

Damiano Anselmi

RENORMALIZATION

Copyright © 2019 D. Anselmi
Sixth Edition, Pisa, February 2019
D. Anselmi, Editor

Contents

Preface	3
1 Functional integral	7
1.1 Path integral	10
1.2 Free field theory	15
1.3 Perturbative expansion	19
1.4 Generating functionals, Schwinger-Dyson equations	29
1.5 Advanced generating functionals	40
1.6 Massive vector fields	44
1.7 Fermions	45
2 Renormalization	57
2.1 Dimensional regularization	60
2.1.1 Limits and other operations in D dimensions	65
2.1.2 Functional integration measure	66
2.1.3 Dimensional regularization for vectors and fermions	67
2.2 Divergences and counterterms	69
2.3 Renormalization to all orders	90
2.4 Locality of counterterms	92
2.5 Power counting	99
2.6 Renormalizable theories	104
2.7 Composite fields	110
2.8 Maximum poles of diagrams	116
2.9 Subtraction prescription	118
2.10 Regularization prescription	119
2.11 Comments about the dimensional regularization	122
2.12 About the series resummation	124

2	<i>CONTENTS</i>
3 Renormalization group	129
3.1 The Callan-Symanzik equation	134
3.2 Finiteness of the beta function and the anomalous dimensions	138
3.3 Fixed points of the RG flow	139
3.4 Scheme (in)dependence	145
3.5 A deeper look into the renormalization group	146
4 Gauge symmetry	153
4.1 Abelian gauge symmetry	153
4.2 Gauge fixing	154
4.3 Non-Abelian global symmetry	160
4.4 Non-Abelian gauge symmetry	165
5 Canonical gauge formalism	169
5.1 General idea behind the canonical gauge formalism	170
5.2 Systematics of the canonical gauge formalism	173
5.3 Canonical transformations	183
5.4 Gauge fixing	185
5.5 Generating functionals	189
5.6 Ward identities	194
6 Quantum electrodynamics	197
6.1 Ward identities	201
6.2 Renormalizability of QED to all orders	205
7 Non-Abelian gauge field theories	211
7.1 Renormalizability of non-Abelian gauge theories to all orders	211
A Notation and useful formulas	219

Preface

Most modern high-energy physics, including the standard model of particle physics, is formulated by means of perturbative quantum field theory.

When the perturbative expansion is performed in a naïve way, it generates “divergences”, that is to say quantities that appear to be “infinite”, instead of being small. Typically, they are due to diverging improper integrals. The presence of divergences suggests that it should be possible to define the perturbative expansion in a smarter way.

With the help of a cutoff, divergences become temporarily finite. Then they can be classified and moved around. Clearly, if a divergence disappears by changing the parametrization of the theory, it is not a true divergence, but just a blunder due to an unfortunate choice of variables. If there exists a reparametrization that makes all the divergences disappear, then the theory is actually convergent.

Divergences can be relocated by performing all sorts of operations that in normal circumstances leave the physics unchanged, such as changes of field variables, as well as redefinitions of the parameters, in particular the coupling constants. Renormalization is the reparametrization that moves the divergences “to the right places”, assuming that such places do exist. In simple theories, fields and couplings just get multiplied by constants, whence the name *re-normalization*. In more complicated situations the redefinitions can even be nonpolynomial. Once the theory is renormalized, the cutoff can be safely removed, and the physical quantities become meaningful.

The reparametrization solves the problem of divergences, and allows us to define the correct perturbative expansion. Under certain, rather general, assumptions it is *always* possible to absorb the divergences into reparametrizations. However, the price can be considerably high: the introduction of infinitely many new independent parameters. If the divergences can be can-

celled by keeping the number of independent parameters finite the theory is called renormalizable. The renormalizable theories acquire a very special status among all theories.

Renormalizability provides a way to select the theories. This selection is actually welcome, since it gives us a reason to discard a huge set of theories that otherwise would have to be *a priori* included. Among other things, no physical theory in more than four spacetime dimensions survives the selection, which makes renormalization a good candidate to explain why we live in four dimensions. The set of renormalizable theories contains the standard model in flat space. Therefore, it allows us to explain three interactions of nature out of four. Unfortunately, there is no known way to formulate quantum gravity so as to include it in the set of renormalizable theories.

Inserting a parameter (the cutoff) to remove it later is a mathematical trick like many. In some sense, it is just a “technicality”, and most of renormalization appears to be a rather technical issue. However, technicalities like this may have extremely important and unforeseen consequences, and considerably affect the physical predictions of the theory. Examples are given by the renormalization-group flow and the anomalies: scale invariant theories can become scale dependent, coupling “constants” can become energy dependent, strong interactions can become weak, eternal particles can decay. The reason why the reparametrizations used to eliminate the divergences do not leave the physics completely unchanged is precisely that they are divergent.

Ironically, the “divergences” are the best known quantities of quantum field theory, to the extent that certain physical amplitudes can be calculated exactly to all orders, because of the intimate relation they have with divergences. At present, perturbative quantum field theory is the most successful theoretical achievement of elementary particle physics. Some of its aspects are so deep that most physicists need years and years to capture their true meanings. In some sense, the conceptual gap between quantum field theory and quantum mechanics can be compared to the one between quantum mechanics and classical mechanics. Several physicists have been puzzled by the indeterminacy principle, and have never accepted that it could be part of the ultimate description of nature. Nowadays, some physicists still view divergences as “pathologies” and think that “renormalization is a way to hide what *we* do not understand under the carpet”. More probably, *they* do not understand what they are talking about.

Removing divergences is just a more sophisticated way to define improper integrals. Following Riemann, we can insert a cutoff, calculate an integral for finite values of the cutoff, and remove the cutoff at the end. If the procedure works, the integral is called convergent. If the procedure does not work, the integral is called divergent. Different prescriptions may lead to different results. For example, it is well known that the Riemann and the Lebesgue approaches are not equivalent.

Quantum field theory requires just one step more. There, we do not have one integral at a time, but a whole *theory*, which is a huge collection of integrals, related to one another. We insert a cutoff, and make calculations for finite values of the cutoff. However, before removing it, we have the freedom to make a variety of operations that *normally* do not change the physics. If the procedure is successful, the theory is actually convergent, otherwise it is divergent. In the end, we discover that the operations we make affect some physical predictions in crucial ways with respect to what we naïvely expected at the beginning. However, there is no source of embarrassment in that, because the Riemann and the Lebesgue approaches may also give results that differ from what one naïvely expects.

Ultimately, renormalization is one of the concepts we understand better, at present, in high-energy theoretical physics. We bet that all the future developments of high-energy physics will emerge more or less directly from it. At the same time, there is no doubt that quantum field theory is still formulated in a rather primitive way. A complete reformulation is desirable. One purpose of this book is to collect the present knowledge about renormalization and stimulate people to start from that point and make an effort to upgrade the formulation of quantum field theory as much as it takes to achieve substantial progress and trigger a renaissance of the topic.

We are aware that in the past decades several approaches alternative to quantum field theory have been proposed, but we remain skeptical about their claimed virtues. Although they are often presented as “beyond quantum field theory”, we do not see any justification to the artificial enthusiasm that has surrounded them for too long. For example, there is little doubt that, conceptually speaking, string theory is a huge step backwards with respect to quantum field theory. We can only wish good luck to those who still do not see that all the alternatives to quantum field theory are doomed to sink into anonymity.

The book contains the basic notions of renormalization. The main goals are to construct perturbative quantum field theory, study the consequences of renormalization, and show that the perturbative formulation of a wide class of quantum field theories, which includes the standard model coupled to quantum gravity, is consistent to all orders. Most issues are treated using modern techniques, privileging the most economical and powerful tools. On the contrary, not much effort is devoted to explain how such a successful theoretical framework has emerged historically. Some aspects of quantum field theory are very involved, and those who study the matter for the first time can greatly benefit from the rational, non historical approach of this book.

Although self-consistent, this book is not meant to replace the existing books on quantum field theory. Since its main focus is renormalization, several basic notions of quantum field theory are just taken for granted. Quantum field theory is formulated using the functional integral and the dimensional regularization technique. Algebraic aspects are covered to the extent that is necessary to treat renormalization. In particular, issues such as the topological properties of anomalies, the geometric aspects of gauge fields, and so on, are not discussed. A number of exercises, with solutions, are distributed along the book to help the layman familiarize with the most important tools of renormalization.

Chapter 1

Functional integral

The *functional integral* is an integral over a space of functions. It is one of the basic tools that can be used to formulate the perturbative expansion of quantum field theory. It also provides an alternative formulation of quantum mechanics, which is equivalent to the Schrödinger and the Heisenberg ones.

The functional integral is defined as a limit of an ordinary multiple integral, when the number of integrated variables tends to infinity. Imagine that spacetime is discretized, with elementary cubic cells of size a , and put into a box of finite size $L = Na$. The discretized pattern is called “lattice” and the distance a between two vertices of the lattice is called “lattice space”. For the moment, we work at finite values of a and N , but at a second stage we take the limits $a \rightarrow 0$ and $N \rightarrow \infty$. At finite a and N , the set of spacetime points x_i is finite and the discretized version of a function $f(x)$ is a finite set of values $f_i = f(x_i)$, with $i = 1, 2, \dots, N$. The f_i are the variables over which we integrate.

Consider the finite-dimensional ordinary integral

$$c(a, N) \int \prod_{i=1}^N df_i \hat{G}(f_i), \quad (1.1)$$

where $c(a, N)$ is a normalization factor, which can depend on a and N , and $\hat{G}(f_i)$ is the discretized version of a generic functional $G(f)$. When a tends to zero and N tends to infinity, the number of integrated variables tends to infinity. Assume that there exists a normalization factor $c(a, N)$ such that the limits $a \rightarrow 0$, $N \rightarrow \infty$ exist. Then, the functional integral over the space

of functions $f(x)$ is defined as

$$\int [df] G(f) = \lim_{\substack{a \rightarrow 0 \\ N \rightarrow \infty}} c(a, N) \int \prod_{i=1}^N df_i \hat{G}(f_i).$$

The simplest integrals we need are Gaussian. The basic Gaussian multiple integral reads

$$\int_{-\infty}^{+\infty} \prod_{i=1}^N dx_i \exp \left(-\frac{1}{2} \sum_{i,j=1}^N x_i M_{ij} x_j \right) = \frac{(2\pi)^{N/2}}{\sqrt{\det M}}, \quad (1.2)$$

where M is a positive-definite symmetric matrix. Formula (1.2) can be proved by diagonalizing M with an orthogonal matrix \mathcal{N} . Write $M = \mathcal{N} D \mathcal{N}^t$, where $D = \text{diag}(m_1, \dots, m_n)$ and m_i are the eigenvalues of M . Perform the change of variables $x = \mathcal{N}y$ and recall that the integration measure is invariant, since $\det \mathcal{N} = 1$. Then, the integral becomes the product of the one-dimensional Gaussian integrals

$$\int_{-\infty}^{+\infty} dy_i \exp \left(-\frac{1}{2} m_i y_i^2 \right) = \sqrt{\frac{2\pi}{m_i}},$$

whence (1.2) follows. We also have the formula

$$\begin{aligned} Z(a) &= \int_{-\infty}^{+\infty} \prod_{i=1}^N dx_i \exp \left(-\frac{1}{2} \sum_{i,j=1}^N x_i M_{ij} x_j + \sum_{i=1}^N x_i a_i \right) \\ &= \frac{(2\pi)^{N/2}}{\sqrt{\det M}} \exp \left(\frac{1}{2} \sum_{i,j=1}^N a_i M_{ij}^{-1} a_j \right), \end{aligned} \quad (1.3)$$

which can be easily proved from (1.2) by means of the translation $x = y + M^{-1}a$.

We can define correlation functions

$$\begin{aligned} \langle x_{i_1} \cdots x_{i_n} \rangle &= \frac{1}{Z(0)} \int_{-\infty}^{+\infty} \prod_{i=1}^N dx_i x_{i_1} \cdots x_{i_n} \exp \left(-\frac{1}{2} \sum_{i,j=1}^N x_i M_{ij} x_j \right) \\ &= \frac{1}{Z(a)} \frac{\partial^n Z(a)}{\partial a_{i_1} \cdots \partial a_{i_n}} \Big|_{a=0}. \end{aligned} \quad (1.4)$$

For example, we find

$$\begin{aligned}\langle x_j x_k \rangle &= \frac{1}{Z(a)} \frac{\partial^2 Z(a)}{\partial a_j \partial a_k} \Big|_{a=0} = M_{jk}^{-1}, \\ \langle x_j x_k x_m x_n \rangle &= M_{jk}^{-1} M_{mn}^{-1} + M_{jm}^{-1} M_{kn}^{-1} + M_{jn}^{-1} M_{km}^{-1}.\end{aligned}\quad (1.5)$$

Every correlation function that contains an odd number of insertion vanishes: $\langle x_{i_1} \cdots x_{i_{2n+1}} \rangle = 0 \ \forall n$. Instead, the correlation functions that contain even numbers of insertions are determined by a simple recursion relation, which reads

$$\langle x_{i_1} \cdots x_{i_{2n}} \rangle = \sum_{k=2}^{2n} M_{i_1 i_k}^{-1} \langle x_{i_2} \cdots \widehat{x_{i_k}} \cdots x_{i_{2n}} \rangle. \quad (1.6)$$

where the hat denotes a missing insertion. This formula is proved by noting that

$$\begin{aligned}\langle x_{i_1} \cdots x_{i_{2n}} \rangle &= \frac{1}{n!} \frac{\partial^{2n}}{\partial a_{i_1} \cdots \partial a_{i_{2n}}} \left(\frac{1}{2} a^t M^{-1} a \right)^n \\ &= \frac{1}{2^{n-1} (n-1)!} \frac{\partial^{2n-1}}{\partial a_{i_2} \cdots \partial a_{i_{2n}}} \left[(M^{-1} a)_{i_1} (a^t M^{-1} a)^{n-1} \right] \\ &= \frac{1}{2^{n-1} (n-1)!} \sum_{k=2}^{2n} M_{i_1 i_k}^{-1} \frac{\partial^{2(n-1)}}{\partial a_{i_2} \cdots \widehat{\partial a_{i_k}} \cdots \partial a_{i_{2n}}} (a^t M^{-1} a)^{n-1} \\ &= \sum_{k=2}^{2n} M_{i_1 i_k}^{-1} \langle x_{i_2} \cdots \widehat{x_{i_k}} \cdots x_{i_{2n}} \rangle.\end{aligned}$$

In the third line the hat on ∂a_{i_k} denotes a missing derivative. The recurrence relation (1.6) gives

$$\langle x_{i_1} \cdots x_{i_{2n}} \rangle = \sum_P M_{P(i_1)P(i_2)}^{-1} \cdots M_{P(i_{2n-1})P(i_{2n})}^{-1}, \quad (1.7)$$

where the sum is over the inequivalent permutations P of $\{i_1, \dots, i_{2n}\}$. By this we mean that identical contributions are counted only once.

Our first goal is to define the $N \rightarrow \infty$ limits of the multiple integrals just met, and others of similar types, and use them to formulate quantum mechanics and perturbative quantum field theory. We begin with quantum mechanics.

1.1 Path integral

Consider a non relativistic particle of mass m , potential $V(q)$ and Lagrangian

$$\mathcal{L}(q, \dot{q}) = \frac{m}{2} \dot{q}^2 - V(q). \quad (1.8)$$

Suppose that the particle is observed in the locations q_{in} at time t_{in} and q_{f} at time t_{f} and that it is not observed in the time interval $t_{\text{in}} < t < t_{\text{f}}$. Quantum mechanics teaches us that it is meaningless to tell “where” the particle is while it is not observed, or even assume that it is somewhere. More or less equivalently, we can imagine that it is anywhere, or everywhere. In particular, it does not make sense to say that the particle moves from q_{in} to q_{f} along a particular trajectory $q(t)$, such as the classical trajectory that extremizes the action

$$S(q_{\text{f}}, t_{\text{f}}; q_{\text{in}}, t_{\text{in}}) = \int_{t_{\text{in}}}^{t_{\text{f}}} dt \mathcal{L}(q(t), \dot{q}(t)). \quad (1.9)$$

A possible way out is to imagine that it moves from q_{in} to q_{f} along all possible *paths*

$$q(t), \quad t_{\text{in}} \leq t \leq t_{\text{f}}, \quad q(t_{\text{in}}) = q_{\text{in}}, \quad q(t_{\text{f}}) = q_{\text{f}},$$

at the same time. Then, each path must contribute to physical quantities, with a suitable (complex) weight. Clearly, if this idea is right we have to “integrate” over the paths.

In some sense, we replace the principle of minimum action with a new principle, which is able to account for the quantum effects. The semiclassical approximation suggests that each path should be weighted by the factor

$$\exp \left(\frac{i}{\hbar} S(q_{\text{f}}, t_{\text{f}}; q_{\text{in}}, t_{\text{in}}) \right). \quad (1.10)$$

Indeed, in the limit $\hbar \rightarrow 0$ the strongly oscillating exponent singles out the trajectory of minimum action as the only one that survives.

These considerations, although inspiring, are still vague. We do not know how to define the integral over the paths. As mentioned before, one possibility is to discretize the problem and define the path integral as the limit of an ordinary multiple integral, when the number of integrated variables tends to infinity. Thus, let us discretize the time interval $t_{\text{in}} \leq t \leq t_{\text{f}}$ by dividing it in N subintervals

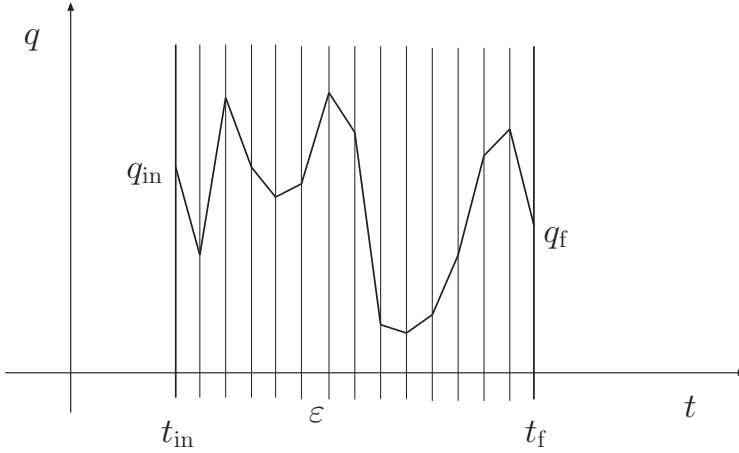
$$t_{i-1} \leq t \leq t_i, \quad t_i = t_{i-1} + \varepsilon, \quad \varepsilon = \frac{t_{\text{f}} - t_{\text{in}}}{N},$$

$i = 1, \dots, N$, with $t_0 = t_{\text{in}}$ and $t_N = t_{\text{f}}$. The path $q(t)$ is then replaced by the set of positions $q_i = q(t_i)$ at times t_i .

The trajectory of the i -th subinterval can be taken to be the one that extremizes the action. However, in many cases simpler subtrajectories are equally good approximations. For example, we can take the straight lines

$$\bar{q}(t) = \frac{q_i - q_{i-1}}{\varepsilon}(t - t_{i-1}) + q_{i-1}. \quad (1.11)$$

This choice produces a picture like



In the limit $\varepsilon \rightarrow 0$ the approximate path can tend to any function $q(t)$, including the ones that are not differentiable and not continuous. From the physical point of view there is no reason why the unobservable trajectory $q(t)$ should be continuous and/or differentiable, so the path integral should sum over *all* functions $q(t)$.

In the i -th subinterval we have the constant velocity

$$\frac{q_i - q_{i-1}}{\varepsilon},$$

so the action (1.9) can be approximated by

$$\sum_{i=1}^N \bar{S}(q_i, t_i; q_{i-1}, t_{i-1}) = \sum_{i=1}^N \left\{ \frac{m(q_i - q_{i-1})^2}{2\varepsilon} - \varepsilon V(q_i) \right\} + \mathcal{O}(\varepsilon^{3/2}), \quad (1.12)$$

where the bar over S is there to remember that we have chosen the special subtrajectories (1.11). Below we prove that $|q_i - q_{i-1}| \sim \mathcal{O}(\varepsilon^{1/2})$ and that

the corrections $\mathcal{O}(\varepsilon^{3/2})$ appearing in formula (1.12) can be neglected in the limit $\varepsilon \rightarrow 0$.

Inspired by (1.10), we weigh each infinitesimal portion of the trajectory by the factor

$$\frac{1}{A} \exp \left(\frac{i}{\hbar} \bar{S}(q_i, t_i; q_{i-1}, t_{i-1}) \right),$$

where A is some normalization constant, to be determined. This means that during a time subinterval the wave function $\psi(q, t)$ evolves into

$$\psi(q, t + \varepsilon) = \frac{1}{A} \int_{-\infty}^{+\infty} dq' \exp \left(\frac{i}{\hbar} \bar{S}(q, t + \varepsilon; q', t) \right) \psi(q', t). \quad (1.13)$$

Consequently, during the finite interval $t_{\text{in}} \leq t \leq t_{\text{f}}$ the evolution of the wave function $\psi(q, t)$ is given by the formula

$$\psi(q, t) = \int_{-\infty}^{+\infty} dq' K(q, t; q', t') \psi(q', t'), \quad (1.14)$$

where $K(q, t; q', t')$, called *kernel* of the time evolution, has the path-integral expression

$$\begin{aligned} K(q, t; q', t') &= \lim_{N \rightarrow \infty} A^{-N} \int \prod_{i=1}^{N-1} dq_i e^{\frac{i}{\hbar} \sum_{i=1}^N \bar{S}(q_i, t_i; q_{i-1}, t_{i-1})} \\ &\equiv \int [dq] \exp \left(\frac{i}{\hbar} S(q) \right), \end{aligned} \quad (1.15)$$

and now $t_0 = t'$, $t_N = t$, $q_0 = q'$, $q_N = q$, $\varepsilon = (t - t')/N$. The last line is the common short-hand notation used to denote the functional integral.

Observe that, in particular, we must have

$$K(q, t; q', t) = \delta(q - q'). \quad (1.16)$$

Schrödinger equation

Now we prove that the time evolution encoded in the path-integral formulas (1.13) and (1.15) is equivalent to the one predicted by quantum mechanics. In particular, we show that the wave function (1.14), with the kernel defined by (1.15), satisfies the Schrödinger equation. We discretize time as explained above, and compare $\psi(q, t + \varepsilon)$ and $\psi(q', t)$ by means of (1.13). We have

$$\psi(q, t + \varepsilon) = \frac{1}{A} \int_{-\infty}^{+\infty} d\Delta e^{\frac{im\Delta^2}{2\hbar\varepsilon} - \frac{i\varepsilon}{\hbar} V(q) + \mathcal{O}(\varepsilon^{3/2})} \psi(q - \Delta, t), \quad (1.17)$$

after a translation $q' = q - \Delta$. Recall that

$$\lim_{\varepsilon \rightarrow 0} \sqrt{\frac{m}{2\pi i \hbar \varepsilon}} e^{\frac{im\Delta^2}{2\hbar\varepsilon}} = \delta(\Delta).$$

This formula can be proved assuming that the mass has a small positive imaginary part. Thus, the two sides of (1.17) match in the limit $\varepsilon \rightarrow 0$ if we take

$$\frac{1}{A} = \sqrt{\frac{m}{2\pi i \hbar \varepsilon}}.$$

Observe that this choice also ensures that (1.16) holds. We are still assuming that $|\Delta| \sim \varepsilon^{1/2}$, which allows us to neglect the Δ dependence contained inside $\mathcal{O}(\varepsilon^{3/2})$. This assumption is justified by the calculations that follow.

Expanding the integrand of (1.17) in powers of Δ , we obtain

$$\begin{aligned} \sqrt{\frac{m}{2\pi i \hbar \varepsilon}} \int_{-\infty}^{+\infty} d\Delta e^{\frac{im\Delta^2}{2\hbar\varepsilon}} \left(1 - \Delta \frac{\partial}{\partial q} + \frac{\Delta^2}{2} \frac{\partial^2}{\partial q^2} + \mathcal{O}(\Delta^3) \right. \\ \left. - \frac{i\varepsilon}{\hbar} V(q) + \mathcal{O}(\varepsilon^{3/2}) \right) \psi(q, t). \end{aligned} \quad (1.18)$$

Defining the integrals

$$I_n = \int_{-\infty}^{+\infty} d\Delta \Delta^n e^{\frac{im\Delta^2}{2\hbar\varepsilon}},$$

we find $I_{2k+1} = 0$ and

$$I_{2k} = -2i\varepsilon\hbar \frac{\partial I_{2k-2}}{\partial m}, \quad I_0 = \sqrt{\frac{2\pi i \hbar \varepsilon}{m}},$$

which gives

$$I_{2k} = \Gamma\left(k + \frac{1}{2}\right) \left(\frac{2i\hbar\varepsilon}{m}\right)^{k+\frac{1}{2}}.$$

We find $I_{2k}/I_0 \sim \varepsilon^k$, which also proves $|\Delta| \sim \varepsilon^{1/2}$, as claimed before. Finally, rearranging (1.17), dividing by ε and taking the limit $\varepsilon \rightarrow 0$, we find the Schrödinger equation

$$i\hbar \frac{\partial \psi}{\partial t} = -\frac{\hbar^2}{2m} \frac{\partial^2 \psi}{\partial q^2} + V\psi.$$

The outcome is independent of the approximation we have used to expand $\bar{S}(q_i, t_i; q_{i-1}, t_{i-1})$. For example, we could have written $V(q_{i-1})$ in (1.12),

instead of $V(q_i)$, or $(V(q_i) + V(q_{i-1}))/2$. The difference is always made of terms that are $\mathcal{O}(\varepsilon^{3/2})$ in the integrand of (1.18), which are negligible in the limit $\varepsilon \rightarrow 0$.

Thus, we have proved that the path integral provides a formulation of quantum mechanics that is equivalent to the Schrödinger and the Heisenberg ones.

Free particle

We explicitly calculate the kernel in the case of the free particle. There,

$$\bar{S}(q_i, t_i; q_{i-1}, t_{i-1}) = \frac{m(q_i - q_{i-1})^2}{2\varepsilon},$$

so we have

$$K_{\text{free}}(q, t; q', t') = \lim_{N \rightarrow \infty} \left(\frac{m}{2\pi i \hbar \varepsilon} \right)^{N/2} \int \prod_{i=1}^{N-1} dq_i e^{\frac{im}{2\hbar\varepsilon} \sum_{i=1}^N (q_i - q_{i-1})^2}.$$

Changing variables to $\tilde{q}_i = q_i - q$, we can rewrite the integral as

$$e^{\frac{im}{2\hbar\varepsilon} (q - q')^2} \int \prod_{i=1}^{N-1} d\tilde{q}_i e^{\frac{im}{2\hbar\varepsilon} (\tilde{q}^t \tilde{M} \tilde{q} + 2\tilde{q}_1 (q - q'))},$$

where

$$\tilde{M} = \begin{pmatrix} 2 & -1 & 0 & 0 & \cdots \\ -1 & 2 & -1 & \cdots & 0 \\ 0 & -1 & \cdots & -1 & 0 \\ 0 & \cdots & -1 & 2 & -1 \\ \cdots & 0 & 0 & -1 & 2 \end{pmatrix} \quad (1.19)$$

is an $(N - 1) \times (N - 1)$ matrix. Now the integral is of the Gaussian form (1.3) with

$$M = -\frac{im}{\hbar\varepsilon} \tilde{M}, \quad a = \frac{im(q - q')}{\hbar\varepsilon} (0, \dots, 0, 1)$$

and $N \rightarrow N - 1$. Again, we assume that the mass has a small positive imaginary part. We have

$$\det \tilde{M} = N, \quad (\tilde{M}^{-1})_{N-1, N-1} = \frac{N-1}{N}. \quad (1.20)$$

The first formula can be proved recursively. Indeed, denoting the $I \times I$ matrix \tilde{M} of (1.19) with \tilde{M}_I , we have

$$\det \tilde{M}_{N-1} = 2 \det \tilde{M}_{N-2} - \det \tilde{M}_{N-3}, \quad \det \tilde{M}_1 = 2 \quad \det \tilde{M}_2 = 3.$$

The second formula of (1.20) gives the last entry of the inverse matrix \tilde{M}_{N-1}^{-1} , and is just the determinant of the associated minor (which coincides with \tilde{M}_{N-2}), divided by the determinant of \tilde{M}_{N-1} .

Finally, using formula (1.3), with the appropriate substitutions, and recalling that $\varepsilon N = t - t'$, we find

$$K_{\text{free}}(q, t; q', t') = \sqrt{\frac{m}{2\pi i \hbar (t - t')}} e^{\frac{im(q-q')^2}{2\hbar(t-t')}},$$

which is the known result.

1.2 Free field theory

Given a classical field theory, described by the action $S(\varphi)$, we want to define the functional integral

$$\int [d\varphi] \exp\left(\frac{i}{\hbar} S(\varphi)\right). \quad (1.21)$$

At present, we can do this only perturbatively, by expanding around the free-field limit.

From now on, we work in Euclidean space, where some complications are avoided. For simplicity, we also set $\hbar = 1$.

Free field theories are described by Gaussian functional integrals. We start from the scalar field in four dimensions. Its action in Euclidean space is

$$S(\varphi) = \frac{1}{2} \int d^4x \left((\partial_\mu \varphi)^2 + m^2 \varphi^2 \right). \quad (1.22)$$

We want to define the generating functional

$$Z(J) \equiv e^{W(J)} = \int [d\varphi] \exp\left(-S(\varphi) + \int \varphi J\right) \quad (1.23)$$

where J are external sources, $\int \varphi J \equiv \int d^4x \varphi(x) J(x)$ and W is the logarithm of Z . First, we discretize the Euclidean space. Each coordinate x^μ is replaced

by an index i_μ , and the field $\varphi(x)$ becomes $\varphi_{\{i_\mu\}}$. The discretized form of the action reads

$$S_{\text{discr}}(\varphi_{\{i_\mu\}}) = \frac{1}{2} \sum_{\{i_\mu\}, \{j_\nu\}} \varphi_{\{i_\mu\}} M_{\{i_\mu\}\{j_\nu\}} \varphi_{\{j_\nu\}}, \quad (1.24)$$

for some matrix $M_{\{i_\mu\}\{j_\nu\}}$ that we do not need to specify here. It suffices to recall that M is the discretized version of the kinetic operator $-\square + m^2$:

$$M_{\{i_\mu\}\{j_\nu\}} = (-\square + m^2)|_{\text{discr}}. \quad (1.25)$$

The discretized version of the functional integral reads

$$Z(J_{\{i_\mu\}}) = \int \prod_{\{i_\mu\}} [d\varphi_{\{i_\mu\}}] \exp \left(-S_{\text{discr}}(\varphi_{\{i_\mu\}}) + \sum_{\{i_\mu\}} \varphi_{\{i_\mu\}} J_{\{i_\mu\}} \right).$$

Using formula (1.3), we find

$$W(J_{\{i_\mu\}}) = \ln Z(J_{\{i_\mu\}}) = \frac{1}{2} \sum_{\{i_\mu\}, \{j_\nu\}} J_{\{i_\mu\}} M_{\{i_\mu\}\{j_\nu\}}^{-1} J_{\{j_\nu\}} - \frac{1}{2} \ln \det M + C, \quad (1.26)$$

where C is a constant that collects also the normalization factor $c(a, L)$ of formula (1.1).

To define the continuum limit, we basically need to define the inverse of M and its determinant. However, note that the determinant, as well as the constant C , appear only in Z and W , but not in the correlation functions (1.4). Therefore, we actually just need to define M^{-1} . This is not difficult, because the inverse of $-\square + m^2$ is by definition the Green function $G_B(x, y)$, that is to say the solution of the equation

$$(-\square_x + m^2)G_B(x, y) = \delta^{(4)}(x - y). \quad (1.27)$$

Normalizing the functional integral conveniently and using (1.3) and (1.26), we can write

$$Z(J) = e^{W(J)}, \quad W(J) = \frac{1}{2} \int d^4x J(x) G_B(x, y) J(y) d^4y. \quad (1.28)$$

We can define the J -dependent correlation functions

$$\begin{aligned} \langle \varphi(x_1) \cdots \varphi(x_n) \rangle_J &= \frac{\int [d\varphi] \varphi(x_1) \cdots \varphi(x_n) \exp(-S(\varphi) + \int \varphi J)}{\int [d\varphi] \exp(-S(\varphi) + \int \varphi J)} \\ &= \frac{1}{Z(J)} \frac{\delta^n Z(J)}{\delta J(x_1) \cdots \delta J(x_n)}, \end{aligned} \quad (1.29)$$

where the subscript J means that sources are non-vanishing. In particular, formulas (1.5) give

$$\langle \varphi(x)\varphi(y) \rangle = \frac{1}{Z(J)} \frac{\partial^2 Z(J)}{\partial J(x) \partial J(y)} \Big|_{J=0} = G_B(x, y), \quad (1.30)$$

$$\begin{aligned} \langle \varphi(x)\varphi(y)\varphi(z)\varphi(w) \rangle &= G_B(x, y)G_B(z, w) + G_B(x, z)G_B(y, w) \\ &\quad + G_B(x, w)G_B(y, z), \end{aligned} \quad (1.31)$$

and so on. We see that, in practice, the free theory contains just one piece of information, which is the Green function.

Formulas (1.6) and (1.7) can be generalized following the same steps. We have

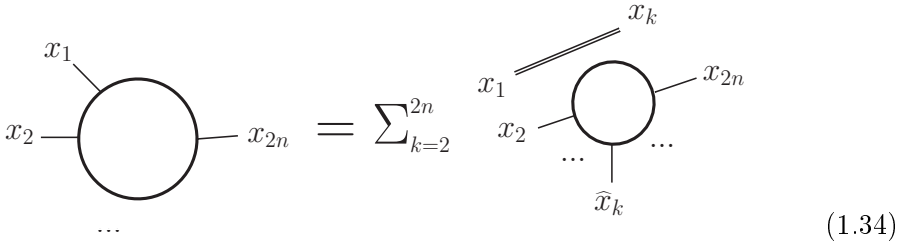
$$\langle \varphi(x_1) \cdots \varphi(x_{2n}) \rangle = \sum_{k=2}^{2n} G_B(x_1, x_k) \langle \varphi(x_2) \cdots \widehat{\varphi(x_k)} \cdots \varphi(x_{2n}) \rangle \quad (1.32)$$

$$= \sum_P G_B(x_{P(1)}, x_{P(2)}) \cdots G_B(x_{P(2n-1)}, x_{P(2n)}), \quad (1.33)$$

while the correlation functions that contain an odd number of insertions vanish.

Equation (1.33) is known as *Wick's theorem*. It says that *i*) the external points $x_1 \cdots x_{2n}$ must be connected pairwise in all inequivalent ways, *ii*) each connection is a Green function and *iii*) each inequivalent set of connections is multiplied by the coefficient 1.

It is natural express Wick's theorem graphically. A Green function is drawn as a double line connecting a pair of points. Then formula (1.32) reads



$$\text{Diagram with legs } x_1, x_2, \dots, x_{2n} = \sum_{k=2}^{2n} \text{Diagram with legs } x_1, x_2, \dots, x_{2n} \text{ and a double line connecting } x_1 \text{ and } x_k \quad (1.34)$$

where the legs attached to the discs denote the insertions of the correlation functions.

The Euclidean Green functions can be computed by switching to momentum space. We define the Fourier transform as

$$\varphi(x) = \int \frac{d^4 p}{(2\pi)^4} e^{ip \cdot x} \tilde{\varphi}(p). \quad (1.35)$$

Then we find

$$G_B(x, y) = \langle \varphi(x) \varphi(y) \rangle = \int \frac{d^4 p}{(2\pi)^4} \frac{e^{ip \cdot (x-y)}}{p^2 + m^2} = \frac{m}{4\pi^2 |x - y|} K_1(m|x - y|), \quad (1.36)$$

where K_1 denotes the modified Bessel function of the second kind.

This result is proved as follows. We must assume that x and y do not coincide, and use a trick to dump the oscillating behavior at infinity. Let us start from the massless limit. To calculate the integral at $m = 0$, we multiply the integrand by $e^{-\delta|p|}$, where $\delta > 0$, and take the limit $\delta \rightarrow 0$ at the end. Switching to spherical coordinates, we first integrate over $|p|$ and later over the angles. The basic steps are

$$\begin{aligned} G_B(x, 0)|_{m=0} &= \lim_{\delta \rightarrow 0^+} \int \frac{d^4 p}{(2\pi)^4} \frac{e^{ip \cdot x - \delta|p|}}{p^2} = \lim_{\delta \rightarrow 0^+} \frac{1}{4\pi^3} \int_0^\pi \frac{d\theta \sin^2 \theta}{(\delta - i|x| \cos \theta)^2} \\ &= \lim_{\delta \rightarrow 0^+} \frac{1}{4\pi^2 x^2} \left(1 - \frac{\delta}{\sqrt{\delta^2 + x^2}} \right) = \frac{1}{4\pi^2 x^2}. \end{aligned} \quad (1.37)$$

To calculate the integral at $m \neq 0$, we make it convergent in a different way, at $x \neq 0$, that is to say by differentiating with respect to m . Then, after switching to spherical coordinates, we first integrate over the angles, and later over $|p|$. We find

$$\begin{aligned} \frac{\partial}{\partial m} G_B(x, 0) &= - \int \frac{d^4 p}{(2\pi)^4} \frac{2m e^{ip \cdot x}}{(p^2 + m^2)^2} \\ &= - \frac{m}{2\pi^2 |x|} \int_0^\infty \frac{p^2 J_1(p|x|) dp}{(p^2 + m^2)^2} = - \frac{m}{4\pi^2} K_0(m|x|), \end{aligned}$$

where $J_1(x)$ is the Bessel function of the first kind. Integrating over m and requiring (1.37) at $m = 0$, we obtain (1.36).

The correlations functions can be mathematically interpreted as distributions. Then, the sources J should be viewed as test functions. Indeed, the Green function $G_B(x, y)$, which appears to be singular at $x = y$, is actually

regular as a distribution. To see this, it is sufficient to observe that when $G_B(x, y)$ acts on a test function $J(y)$, the behavior of the integral

$$\int_{|x-y|\sim 0} d^4y G_B(x, y) J(y) \sim \frac{1}{4\pi^2} \int d^4y \frac{J(y)}{(x-y)^2},$$

around $x \sim y$ shows no singularity. We have used the result (1.37), since the behavior of (1.36) at $x \sim y$ coincides with the behavior at $m \rightarrow \infty$.

1.3 Perturbative expansion

Interacting theories are defined by expanding perturbatively around their free field limits. Although this sounds like a straightforward process, the perturbative expansion of quantum field theory actually underlies a huge conceptual advancement with respect to the notions we are accustomed to. To clarify this point, it is worth to pay attention to what we do when we normally approximate. We have, say, difficult differential equations, which we want to solve. We know some classes of exact solutions, which typically do not cover the cases of physical interest. We realize that some physical situations are only slightly different from those described by the exact solutions, so we work out other solutions by expanding perturbatively around the exact ones. What is important for our present discussion, is that we are talking about a well defined problem, described by difficult, but well defined, equations. *Then*, we approximate. We approximate something that does exist, something that exists *before* the approximation.

In quantum field theory, instead, we must really start from nothing, apart from the free field limit. There are no equations, and no theory, before we make approximations. Thus, when we say that we perturbatively expand around the free field theory, we are actually lying: we are not expanding at all. The truth is that we are perturbatively *building* the interactive theory, piece by piece, out of the free field one. The enterprise we are going to undertake is a creative one, not just a deductive process. Therefore, if something goes wrong along the way, it will be no real surprise. To solve the problems that emerge, we have to be more and more creative. In particular, we have to build the mathematics that we need by ourselves. Moreover, every time we find a difficulty, and guess a possible solution, we must start over, implement the proposed solution from the very beginning, and rederive everything up to the

point where we found the problem, check that the problem does disappear as expected and ensure that no collateral difficulties emerge.

Another crucial point is that the perturbative expansion should be considered as a formal power series. In mathematics, a formal power series is a power series that is just viewed as a list of addends, disregarding completely whether the sum converges or not. Perturbative quantum field theory investigates the consistency of the perturbative expansion as a formal power series. It studies the properties of the addends (e.g. their consistency with gauge invariance, unitarity, causality, etc.) and the relations among them. Proving that, for example, the Standard Model is consistent to all orders, as a perturbative quantum field theory, which is one of the main objectives of this book, is already a nontrivial task. The great advantage of working with formal power series is that it allows us to freely exchange the sum with derivative operations, as well as integral operations. Only at the very end we will inquire whether the sum converges or not. Indeed, it is meaningless to demand that a power series be convergent before having shown that it obeys all the desirable physical and mathematical requirements as a formal powers series. The renormalization group and the anomalies of quantum field theory provide well-known cases where the power series is in the end convergent.

Having to build some of the mathematics anew is not surprising either. If we take for granted that the mathematics we already have is good enough to formulate the physical laws of so far unexplored research domains, we may be making a too restrictive assumption. More reasonably, our mathematics is a product of our interaction with the environment in which we are placed as human beings. When we explore energy domains that are very far from those we are accustomed to, the mathematics we have previously developed may be unsatisfactory. In the study of cosmologic and astrophysical phenomena, for example, it has so far proved to be exhaustive. However, in some cases, such as in the study of microscopic phenomena, it has already shown its limitations. In the case of quantum mechanics, we could fill the gap by means of a “correspondence principle”. The idea was that, although there was a huge difference between the classical and the quantum phenomena, at least there was a sort of correspondence between the two. Clearly, we cannot expect to go on forever relying on lucky correspondences, to the extent that quantum field theory forces us to abandon that idea. For example, the “classical” Lagrangian of quantum chromodynamics, which is

the theory that describes what the strong interactions become at high energies, has no correspondence with classical phenomena. Exploring smaller and smaller distances, the problem becomes harder and harder, and we may be forced to give up every correspondence with what we know, and even renounce common sense and intuition, to develop a completely new mathematics by plunging into pure abstraction and technicalism. Quantum field theory, renormalization, with all the problems we find along the way and the partial solutions we work out, give us hints of what the new mathematics will have to be.

That said, the only thing we can do in this moment is pretend there is nothing to worry about, and make a step forward along the process of “creative approximation”.

Consider a theory of interacting scalar fields with action $S(\varphi) = S_0(\varphi) + S_I(\varphi)$, where $S_0(\varphi)$ is (1.22). For concreteness, we can take the φ^4 theory in four dimensions, which has

$$S(\varphi) = \int d^4x \left(\frac{1}{2}(\partial_\mu \varphi)^2 + \frac{m^2}{2}\varphi^2 + \frac{\lambda}{4!}\varphi^4 \right). \quad (1.38)$$

Defining $Z(J)$ as in (1.23) we can write

$$\begin{aligned} Z(J) &= \int [d\varphi] \exp \left(-S(\varphi) + \int J\varphi \right) \\ &= \int [d\varphi] \exp(-S_I(\varphi)) \exp \left(-S_0(\varphi) + \int J\varphi \right) \\ &= \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} \int [d\varphi] S_I^n(\varphi) \exp \left(-S_0(\varphi) + \int J\varphi \right) \\ &= Z_0(J) \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} \langle S_I^n(\varphi) \rangle_{0,J}, \end{aligned} \quad (1.39)$$

where $Z_0(J)$ is given by (1.28). We use the subscript 0 to denotes quantities at $\lambda = 0$. In particular, $\langle \cdots \rangle_{0,J}$ are free-field correlation functions at nonvanishing sources. We have

$$\langle S_I^n(\varphi) \rangle_{0,J} = \left(\frac{\lambda}{4!} \right)^n \int \prod_{i=1}^n d^4x_i \langle \varphi^4(x_1) \cdots \varphi^4(x_n) \rangle_{0,J}.$$

Now, by (1.29) every φ -insertion can be expressed as a functional derivative

with respect to J . Therefore,

$$\langle S_I^n(\varphi) \rangle_{0,J} = \left(\frac{\lambda}{4!} \right)^n \frac{1}{Z_0(J)} \int \prod_{i=1}^n d^4 x_i \frac{\delta^{4n} Z_0(J)}{\delta J^4(x_1) \cdots \delta J^4(x_n)}.$$

Inserting this formula into (1.39), we get

$$\begin{aligned} Z(J) &= \sum_{n=0}^{\infty} \frac{1}{n!} \left(-\frac{\lambda}{4!} \right)^n \prod_{i=1}^n \left(\int d^4 x_i \frac{\delta^4}{\delta J^4(x_i)} \right) Z_0(J) \\ &= \exp \left(-\frac{\lambda}{4!} \int d^4 x \frac{\delta^4}{\delta J^4(x)} \right) Z_0(J). \end{aligned}$$

More generally, we have

$$Z(J) = e^{W(J)} = \exp \left(-S_I \left(\frac{\delta}{\delta J} \right) \right) Z_0(J). \quad (1.40)$$

The scalar field inside S_I is formally replaced by the functional derivative $\delta/\delta J$, which acts on the free-field generating functional $Z_0(J)$.

Formula (1.40) expresses the generating functional of the interacting theory as an infinite sum of terms, each of which involves just functional derivatives of the generating functional of the free theory (which, as we know, contains only the Green function) and integrals over coordinates. Some functional derivatives are taken at the same point, which is called “vertex”. Moreover, the Green functions connect pairs of points, as we see from the Wick theorem (1.34). Formula (1.40) can be efficiently expressed diagrammatically. Diagrams are made of vertices and lines, and are drawn following a simple set of rules, which we now derive.

Feynman rules

The correlation functions can be defined from the expansion of the generating functional $Z(J)$ in powers of J :

$$Z(J) = Z(0) \sum_{n=0}^{\infty} \frac{1}{n!} \int \left(\prod_{i=1}^n d^4 x_i \right) \langle \varphi(x_1) \cdots \varphi(x_n) \rangle J(x_1) \cdots J(x_n).$$

For some practical purposes, it is also useful to define correlation functions that have a different normalization. At $J \neq 0$ we define

$$\langle \varphi(x_1) \cdots \varphi(x_n) \rangle'_J = \frac{1}{Z_0(0)} \frac{\delta^n Z(J)}{\delta J(x_1) \cdots \delta J(x_n)}, \quad (1.41)$$

while at $J = 0$ we write them as $\langle \varphi(x_1) \cdots \varphi(x_n) \rangle'$. In particular, we have

$$\langle \varphi(x_1) \cdots \varphi(x_n) \rangle = \frac{Z_0(0)}{Z(0)} \langle \varphi(x_1) \cdots \varphi(x_n) \rangle' = \frac{\langle \varphi(x_1) \cdots \varphi(x_n) \rangle'}{\langle 1 \rangle'}. \quad (1.42)$$

Since we have normalized $Z_0(0)$ to 1, we could omit this factor. However, the formulas are more explicit if we keep it, which also emphasizes that $Z(0)$ is not equal to one.

Observe that $Z(J)$ can be viewed as the generating functional $Z'(J)$ of the correlation functions (1.41):

$$Z'(J) = Z_0(0) \sum_{n=0}^{\infty} \frac{1}{n!} \int \left(\prod_{i=1}^n d^4 x_i \right) \langle \varphi(x_1) \cdots \varphi(x_n) \rangle' J(x_1) \cdots J(x_n) = Z(J). \quad (1.43)$$

Consider a generic correlation function (1.41) in the φ^4 theory. Writing

$$\begin{aligned} \langle \varphi(x_1) \cdots \varphi(x_n) \rangle' &= \frac{1}{Z_0(0)} \int [d\varphi] \varphi(x_1) \cdots \varphi(x_n) e^{-S_0(\varphi) - \frac{\lambda}{4!} \int d^4 x \varphi^4(x)} \\ &= \sum_{k=0}^{\infty} \frac{(-\lambda)^k}{(4!)^k k!} \int \left(\prod_{j=1}^k d^4 y_j \right) \langle \varphi(x_1) \cdots \varphi(x_n) \prod_{j=1}^k \varphi^4(y_j) \rangle_0, \end{aligned} \quad (1.44)$$

we obtain a sum of contributions that are due to free-field correlation functions with $n + 4k$ insertions. We call the points x_1, \dots, x_n “external” and the points y_1, \dots, y_k “internal”. Each internal point carries four φ insertions. We refer to it as a vertex with four legs.

The free-field correlation functions of (1.44) can be worked out by means of Wick’s theorem. Let us consider the graphical version (1.34) of that theorem. We see that each point is connected once to every other point. Moreover, each contribution is multiplied by the coefficient one. Thus, the interacting correlation function (1.44) is expressed as a sum of diagrams that are constructed by applying the following rules:

- 1) the diagrams have n external points x_1, \dots, x_n and an arbitrary number k of internal points y_1, \dots, y_k ; the latter are called vertices;
- 2) lines connect pairs of points; a line is called internal if it connects two internal points, otherwise it is called external;
- 3) the line that connects two points z and w is associated with the Green function $G_B(z, w)$;
- 4) four legs are attached to each internal point, one leg to each external point;

5) each diagram with k vertices is multiplied by the factor

$$\frac{1}{k!} \left(\frac{-\lambda}{4!} \right)^k; \quad (1.45)$$

6) the positions y of the vertices are integrated with measure d^4y .

For example, consider the case $n = 2$, $k = 0, 1$. We have

$$\begin{aligned} \langle \varphi(x_1) \varphi(x_2) \rangle' &= G_B(x_1, x_2) - \frac{\lambda}{4!} \int d^4y \langle \varphi(x_1) \varphi(x_2) \varphi^4(y) \rangle_0 + \mathcal{O}(\lambda^2) \\ &= G_B(x_1, x_2) - \frac{\lambda}{2} \int d^4y G_B(x_1, y) G_B(y, y) G_B(y, x_2) \\ &\quad - \frac{\lambda}{8} G_B(x_1, x_2) \int d^4y G_B^2(y, y) + \mathcal{O}(\lambda^2), \end{aligned}$$

which graphically reads

$$\begin{array}{c} \text{---} x_1 \text{---} \bigcirc \text{---} x_2 \\ \text{---} x_1 \text{---} \text{---} x_2 \end{array} = \text{---} x_1 \text{---} \text{---} x_2 - \frac{\lambda}{2} \begin{array}{c} \bigcirc \\ \text{---} x_1 \text{---} y \text{---} x_2 \end{array} - \frac{\lambda}{8} \begin{array}{c} x_1 \text{---} \text{---} x_2 \\ \bigcirc \text{---} y \end{array} \quad (1.46)$$

plus $\mathcal{O}(\lambda^2)$. Different contributions originated by the right-hand side of (1.34) can give the same diagram, that is to say the same integral. For example, the second diagram on the right-hand side of (1.46) appears 12 times, which is why its coefficient is in the end $1/2$. Instead, the third diagram appears 3 times, so its coefficient is $1/8$.

We can collect the arrangements that give the same diagram into a single contribution, provided that we multiply it by a suitable combinatorial factor. Then, the perturbative expansion is organized as a sum over inequivalent diagrams G , which are multiplied by (1.45) and an extra factor s_G that counts how many contributions of Wick's theorem give the same G .

It is also convenient to move to momentum space, where some further simplifications occur. For example, consider the last-but-one diagram of (1.46). We find

$$\begin{aligned} &\int d^4u G_B(x, u) G_B(u, u) G_B(u, y) \\ &= \int d^4u \int \frac{d^4p}{(2\pi)^4} \frac{d^4k}{(2\pi)^4} \frac{d^4q}{(2\pi)^4} \frac{e^{ip(x-u)+iq(u-y)}}{(p^2 + m^2)(k^2 + m^2)(q^2 + m^2)}. \quad (1.47) \end{aligned}$$

The u -integral can be evaluated immediately, and gives $(2\pi)^4 \delta^{(4)}(p - q)$. Thus, (1.47) is the Fourier transform of

$$(2\pi)^4 \delta^{(4)}(p - q) \frac{1}{p^2 + m^2} \left(\int \frac{d^4 k}{(2\pi)^4} \frac{1}{k^2 + m^2} \right) \frac{1}{q^2 + m^2} \quad (1.48)$$

on p and q . This formula illustrates some properties that are actually valid for all graphs. First, we learn that it is much more convenient to work in momentum space, rather than in coordinate space. Indeed, (1.48) looks much simpler than the left-hand side of (1.47). Second, the theory is invariant under translations, so the total momentum is conserved. As a consequence, each correlation function is multiplied by a delta function like the one appearing in (1.48), which ensures that the momentum that enters the graph equals the momentum that exits from it, or, equivalently, that the total momentum that enters the graph vanishes. We do not need to write this delta function down every time, and from now on we will simply omit it. Third, the factors $1/(p^2 + m^2)$ and $1/(q^2 + m^2)$ are just the Green functions attached to the external legs: they do not enter the surviving integral. Thus, in momentum space we can “amputate” the diagram, which means omit the Green functions attached to the external legs. Note that the factorization (1.48) does not occur in coordinate space.

What remains is the “core” of our diagram, that is to say its truly non-trivial part, which is, in the case at hand,

$$\int \frac{d^4 k}{(2\pi)^4} \frac{1}{k^2 + m^2}. \quad (1.49)$$

Unfortunately, the integral (1.49) is infinite, as are many integrals that we are going to work with. However, this kind of problem, which is the main topic of this book, does not concern us right now. What is important here is that we have identified a few tricks that can help us save a lot of effort, by working in momentum space and concentrating on what occurs *inside* the diagram, since what happens outside is not new. From a certain point onwards, we will not need to use double lines to denote Green functions anymore, apart from the situations where it is really necessary: it will be understood that internal lines carry Green functions, while external lines do not.

Focusing on the cores of diagrams, we can now formulate the Feynman rules in momentum space, in arbitrary d dimensions, for the correlation functions (1.41) of a scalar field theory with arbitrary interactions.

The Fourier transform $\tilde{G}_B(p)$ of the two-point function $\langle \varphi(x)\varphi(y) \rangle'$ is called *propagator*. We have

$$\langle \varphi(x)\varphi(y) \rangle' = \int \frac{d^d p}{(2\pi)^d} e^{ip(x-y)} \tilde{G}_B(p).$$

The propagator is graphically denoted with a line that connects two points. We associate a *vertex* with each interaction term of the Lagrangian. A vertex is graphically denoted with lines ending at the same point, also called *legs*. Each leg is a field φ . The value of the vertex is equal to minus the coefficient of the associated Lagrangian term, summed over the permutations of identical legs. In momentum space, the momentum p of the Fourier transform $\tilde{\varphi}(p)$ is conventionally oriented towards the vertex.

For example, in the φ^4 theory we have (in arbitrary dimensions)

$$\text{---}\frac{\text{---}}{p}\text{---} = \frac{1}{p^2 + m^2} \quad \text{---}\times\text{---} = -\lambda \quad (1.50)$$

Consider a correlation function (1.41) at $J = 0$ with $n = E$ external legs and assume that we want to calculate its $\mathcal{O}(\lambda^k)$ -corrections. To achieve this goal,

- 1) assign a momentum p to every external leg, imposing overall momentum conservation;
- 2) draw all different diagrams G that have k vertices and E external legs;
- 3) assign a momentum q to every internal leg, imposing momentum conservation at every vertex.

Next, associate an integral \mathcal{I}_G with each diagram G as follows:

- a) write the propagator associated with every internal leg;
- b) multiply by the value of every vertex;
- c) multiply by the combinatorial factor c_G explained below;
- d) integrate over the surviving independent internal momenta q , with the measures $d^d q / (2\pi)^d$.

The combinatorial factor is given by the formula

$$c_G = \frac{s_G}{\prod_i n_i! c_i^{n_i}}. \quad (1.51)$$

Here, n_i is the number of vertices of type i contained in G , and c_i^{-1} is the combinatorial factor that multiplies the vertex of type i . For example, $c_i =$

$N!$, if the i th vertex has N identical legs, such as φ^N . Instead, $c_i = N_1!N_2!$, if the vertex is $\varphi^{N_1}\varphi^{N_2}$, and so on. Finally, the numerator s_G in the number of contributions of Wick's theorem that lead to the same diagram G .

The safest way to compute s_G is by drawing the vertices of G on a piece of paper, together with E points associated with the external legs. Then, count how many ways to connect the external legs to the legs attached to the vertices give the diagram G . The result of this counting is s_G . It is not advisable to avoid the counting and compute s_G by means of shortcuts (typically based on the symmetry properties of the diagram, which may be very difficult to spot), although some textbooks suggest to do so.

Normally, s_G is a huge number, to the extent that it almost simplifies the factors appearing in the denominator of (1.51). This is one reason why it is convenient to arrange the expansion in terms of diagrams. Nevertheless, sometimes it can be better, for theoretical purposes, to forget about diagrams and write the expansion as a sum over the sets of contributions coming from Wick's theorem, each of which has $s = 1$. So doing, it is much easier to have control over the combinatorial factors. We will use this kind of expansion in some proofs later on.

Diagrams can also be classified according to the number L of their “loops”. The precise definition of L is the number of independent internal momenta q , those on which we must integrate. Thus, formula (1.46) contains a one-loop diagram and a two-loop one. We will see later that the expansion in powers of λ coincides with the expansion in the number of loops. Graphically, loops appear as closed internal lines. However, it is not always easy to count them as such.

Basically, the combinatorial factors are due to identical legs. This is the reason why, to simplify some formulas, it is common to divide each Lagrangian term by the permutations of its identical legs. For example, in the φ^4 theory we have multiplied the quadratic part of the Lagrangian by $1/2!$ and the vertex by $1/4!$. With a different normalization, the propagators and the vertices get multiplied by extra coefficients. Apart from that, the rules to construct graphs and the formula for the combinatorial factors remain the same.

Finally, observe that the factors $1/(\prod_i n_i!)$ in c_G are brought by the

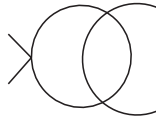
expansion of the exponential in power series, e.g.

$$\exp\left(-\frac{\lambda_4}{4!} \int \varphi^4 - \frac{\lambda_6}{6!} \int \varphi^6\right) = \sum_{n,m=0}^{\infty} \frac{(-\lambda_4)^n (-\lambda_6)^m}{n!m!(4!)^n (6!)^m} \left(\int \varphi^4\right)^n \left(\int \varphi^6\right)^m,$$

and correspond to the permutations of identical vertices.

We illustrate the calculation of combinatorial factors with a couple of examples. Consider the one-loop diagram of (1.46). It contains just one vertex with $c = 4!$. Moreover, we can easily verify that $s = 4 \cdot 3$, since the left external leg can be connected to the vertex in four ways, and then the right external leg can be connected to the vertex in three ways. In this particular case, the diagram is uniquely determined once the external legs are assigned. Thus, $c_G = (4 \cdot 3)/4! = 1/2$, which is indeed the factor that multiplies the diagram in formula (1.46), together with the value of the vertex, which is $-\lambda$.

Next, consider the diagram



(1.52)

It is made of three identical vertices, so we have a factor $1/(4!)^3$ and a factor $1/3!$. The coefficient s is equal to $3(4!)^3$. To calculate it, let us first draw three vertices with four legs each, and four external points. Then we connect the points in all the ways that lead to the graph we want. We begin from the up-left external leg, which can be arranged in $4 \cdot 3$ ways, where 3 is the number of vertices we can choose, and 4 is the number of legs of each vertex. Once that is done, the down-left external leg can be chosen in just 3 ways, because its vertex is already determined. Next, the up-right external leg can be arranged in $4 \cdot 2$ ways, after which the down-right external leg can be connected in 4 ways. Then, consider an internal leg of the left vertex: it can be attached to other internal legs in 6 ways. When this is done, the remaining internal leg of the left vertex can be attached to 3 internal legs. Finally, the remaining internal legs can be connected in 2 ways. In total

$$c_G = \frac{4 \cdot 3 \cdot 3 \cdot 4 \cdot 2 \cdot 4 \cdot 6 \cdot 3 \cdot 2}{3!(4!)^3} = \frac{3(4!)^3}{3!(4!)^3} = \frac{1}{2}.$$

Because of (1.42), the information just given is also sufficient to determine the correlation functions (1.29) at $J = 0$. In particular, $Z(0) = \langle 1 \rangle' Z_0(0)$

is a sum over diagrams with no external legs. There is a simple way to characterize the correlation functions without primes. Indeed, they differ from the correlation functions with primes just because they do not receive contributions from the diagrams that contain subdiagrams with no external legs. This statement will be proved at the end of the next section. Here we just give a simple example: the two-point function without primes at $\mathcal{O}(\lambda)$ simply loses the last term of (1.46), so

$$\langle \varphi(x_1) \varphi(x_2) \rangle = G_B(x_1, x_2) - \frac{\lambda}{2} \int d^4 y G_B(x_1, y) G_B(y, y) G_B(y, x_2) + \mathcal{O}(\lambda^2).$$

1.4 Generating functionals, Schwinger-Dyson equations

The rules given in the previous section determine the correlation functions with primes and the generating functional $Z(J)$. It turns out that $Z(J)$ contains redundant information. For example, working with $W(J)$, instead, of $Z(J)$, it is possible to reduce a lot of effort. A third functional, which is the Legendre transform of $W(J)$ and is denoted with $\Gamma(\Phi)$, allows us to further simplify the calculations. In this section we study the generating functionals and their properties. We start by deriving a functional equation for $Z(J)$, called *Schwinger-Dyson equation*.

We begin by noting that the functional integral of a total functional derivative is zero. We have

$$\begin{aligned} 0 &= \int [d\varphi] \frac{\delta}{\delta \varphi(x)} \exp \left(-S(\varphi) + \int J\varphi \right) \\ &= \int [d\varphi] \left(-\frac{\delta S}{\delta \varphi(x)} + J(x) \right) \exp \left(-S(\varphi) + \int J\varphi \right). \end{aligned} \quad (1.53)$$

Using the perturbative expansion, it is sufficient to prove this formula for the free field theory, but with an arbitrary set of φ -insertions. Consider a massive field. In the discretized version of the functional integral, where we have a finite number of ordinary integrals, we obviously have the identity

$$0 = \int_{-\infty}^{+\infty} \prod_{\{i_\mu\}} [d\varphi_{\{i_\mu\}}] \frac{\partial}{\partial \varphi_{\{k_\rho\}}} \left\{ \left(\prod_{\{n_\sigma\} \subset I} \varphi_{\{n_\sigma\}} \right) \exp \left(-S_{\text{discr}}(\varphi_{\{i_\mu\}}) \right) \right\}, \quad (1.54)$$

for every $\{k_\rho\}$ and every set I of insertions $\varphi_{\{n_\sigma\}}$, where $S_{\text{discr}}(\varphi_{\{i_\mu\}})$ is the free discretized action (1.24). Indeed, one integral, the one over $\varphi_{\{k_\rho\}}$, vanishes, because the exponential contains

$$\exp\left(-\frac{m^2}{2}\varphi_{\{k_\rho\}}^2\right),$$

which is sufficient to kill all contributions of the boundary $\varphi_{\{k_\rho\}} \rightarrow \pm\infty$. Since (1.54) holds for every lattice space a and size L , it also holds in the continuum limit.

The result is actually much more general, to the extent that it also holds when the mass vanishes and the free-field action is not positive definite in the Euclidean framework (which is the case, among others, of gravity). Indeed, we should not forget that, although we are temporarily working in the Euclidean framework, the correct theory is the one in Minkowski spacetime. There, the functional integral (1.21) contains an oscillating integrand, which can be always damped at infinity by assuming that the field has a mass with a small positive imaginary part $+i\varepsilon$, which is later sent to zero. So doing, we can prove that the identity (1.53) is always true in perturbative quantum field theory. The reader who is familiar with the operatorial formulation of quantum field theory will notice that this prescription is also the one that defines the correlation functions as T-ordered products. In other words, the functional integral automatically selects the T-ordered correlation functions.

Then, formula (1.53) gives

$$J(x)Z(J) = \int [d\varphi] \left[(-\square + m^2)\varphi(x) + \frac{\lambda}{3!}\varphi^3(x) \right] \exp\left(-S(\varphi) + \int J\varphi\right), \quad (1.55)$$

which can be graphically represented as

$$\bullet_x \quad \text{disc} = (-\square + m^2) \text{--- disc} + \frac{\lambda}{3!} \text{disc} \text{--- disc} \quad (1.56)$$

Here the disc stands for $Z(J)$ and the dot for J . A leg attached to the disc is a functional derivative with respect to J , i.e. a φ -insertion. Three legs meeting at the same point x denote three functional derivatives with respect to $J(x)$. To write (1.55), we have exchanged the functional integral with the derivatives contained in \square . In general, we have the identity $\partial_\mu \langle \varphi(x) \cdots \rangle =$

$\langle \partial_\mu \varphi(x) \cdots \rangle$, where the dots stand for any insertions at points different from x . We can prove this identity as follows. Consider the generating functional $Z(J)$ (1.23) and (for definiteness) the two-point function

$$\langle \varphi(x) \varphi(y) \rangle_J = Z(J)^{-1} \frac{\delta^2 Z(J)}{\delta J(x) \delta J(y)}.$$

If we write $J(x) = J_1(x) - \partial_\mu J_2^\mu(x)$ inside (1.23), where J_1 and J_2^μ are arbitrary, the functional derivative with respect to J_2^μ originates an insertion of $\partial_\mu \varphi(y)$. To see this, we must use

$$\int \varphi(J_1 - \partial_\mu J_2^\mu) = \int \varphi J_1 + \int J_2^\mu (\partial_\mu \varphi),$$

where the integration by parts can be justified by assuming that J_2^μ decreases rapidly enough at infinity. Indeed, since the sources J are test functions, we can choose them as smooth as we want and, if needed, with compact support. Thus, we find

$$\begin{aligned} \langle \varphi(x) \partial_\mu \varphi(y) \rangle_J &= Z(J)^{-1} \frac{\delta^2 Z(J)}{\delta J_1(x) \delta J_2^\mu(y)} = Z(J)^{-1} \frac{\delta^2 Z(J_1 - \partial_\rho J_2^\rho)}{\delta J_1(x) \delta J_2^\mu(y)} \\ &= Z(J)^{-1} \partial_\rho^{(y)} \frac{\delta^2 Z(J)}{\delta J(x) \delta J(y)} = \partial_\rho^{(y)} \langle \varphi(x) \varphi(y) \rangle_J. \end{aligned}$$

Multiplying both sides of (1.55) by $G_B(y, x)$, integrating over x and re-labeling $y \rightarrow x$ we obtain

$$x \text{ --- } \text{[shaded circle]} = x \text{ --- } \bullet \text{ --- } \text{[shaded circle]} - \frac{\lambda}{3!} \text{ --- } \text{[double circle with one shaded]} \quad (1.57)$$

where, as before, the double line stands for the Green function.

We can derive an alternative equation,

$$(-\square + m^2) \langle \varphi(x) \rangle_J = J(x) - \frac{\lambda}{3!} \langle \varphi^3(x) \rangle_J, \quad (1.58)$$

if we divide both sides of (1.55) by $Z(J)$. Again, if we insert the Green function in (1.58), we get

$$\langle \varphi(x) \rangle_J = \int d^4 y \, G_B(x, y) J(y) - \frac{\lambda}{3!} \int d^4 y \, G_B(x, y) \langle \varphi^3(y) \rangle_J. \quad (1.59)$$

Now, recalling that $Z = \exp(W)$, observe that

$$\langle \varphi^3 \rangle_J = e^{-W(J)} \frac{\delta^3}{\delta J^3} e^{W(J)} = W''' + 3W'W'' + W'^3, \quad (1.60)$$

each apex denoting a J derivative. Then equation (1.59) can be graphically represented as

$$\begin{aligned}
 & \text{---} \bigcirc_W = \text{---} \bullet \text{---} -\frac{\lambda}{3!} \text{---} \bigcirc_W \\
 & -\frac{\lambda}{2} \text{---} \bigcirc_W = -\frac{\lambda}{3!} \text{---} \bigcirc_W \text{---} \bigcirc_W \text{---} \bigcirc_W
 \end{aligned} \quad (1.61)$$

where now the disc denotes W . Again, the legs attached to the disc denote functional derivatives with respect to J .

The third generating functional $\Gamma(\Phi)$ is the Legendre transform of W . Define the functional $\Phi(J)$ as

$$\Phi(J)_x = \frac{\delta W(J)}{\delta J(x)} = \langle \varphi(x) \rangle_J. \quad (1.62)$$

From (1.28) we have

$$\Phi(J)_x = \int d^4y G_B(x, y) J(y) + \mathcal{O}(\lambda).$$

We can perturbatively invert $\Phi(J)$ and define the functional $J(\Phi)(x)$ such that $J(\Phi(J))_x = 1$. We have

$$J(\Phi)_x = (-\square + m^2)\Phi(x) + \mathcal{O}(\lambda). \quad (1.63)$$

Now, the functional $\Gamma(\Phi)$ is defined as

$$\Gamma(\Phi) = -W(J(\Phi)) + \int d^4x J(\Phi)_x \Phi(x). \quad (1.64)$$

We easily find

$$\Gamma(\Phi) = \frac{1}{2} \int d^4x ((\partial_\mu \Phi)^2 + m^2 \Phi^2) + \mathcal{O}(\lambda),$$

so Γ looks like a sort of “quantum action”. In the literature it is often called *effective action*. Note, however, that in Minkowski spacetime Γ is not even real. The Γ functional collects the amplitudes that are necessary to calculate the S matrix.

Let us work out the Schwinger-Dyson equation satisfied by Γ . First, observe that since Γ is a Legendre transform we have

$$\frac{\delta\Gamma(\Phi)}{\delta\Phi(x)} = J(\Phi)_x. \quad (1.65)$$

This relation can be easily verified by explicit differentiation. Second, using the formula for the derivative of the inverse function, we also have

$$\frac{\delta^2 W}{\delta J(x)\delta J(y)} = \frac{\delta\Phi(J)_y}{\delta J(x)} = \left(\frac{\delta J(\Phi)_x}{\delta\Phi(y)} \right)^{-1} = \left(\frac{\delta^2\Gamma(\Phi)}{\delta\Phi(x)\delta\Phi(y)} \right)^{-1}. \quad (1.66)$$

We write this formula symbolically as $W_{xy} = 1/\Gamma_{xy}$, where the subscripts denote derivatives with respect to the arguments (J for W , Φ for Γ) at the specified points. Third,

$$W_{xyz} = - \int \frac{1}{\Gamma_{xs}} \frac{1}{\Gamma_{yt}} \frac{1}{\Gamma_{zu}} \Gamma_{stu}, \quad (1.67)$$

where the integral is over the repeated subscripts. Using (1.60)-(1.67), equation (1.58) becomes

$$-\frac{\delta\Gamma(\Phi)}{\delta\Phi(x)} = -(-\square + m^2)\Phi(x) - \frac{\lambda}{3!} \left(\Phi^3(x) - \int \frac{1}{\Gamma_{xs}} \frac{1}{\Gamma_{xt}} \frac{1}{\Gamma_{xu}} \Gamma_{stu} + \frac{3}{\Gamma_{xx}} \Phi(x) \right). \quad (1.68)$$

Graphically, this formula reads

$$\begin{aligned} \text{---}_x \bigcirc_{-\Gamma} &= - \left(-\square + m^2 \right) \Phi(x) - \frac{\lambda}{3!} \Phi^3(x) \\ &\quad - \frac{\lambda}{3!} \text{---}_x \bigcirc_{-\Gamma} \text{---}_x + \frac{\lambda}{2!} \bigcirc_{-\Gamma} \text{---}_x \end{aligned} \quad (1.69)$$

where the line with a cut denotes $1/\Gamma''$.

We know that the correlation functions $\langle \varphi \cdots \varphi \rangle$ can be expressed as functional derivatives of $Z(J)$ with respect to J , calculated at $J = 0$, and divided by $Z(0)$. Similarly, the functional derivatives of W with respect to J , calculated at $J = 0$, and the functional derivatives of Γ with respect to Φ , calculated at $\Phi = 0$, define W and Γ correlation functions, respectively. Our purpose is to characterize the correlation functions of Z , W and Γ more precisely and find the relations among them.

The functional Z is the generator of all correlation functions. We prove that W is the generating functional of the *connected* correlation functions. That is to say, W contains precisely the contributions to Z originated by connected diagrams. We then write

$$W(J) = \sum_{n=0}^{\infty} \frac{1}{n!} \int \left(\prod_{i=1}^n d^4 x_i \right) \langle \varphi(x_1) \cdots \varphi(x_n) \rangle_c J(x_1) \cdots J(x_n),$$

where the subscript c stands for “connected”.

Moreover, we prove that Γ is the generating functional of the connected, amputated (which means that the external legs carry no Green functions G_B) *one-particle irreducible* (commonly abbreviated as 1PI) correlation functions, which we simply call “irreducible”. Irreducible diagrams are those that do not become disconnected by cutting one internal line. Precisely, we prove that $-\Gamma$ exactly contains the (amputated) contributions to Z and W that are due to irreducible diagrams, with only one exception: the free two-point function, which has an extra minus sign. We then write

$$-\Gamma(\Phi) = \sum_{n=0}^{\infty} \frac{1}{n!} \int \left(\prod_{i=1}^n d^4 x_i \right) \langle \varphi(x_1) \cdots \varphi(x_n) \rangle_{1\text{PI}} \Phi(x_1) \cdots \Phi(x_n).$$

To prove that the W and Γ correlation functions are connected and irreducible, respectively, it is sufficient to note that

- i) W and Γ are connected and irreducible, respectively, at the free-field level;
- ii) the W equation (1.61) and the Γ equation (1.69) are connected and irreducible, respectively;
- iii) equations (1.61) and (1.69) can be solved algorithmically from the free-field theory.

Property i) is obvious. We now prove that equations (1.61) and (1.69) are connected and irreducible, respectively. Observe that equation (1.57), instead, is neither of the two. Indeed, (1.57) contains the product $J(x)Z(J)$,

and generates disconnected contributions when we differentiate with respect to J . Equation (1.61) contains no products of functionals, which means that it is connected. On the other hand, it is clearly reducible. Finally, equation (1.69) is connected and irreducible. Indeed, the first three terms of (1.69) are the classical field equations. Being local, they are just vertices, rather than diagrams, so they are irreducible. The other terms of (1.69) are clearly irreducible.

Next, we prove that equations (1.57), (1.61) and (1.69) can be solved algorithmically starting from the free-field theory. Observe that by repeatedly differentiating those equations with respect to the sources, J or Φ , and later setting J or Φ to zero, we obtain relations among the correlation functions of Z , W and Γ . Each differentiation amounts to add a leg to a disc and sum appropriately. The right-hand sides of equations (1.57), (1.61) and (1.69) are the sums of two sets of contributions, which we call U_1 and U_2 . The set U_1 is the one that does not carry a factor of λ . It contains no disc, or a disc with no leg. The set U_2 is the one that carries a factor of λ and contains discs with at most three legs. An analogous decomposition holds for the differentiated equations and is the crucial property to prove our construction. If we take n functional derivatives, the left-hand sides become discs with $n + 1$ legs, which stand for the $(n + 1)$ -point correlation functions. The right-hand sides are, again, the sums of two types of contributions, U_1 and U_2 . The set U_1 contains no factor of λ and discs with at most n legs. In the cases of W and Γ such a set vanishes after a sufficient number of functional derivatives. The set U_2 contains a factor of λ and discs with at most $n + 3$ legs.

Equations (1.57), (1.61) and (1.69) ensure that to determine the $(n + 1)$ -point function to the order λ^k , it is sufficient to know the m -point functions, $m \leq n$, up to the order λ^k and the m' -point functions $m' \leq n + 3$, up to the order λ^{k-1} . Iterating the argument r times, we find that to determine the $(n + 1)$ -point function up to the order λ^k , we need to know the m -point functions, $m \leq n - r + 3h$, to the orders λ^{k-h} , with $h = 0, 1, \dots, r + 1$. Taking $r = n + 3k$, we need to know the m -point functions, $m \leq 3(h - k)$, up to the order λ^{k-h} : if $H \neq K$ we have zero, if $H = K$ we have $Z_0(0)$, which can be normalized to 1. This proves that equations (1.57), (1.61) and (1.69) can be solved algorithmically, as claimed.

We have considered, for simplicity, the φ^4 theory, but the results clearly extend to any polynomial theory in arbitrary spacetime dimensions.

Clearly, disconnected diagrams are products of connected ones, so $W(J)$ and $Z(J)$ contain the same amount of information. However, working with $W(J)$ instead of $Z(J)$ saves us some effort. In the free-field limit, for example, only the two-point function is connected, so $W(J)$ contains just one term [see (1.28)], while $Z(J)$ contains infinitely many, because it is the exponential of $W(J)$.

The simplification due to Γ is more clearly visible in momentum space, rather than in coordinate space. Observe that a convolution becomes a product after Fourier transform. The reducible diagrams are those that can be split into two parts, connected by a single leg. In momentum space they factorize, so they “disconnect”. Clearly, we loose no information if we concentrate on the “minimal” factors of such products. Working with Γ we take advantage of this simplification.

So far, we have proved that all the diagrams that contribute to W (respectively, Γ) are connected (irreducible). We still have to prove the converse, i.e. that all the connected (irreducible) diagrams do contribute to W (Γ). To show this, we proceed as follows.

Let us begin with W . Write

$$Z(J) = 1 + W(J) + \frac{1}{2!}W^2(J) + \frac{1}{3!}W^3(J) \cdots \quad (1.70)$$

Since $Z(J)$ contains all the diagrams, and $W(J)$ contains only connected diagrams, $W(J)$ contains all the connected diagrams of $Z(J)$. Now, take the connected part of equation (1.70), and note that the powers $W^n(J)$, $n > 1$, can give only disconnected contributions. Using (1.43), we get

$$Z'(J)|_c = Z(J)|_c = 1 + W(J)|_c = 1 + W(J). \quad (1.71)$$

Thus, the connected diagrams contained in $Z'(J) = Z(J)$ and $W(J)$ coincide. Moreover, in these two functionals they appear with the same coefficients. This property ensures that the Feynman rules we have determined for Z can be used also for W : we just have to discard the diagrams that are disconnected.

Comparing the two sides of (1.70), we get, in the first few cases,

$$\begin{aligned} \langle \varphi(x) \rangle_c &= \langle \varphi(x) \rangle, & \langle \varphi(x)\varphi(y) \rangle_c &= \langle \varphi(x)\varphi(y) \rangle - \langle \varphi(x) \rangle \langle \varphi(y) \rangle, \\ \langle \varphi(x)\varphi(y)\varphi(z) \rangle_c &= \langle \varphi(x)\varphi(y)\varphi(z) \rangle - \langle \varphi(x)\varphi(y) \rangle \langle \varphi(z) \rangle - \langle \varphi(y)\varphi(z) \rangle \langle \varphi(x) \rangle \\ &\quad - \langle \varphi(z)\varphi(x) \rangle \langle \varphi(y) \rangle + 2\langle \varphi(x) \rangle \langle \varphi(y) \rangle \langle \varphi(z) \rangle. \end{aligned}$$

Observe that $W(0)$ is the sum of connected diagrams that have no external legs. Consider a correlation function $\langle \varphi(x_1) \cdots \varphi(x_n) \rangle$ and write it in terms of W derivatives [see (1.60) for an example]. It is easy to check that $W(0)$ never appears: only the derivatives $W^{(n)}$ with $n > 0$ are involved. Thus, the correlation function $\langle \varphi(x_1) \cdots \varphi(x_n) \rangle$ can be expressed as the sum of products of connected diagrams that have a nonvanishing number of external legs. This statement was left without proof at the end of the previous section. Instead, the correlation function $\langle \varphi(x_1) \cdots \varphi(x_n) \rangle' = \langle \varphi(x_1) \cdots \varphi(x_n) \rangle e^{W(0) - W_0(0)}$ contains products of all the connected diagrams, including those that have no external legs. The diagrams that appear in both correlations functions are multiplied by the same coefficients.

It remains to study the correlation functions of $-\Gamma$. From (1.65) and (1.69) we see that $J(\Phi)$ is a sum of irreducible diagrams. Consider (1.64) and restrict it to the irreducible diagrams. We have

$$-\Gamma(\Phi) = -\Gamma(\Phi)|_{1\text{PI}} = W(J(\Phi))|_{1\text{PI}} - \int J(\Phi)\Phi|_{1\text{PI}}. \quad (1.72)$$

To manipulate this formula, it is convenient to write $J = (-\square + m^2)\Phi + \Delta J$ and expand in powers of ΔJ , where $\Delta J = \mathcal{O}(\lambda)$ can be read from the right-hand sides of (1.68) and (1.69). We find

$$W((-\square + m^2)\Phi) = W(J - \Delta J) = W(J) - \int \Delta J \frac{\delta W}{\delta J} + \frac{1}{2} \int \Delta J \frac{\delta^2 W}{\delta J \delta J} \Delta J + \cdots$$

Turning this expansion around, we can also write

$$\begin{aligned} W(J) - \int J\Phi &= W((-\square + m^2)\Phi) + \int (\Delta J - J) \frac{\delta W}{\delta J} - \frac{1}{2} \int \Delta J \frac{\delta^2 W}{\delta J \delta J} \Delta J \\ &+ \cdots = W((-\square + m^2)\Phi) - \int \Phi(-\square + m^2)\Phi - \frac{1}{2} \int \Delta J \frac{\delta^2 W}{\delta J \delta J} \Delta J + \cdots \end{aligned}$$

Now we take the one-particle irreducible contributions of both sides of this equation. Note that the last term, as well as the higher-order corrections collected inside the dots, always give reducible diagrams, since ΔJ contains vertices. Thus, we get

$$-\Gamma(\Phi) = W((-\square + m^2)\Phi)|_{1\text{PI}} - \int \Phi(-\square + m^2)\Phi. \quad (1.73)$$

Replacing J by $(-\square + m^2)\Phi$ inside $W(J)$ is equivalent to amputate the external legs and attach a field Φ to them. Formula (1.73) tells us that $-\Gamma$

contains the amputated irreducible diagrams of W , with exactly the same coefficient they have in W , apart from the free two-point function, which has an extra minus sign because of the last term of (1.73). Indeed, at the free-field level we have

$$\Gamma_0 = \frac{1}{2} \int \Phi(-\square + m^2)\Phi = \frac{1}{2} \int J(-\square + m^2)^{-1}J = W_0,$$

so $+\Gamma$ is the amputated W (instead of $-\Gamma$). Finally, the Feynman rules worked out for Z and W also work for $-\Gamma$ (apart from the free two-point function), provided we discard the reducible diagrams.

It is easy to see that the results of this section do not depend on the form of the vertex, nor on the free-field action around which we perturb, nor on the type of the fields. For example, if we replace the interaction $\sim \int \varphi^4$ by $\sim \int \varphi^6$, or by the sum of $\int \varphi^4$ and $\int \varphi^6$, or even by interactions that contain derivatives, such as $\sim \int \varphi^2(\partial_\mu \varphi)^2$, etc., all the arguments given above can be generalized with obvious modifications. The only assumption that is crucial for the derivation is that the interactions be local, which means that each vertex should be the integral of a monomial constructed with the fields and their derivatives. In the end, we find that in every local perturbative quantum field theory the generating functional Z contains all the correlation functions, while W and Γ contain only the connected and one-particle irreducible correlation functions, respectively. Moreover, the correlation functions appear in Z , W and Γ with the same coefficients, apart from the Γ free-field two-point function.

Exercise 1 *Integrating (1.69), calculate $\Gamma(\Phi)$ at the tree level and at one loop.*

Solution. The first line of (1.69) can be integrated straightforwardly, and gives $S(\Phi)$. The second line is made of two terms. The first of them can generate only two-loop diagrams, so we can neglect it. The second term gives diagrams that contain at least one loop. Thus, at the tree level the Γ functional coincides with the classical action: $\Gamma(\Phi) = S(\Phi)$.

To calculate the one-loop corrections it is sufficient to calculate Γ_{xx} at the tree level, which is just

$$S_{xy} \equiv \frac{\delta^2 S(\Phi)}{\delta \Phi(x) \delta \Phi(y)} = (-\square + m^2 + \frac{\lambda}{2} \Phi^2(x)) \delta(x - y),$$

in the limit $y \rightarrow x$. Then we insert it into the last term of equation (1.69), which becomes

$$\frac{\lambda}{2}\Phi(x)\frac{1}{\Gamma_{xx}} = \frac{\lambda}{2}\Phi(x)\frac{1}{S_{xx}} = \frac{1}{2}\frac{\delta}{\Phi(x)} \int d^4y \ln S_{yz}|_{z \rightarrow y}, \quad (1.74)$$

having used $\lambda\Phi = S'''$. Finally, the Γ functional reads

$$\Gamma(\Phi) = S(\Phi) + \frac{1}{2} \int d^4x \ln S_{xy}|_{y \rightarrow x} \equiv S(\Phi) + \frac{1}{2} \text{tr} \left[\ln \frac{\delta^2 S(\Phi)}{\delta\Phi(x)\delta\Phi(y)} \right], \quad (1.75)$$

plus two-loop corrections, plus unimportant constants. Although for clarity we have used the φ^4 theory to derive this result, it can be easily checked that formula (1.75) holds for an arbitrary action $S(\Phi)$, because the specific form of the action is actually not necessary for the derivation. \square

The classical action $S(\varphi)$ and the functional $-W(J)$ satisfy an interesting duality relation. Consider iJ as the “fields”, $S_J(iJ) \equiv -W(-iJ)$ as their classical action and φ as the sources coupled to iJ . Then, the W functional is equal to $-S(\varphi)$ itself. Precisely,

$$\int [dJ] \exp \left(W(-iJ) + \int iJ\varphi \right) = \exp(-S(\varphi)). \quad (1.76)$$

Indeed, using (1.23) the left-hand side can be written as

$$\int [dJ d\varphi'] \exp \left(-S(\varphi') + i \int \varphi J - i \int \varphi' J \right).$$

Integrating over J we get the “functional δ function”

$$\delta_F(\varphi - \varphi') = \prod_x \delta(\varphi(x) - \varphi'(x)),$$

whose meaning can be easily understood from the discretized version of the functional integral. Finally, integrating over φ' we get the right-hand side of (1.76).

From the perturbative point of view, it does not really matter whether J is multiplied by i or not. Thus, we can also write

$$\int [dJ] \exp \left(W(J) - \int J\varphi \right) = \exp(-S(\varphi)).$$

The meaning of this identity is that if we take the diagrams that contribute to the connected correlation functions, replace their vertices by minus the connected diagrams themselves, and the propagators by minus their inverses, the results we obtain are minus the vertices again.

1.5 Advanced generating functionals

We can also define generating functionals for n -particle irreducible connected Green functions, that is to say connected Green functions that become disconnected when n or fewer internal lines are cut into two. In this section we explain how. Although the new functionals are rarely met in the literature, they can help us gain a more complete picture of what we are doing. Moreover, some generalizations of these functionals are useful treat some topic of the next chapters.

We first study the generating functional of two-particle irreducible Green functions. We introduce a new source $K(x, y)$ coupled to the bilinear $\varphi(x)\varphi(y)$ and define

$$Z(J, K) = \int [d\varphi] \exp \left(-S(\varphi) + \int J\varphi + \frac{1}{2} \int \varphi K \varphi \right) = e^{W(J, K)},$$

where $\int \varphi K \varphi = \int dx \varphi(x) K(x, y) \varphi(y) dy$. Then, define

$$\Phi(x) = \frac{\delta W}{\delta J(x)} = \langle \varphi(x) \rangle, \quad N(x, y) = \frac{\delta^2 W}{\delta J(x) \delta J(y)} = \langle \varphi(x) \varphi(y) \rangle_c, \quad (1.77)$$

at nonzero J and K . Observe that

$$\frac{\delta W}{\delta K(x, y)} = \frac{1}{2} (N(x, y) + \Phi(x)\Phi(y)) = \frac{1}{2} \frac{\delta^2 W}{\delta J(x) \delta J(y)} + \frac{1}{2} \frac{\delta W}{\delta J(x)} \frac{\delta W}{\delta J(y)}. \quad (1.78)$$

This is a functional differential equation for $W(J, K)$. It shows that the K dependence is not unrelated to the J dependence, so the advanced functional $W(J, K)$ does not contain new information, but just the information already known, expressed in a different way.

Now, call $\Gamma_2(\Phi, N)$ the Legendre transform of $W(J, K)$ with respect to both J and K , that is to say

$$\begin{aligned} \Gamma_2(\Phi, N) &= -W(J, K) + \int \frac{\delta W}{\delta J} J + \int \frac{\delta W}{\delta K} K \\ &= -W(J, K) + \int J\Phi + \frac{1}{2} \int (NK + \Phi K \Phi), \end{aligned}$$

where NK stands for $N(x, y)K(x, y)$ and J and K are meant to be functions of Φ and N , obtained by inverting (1.77). That this transform is well defined will become evident soon. Differentiating Γ_2 we get

$$\frac{\delta \Gamma_2}{\delta \Phi(x)} = J(x) + \int K(x, y) \Phi(y) dy, \quad \frac{\delta \Gamma_2}{\delta N(x, y)} = \frac{1}{2} K(x, y). \quad (1.79)$$

To retrieve $\Gamma(\Phi)$ from $\Gamma_2(\Phi, N)$ it is sufficient to set $K = 0$, because then $W(J, K)$ becomes precisely the functional $W(J)$ encountered before. Inverting (1.79) we obtain Φ and N as functions of J and K . Once K is set to zero, the relations $\Phi = \Phi(J, 0)$ and $N = N(J, 0)$ allow us to express J as a function $J(\Phi)$ of Φ , which coincides with the relation found in the previous sections, but also N as a function $N(\Phi)$ of Φ . Finally,

$$\Gamma(\Phi) = \Gamma_2(\Phi, N(\Phi)).$$

At $J = K = 0$ we have that Φ is the expectation value of the field and N is the propagator.

The functional $\Gamma_2(\Phi, N)$ can also be seen as the Legendre transform

$$\Gamma_2(\Phi, N) = \Gamma(\Phi, K) + \frac{1}{2} \int K(N + \Phi\Phi) \quad (1.80)$$

of $\Gamma(\Phi, K)$, which is the usual Γ functional for the modified classical action

$$S(\varphi, K) = S(\varphi) - \frac{1}{2} \int \varphi K \varphi. \quad (1.81)$$

Exercise 2 Calculate $\Gamma_2(\Phi, N)$ for a free scalar field and rederive $\Gamma(\Phi)$.

Solution. The source $K(x, y)$ is like a non-local squared mass, so $W(J, K)$ can be obtained from the usual functional, replacing the mass m^2 with $m^2 - K$. From (1.26) we get

$$W(J, K) = \frac{1}{2} \int J(-\square + m^2 - K)^{-1} J - \frac{1}{2} \text{tr} \ln [-\square + m^2 - K].$$

We immediately find

$$\Phi = (-\square + m^2 - K)^{-1} J, \quad N = (-\square + m^2 - K)^{-1},$$

thus

$$\Gamma_2(\Phi, N) = \frac{1}{2} \int [(\partial_\mu \Phi)^2 + m^2 \Phi^2] - \frac{1}{2} \text{tr} \ln N + \frac{1}{2} \text{tr} [(-\square + m^2)N - 1].$$

Observe that objects such as $\ln N$ and N^{-1} are meaningful, since by (1.77) N^{-1} is just the scalar propagator. Setting $K = 0$ we find $N = (-\square + m^2)^{-1}$ and the usual free-field Γ -functional

$$\Gamma_2(\Phi, (-\square + m^2)^{-1}) = \frac{1}{2} \int [(\partial_\mu \Phi)^2 + m^2 \Phi^2] + \frac{1}{2} \text{tr} \ln (-\square + m^2) = \Gamma(\Phi),$$

which agrees with (1.75). In an interacting theory we obtain this expression plus corrections proportional to the couplings. As said before, all such corrections are made of tree-level expressions plus two-particle irreducible diagrams.

Exercise 3 Calculate $\Gamma_2(\Phi, N)$ at one loop for a generic theory $S(\varphi)$.

Solution. We start from formula (1.75), which gives the most general one-loop Γ -functional, and apply it to a classical theory with modified action (1.81). We obtain the one-loop Γ -functional

$$\Gamma(\Phi, K) = S(\Phi) - \frac{1}{2} \int \Phi K \Phi + \frac{1}{2} \text{tr} \ln(S'' - K),$$

where S'' stands for S_{xy} . Now we further Legendre-transform with respect to K . Differentiating we get

$$\frac{\delta \Gamma}{\delta K} = -\frac{1}{2} \frac{1}{S'' - K} - \frac{1}{2} \Phi \Phi = -\frac{\delta W}{\delta K},$$

which gives

$$N = \frac{1}{S'' - K}.$$

Finally, using (1.80) the one-loop functional Γ_2 is

$$\Gamma_2(\Phi, N) = S(\Phi) - \frac{1}{2} \text{tr} \ln N + \frac{1}{2} \text{tr} [N S''(\Phi) - 1]. \quad (1.82)$$

□

Now we study the diagrammatics of $\Gamma_2(\Phi, N)$. Since every one-loop diagrams are two-particle reducible, unless they contain just one vertex (in which case they are called “tadpoles”), it is useful to consider the difference $\tilde{\Gamma}_2(\Phi, N)$ between $\Gamma_2(\Phi, N)$ and its one-loop expression (1.82):

$$\tilde{\Gamma}_2(\Phi, N) = \Gamma_2(\Phi, N) - S(\Phi) + \frac{1}{2} \text{tr} \ln N - \frac{1}{2} \text{tr} [N S''(\Phi) - 1]. \quad (1.83)$$

Now, the functional $\Gamma(\Phi, K)$ is the set of one-particle irreducible diagrams of the theory $S(\varphi, K)$, namely the set of one-particle irreducible diagrams of $S(\varphi)$ with inverse propagator shifted by $-K$. Separate the tree-level contribution $S(\Phi, K)$ of $\Gamma(\Phi, K)$ from the rest, by writing

$$\Gamma(\Phi, K) = S(\Phi) - \frac{1}{2} \int \Phi K \Phi + \tilde{\Gamma}(\Phi, K). \quad (1.84)$$

The two-point function of $\Gamma(\Phi, K)$ is

$$\frac{\delta^2 \Gamma(\Phi, K)}{\delta \Phi \delta \Phi} = S''(\Phi) - K - \frac{\delta^2 \tilde{\Gamma}(\Phi, K)}{\delta \Phi \delta \Phi} = \left(\frac{\delta^2 W(J, K)}{\delta J \delta J} \right)^{-1} = \frac{1}{N}. \quad (1.85)$$

The last two equalities follow from (1.66) and the second of (1.77).

Take formula (1.80). Using (1.84) and (1.85) we have

$$\Gamma_2(\Phi, N) = S(\Phi) + \tilde{\Gamma}(\Phi, K) + \frac{1}{2} \text{tr} \left[N S''(\Phi) - N \frac{\delta^2 \tilde{\Gamma}(\Phi, K)}{\delta \Phi \delta \Phi} - 1 \right].$$

Now we must re-express K as a function of Φ and N on the left-hand side. Formula (1.85) tells us that all propagators just become N . Then we just have to replace the sources K that appear in the vertices. Observe that each K is attached to two φ legs, so also two propagators N . Thus, we have to consider the product NKN . Using (1.85) we see that

$$NKN \rightarrow N S''(\Phi) N - N - N \frac{\delta^2 \tilde{\Gamma}(\Phi, K)}{\delta \Phi \delta \Phi} N.$$

The sources K on the right-hand side can be treated recursively. Then it is easy to see that the diagrams of $\Gamma_2(\Phi, N)$, and also those of $\tilde{\Gamma}_2(\Phi, N)$, are one-particle irreducible.

Working out the N derivative of $\tilde{\Gamma}_2$ and using (1.85), we get

$$\frac{\delta \tilde{\Gamma}_2}{\delta N} = \frac{\delta \Gamma_2}{\delta N} + \frac{1}{2N} - \frac{1}{2} S''(\Phi) = -\frac{1}{2} \frac{\delta^2 \tilde{\Gamma}(\Phi, K)}{\delta \Phi \delta \Phi}.$$

Repeating the argument above, we find that the diagrams of $\delta \tilde{\Gamma}_2 / \delta N$ are also one-particle irreducible. Then the diagrams of $\tilde{\Gamma}_2$ are two-particle irreducible, because taking an N derivative is equivalent to cutting one internal line.

The functional Γ_∞ is defined by coupling sources $K_n(x_1, \dots, x_n)$ to arbitrary strings $\varphi(x_1) \cdots \varphi(x_n)$ of φ -insertions:

$$Z(J, K) = \int [d\varphi] \exp \left(-S(\varphi) + \int J\varphi + \sum_{n=2}^{\infty} \frac{1}{n!} \int K_n \overbrace{\varphi \cdots \varphi}^n \right).$$

Then $W(J, K) = \ln Z(J, K)$ as usual and

$$\Phi = \frac{\delta W}{\delta J} = \langle \varphi \rangle, \quad N_n = \frac{\delta^n W}{\underbrace{\delta J \cdots \delta J}_n} = \langle \overbrace{\varphi \cdots \varphi}^n \rangle_c. \quad (1.86)$$

We have, in compact notation,

$$\frac{\delta W}{\delta K_n} = \frac{1}{n!} \langle \overbrace{\varphi \cdots \varphi}^n \rangle = \frac{1}{n!} e^{-W} \frac{\delta^n}{\delta J^n} e^W \Big|_{W' \rightarrow \Phi, W^{(k)} \rightarrow N_k}.$$

Performing the Legendre transform with respect to all the sources, we obtain the functional

$$\Gamma_\infty(\Phi, N) = -W(J, K) + \int \frac{\delta W}{\delta J} J + \sum_{n=2}^{\infty} \int \frac{\delta W}{\delta K_n} K_n,$$

where J and the sources K_n need to be expressed as functions of Φ and N_k by inverting (1.86). The functional $\Gamma(\Phi)$ is retrieved by setting all the sources K_n to zero. The functional $\Gamma_2(\Phi, N)$ is obtained by setting all of them to zero but K_2 , and so on.

1.6 Massive vector fields

So far, we have just considered scalar fields. Massive vector fields can be treated in a similar way, while fermions of spin 1/2 require that we extend the definition of functional integral to anticommuting variables. Finally, gauge fields need a separate treatment, since the definition of the functional integral in the presence of gauge symmetries is not obvious, even in the Gaussian limit.

In the case of massive vector fields, we start from the free Proca action

$$S_{\text{free}}(W) = \int d^4x \left(\frac{1}{4} W_{\mu\nu}^2 + \frac{m^2}{2} W_\mu^2 \right), \quad (1.87)$$

where $W_{\mu\nu} \equiv \partial_\mu W_\nu - \partial_\nu W_\mu$. The field equations

$$-\square W_\mu + \partial_\mu \partial_\nu W_\nu + m^2 W_\mu = 0 \quad (1.88)$$

ensure that the theory propagates only three degrees of freedom at the classical level, since the divergence of (1.88) gives $m^2 \partial_\mu W_\mu = 0$. The propagator $G_{\mu\nu}(x, y) = \langle W_\mu(x) W_\nu(y) \rangle$ is the solution of the differential equation

$$(-\square \delta_{\mu\nu} + \partial_\mu \partial_\nu + m^2 \delta_{\mu\nu}) G_{\nu\rho}(x, y) = \delta_{\mu\rho} \delta^{(4)}(x - y)$$

and can be easily expressed by means of the Green function G_B of the scalar field. Indeed, recalling (1.27), we find

$$G_{\mu\nu}(x, y) = \left(\delta_{\mu\nu} - \frac{\partial_\mu \partial_\nu}{m^2} \right) G_B(x, y) = \int \frac{d^4 p}{(2\pi)^4} e^{ip \cdot (x-y)} \frac{\delta_{\mu\nu} + \frac{p_\mu p_\nu}{m^2}}{p^2 + m^2}. \quad (1.89)$$

At the quantum level the degrees of freedom can be counted by counting the poles of the propagator in momentum space, after switching to the Minkowskian framework. Basically, on the pole, which can be reached from the Euclidean notation by writing $p = (\pm im, 0, 0, 0)$ in the rest frame, the numerator is the matrix $\text{diag}(0, 1, 1, 1)$. The three positive eigenvalues are the propagating degrees of freedom, while the eigenvalue zero corresponds to the nonpropagating degree of freedom $\partial_\mu W_\mu$.

When we add interactions, the Feynman rules and the diagrammatics follow straightforwardly, as well as the definitions of the generating functionals.

Massless vectors are the fields of gauge theories. It is clear that the massless limit of (1.89) is singular. Thus, gauge fields need a separate discussion. For the same reason, the ultraviolet limit of a theory that contains massive vector fields is singular, because there the mass becomes negligible. Another way to see the problem of massive vector fields at high-energies is that the propagator in momentum space behaves like $\sim p_\mu p_\nu / (m^2 p^2)$ for large, p instead of $\sim 1/p^2$. We will see that this behavior does not decrease fast enough to have renormalizability. In general, an interacting quantum field theory that contains massive vector fields is nonrenormalizable. The same conclusion applies to the theories that contain massive fields of higher spins, which we do not treat here.

1.7 Fermions

The functional integral provides a formulation of quantum mechanics that is equivalent to the orthodox ones. Its main virtue is that it allows us to work with functions, instead of operators. In practice, summing over all paths that connect the initial point to the final one has the same effect as working with objects that have nontrivial commutators. In some sense, the right-hand sides of the commutators $[\hat{q}, \hat{p}] = i$, $[\hat{a}, \hat{a}^\dagger] = 1$, where $[A, B] = AB - BA$, are replaced by the functional integration.

We know that, to be consistency with the Fermi statistics, the second quantization of fermions is done, in the operator approach, by assuming that

there exist annihilation and creation operators \hat{a}_f and \hat{a}_f^\dagger that satisfy the anticommutation relation $\{\hat{a}_f, \hat{a}_f^\dagger\} = 1$, where $\{A, B\} = AB + BA$. We expect that a functional integral over fermions can replace the right-hand side of this anticommutator. We do not expect, however, that it can do more than that, for example allow us to work with commuting objects, instead of anticommuting ones. Indeed, the Pauli exclusion principle, which is the origin of anticommutators, survives the classical limit $\hbar \rightarrow 0$. The right-hand sides of commutators and anticommutators vanish when \hbar tends to zero, but the left-hand sides remain unchanged. Thus, to describe fermions we need to work with anticommuting objects, and define a suitable integral over them.

Such objects are called Grassmann variables, and for the time being we denote them with $\theta_i, \bar{\theta}_i$. They satisfy

$$\{\theta_i, \theta_j\} = \{\theta_i, \bar{\theta}_j\} = \{\bar{\theta}_i, \bar{\theta}_j\} = 0.$$

We also need to define functions of such variables, then the “ordinary” integral over them and finally the functional integral. These concepts will sound a bit formal, however we know that we must be prepared to upgrade the mathematics and include notions that may not sound familiar at first. Quantum mechanics already thought us a lot on this. In the derivations of physical predictions we may need to work with quantities, such as the wave function, which do not have a direct connection with reality. Ultimately, we just need to retrieve ordinary real numbers at the very end, when we work out cross sections or any other physical quantity. In the intermediate steps, we are free to introduce any objects we want, not matter how awkward they may look at first sight, as long as they are equipped with a set of consistent axioms that allow us to manipulate them, and are such that the physical quantities we get at the end are real.

Consider a generic function of a single Grassmann variable θ . Making a Taylor expansion around $\theta = 0$, we just have

$$f(\theta) = a + \theta b, \quad a = f(0), \quad b = f'(0). \quad (1.90)$$

Here a and b are constants. Every other term of the Taylor expansion disappears, since $\theta^2 = (1/2)\{\theta, \theta\} = 0$.

Similarly, a function of two variables $\theta, \bar{\theta}$ reads

$$g(\theta, \bar{\theta}) = c + \theta d + \bar{\theta} e + \theta \bar{\theta} f,$$

c, d, e and f being other constants.

Ordinary commuting variables are normally called “ c -numbers”, to distinguish them from the Grassmann variables. If the function f of formula (1.90) is a c -number, then a is also a c -number, while b is an anticommuting constant. If f is anticommuting, then a also is anticommuting, while b is a c -number. We also say that c -numbers have bosonic statistics, while anticommuting objects have fermionic statistics.

Now we define the integral of $f(\theta)$ in $d\theta$. We introduce differentials $d\theta$, which are also anticommuting objects, and assume that the integral is linear and translational invariant. By linearity

$$\int d\theta f(\theta) = \left(\int d\theta 1 \right) a + \left(\int d\theta \theta \right) b,$$

so it is sufficient to define the integrals of 1 and θ . Let us perform the change of variables $\theta = \theta' + \xi$, where ξ is constant and anticommuting. Then, by translational invariance $d\theta = d\theta'$, so

$$\int d\theta \theta = \int d\theta' (\theta' + \xi) = \int d\theta' \theta' + \left(\int d\theta' 1 \right) \xi = \int d\theta \theta + \left(\int d\theta 1 \right) \xi.$$

We conclude that the integral of 1 in $d\theta$ must vanish. Then, the integral of θ must not be zero, otherwise our integral would identically vanish. Normalizing the integral of θ to 1, we have the formal rules

$$\int d\theta 1 = 0, \quad \int d\theta \theta = 1,$$

which define the *Berezin integral*.

In practice, the Berezin integral behaves like a derivative. For example, under a rescaling $\theta' = c\theta$ we have

$$1 = \int d\theta' \theta' = c \int d(c\theta) \theta = \int d\theta \theta,$$

whence, differently from usual,

$$d(c\theta) = \frac{1}{c}d\theta.$$

This rule coincides with the one of the derivative with respect to θ .

The basic Gaussian integral reads

$$\int d\bar{\theta}d\theta e^{-m\bar{\theta}\theta} = \int d\bar{\theta}d\theta (1 - m\bar{\theta}\theta) = m. \quad (1.91)$$

The minus sign disappears when we anticommute $\bar{\theta}$ with $d\theta$.

With more variables it is easy to prove that

$$\int \prod_{i=1}^N d\bar{\theta}_i d\theta_i \bar{\theta}_{i_1} \theta_{j_1} \cdots \bar{\theta}_{i_N} \theta_{j_N} = (-1)^N \varepsilon_{i_1 \dots i_N} \varepsilon_{j_1 \dots j_N}. \quad (1.92)$$

Indeed, the result must be completely antisymmetric in $i_1 \cdots i_N$ and $j_1 \cdots j_N$. Taking $i_k = j_k = k$ and using (1.91) we correctly get $(-1)^N$.

Then, defining the action

$$S(\bar{\theta}, \theta) = \sum_{i,j=1}^N \bar{\theta}_i M_{ij} \theta_j,$$

where M_{ij} is some matrix, we get

$$\begin{aligned} \int \prod_{i=1}^N d\bar{\theta}_i d\theta_i e^{-S(\bar{\theta}, \theta)} &= \frac{(-1)^N}{N!} \int \prod_{i=1}^N d\bar{\theta}_i d\theta_i S^N(\bar{\theta}, \theta) \\ &= \frac{1}{N!} \varepsilon_{i_1 \dots i_N} \varepsilon_{j_1 \dots j_N} M_{i_1 j_1} \cdots M_{i_N j_N} = \det M. \end{aligned}$$

Every other contribution coming from the exponential integrates to zero, because it cannot saturate the Grassmann variables θ and $\bar{\theta}$. We can easily generalize this formula to

$$\begin{aligned} Z(\bar{\xi}, \xi) &\equiv \int \prod_{i=1}^N d\bar{\theta}_i d\theta_i \exp \left(-S(\bar{\theta}, \theta) + \sum_{i=1}^N (\bar{\xi}_i \theta_i + \bar{\theta}_i \xi_i) \right) \\ &= \exp \left(\sum_{i,j=1}^N \bar{\xi}_i M_{ij}^{-1} \xi_j \right) \det M, \end{aligned} \quad (1.93)$$

with the help of the translation $\theta = \theta' + M^{-1}\xi$, $\bar{\theta} = \bar{\theta}' + \bar{\xi}M^{-1}$.

Finally, a generic change of variables $\theta = \theta(\theta')$ produces the reciprocal of the usual Jacobian determinant,

$$\prod_{i=1}^N d\theta_i = \left(\det \frac{\partial \theta}{\partial \theta'} \right)^{-1} \prod_{i=1}^N d\theta'_i. \quad (1.94)$$

The derivative with respect to Grassmann variables can be placed at the left of the differentials $d\bar{\theta}$, $d\theta$, or at the right of them, so we define left- and right-derivatives ∂_l and ∂_r , which can at most differ by a minus sign. Precisely, the differential of a function reads

$$df(\bar{\theta}, \theta) = \frac{\partial_r f}{\partial \bar{\theta}_i} d\bar{\theta}_i + \frac{\partial_r f}{\partial \theta_i} d\theta_i = d\bar{\theta}_i \frac{\partial_l f}{\partial \bar{\theta}_i} + d\theta_i \frac{\partial_l f}{\partial \theta_i}.$$

Of course, $\partial_l/\partial \bar{\theta}$ and $\partial_l/\partial \theta$ are anticommuting objects, as well as $\partial_r/\partial \bar{\theta}$ and $\partial_r/\partial \theta$. However, observe that

$$\frac{\partial_r}{\partial \bar{\theta}_i} \frac{\partial_l}{\partial \theta_j} = \frac{\partial_l}{\partial \theta_j} \frac{\partial_r}{\partial \bar{\theta}_i}.$$

We can define averages

$$\langle \theta_{i_1} \cdots \theta_{i_n} \bar{\theta}_{j_1} \cdots \bar{\theta}_{j_n} \rangle = \frac{1}{Z(\bar{\xi}, \xi)} \frac{\partial_l}{\partial \xi_{i_1}} \cdots \frac{\partial_l}{\partial \xi_{i_n}} \frac{\partial_r}{\partial \xi_{j_n}} \cdots \frac{\partial_r}{\partial \xi_{j_1}} Z(\bar{\xi}, \xi) \Big|_{\xi=\bar{\xi}=0}.$$

In particular,

$$\langle \theta_i \bar{\theta}_j \rangle = M_{ij}^{-1}, \quad \langle \theta_i \theta_j \bar{\theta}_k \bar{\theta}_l \rangle = M_{il}^{-1} M_{jk}^{-1} - M_{ik}^{-1} M_{jl}^{-1}. \quad (1.95)$$

We can also have integrals over both commuting variables x and anti-commuting variables θ . Writing $\bar{z} = (\bar{x}, \bar{\theta})$ and $z = (x, \theta)$, we define the *superdeterminant* as

$$\begin{aligned} (\text{sdet} M)^{-1} &\equiv \frac{1}{(2\pi)^N} \int d\bar{z} dz \exp(-\bar{z}^t M z) \\ &= \frac{1}{(2\pi)^N} \int d\bar{x} dx d\bar{\theta} d\theta \exp(-\bar{x}^t A x - \bar{x}^t B \theta - \bar{\theta} C x - \bar{\theta} D \theta), \end{aligned}$$

where the normalization factor is chosen so that $\text{sdet} \mathbf{1} = 1$,

$$M = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

and each block A , B , C and D is a $N \times N$ matrix, where A , D contain commuting entries and B , C contain anticommuting entries. To compute the superdeterminant, we perform the translations $\bar{y}^t = \bar{x}^t + \bar{\theta} C A^{-1}$ and $y = x + A^{-1} B \theta$ and observe that in the variables $\bar{\zeta} = (\bar{y}, \bar{\theta})$, $\zeta = (y, \theta)$, we

have $\bar{z}^t M z = \bar{y}^t A y + \bar{\theta}(D - C A^{-1} B)\theta$, so the integrals over commuting and anticommuting variables factorize. At the end, we find

$$\text{sdet} M = \frac{\det A}{\det(D - C A^{-1} B)}. \quad (1.96)$$

A useful property, which we do not prove here, is

$$\ln \text{sdet}(\exp M) = \text{str} M \equiv \text{tr} A - \text{tr} D, \quad (1.97)$$

where “str” denotes the so-called supertrace. Moreover, the infinitesimal variation is

$$\delta \text{sdet} M = \delta \exp(\text{str} \ln M) = (\text{sdet} M) \text{str}(M^{-1} \delta M). \quad (1.98)$$

In this book we just need a special case of this formula, when $M = 1 + \delta M$ and δM is small. Then

$$\text{sdet}(1 + \delta M) \sim 1 + \text{tr}[\delta A] - \text{tr}[\delta D] = 1 + \text{str}[\delta M].$$

This result can also be proved by expanding formula (1.96) to the first order in δM .

Finally, for a generic change of variables $(\bar{z}, z) \rightarrow (\bar{\zeta}, \zeta)$ we have

$$d\bar{z}dz = d\bar{\zeta}d\zeta \text{sdet} \frac{\partial(\bar{z}, z)}{\partial(\bar{\zeta}, \zeta)}.$$

Again, we leave this formula without proof, but it is easy to derive the infinitesimal version that we need later. For $(\bar{z}, z) = (\bar{\zeta} + \delta\bar{\zeta}, \zeta + \delta\zeta)$ we have, to the first order,

$$d\bar{z}dz \sim d\bar{\zeta}d\zeta \left(1 + \text{str} \frac{\partial(\delta\bar{\zeta}, \delta\zeta)}{\partial(\bar{\zeta}, \zeta)} \right). \quad (1.99)$$

The minus sign inside the supertrace is due to the exponent -1 of the Jacobian determinant in (1.94).

The continuum limit is now straightforward. Consider for example free Dirac fermions, which have the action

$$S(\bar{\psi}, \psi) = \int d^4x \bar{\psi}(\not{\partial} + m)\psi. \quad (1.100)$$

Here $\not{\partial} = \gamma^\mu \partial_\mu$ and γ^μ are the γ matrices in Euclidean space, which satisfy $\{\gamma^\mu, \gamma^\nu\} = 2\delta^{\mu\nu}$ and $(\gamma^\mu)^\dagger = \gamma^\mu$. The Green function $G_F(x, y) = \langle \psi(x) \bar{\psi}(y) \rangle$ is the solution of the equation

$$(\not{\partial}_x + m)G_F(x, y) = \delta^{(4)}(x - y).$$

We have

$$G_F(x, y) = (-\not{\partial}_x + m)G_B(x, y) = \int \frac{d^4 p}{(2\pi)^4} \frac{-i\not{p} + m}{p^2 + m^2} e^{ip \cdot (x-y)}. \quad (1.101)$$

Define the generating functionals

$$Z(\bar{\xi}, \xi) = \int [d\bar{\psi} d\psi] \exp \left(-S(\bar{\psi}, \psi) + \int \bar{\xi} \psi + \int \bar{\psi} \xi \right) = e^{W[\bar{\xi}, \xi]},$$

where $\int \bar{\xi} \psi$ and $\int \bar{\psi} \xi$ stand for $\int d^4 x \bar{\xi}(x) \psi(x)$ and $\int d^4 x \bar{\psi}(x) \xi(x)$, respectively. Using (1.93) we find

$$W(\bar{\xi}, \xi) = \int d^4 x \bar{\xi}(x) G_F(x, y) \xi(y) d^4 y$$

plus an irrelevant constant.

Wick's theorem reads

$$\langle \chi_1 \cdots \chi_{2n} \rangle = \sum_P \varepsilon_P \langle \chi_{P(1)} \chi_{P(2)} \rangle \cdots \langle \chi_{P(2n-1)} \chi_{P(2n)} \rangle,$$

where χ_i can either be $\psi(x_i)$ or $\bar{\psi}(x_i)$, while ε_P is the signature of the permutation P . Precisely, ε_P is equal to 1 or -1 , depending on whether $\{P(1), P(2), \dots, P(2n)\}$ is obtained from $\{1, 2, \dots, 2n\}$ by means of an even or odd number of permutations of two nearby elements. The free correlation functions with an odd number of insertions still vanish.

The perturbative expansion around the free theory is defined by following the guideline of scalar fields. We can consider, for example, the four-fermion model

$$S_4(\bar{\psi}, \psi) = \int d^4 x \left(\bar{\psi} (\not{\partial} + m) \psi - \frac{\lambda}{4} (\bar{\psi} \psi)^2 \right) \quad (1.102)$$

The Feynman rules are

$$\begin{array}{c} \alpha \\ \hline \xrightarrow{p} \beta \end{array} = \left(\frac{1}{i\not{p} + m} \right)_{\alpha\beta} \quad \begin{array}{c} \alpha \quad \beta \\ \diagdown \quad \diagup \\ \diagup \quad \diagdown \\ \gamma \quad \delta \end{array} = \frac{\lambda}{2} (\delta^{\alpha\beta} \delta^{\gamma\delta} - \delta^{\alpha\delta} \delta^{\gamma\beta}) \quad (1.103)$$

where α, β , etc. are spinor indices. Observe that:

- 1) the vertex and the Lagrangian term have opposite signs;
- 2) each incoming line of the vertex is a ψ and each outgoing line is a $\bar{\psi}$;
- 3) fermion lines are drawn with arrows pointing from the right to the left; then their Lorentz indices are ordered from the left to the right;
- 4) if the Lagrangian term is ordered by putting each $\bar{\psi}$ to the left and each ψ to the right, the vertex is drawn by putting the $\bar{\psi}$ legs to the left, and the ψ legs to the right;
- 5) if we order the fields $\bar{\psi}$ (respectively, ψ) from the left to the right, the legs associated with them are ordered from the top to the bottom (resp., from the bottom to the top);
- 6) the exchange of two identical fermion lines flips the overall sign;
- 7) the vertices must include all the permutations of identical lines.

Point 1) is due to the minus sign that appears in front of the action in the exponential factor e^{-S} . Point 7) is why the factor 4 of $\lambda/4$ drops out.

Diagrams are constructed with the previous rules, plus the following one:

- 8) every fermion loop must be multiplied by a factor (-1) .

Finally, in evaluating the diagram it must be remembered that, because of 3),

- 9) fermion lines must be followed in the sense opposite to the arrow.

The minus sign in front of fermion loops is due to the Berezin integral. Consider for example

$$\int \prod_i d\bar{\theta}_i d\theta_i (\bar{\theta} V_1 \theta) (\bar{\theta} V_2 \theta) \exp \left(-\bar{\theta}^t M \theta \right),$$

where the V_i 's are the matrices that appear in the vertices, possibly depending on other fields. Using (1.95) we obtain

$$-\text{tr}[V_1 M^{-1} V_2 M^{-1}] + (-\text{tr}[V_1 M^{-1}])(-\text{tr}[V_2 M^{-1}]).$$

The first contribution corresponds to a one-loop diagram that contains both vertices, and is indeed multiplied by -1 . The second contribution is instead the product of two diagrams, each of which has one loop and contains a single vertex.

The result is easily generalized to diagrams with an arbitrary number of loops. The minus sign appears when we move the first $\bar{\theta}$ to the far right:

$$\prod_{i=1}^N (\bar{\theta} V_i \theta) \rightarrow -\text{tr}[V_1 \langle \theta \bar{\theta} \rangle V_2 \langle \theta \bar{\theta} \rangle V_3 \langle \theta \bar{\theta} \rangle V_4 \theta \cdots V_N \langle \theta \bar{\theta} \rangle].$$

The simplest example of scalar-fermion theory is the (massless) Yukawa model

$$S(\varphi, \psi) = \int d^4x \left(\frac{1}{2} (\partial_\mu \varphi)^2 + \bar{\psi} (\not{\partial} + g\varphi) \psi + \frac{\lambda}{4!} \varphi^4 \right) \quad (1.104)$$

with Feynman rules

$$\begin{array}{ccc} \text{---} k \text{---} & = \frac{1}{k^2} & \xrightarrow[\beta]{\alpha} p \text{---} = \left(\frac{1}{i\not{p}} \right)_{\alpha\beta} \\ \begin{array}{c} \diagup \text{---} \\ \diagdown \text{---} \end{array} & = -\lambda & \begin{array}{c} \diagup \text{---} \\ \diagdown \text{---} \end{array} = -g \end{array} \quad (1.105)$$

The rules to construct the diagrams are the same as before.

The functional Γ is defined as the Legendre transform

$$\Gamma(\Phi, \bar{\Psi}, \Psi) = -W(J, \bar{\xi}, \xi) + \int J\Phi + \int \bar{\xi}\Psi + \int \bar{\Psi}\xi,$$

where

$$\Phi = \frac{\delta W}{\delta J}, \quad \Psi = \frac{\delta_l W}{\delta \bar{\xi}}, \quad \bar{\Psi} = \frac{\delta_r W}{\delta \xi}.$$

All the arguments applied before to prove that W and Γ are the generating functionals of the connected and one-particle irreducible diagrams, respectively, can be repeated here with obvious modifications. Actually, we have already remarked that the derivation can be extended to the most general local perturbative quantum field theory. The Feynman rules for Z , W and Γ are the same, since the diagrams appear in each functional with the same coefficients (apart from the free two-point function of Γ).

Working on Γ makes the study of renormalization much simpler. For this reason, from now on we mostly concentrate on the irreducible diagrams.

The locality assumption, which is crucial for perturbative quantum field theory, has intriguing aspects. It requires that the action $S(\varphi, \bar{\psi}, \psi, V_\mu, \dots)$ be a local functional of the fields. It should be noted, however, that the action S does not contain the true interactions, which are encoded in the correlation functions. As we will see, the correlation functions are most of the times nonlocal. So, why should we require that the classical action be local? Even more, why should we require that there exist a classical action, and the theory be built on it? Why not investigate all the conceivable Γ functionals directly?

An attempt like this has been made, a few decades ago, but did not lead to substantial progress. The point is that if we do not have a sufficiently constrained starting point, such as a local (and renormalizable, as we will see) classical action, what we can say is so arbitrary that making predictions becomes almost impossible. We have to remember that when we explore the quantum world, we are not in the same situation as when we explore the classical world. We can make only sporadic experiments, and just collect data here and there. Instead, when we observe the world around us, the observed object emits a practically infinite number of photons, which are collected by our eyes, or instruments, in a finite amount of time, and each photon is like an individual experiment. Because of this, we do not worry so much about constraining the physical laws *a priori*, because the experimental observation is so powerful that it constrains them for us *a posteriori*. On the other hand, if we did not have a way to select classes of theories and interactions *a priori* in quantum field theory, we would not be able to get anywhere.

All this is fine, but prompts a dilemma: why should nature arrange itself so as to make us capable of investigating it? And isn't it a really twisted assumption to require that the observable interaction be built starting from a local "classical" action that may have no direct connection with the experimental observation of the classical world?

Well, such an involved principle is all that remains of the correspondence principle. We call S the "classical action" not because it has something to do with classical phenomena, but because it is the starting point of a process of *quantization*. Since we cannot have a direct intuition of the quantum world, the best we can hope is to be able to quant-ize a phantom of the classical

world. If we did not even have this chance, we would probably have no way to make progress in high-energy physics.

After properly formulating local, perturbative, renormalizable quantum field theory, we will be ready to explore more general quantum field theories, including the nonrenormalizable and the nonlocal ones. What we stress here is that if we make a too long conceptual jump at the beginning, we risk to plunge into the domain of absolute arbitrariness. We have to start from what is working for sure, or has most chances to work, and depart from that little by little.

Chapter 2

Renormalization

We have seen that the perturbative expansion produces ill defined integrals, such as (1.49). This is the first serious problem of the “creative approximation” we undertook in the previous chapter. Despite what may appear at first sight, it is not so difficult to overcome this difficulty. It is useful to compare this situation to the situation of a mathematician in front of an integral over the real line,

$$\int_{-\infty}^{+\infty} dx f(x). \quad (2.1)$$

Written like this, this expression has no intrinsic meaning, and needs to be defined. Riemann gives us a natural attempt to define it by means of the improper integral

$$\lim_{\Lambda \rightarrow \infty} \int_{-\Lambda}^{+\Lambda} dx f(x). \quad (2.2)$$

Precisely, a “cutoff” Λ is inserted, to replace the original integral (2.1) into a definite one. After calculating the definite integral, the limit $\Lambda \rightarrow \infty$ is studied. If the limit exists, the integral is convergent. If the limit does not exist, the integral is divergent.

In quantum field theory we do not have to define *one* integral, but a *theory*, which contains an infinite number of integrals, one for each diagram. Different diagrams may be related to one another by certain identities. Physical quantities involve, in general, sums, products and convolutions of integrals. If a single integral does not converge, the reason may simply be that we have isolated that integral from the rest of the theory in an inconvenient

way. This happens, for example, when the “divergence” disappears by changing variables (fields, spacetime coordinates or momenta, couplings and any other parameter of the theory), i.e. by performing all sorts of operations that normally do not change the physics. When that is the case, the divergence is not a problem, but just a blunder due to an unfortunate parametrization of the theory.

Instead of taking the limit $\Lambda \rightarrow \infty$ integral by integral, right after inserting the cutoff, we postpone this operation till the physical quantities have been fully worked out. In the meantime, we take the liberty to perform a number of “almost innocuous” operations, which means move the Λ divergences around, from one quantity to another, by performing changes of field variables and reparametrizations. Before claiming that our theory is ill defined, we want to take full advantage of the freedom we have. It is not correct to view the single integral as an improper integral: it is correct to view the whole theory as an “improper theory”.

Thus, we have to answer the following question: is there a rearrangement based on reparametrizations and field redefinitions after which the theory admits the $\Lambda \rightarrow \infty$ limit?

The insertion of a cutoff is called *regularization*, the rearrangement of the Λ divergences that allows us to achieve our goal is called *renormalization*. Of course, we will have to prove that the physical results do not depend on the way we regularize and renormalize our theory.

The cutoff is a useful tool to classify the divergences. In principle, it may not be strictly necessary to introduce it, and in the literature there exist several regularization-independent approaches that do not make explicit use of a cutoff. On the other hand, working with a cutoff is very convenient, because it helps us keep track of what we do when we move the divergences around. The goal of the rearrangement is to identify the right places of the divergences, so that, after moving those “infinities” to their right destinations, the limit $\Lambda \rightarrow \infty$ makes sense *in all the physical quantities*, but not necessarily in the single integrals and the quantities that are physically meaningless. If this program works, we obtain a consistent (perturbative) definition of the local quantum field theory.

Definition 1 *A theory is called convergent if, possibly after a reparametrization, all the physical quantities admit the limit $\Lambda \rightarrow \infty$. Otherwise it is called divergent.*

The definition of convergent theory is not equivalent to the definition of “renormalizable” theory. We will appreciate the difference later.

The cutoff (2.2) is the simplest and most intuitive way to smooth out the singularities. It amounts to state that the domain of integration is bounded to momenta that have a modulus smaller than Λ . Clearly, this trick makes every integral convergent at finite Λ . For example, the two-point function $G_B(x, y)$ is divergent at coinciding points. At finite Λ we find

$$\begin{aligned} G_B(x, x) &= \int_{|p| \leq \Lambda} \frac{d^4 p}{(2\pi)^4} \frac{1}{p^2 + m^2} = \frac{1}{16\pi^2} \left[\Lambda^2 - m^2 \ln \left(1 + \frac{\Lambda^2}{m^2} \right) \right] \\ &= \frac{1}{16\pi^2} \left[\Lambda^2 - m^2 \ln \frac{\Lambda^2}{m^2} + m^2 \mathcal{O} \left(\frac{m^2}{\Lambda^2} \right) \right]. \end{aligned} \quad (2.3)$$

When Λ is sent to infinity, we have a quadratic divergence, which is the term proportional to Λ^2 , plus a logarithmic divergence, which is the term proportional to $\ln \Lambda$, plus finite contributions.

Divergences occur at large momenta, or, equivalently, at coinciding points. They are basically due to the locality of our theories. If we were satisfied with nonlocal, rather than pointlike, interactions, then we could easily construct theories with no divergences. However, that is not our purpose, because, as we have remarked at the end of the previous chapter, nonlocalities may open that door to a huge arbitrariness. It is better to first deal with divergences in local theories, then investigate nonlocal theories. Besides, we have already said that the divergences of isolated integrals are not the true problem: it would be a mistake to throw away theories just because they look divergent at first sight.

Definition 2 *Given a theory having Feynman rules F , a regularization is any deformation F_Λ of the Feynman rules that gives sense to all the individual integrals generated by the perturbative expansion, and is such that F_Λ gives back F when the deformation is switched off.*

We stress that the regularization does not need to be physical, because the cutoff must be eventually removed. Actually, the most common regularization techniques are unphysical, in the sense that regularized theories are not physically acceptable as quantum field theories *per se*, because they violate some physical principle. The cutoff is an example of unphysical regularization, since it violates unitarity. Indeed, it excludes the contributions

of high frequencies from the integrals, while unitarity says (loosely speaking) that the set of particles that circulate in the loop must coincide with the set of ingoing and outgoing particles.

On the other hand, the violation of locality does not sound like the violation of a physical principle, so a theory regularized in a nonlocal way might well be physical in its own right. Yet, we stress again that the intrinsic arbitrariness of nonlocal theories makes us postpone their investigation to the very end. For the moment, the problems we find in local theories are rather welcome, because they give us hope to select the set of theories that are admitted a priori. If the selection is powerful enough we might be able to make predictions that can be successfully compared with experiments.

It may be objected that inserting a cutoff *à la* Riemann may not be the smartest choice. The Lebesgue integral is known to be an excellent generalization of the Riemann integral, and supersedes it in many ways. So, the natural question is: can we regularize quantum field theory *à la* Lebesgue? Unfortunately, nobody has pursued this direction, so far. Nevertheless, there exists a regularization technique that very well fits the needs of perturbative quantum field theory. This is the dimensional regularization.

2.1 Dimensional regularization

The *dimensional regularization* is a regularization technique based on the continuation of the dimension of spacetime to complex values. We recall that, as awkward as this concept may sound at first, we just need to provide a consistent formal construction, and equip it with a set of axioms that allow us to make manipulations and get back to real numbers in the physical predictions.

Consider an integral \mathcal{I}_4 in four dimensions, in momentum space. Call the integrated momentum p and the external momenta k . Assume that the integrand is Lorentz invariant in Minkowski spacetime, and a rational function. To dodge a number of nuisances that are not important for the present discussion, we still choose to work in the Euclidean framework. There, the integrand is invariant under rotations, and can be expressed as a function f of p^2 and the scalar products $p \cdot k$:

$$\mathcal{I}_4(k) = \int \frac{d^4 p}{(2\pi)^4} f(p^2, p \cdot k).$$

An *analytic integral* $\mathcal{I}_D(k)$ in complex D dimensions can be associated with \mathcal{I}_4 as follows. Replace the four-dimensional integration measure d^4p with a formal D -dimensional measure d^Dp , and include a $(2\pi)^D$ in the denominator for convenience, instead of $(2\pi)^4$. Replace p_μ and k_μ with formal D -dimensional vectors inside the integrand. This gives

$$\mathcal{I}_D(k) = \int \frac{d^Dp}{(2\pi)^D} f(p^2, p \cdot k). \quad (2.4)$$

We want to define the analytic integral in D dimensions so that it coincides with the ordinary integral $\mathcal{I}_d(k)$ when D takes integer values d and $\mathcal{I}_d(k)$ is convergent. When $\mathcal{I}_d(k)$ is not convergent, we want to use $\mathcal{I}_D(k)$ to classify its divergence.

To achieve this goal, we start by writing the analytic integral $\mathcal{I}_D(k)$ in spherical coordinates. The measure reads

$$\begin{aligned} \int d^Dp &= \int_0^\infty p^{D-1} dp \times \\ &\times \int_0^{2\pi} d\theta_1 \int_0^\pi d\theta_2 \sin \theta_2 \cdots \int_0^\pi d\theta_{D-1} \sin^{D-2} \theta_{D-1}, \end{aligned}$$

any time D is integer. When L is integer and greater than one, we also have

$$\int_0^{2\pi} d\theta_1 \int_0^\pi d\theta_2 \sin \theta_2 \cdots \int_0^\pi d\theta_{L-1} \sin^{L-2} \theta_{L-1} = \frac{2\pi^{L/2}}{\Gamma(\frac{L}{2})},$$

which is the total solid angle in L dimensions.

Since the external momenta k are finitely many, because a Feynman diagram has a finite number of external legs, the integrand of (2.4) depends on finitely many angles $\theta_{D-L}, \dots, \theta_{D-1}$. The number D is still unspecified and for the time being we can imagine that it is integer and sufficiently large, in any case larger than L . Then we can write

$$\begin{aligned} \mathcal{I}_D(k) &= \frac{1}{2^{D-1} \pi^{(D+L)/2} \Gamma(\frac{D-L}{2})} \int_0^\infty dp \times \\ &\times \int_0^\pi d\theta_{D-L} \int_0^\pi d\theta_2 \cdots \int_0^\pi d\theta_{D-1} p^{D-1} \bar{f}(p, \theta_1 \cdots \theta_L, D). \end{aligned} \quad (2.5)$$

The function \bar{f} also includes the factors $\sin^{i-1} \theta_i$, $i = D-L, \dots, D-1$.

Now, the expression on the right-hand side of (2.5) is meaningful for generic complex D . Assume that there is an open domain \mathcal{D} in the complex

plane where the integral $\mathcal{I}_D(k)$, written as in (2.5), is well-defined. Evaluate $\mathcal{I}_D(k)$ in \mathcal{D} . Then, analytically continue the function $\mathcal{I}_D(k)$ from \mathcal{D} to the rest of the complex plane. The value of this function at $D = 4$, if it exists, is the physical value of the integral $\mathcal{I}_4(k)$. If it does not exist, the function $\mathcal{I}_D(k)$ has poles around $D = 4$. Such poles classify its divergences.

For example,

$$\mathcal{I}_D(m) \equiv \int \frac{d^D p}{(2\pi)^D} \frac{1}{p^2 + m^2} = \frac{1}{2^{D-1} \pi^{D/2} \Gamma(\frac{D}{2})} \int_0^\infty dp \frac{p^{D-1}}{p^2 + m^2}. \quad (2.6)$$

The integral is well-defined in the strip $0 < \text{Re } D < 2$. The analytic continuation gives (see Appendix A, formula (A.5))

$$\frac{\Gamma(1 - \frac{D}{2}) m^{D-2}}{(4\pi)^{D/2}} = \frac{1}{16\pi^2} \left[-\frac{2m^2}{\varepsilon} + m^2 \left(\ln \frac{m^2}{4\pi} - 1 + \gamma_E \right) + \mathcal{O}(\varepsilon) \right], \quad (2.7)$$

where $\gamma_E = 0.5772\dots$ is the Euler-Mascheroni constant. The right-hand side of formula (2.7) is the expansion around four dimensions, having written $D = 4 - \varepsilon$ and used formula (A.8).

Observe that the term $m^2 \ln m^2$ coincides with the one of (2.3). The logarithmic divergences of (2.3) and (2.7) coincide after identifying $\ln \Lambda$ with $1/\varepsilon$. Indeed, for large momenta we have

$$\int_{|p| \geq \delta} \frac{d^D p}{(2\pi)^D (p^2)^2} \sim \frac{1}{8\pi^2 \varepsilon} + \text{finite}, \quad \int_{\delta \leq |p| \leq \Lambda} \frac{d^4 p}{(2\pi)^4 (p^2)^2} \sim \frac{1}{8\pi^2} \ln \Lambda,$$

where δ is an infrared cutoff. The other contributions to (2.3) and (2.7) differ from each other. In particular, (2.7) contains no analogue of the quadratic divergence Λ^2 . Differences and similarities will become clearer later.

What happens when the integral, expressed in the form (2.5) does not admit a domain of convergence \mathcal{D} ? Or when it admits more disconnected domains of convergence?

First, observe that the Feynman rules of a local quantum field theory can only give rational integrands. Then, if the domain of convergence \mathcal{D} exists, it is always unique (a strip $X < \text{Re } D < Y$), which ensures that the analytic continuation is also unique, as well as the value of the integral in D dimensions. The situation where an integral admits two disconnected convergence domains cannot occur.

If an integral does not admit a convergence domain, assume that we can decompose the integrand f into a finite sum of integrands f_i , such that each of

them admits its own domain of convergence \mathcal{D}_i . Then we define the integral of f as the sum of the integrals of each f_i . For example, the integrand $f \equiv 1$ does not admit a domain of convergence. However, writing

$$1 = \frac{p^2 + m^2}{p^2 + m^2} = f_1 + f_2, \quad f_1 = \frac{p^2}{p^2 + m^2}, \quad f_2 = \frac{m^2}{p^2 + m^2},$$

we see that f_1 and f_2 admit the domains of convergence $-2 < \text{Re } D < 0$ and $0 < \text{Re } D < 2$, respectively. We thus find

$$\int \frac{d^D p}{(2\pi)^D} f_1 = \frac{D m^D \Gamma(-\frac{D}{2})}{2^{D+1} \pi^{D/2}}, \quad \int \frac{d^D p}{(2\pi)^D} f_2 = m^2 \mathcal{I}_D(m).$$

Summing the two contributions, we discover that the analytic integral of one is actually zero. The same integral, treated with the cutoff method, behaves like Λ^4 . We learn that the dimensional regularization kills every powerlike divergence. It is sensitive only to the logarithmic divergences, which manifest themselves as poles in $1/\varepsilon$.

With exactly the same procedure we can calculate the analytic integral of $(p^2)^\alpha$, for every complex α : we find again 0. More generally, let $f(p)$ be a rational function of p . Let α_{IR} and α_{UV} denote the exponents such that

$$f(p) \sim (p^2)^{\alpha_{\text{IR}}}, \quad f(p) \sim (p^2)^{\alpha_{\text{UV}}},$$

for $p \rightarrow 0$ and $p \rightarrow \infty$, respectively. Decompose the integrand as

$$f(p) \left(\frac{p^2 + m^2}{p^2 + m^2} \right)^n = \sum_{k=0}^n \binom{n}{k} (m^2)^{n-k} \frac{f(p)(p^2)^k}{(p^2 + m^2)^n}.$$

The integral of the k -th term of the sum is convergent in the strip $-2\alpha_{\text{IR}} - 2k < \text{Re } D < 2n - 2\alpha_{\text{UV}} - 2k$, which is non-trivial if its width $2n - 2\alpha_{\text{UV}} + 2\alpha_{\text{IR}}$ is strictly positive. Note that the width is k independent. Thus, if we choose n sufficiently large, in particular larger than $\alpha_{\text{UV}} - \alpha_{\text{IR}}$, all the terms of the sum can be integrated.

Concluding, we can always decompose the analytic integral of a rational function as a finite sum of integrals admitting nontrivial convergence domains. The construction easily extends to multiple integrals. Since a local quantum field theory can only generate rational integrands, our arguments prove that the dimensional-regularization technique is able to define every integral we need.

It remains to prove that our definition is consistent. We do not provide a complete proof here, but collect the basic arguments and mention the key properties of the integral.

First, the analytic integral is linear, and invariant under translations and rotations. In particular, the result does not depend on the center of the polar coordinates used to write (2.5). Moreover, the usual formulas for the multiple integration and the change of variables hold.

The rules of multiple integration deserve some comment. It is always safe to split an analytic integral in D dimensions as the sequence of two analytic integrals in D_1 and D_2 dimensions, with $D = D_1 + D_2$, which are defined as explained above:

$$\int \frac{d^D p}{(2\pi)^D} = \int \frac{d^{D_1} p_1}{(2\pi)^{D_1}} \int \frac{d^{D_2} p_2}{(2\pi)^{D_2}}.$$

Sometimes, however, it is convenient to split the integral as an analytic integral followed by an ordinary integral. For example,

$$\int \frac{dp_1}{2\pi} \int \frac{d^{D-1} p_2}{(2\pi)^{D-1}}, \quad \int \frac{d^4 p_1}{(2\pi)^2} \int \frac{d^{-\varepsilon} p_2}{(2\pi)^{-\varepsilon}}, \quad (2.8)$$

and so on. This kind of decomposition also works. However, the outside integral is still to be meant in the analytic sense. Precisely, after evaluating the inside integral, we obtain the ordinary integral of a function f that depends on D . That integral must be evaluated in a domain \mathcal{D} where it converges, and analytically continued to the rest of the complex plane, as explained above. If a domain \mathcal{D} does not exist, it must be written as a finite linear combination of ordinary integrals that separately admit domains of convergence \mathcal{D}_i . For example, if we use the second split of (2.8) on $\mathcal{I}_D(m)$, we can represent it as a four-dimensional integral:

$$\mathcal{I}_D(m) = \frac{\Gamma(1 + \varepsilon/2)}{(4\pi)^{-\varepsilon/2}} \int \frac{d^4 p_1}{(2\pi)^2} \frac{1}{(p_1^2 + m^2)^{1+\varepsilon/2}}.$$

Neglecting the prefactor, which tends to 1 when ε tends to zero, this formula can be viewed as an alternative regularization of the integral. It does not change the integration per se and does not introduce a cutoff for the large momenta. Instead, it replaces the propagator by

$$\frac{1}{(p^2 + m^2)^{1+\varepsilon/2}},$$

where ε is a complex number. The integrals have to be calculated in a complex domain of ε -values where they converge, and then analytically continued to the rest of the complex plane. In the literature, this procedure is known as *analytic regularization*. The good feature of the analytic regularization is that it deals with ordinary integrals all the time, so its consistency is easier to prove. We anticipate that, however, it breaks gauge invariance, while the dimensional regularization manifestly preserves it. Using the analytic regularization (or the cutoff one), gauge invariance has to be recovered by hand, which is possible, but requires a lot of effort. The dimensional regularization is a sort of rationalized analytic regularization, which knows how to rearrange itself so as to preserve gauge invariance at no cost.

Finally, it is normally not safe to split an analytic integral as an ordinary integral followed by an analytic integral, e.g.

$$\int \frac{d^{-\varepsilon}p_1}{(2\pi)^{-\varepsilon}} \int \frac{d^4p_2}{(2\pi)^2}$$

because the ordinary integral might not converge. Check it on $\mathcal{I}_D(m)$.

2.1.1 Limits and other operations in D dimensions

Limits can be taken applying similar steps. Consider a function $f(D, x)$. Its limit $f(D, x_0)$ for $x \rightarrow x_0$ is defined by applying the following two rules:

a) search for an open set \mathcal{D} of the complex plane where the limit exists, calculate it there, and analytically continue the result to the complex plane;

b) if $f(D, x)$ admits no such \mathcal{D} , search for a decomposition of $f(D, x)$ into a finite sum $\sum_i f_i(D, x)$, such that each $f_i(D, x)$ admits a complex domain \mathcal{D}_i where the limit exists, proceed as in point *a)* for each $f_i(D, x)$ and sum the analytic continuations $f_i(D, x_0)$.

As an example, consider the integral

$$\int \frac{d^D p}{(2\pi)^D} \frac{\Lambda^2}{(p^2 + m^2)(p^2 + m^2 + \Lambda^2)}. \quad (2.9)$$

It can be evaluated by means of formula (A.2) of Appendix A, which allows us to express it as

$$\int_0^1 dx \int \frac{d^D p}{(2\pi)^D} \frac{\Lambda^2}{(p^2 + m^2 + x\Lambda^2)^2}.$$

Then formula (A.4) gives

$$\begin{aligned} & \Lambda^2 \frac{\Gamma(2 - \frac{D}{2})}{(4\pi)^{D/2}} \int_0^1 dx (m^2 + x\Lambda^2)^{D/2-2} \\ &= \frac{\Gamma(1 - \frac{D}{2}) m^{D-2}}{(4\pi)^{D/2}} \left[1 - \left(1 + \frac{\Lambda^2}{m^2} \right)^{D/2-1} \right]. \end{aligned} \quad (2.10)$$

If we take Λ to infinity in the integrand of (2.9) we get (2.7). Now, consider the final result (2.10). It admits a regular limit only in the domain $\text{Re}D < 2$. The analytic continuation of the limit in such a domain gives again (2.7).

To interchange derivatives and integrals, derivatives and limits, and perform all sorts of similar operations, we must follow the same guideline, namely *a)* decompose the function f into a finite sum of functions f_i each of which admits a domain \mathcal{D}_i of the complex plane where the operation can be performed ordinarily, once the integral is expressed in the form (2.5), *b)* analytically continue each result to the complex plane, and *c)* sum the analytic continuations.

2.1.2 Functional integration measure

Now we prove an important property that is going to be useful in many contexts. We say that a function of the fields and their derivatives, evaluated at the same spacetime point, is ultralocal if it depends polynomially on the derivatives of the fields. It does not need to be polynomial in the fields themselves. We prove that

Theorem 1 *In dimensional regularization, the functional integration measure is invariant under every ultralocal change of field variables.*

Proof. Let φ^i denote the fields and $\varphi^i \rightarrow \varphi^{i'}$ the change of field variables. If the field redefinition is ultralocal, then there exists a finite number of local functions $F_{ij}^{\mu_1 \dots \mu_n}$ such that

$$\frac{\delta \varphi^{i'}(x)}{\delta \varphi^j(y)} = \sum_{n=0}^N F_{ij}^{\mu_1 \dots \mu_n}(\varphi(x)) \partial_{\mu_1} \dots \partial_{\mu_n} \delta^{(D)}(x - y) \quad (2.11)$$

and the Jacobian determinant can be written as

$$\begin{aligned}\mathcal{J} &= \det \frac{\delta \varphi^{i'}(x)}{\delta \varphi^j(y)} = \exp \left(\text{tr} \frac{\delta \varphi^{i'}(x)}{\delta \varphi^j(y)} \right) = \exp \left(\int d^D x \frac{\delta \varphi^{i'}(x)}{\delta \varphi^i(x)} \right) \\ &= \exp \left(\sum_{n=0}^N \partial_{\mu_1} \cdots \partial_{\mu_n} \delta^{(D)}(0) \int d^D x F_{ij}^{\mu_1 \cdots \mu_n}(\varphi(x)) \right)\end{aligned}$$

Because of (2.11), the exponent is a finite sum of local functionals multiplied by $\delta^{(D)}(0)$ or derivatives of $\delta^{(D)}(0)$. Such expressions vanish using the dimensional regularization, because in momentum space they read

$$\partial_{\mu_1} \cdots \partial_{\mu_n} \delta^{(D)}(0) = i^n \int \frac{d^D p}{(2\pi)^D} p_{\mu_1} \cdots p_{\mu_n}. \quad (2.12)$$

Recalling that the analytic integral is invariant under rotations, we obtain zero when n is odd, but also zero when n is even. Indeed,

$$\int \frac{d^D p}{(2\pi)^D} p_{\mu_1} \cdots p_{\mu_{2k}} \propto (\delta_{\mu_1 \mu_2} \cdots \delta_{\mu_{2k-1} \mu_{2k}} + \text{perms.}) \int \frac{d^D p}{(2\pi)^D} (p^2)^k = 0. \quad (2.13)$$

□

The theorem we just proved is very general. It also holds when the change of variables is not polynomial in the derivatives of the fields, but can be treated as a perturbative series of ultralocal terms. Moreover, it holds for all types of fields: scalars, fermions, vectors, tensors, as well as fields of higher spins. To include fields of different statistics in the proof, it is sufficient to replace the determinant by the superdeterminant and the trace by the supertrace.

We say that a functional is perturbatively local if it can be perturbatively expanded as a series of terms that are polynomial in the fields and their derivatives, evaluated at the same spacetime point. It is perturbatively ultralocal if it can be perturbatively expanded in a series of terms that are ultralocal. In some situations we may just use the terms “local” and “ultralocal” in this extended sense.

2.1.3 Dimensional regularization for vectors and fermions

In the dimensional regularization the coordinates x^μ , the momenta p_μ , the Kronecker tensor $\delta_{\mu\nu}$, and so on, have to be viewed as purely formal objects.

We need to give a consistent set of operations to manipulate such objects, so that the four dimensional results are retrieved when $D = 4$. Similarly, vector fields A_μ , the gamma matrices γ_μ and spinors ψ^α also have to be considered as formal objects. In particular, the gamma “matrices” should not be viewed as true matrices, although we keep calling them with their usual name.

We define the D -dimensional Dirac algebra as a set of formal objects γ_μ that are equipped with a formal trace operation and satisfy the following axioms:

$$\begin{aligned} \{\gamma_\mu, \gamma_\nu\} &= 2\delta_{\mu\nu}\mathbb{1}, & \gamma_\mu^\dagger &= \gamma_\mu, & \text{tr}[\gamma_{\mu_1} \cdots \gamma_{\mu_{2n+1}}] &= 0, \\ \text{tr}[AB] &= \text{tr}[BA], & \text{tr}[\mathbb{1}] &= f(D), & f(4) &= 4. \end{aligned} \quad (2.14)$$

In particular, the formal trace is cyclic and vanishes on an odd product of gamma matrices. Using the formal Dirac algebra, that is to say the first axiom of (2.14), we can reduce every trace to the trace of the identity, which we call $f(D)$. The function $f(D)$ must be equal to 4 in four dimensions, but is otherwise arbitrary.

Specifically, the axioms (2.14) imply

$$\text{tr}[\gamma_{\mu_1} \cdots \gamma_{\mu_{2n}}] = \sum_{i=2}^{2n} (-1)^i \delta_{\mu_1 \mu_i} \text{tr}[\gamma_{\mu_2} \cdots \hat{\gamma}_{\mu_i} \cdots \gamma_{\mu_{2n}}], \quad (2.15)$$

where $\hat{\gamma}_{\mu_i}$ means that the matrix γ_{μ_i} is dropped. The proof is identical to the one in four dimensions. In particular,

$$\begin{aligned} \text{tr}[\gamma_\mu \gamma_\nu] &= f(D) \delta_{\mu\nu}, \\ \text{tr}[\gamma_\mu \gamma_\nu \gamma_\rho \gamma_\sigma] &= f(D) (\delta_{\mu\nu} \delta_{\rho\sigma} - \delta_{\mu\rho} \delta_{\nu\sigma} + \delta_{\mu\sigma} \delta_{\nu\rho}). \end{aligned}$$

We also have the identities

$$\gamma_\mu \gamma_\mu = D\mathbb{1}, \quad \gamma_\mu \gamma_\rho \gamma_\mu = (2 - D)\gamma_\rho.$$

It seems that in D dimensions everything proceeds smoothly, with minor modifications with respect to the usual formulas, but it is actually not true. In four dimensions we can define also a matrix γ_5 that satisfies $\{\gamma_\mu, \gamma_5\} = 0$. A matrix with such properties does not exist in complex D dimensions. Another object that cannot be extended to D dimensions is the tensor $\varepsilon_{\mu\nu\rho\sigma}$, because it would have a complex number of indices! For the moment we

ignore these problems and limit ourselves to nonchiral theories, where γ_5 and $\varepsilon_{\mu\nu\rho\sigma}$ do not appear in the Lagrangian and the Feynman diagrams. Later, we will see that the problem is related to the appearance of an important “anomaly”. Another fact that is worth mentioning is that in odd dimensions it can be inconsistent to assume that the trace of an odd product of gamma matrices vanishes. For example, in three dimensions the trace $\text{tr}[\sigma_i \sigma_j \sigma_k]$, where σ_i are the Pauli matrices, is not zero, but proportional to the tensor ε_{ijk} . There exist modified versions of the dimensional regularization that bypass these difficulties. We will introduce them when needed.

The dimensionally regularized versions of the models studied so far have formally identical Feynman rules (1.50), (1.103) and (1.105). However, for $D \neq 4$ the couplings are dimensionful even when they are dimensionless in $D = 4$. It is convenient to redefine them in a dimensionless way, by isolating suitable powers of an energy scale μ . For example, the Lagrangians (1.38) and (1.104) become

$$S(\varphi) = \int d^D x \left(\frac{1}{2} (\partial_\mu \varphi)^2 + \frac{m^2}{2} \varphi^2 + \frac{\lambda \mu^\varepsilon}{4!} \varphi^4 \right) \quad (2.16)$$

and

$$S(\varphi, \psi) = \int d^D x \left(\frac{1}{2} (\partial_\mu \varphi)^2 + \bar{\psi} \left(\not{\partial} + g \mu^{\varepsilon/2} \varphi \right) \psi + \frac{\lambda \mu^\varepsilon}{4!} \varphi^4 \right), \quad (2.17)$$

respectively. In the new parametrization, both g and λ are dimensionless in arbitrary D , and the Feynman rules are (1.50) and (1.105) with the replacements $g \rightarrow g \mu^{\varepsilon/2}$ and $\lambda \rightarrow \lambda \mu^\varepsilon$.

2.2 Divergences and counterterms

Now that we know that each diagram is associated with a well regularized integral, we can study the general properties of the diagrammatics.

Consider a diagram G with V vertices, E external legs and I internal legs. Assign an independent momentum to each leg, internal and external. In total, this gives $I + E$ momenta. Once we impose the momentum conservation at each vertex, we remain with $I + E - V$ independent momenta. Now, observe that the external legs contain $E - 1$ independent momenta, because the E th momentum is determined by the global momentum conservation. Therefore,

the diagram G contains $I + E - V - (E - 1) = I - V + 1 = L$ independent internal momenta, and the integral associated with G is performed over them. We call L the number of *loops* of the diagram. It satisfies topological formula

$$L - I + V = 1, \quad (2.18)$$

which holds for every diagram, in every theory. It is called topological, because it coincides with Euler's formula for simple polyhedra, namely

$$v - e + f = 2$$

where v is the number of vertices, e is the number of edges and f is the number of faces of the polyhedron. The correspondence with (2.18) is $v = V$, $e = I$ and $f = L + 1$. Indeed, dropping the external legs and adding the “loop at infinity”, which is the $(L + 1)$ -th face, a graph becomes a generalized polyhedron, namely a polyhedron whose faces are not necessarily flat (which still satisfies Euler's formula).

Another very general fact is that the expansion in the number of loops coincides with the expansion in powers of \hbar . Although we have set $\hbar = 1$ so far, we can easily restore the \hbar dependence by writing the generating functionals $Z(J)$ and $W(J)$ as

$$Z(J) = \int [d\varphi] \exp \left(-\frac{1}{\hbar} S(\varphi) + \int J\varphi \right) = \exp \left(\frac{1}{\hbar} W(J) \right),$$

while $\Gamma(\Phi)$ is defined as before. In the new Feynman rules a propagator gets a factor \hbar and a vertex gets a factor $1/\hbar$. Therefore, each diagram is multiplied by a factor

$$\hbar^{I-V} = \frac{\hbar^L}{\hbar},$$

having used (2.18). Diagrams contribute to Z in the usual way. If they are connected they contribute to W/\hbar , because $Z = \exp(W/\hbar)$. If they are irreducible they contribute to $-\Gamma/\hbar$. We thus see that the L -loop contributions to W and Γ are multiplied by \hbar^L .

Consider the “ φ_d^N theory”, which is the d -dimensional scalar field theory with interaction φ^N , which has the action

$$S(\varphi) = \int d^d x \left(\frac{1}{2} (\partial_\mu \varphi)^2 + \frac{m^2}{2} \varphi^2 + \lambda \frac{\varphi^N}{N!} \right).$$

For the moment we do not need to continue the physical dimension to complex values. Let $[\mathcal{O}]$ denote the dimension of an object \mathcal{O} in units of mass. Coordinates have dimension -1 , while momenta have dimension 1 . Since the action is dimensionless, the Lagrangian must have dimension d . From the kinetic term, or the mass term, we can read the dimension of φ . Then, we can read the dimension of λ from the vertex. We find

$$[x] = -1, \quad [\partial] = 1, \quad [\varphi] = \frac{d}{2} - 1, \quad [\lambda] = N \left(1 - \frac{d}{2}\right) + d. \quad (2.19)$$

Consider again a diagram G with V vertices, E external legs and I internal legs. Since N legs are attached to each vertex, we have NV legs in total. Of these, E exit the diagram and $2I$, connected in pairs, build the internal legs, each of which is attached to two vertices. Therefore, we have the identity

$$E + 2I = NV. \quad (2.20)$$

Calling the loop momenta p_i , the integral associated with the Feynman diagram has the form

$$\mathcal{I}_G(k, m) = \int \prod_{i=1}^L \frac{d^d p_i}{(2\pi)^d} \prod_{i=1}^L \frac{1}{(p_i + k_i)^2 + m^2} \prod_{j=1}^{V-1} \frac{1}{(\Delta p_j + k'_j)^2 + m^2}, \quad (2.21)$$

where k and k' are linear combinations of external momenta, with coefficients ± 1 . Moreover, the Δp_j s are nontrivial linear combinations of the integrated momenta p with coefficients ± 1 . We have used (2.18) to organize the integrand in the way shown.

We need to check the convergence of the integral in all regions of integration. Since we are in the Euclidean framework, the integral is regular for finite values of the momenta p . We just need to study its behavior when the momenta tend to infinity in all possible ways. It is sufficient to consider the following situations: *i*) let the momenta of all internal legs tend to infinity with the same velocity, or *ii*) keep the momenta of some internal legs fixed. A singularity that occurs in case *i*) is called ultraviolet overall divergence. A singularity that occurs in case *ii*) is called ultraviolet subdivergence. Since in this book we treat only ultraviolet divergences, we omit to specify it from now on.

Overall divergences are studied by rescaling the integrated momenta p with a factor λ ,

$$p_i \rightarrow \lambda p_i, \quad (2.22)$$

and then sending λ to infinity. Subdivergences are studied by performing the rescaling (2.22) with the constraint that the momenta of some internal legs are kept fixed. It can be shown that once the divergences due to the two types of limits *i*) and *ii*) are cured, the integral becomes convergent. In other words, all other ways to send the momenta to infinity are then automatically cured, because they amount to some combinations of the limits *i*) and *ii*). For example, if some momenta p'_i are rescaled by a factor λ , and the other momenta p''_i by a factor λ^2 , then sending λ to infinity is like first rescaling the p''_i s at fixed p'_i s, then rescaling the p'_i s.

The subdivergences are the overall divergences of a suitable subdiagram G_{sub} of G . Precisely, G_{sub} is the irreducible part of the diagram obtained by cutting the G internal legs whose momenta are kept fixed. Clearly, if G is irreducible, as we are going to assume from now on, the subdiagrams G_{sub} have fewer loops, because when we cut one or more G internal lines we necessarily break some loop. Moreover, since the perturbative expansion, namely the expansion in powers of \hbar , coincides with the loop expansion, the divergences can be subtracted algorithmically. In other words, when we deal with an L -loop diagram, we can assume to be already equipped with the set of counterterms that take care of its subdiagrams G_{sub} . For the moment, we ignore the subdivergences and concentrate on the overall divergences.

Let us compute the dimension of $\mathcal{I}_G(k, m)$. The momentum integration measure $d^d p$ has dimension d , while the propagators have dimension -2 . Using (2.18) and (2.20), we have

$$[\mathcal{I}_G(k, m)] = Ld - 2I = V \left[N \left(\frac{d}{2} - 1 \right) - d \right] - E \left(\frac{d}{2} - 1 \right) + d. \quad (2.23)$$

Make the rescaling (2.22) and consider the behavior of the integral when λ tends to infinity. Let $\omega(\mathcal{I}_G)$ denote the power of λ in this limit. In our present case, given the form of the integral (2.21), we have $\omega(\mathcal{I}_G) = [\mathcal{I}_G]$. However, if some external momentum or a mass factorizes, we may have $\omega(\mathcal{I}_G) < [\mathcal{I}_G]$. In general, we have the inequality $\omega(\mathcal{I}_G) \leq [\mathcal{I}_G]$. We call $\omega(\mathcal{I}_G)$ the *degree of divergence* of the diagram G . If $\omega(\mathcal{I}_G) < 0$ and there are no subdivergences, then the integral is ultraviolet convergent, because it is convergent in all the regions of integration. Instead, if $\omega(\mathcal{I}_G) \geq 0$, or $\omega(\mathcal{I}_G) < 0$ but there are subdivergences, the diagram is potentially ultraviolet divergent.

To begin with, consider a one-loop diagram. Since it has no subdiagrams, there can be only an overall divergence, but no subdivergences. Differentiate

the diagram one time with respect to an external momentum k or a mass m , and observe that

$$\begin{aligned}\frac{\partial}{\partial k_\mu} \frac{1}{(p+k)^2 + m^2} &= -\frac{2(p+k)_\mu}{[(p+k)^2 + m^2]^2}, \\ \frac{\partial}{\partial m} \frac{1}{(p+k)^2 + m^2} &= -\frac{2m}{[(p+k)^2 + m^2]^2}.\end{aligned}$$

The differentiated diagram has a smaller degree of divergence:

$$\omega\left(\frac{\partial \mathcal{I}_G}{\partial K}\right) \leq \left[\frac{\partial \mathcal{I}_G}{\partial K}\right] = [\mathcal{I}_G] - 1,$$

where K is k_μ or m . Repeating the argument, we obtain

$$\omega\left(\frac{\partial^{n+r} \mathcal{I}_G}{\partial k_{\mu_1} \cdots \partial k_{\mu_n} \partial m^r}\right) \leq \left[\frac{\partial^{n+r} \mathcal{I}_G}{\partial k_{\mu_1} \cdots \partial k_{\mu_n} \partial m^r}\right] = [\mathcal{I}_G] - n - r.$$

If $n + r$ is sufficiently large, $[\mathcal{I}_G] - n - r$ becomes negative. Thus, if we differentiate the integral a sufficient number of times with respect to its external momenta and/or the masses, the integral becomes overall convergent. Said in equivalent words, the differentiation kills the overall divergent part. When we integrate back the result, we discover that the divergent part must be polynomial in the masses and the external momenta. This is the crucial property of renormalization, and is called *locality* of the counterterms, because the Fourier transform of a polynomial of the momenta is a finite sum of delta functions and derivatives of delta functions, which are distributions localized at a single point.

Now we describe how to subtract the divergent part of a diagram. Call the integrand $f(p, k, m)$ and consider

$$\mathcal{I}_{G_R}(k, m) = \int \frac{d^d p}{(2\pi)^d} \left(f(p, k, m) - \sum_{n=0}^{\bar{\omega}} \frac{1}{n!} k_{\mu_1} \cdots k_{\mu_n} \frac{\partial_0^n f(p, k, m)}{\partial k_{\mu_1} \cdots \partial k_{\mu_n}} \right), \quad (2.24)$$

where $\bar{\omega}$ is to be determined, and the subscript 0 in ∂_0^n means that after taking the n derivatives with respect to k , k is set to zero. The sum in (2.24) collects the “counterterms”. They remove the divergences from the integral. In practice, we subtract the first $\bar{\omega}$ terms of the Taylor expansion of the integrand around vanishing external momenta. The integrand of \mathcal{I}_{G_R} is still

a rational function, and it is proportional to $\bar{\omega} + 1$ powers of the external momenta. Thus, we can write

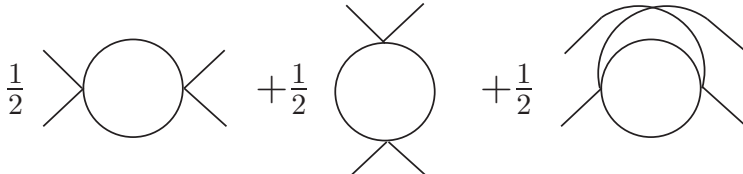
$$\mathcal{I}_{G_R}(k, m) = k_{\mu_1} \cdots k_{\mu_{\bar{\omega}+1}} \int \frac{d^d p}{(2\pi)^d} f_{\mu_1 \cdots \mu_{\bar{\omega}+1}}(p, k, m),$$

for some other rational functions $f_{\mu_1 \cdots \mu_{\bar{\omega}+1}}$. Now,

$$\omega(\mathcal{I}_{G_R}) \leq [\mathcal{I}_{G_R}] - \bar{\omega} - 1 = [\mathcal{I}_G] - \bar{\omega} - 1.$$

If we choose $\bar{\omega} = [\mathcal{I}_G]$, we obtain $\omega(\mathcal{I}_{G_R}) < 0$, which means that \mathcal{I}_{G_R} is overall convergent.

For example, consider the one-loop correction to the four-point function in the theory φ_4^4 . We have the sum of the three diagrams



$$\frac{1}{2} \text{ (diagram 1)} + \frac{1}{2} \text{ (diagram 2)} + \frac{1}{2} \text{ (diagram 3)} \quad (2.25)$$

each of which has the form $(\lambda^2/2)\mathcal{I}(k, m)$, where

$$\mathcal{I}(k, m) = \int \frac{d^4 p}{(2\pi)^4} \frac{1}{p^2 + m^2} \frac{1}{(p+k)^2 + m^2}, \quad (2.26)$$

with different combinations k of the external momenta. The integral (2.26) has $\omega = 0$ and a logarithmic divergence. The subtracted integral reads

$$\mathcal{I}_R(k, m) = -k_\mu \int \frac{d^4 p}{(2\pi)^4} \frac{2p_\mu + k_\mu}{(p^2 + m^2)^2 ((p+k)^2 + m^2)},$$

which is clearly convergent.

We have successfully subtracted the one-loop integrals, but does our procedure make physical sense? Or did we just arbitrarily change the theory we started with? Here enters the crucial property of counterterms, their locality. Formula (2.24) shows that in momentum space the counterterms are polynomial in the external momenta. For example, the counterterm for (2.26) is

$$\mathcal{R}(k, m) \equiv \mathcal{I}_R(k, m) - \mathcal{I}(k, m) = - \int \frac{d^4 p}{(2\pi)^4} \frac{1}{(p^2 + m^2)^2}, \quad (2.27)$$

which is k independent, i.e. just a (divergent) constant.

While a diagram is a nonlocal function of the external momenta, its divergent part is local. Thanks to this property, it looks like a vertex, or an inverse propagator. For this reason, it can be subtracted by adding *ad hoc* local terms to the action. To do this, however, we have to use a specific regularization, because the integral (2.27) is meaningless without a cutoff. If we use a specific regularization, for example the dimensional one, then we can consistently separate \mathcal{I} and \mathcal{R} , and move the counterterms around at will. We stress that the use of an explicit regularization is not necessary to define perturbative quantum field theory. It is however very convenient to keep track of what we are doing.

So, let us switch to the dimensional regularization. The integral of (2.26) is promoted to D dimensions as

$$\mathcal{I}_D(k, m) = \int \frac{d^D p}{(2\pi)^D} \frac{1}{p^2 + m^2} \frac{1}{(p+k)^2 + m^2}. \quad (2.28)$$

and each λ gets multiplied by μ^ε . Using Feynman parameters, namely formula (A.2), we can rewrite the integral as

$$\int_0^1 dx \int \frac{d^D p}{(2\pi)^D} \frac{1}{((p+kx)^2 + m^2 + k^2x(1-x))^2}. \quad (2.29)$$

Then, we can make a translation $p \rightarrow p - kx$ and use (A.4). We get

$$\begin{aligned} \frac{\lambda^2 \mu^{2\varepsilon}}{2} \mathcal{I}_D(k, m) &= \frac{\lambda^2 \mu^{2\varepsilon} \Gamma\left(2 - \frac{D}{2}\right)}{2(4\pi)^{D/2}} \int_0^1 dx (k^2 x(1-x) + m^2)^{D/2-2} \\ &= \frac{\lambda^2 \mu^\varepsilon}{16\pi^2 \varepsilon} + \frac{\lambda^2 \mu^\varepsilon}{32\pi^2} \left(2 - \gamma_E + \ln \frac{4\pi\mu^2}{m^2} \right. \\ &\quad \left. - 2\sqrt{1 + \frac{4m^2}{k^2}} \operatorname{arcsinh} \sqrt{\frac{k^2}{4m^2}} \right) + \mathcal{O}(\varepsilon). \end{aligned} \quad (2.30)$$

The counterterm (2.27) becomes

$$\begin{aligned} -\frac{\lambda^2 \mu^{2\varepsilon}}{2} \int \frac{d^D p}{(2\pi)^D} \frac{1}{(p^2 + m^2)^2} &= -\frac{\lambda^2}{2} \frac{\Gamma\left(2 - \frac{D}{2}\right)}{(4\pi)^{D/2}} \mu^\varepsilon \left(\frac{\mu}{m}\right)^\varepsilon \\ &= -\lambda^2 \mu^\varepsilon \left(\frac{1}{16\pi^2 \varepsilon} + c_1 \right), \end{aligned} \quad (2.31)$$

where the constant c_1 is finite in the limit $\varepsilon \rightarrow 0$. We see that the divergent part does not depend on the external momenta. Since the diagrams of (2.25) are three, and all of them have the same divergent part, we have to modify the action so as to subtract three times the divergent part of (2.31). This result can be achieved by adding

$$\Delta\mathcal{L} = 3\lambda^2\mu^\varepsilon \left(\frac{1}{16\pi^2\varepsilon} + c_1 \right) \frac{\varphi^4}{4!} \quad (2.32)$$

to the Lagrangian. Note that the power $\mu^{2\varepsilon}$ provided by the diagram, shown in the first line of (2.30), has become μ^ε in the counterterm (2.32), to match to dimensions correctly. The other factor μ^ε gets expanded in ε . As a consequence, μ can enter logarithms in the right places to make their arguments dimensionless, such as in the second line of (2.30).

Now we note that the constant c_1 appearing in (2.32) does not actually need to be the one of (2.31), because what is important is to subtract the divergent part. Thus, the c_1 of formula (2.32) can be arbitrary. Roughly speaking, when we subtract infinity, we can as well subtract infinity plus any finite constant. Later on we will see that the physical quantities do not depend on this arbitrariness.

The correction (2.32) produces an extra vertex

$$\begin{array}{c} \diagup \\ \bullet \\ \diagdown \end{array} = -3\lambda^2\mu^\varepsilon \left(\frac{1}{16\pi^2\varepsilon} + c_1 \right) \quad (2.33)$$

that must be added to the Feynman rules. The vertex (2.33) carries an additional (hidden) power of \hbar , since it is of order λ^2 . Diagrammatically, it counts like a one-loop diagram, so it appears in the right place to subtract the divergences of (2.25). The finite value of a single subtracted diagram of (2.25) is thus

$$-\frac{\lambda^2\mu^\varepsilon}{16\pi^2} \left(\sqrt{1 + \frac{4m^2}{k^2}} \operatorname{arcsinh} \sqrt{\frac{k^2}{4m^2}} - \frac{1}{2} \ln \frac{4\pi\mu^2}{m^2} + c \right), \quad (2.34)$$

where c is an arbitrary finite, k -independent constant. We may as well assume that c is independent of μ and m . The result admits a smooth massless limit

$$\frac{\lambda^2\mu^\varepsilon}{32\pi^2} \left(\ln \frac{4\pi\mu^2}{k^2} - 2c \right), \quad (2.35)$$

which can also be easily computed from (2.28). Indeed, using (A.3) to do the integral over x , the first line of (2.30) gives at $m = 0$

$$\frac{\lambda^2 \mu^{2\varepsilon}}{2} \mathcal{I}_D(k, 0) = \frac{\lambda^2 \mu^{2\varepsilon} \Gamma(2 - \frac{D}{2}) [\Gamma(\frac{D}{2} - 1)]^2}{2(4\pi)^{D/2} \Gamma(D - 2)} (k^2)^{D/2-2}. \quad (2.36)$$

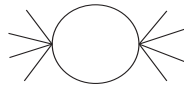
Note again that the factor in front of expression (2.35) is μ^ε instead of $\mu^{2\varepsilon}$, and that the argument of the logarithm contains appropriate factors of μ that make it dimensionless.

The modification (2.32), which subtracts the divergence away, does not look so serious after all. In the end, it just amounts to redefining the coupling constant in front of φ^4 . We are certainly allowed to do that, since we have not attached any physical meaning to λ , so far. This is the idea of *renormalization*, and justifies its name. It is the removal of the divergences by means of redefinitions of fields and parameters. Note that it would not be possible to achieve this goal if the counterterms were not local, since the action is local by assumption. At the same time, locality alone is not sufficient to ensure that the divergences can be renormalized.

Consider for example the theory φ_4^6 . We write the interacting Lagrangian as

$$\mathcal{L}_I = \lambda_6 \mu^{2\varepsilon} \frac{\varphi^6}{6!},$$

where λ_6 is a coupling constant of dimension -2 . At one loop we have divergent diagrams such as



$$(2.37)$$

The corresponding integral is again proportional to $\mathcal{I}_D(k, m)$. However, to subtract this kind of divergence we need to modify the Lagrangian with a counterterm of the form

$$\Delta \mathcal{L} = 35 \lambda_6^2 \mu^{3\varepsilon} \left(\frac{1}{16\pi^2 \varepsilon} + c_1 \right) \frac{\varphi^8}{8!}, \quad (2.38)$$

where $35 = 8!/(2!4!)$ is the number of nontrivial permutations of the external legs. The modified action contains an interaction, φ^8 , that is not present in the action of the theory φ_4^6 . Therefore, (2.38) cannot be absorbed into

a simple redefinition of the fields and the couplings, but demands a radical modification of the theory, from φ_4^6 to $\varphi_4^6 + \varphi_4^8$. Moreover, that modification is not even sufficient. Using two vertices φ^8 we can easily construct a one-loop diagram similar to (2.37), with 6+6 external legs. Again, it is logarithmically divergent and its divergent part can be subtracted only at the price of introducing a vertex φ^{12} . We can go on like this indefinitely: we discover that the renormalization of divergences is possible only at the price of introducing infinitely many new vertices and new independent couplings.

Concluding, the locality of counterterms is necessary, but not sufficient, to have control on the divergences. We need to check that all the counterterms have the form of the terms that are already contained in the initial Lagrangian. When that happens, the divergences can be removed by redefining the fields and the couplings, the subtraction of divergences is a stable procedure and the final Lagrangian is a simple redefinition of the initial one. Otherwise, we can attempt to stabilize the Lagrangian, by adding new *ad hoc* local terms. Next, we must check that a finite number of such new terms is sufficient to stabilize the subtraction of divergences to all orders. If that does not happen, the final Lagrangian contains infinitely many independent couplings and interactions.

The theories that contain finitely many vertices and are stable under renormalization are called renormalizable. The theories that are not stable under renormalization, because they end up containing infinitely many independent terms, are called nonrenormalizable. As we will prove, the theory φ_4^4 is renormalizable. We have already proved that the theory φ_4^6 cannot be stabilized, so it is nonrenormalizable.

Nonrenormalizable theories are described by nonpolynomial Lagrangians, which are the sums of local terms with arbitrarily high powers of the fields and their derivatives. We have

$$\mathcal{L}_{\text{nonren}} = \frac{1}{2}(\partial_\mu \varphi)^2 + \frac{m^2}{2}\varphi^2 + \sum_{\{m,n\}} \frac{\lambda_{\{m,n\}}}{M^{X(m,n)}} \prod_i (\partial^{m_i} \varphi^{n_i}), \quad (2.39)$$

where M is some energy scale and

$$X(m,n) = \sum_i \left(m_i + n_i \frac{d-2}{2} \right) - d$$

is chosen to make the couplings $\lambda_{\{m,n\}}$ dimensionless. Nonrenormalizable theories are problematic from the physical point of view. Their correlation

functions depend on infinitely many parameters, which means, at the practical level, that infinitely many measurements are necessary to determine the theory completely, and make predictions that are valid at arbitrarily high energies.

On the other hand, in most cases nonrenormalizable theories can be used to make predictions at low energies. If a monomial \mathcal{O} in the fields and their derivatives has dimension $d_{\mathcal{O}}$, then its insertion into a correlation function behaves like $E^{d_{\mathcal{O}}}$ at low energies, so the interacting Lagrangian behaves like

$$\mathcal{L}_I \sim E^d \sum_{\{m,n\}} \lambda_{\{m,n\}} \left(\frac{E}{M} \right)^{X(m,n)}.$$

We can have three typical cases.

- 1) If all the dimensionless couplings $\lambda_{\{m,n\}}$ are of comparable orders at energies $E \ll M$, only a finite number of interactions is important. We then say that *almost all* interactions become negligible at low energies. However, the number of interactions that are important grows with the energy and becomes infinite at $E \sim M$.
- 2) A behavior like $\lambda_{\{m,n\}} \sim X(m,n)^{-X(m,n)}$ for large m,n , ensures that almost all the couplings can be neglected in every energy range bounded from above. The number of important couplings grows with the energy and becomes infinite only at $E = \infty$.
- 3) A behavior like $\lambda_{\{m,n\}} \sim X(m,n)^{X(m,n)}$ for large m,n , ensures that the parameter $\lambda_{\{m,n\}}$ is negligible for energies

$$E \ll \frac{M}{X(m,n)}.$$

There exists no energy range where almost all couplings can be neglected.

Intermediate types of behaviors can be traced back to these three cases. The behaviors of the couplings $\lambda_{\{m,n\}}$ are a priori unknown, but comparison with experiments can suggest whether we are in the situations 1), 2) or 3).

Even in the worst case, a nonrenormalizable theory may still have a non-trivial predictive content. Indeed, even if the Lagrangian contains infinitely many independent unknown parameters, there might still exist physical quantities that just depend on a finite subset of them. The hard part is to work out those physical quantities and make experiments that are suitable for them.

Strictly speaking, the difference between renormalizable and nonrenormalizable theories is that the former are always predictive, in an obvious way, while the latter can be predictive, but only in a rather nontrivial way.

It is worth to stress that the nonrenormalizable theories are much less problematic from the mathematical point of view, where it does not really matter whether the number of independent couplings is finite or infinite. Indeed, most renormalization theorems we are going to prove hold both for renormalizable and nonrenormalizable theories

φ_4^4 at one loop

Let us complete the one-loop renormalization of the φ_4^4 theory. Formula (2.23) gives

$$\omega_G \leq 4 - E,$$

so the potentially divergent diagrams are those with $\omega_G \geq 0$, i.e. $E \leq 4$. The renormalization of the four-point function has been discussed above. Since the φ_4^4 theory has a \mathbb{Z}_2 symmetry $\varphi \rightarrow -\varphi$, the correlation functions that contain an odd number of insertions are identically zero. Moreover, the diagrams with zero external legs need not be considered, since they can always be subtracted by adding a constant to the Lagrangian. We remain with the one-loop correction to the two-point function, which is the second term on the right-hand side of (1.46) and gives the integral

$$-\frac{\lambda\mu^\varepsilon}{2} \int \frac{d^D p}{(2\pi)^D} \frac{1}{p^2 + m^2} = \lambda m^2 \left(\frac{1}{16\pi^2\varepsilon} - c_2 \right), \quad (2.40)$$

where the constant c_2 is regular for $\varepsilon \rightarrow 0$. To subtract this divergence we modify the action by adding

$$\Delta' \mathcal{L} = \lambda m^2 \left(\frac{1}{16\pi^2\varepsilon} - c_2 \right) \frac{\varphi^2}{2}. \quad (2.41)$$

Again, we can take an arbitrary finite c_2 here, different from the one appearing in (2.40). Collecting (2.32) and (2.41), the full one-loop renormalized action reads

$$\begin{aligned} S_1(\varphi) = \int d^D x \left(\frac{1}{2} (\partial_\mu \varphi)^2 + m^2 \left(1 + \frac{\lambda}{16\pi^2\varepsilon} - \lambda c_2 \right) \frac{\varphi^2}{2} \right. \\ \left. + \lambda \mu^\varepsilon \left(1 + \frac{3\lambda}{16\pi^2\varepsilon} + 3\lambda c_1 \right) \frac{\varphi^4}{4!} \right). \end{aligned} \quad (2.42)$$

More generally, the renormalized action can be written as

$$S_R(\varphi, \lambda, m, \mu) = \int d^D x \left(\frac{Z_\varphi}{2} (\partial_\mu \varphi)^2 + m^2 Z_{m^2} \frac{Z_\varphi \varphi^2}{2} + \lambda \mu^\varepsilon Z_\lambda \frac{Z_\varphi^2 \varphi^4}{4!} \right), \quad (2.43)$$

where the coefficients

$$\begin{aligned} Z_\varphi &= 1 + \mathcal{O}(\lambda^2), & Z_{m^2} &= 1 + \frac{\lambda}{16\pi^2\varepsilon} - \lambda c_2 + \mathcal{O}(\lambda^2), \\ Z_\lambda &= 1 + \frac{3\lambda}{16\pi^2\varepsilon} + 3\lambda c_1 + \mathcal{O}(\lambda^2), \end{aligned} \quad (2.44)$$

which depend on λ and ε , are called “renormalization constants”. At one loop, S_R coincides with S_1 . If we collect the field and parameter redefinitions into the “bare” quantities

$$\varphi_B = Z_\varphi^{1/2} \varphi, \quad m_B^2 = m^2 Z_{m^2}, \quad \lambda_B = \lambda \mu^\varepsilon Z_\lambda, \quad (2.45)$$

then $S_R(\varphi, \lambda, m, \mu)$ can be rewritten in bare form

$$S_B(\varphi_B, \lambda_B, m_B) \equiv \int d^D x \left(\frac{1}{2} (\partial_\mu \varphi_B)^2 + m_B^2 \frac{\varphi_B^2}{2} + \lambda_B \frac{\varphi_B^4}{4!} \right). \quad (2.46)$$

We see that the bare action is exactly the classical action.

We have already observed that the constants c_1 and c_2 of formula (2.42) are arbitrary. Any time we subtract a pole $1/\varepsilon$, we can equivalently subtract $1/\varepsilon$ plus a finite constant. This arbitrariness amounts to a finite redefinition of the fields and the parameters, which has no physical significance.

A specific prescription to choose such arbitrary constants is called *subtraction scheme*. For example, subtracting the first terms of the Taylor expansion around vanishing external momenta is a scheme prescription. In massless theories this prescription is not convenient, because it can originate spurious infrared divergences. Then it is better, for example, to subtract the first terms of the Taylor expansion around some nontrivial configurations of the external momenta. We can even choose different configurations for different diagrams. A very popular scheme, called *minimal subtraction scheme*, amounts to subtract just the poles in ε , with no finite parts attached.

The constants c_1 and c_2 parametrize the scheme arbitrariness at one loop. The residues of the poles $1/\varepsilon$, on the other hand, are scheme independent. For example, comparing (2.3) and (2.7), we have remarked that the coefficients

of $\ln \Lambda$ and $1/\varepsilon$ coincide, as well as the term $m^2 \ln m^2$. Instead, the quadratic divergences Λ^2 end into the arbitrary constant c_2 .

A few tricks can allow us to compute the divergent parts quite easily, taking advantage of their locality. Consider for example the integral $\mathcal{I}_D(k, m)$ of formula (2.28). We know that its divergent part is a polynomial of degree zero in k and m . Therefore, it just a constant, and can be calculated by setting k and m to the values we like. We cannot put $k = m = 0$, however, because this affects the domain of integration in a nontrivial way: the rules of the dimensional regularization do not allow us to exchange the integral and the limits $k \rightarrow 0, m \rightarrow 0$ in this case. A better choice is to keep $m \neq 0$ and put $k = 0$: since the domain of integration is unaffected, the limit $k \rightarrow 0$ can be safely taken inside the integral. We could also keep $k \neq 0$ and put $m = 0$, but the first choice is more convenient. Then, (2.28) becomes a standard integral (see Appendix A) and its divergent part can be worked out immediately.

More generally, since the divergent part of a diagram is a polynomial of the external momenta k and the masses m , if we differentiate the integral with respect to k and m , we can reduce the degree of the polynomial to zero, and then proceed as above. If we differentiate in all possible ways, we can fully reconstruct the polynomial, i.e. the divergent part of the diagram.

Using these tricks,

Exercise 4 *Compute the one-loop renormalization of the φ_6^3 theory.*

Solution. The renormalized action reads

$$S(\varphi) = \int d^D x \left(\frac{Z_\varphi}{2} (\partial_\mu \varphi)^2 + \frac{m^2}{2} Z_{m^2} Z_\varphi \varphi^2 + \lambda \mu^\varepsilon Z_\lambda Z_\varphi^{3/2} \frac{\varphi^3}{3!} + m^4 \mu^{-\varepsilon} \Delta_1 \varphi \right),$$

where $\varepsilon = 3 - D/2$. At one loop the divergent diagrams are those with one, two and three external legs. The tadpole is

$$-\lambda \mu^\varepsilon \frac{\Gamma(1 - \frac{D}{2}) m^{D-2}}{(4\pi)^{D/2}} = -\frac{\lambda m^4 \mu^{-\varepsilon}}{2(4\pi)^3 \varepsilon} + \text{finite},$$

whence

$$\Delta_1 = -\frac{\lambda}{2(4\pi)^3 \varepsilon}.$$

The self-energy is equal to the first line of (2.30). The difference is that now we have to expand it around $D = 6$ instead of $D = 4$. We obtain the divergent part

$$-\frac{\lambda^2}{12\varepsilon(4\pi)^3} (k^2 + 6m^2),$$

which gives

$$Z_\varphi = 1 - \frac{\lambda^2}{12\varepsilon(4\pi)^3} + \mathcal{O}(\lambda^4), \quad Z_{m^2} = 1 - \frac{5\lambda^2}{12\varepsilon(4\pi)^3} + \mathcal{O}(\lambda^4). \quad (2.47)$$

The divergent part of the correction to the vertex can be calculated at vanishing external momenta. We have

$$-\int \frac{d^D p}{(2\pi)^D} \frac{\lambda^3 \mu^{3\varepsilon}}{(p^2 + m^2)^3} = -\lambda^3 \mu^{3\varepsilon} \frac{\Gamma(3 - D/2)}{2(4\pi)^{D/2}} (m^2)^{D/2-3} = -\frac{\lambda^3 \mu^\varepsilon}{2\varepsilon(4\pi)^3} + \text{finite},$$

so the vertex renormalization constant is

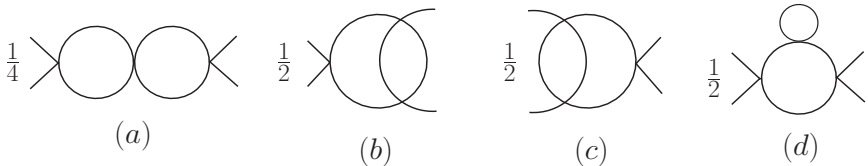
$$Z_\lambda = 1 - \frac{3\lambda^2}{8\varepsilon(4\pi)^3} + \mathcal{O}(\lambda^4). \quad (2.48)$$

□

Most of the properties of the renormalization at one loop generalize to all orders. Now we make some remarks about the renormalization at two loops, which help us introduce the proofs of all-order statements.

φ_4^4 at two loops

We denote the vertices provided by the one-loop counterterms (2.32) and (2.41) with a dot, as in (2.33). At two loops, we have diagrams that contain, in general, both subdivergences and overall divergences. For example, consider the following two-loop corrections to the four-point function, given by the diagrams



$$(2.49)$$

plus two permutations of each. We begin by concentrating on the first three diagrams, since the fourth one is much simpler to deal with. The subdivergences of the diagrams (a), (b) and (c) are subtracted by

$$\begin{array}{ccc}
 \frac{1}{2} \text{ (diagram with circle and dot on right)} & & \frac{1}{2} \text{ (diagram with circle and dot on left)} \\
 (e) & & (f)
 \end{array} \quad (2.50)$$

The subdivergences of each subdiagram of (a), (b) and (c) are given by one third of (2.33). Moreover, (2.33) absorbs also a combinatorial factor $1/2$, which is the combinatorial factor of the diagrams (2.25). It is convenient to define separate counterterms for the three diagrams (2.25), even if they are equal in value. We do this by splitting the counterterm (2.33) into the sum of three equal contributions, and using appropriate labels to remember which diagram they cure. So doing, we obtain

$$\text{(diagram with dot)} = \text{(diagram with square and minus)} + \text{(diagram with square and I)} + \text{(diagram with square and X)} \quad (2.51)$$

Observe that the rules to compute the combinatorial factors remain the same after this splitting. If A is the value of a vertex with N external legs and V is the number of times it appears in a diagram, its contribution is $A^V/(N!^V V!)$. Now, if A is decomposed as a sum $\sum_{i=1}^n a_i$, the multinomial formula ensures that each “subvertex” contributes with the same rule. Indeed we have

$$\frac{A^V}{(N!)^V V!} = \frac{(\sum_{i=1}^n a_i)^V}{(N!)^V V!} = \sum_{\{n_i\}} \prod_{i=1}^n \frac{a_i^{n_i}}{(N!)^{n_i} n_i!}, \quad (2.52)$$

where the sum is taken over sets of non-negative n_i s such that $\sum_{i=1}^n n_i = V$. Note that it is not necessary that each term a_i of the sum be symmetrized under the exchange of its external legs.

Now, consider diagram (a) and its counterterms:

$$\frac{1}{4} \text{ (diagram with two circles)} + \frac{1}{2} \text{ (diagram with circle and square minus)} + \frac{1}{2} \text{ (diagram with circle and square minus)} \quad (2.53)$$

This sum is free of subdivergences. Indeed, the two subdiagrams of (a) are of the first type of the list that appears on the right-hand side of (2.51). Moreover, recall that each subdiagram carries a combinatorial factor $1/2$, which explains why the counterterms in (2.53) are correctly multiplied by $1/2$ instead of $1/4$.

Next, consider diagram (b): the sum

$$\frac{1}{2} \text{diagram 1} + \frac{1}{2} \text{diagram 2} + \frac{1}{2} \text{diagram 3} \quad (2.54)$$

is also free of subdivergences. Observe that this time we use the second and third vertices of (2.51), because they both correspond to the divergent subdiagram of (b). Again the combinatorial factors match, taking into account the factor $1/2$ absorbed by the divergent subdiagram. Diagram (c) is treated symmetrically to (b).

In conclusion, the sums

$$s_1 = (a) + \frac{1}{3}(e) + \frac{1}{3}(f), \quad s_2 = (b) + \frac{2}{3}(e), \quad s_3 = (c) + \frac{2}{3}(f). \quad (2.55)$$

are all free of subdivergences. Therefore, so is the total $t = s_1 + s_2 + s_3 = (a) + (b) + (c) + (e) + (f)$.

Since s_1 , s_2 and s_3 are free of subdivergences, so are their derivatives with respect to the external momenta and the masses. Now, a sufficient number of such derivatives does kill the overall divergences of s_1 , s_2 and s_3 and produces fully convergent integrals. This proves that the divergent parts of the subtracted integrals s_1 , s_2 and s_3 are polynomial in the masses and the external momenta.

Let us explicitly check this fact in s_1 . For simplicity, we work at $m = 0$. Diagram (a) is very easy to calculate, since it is basically the square of any diagram of (2.25). We have

$$(a) = \frac{(-\lambda)^3 \mu^\varepsilon}{4} (\mu^\varepsilon \mathcal{I}_D(k, 0))^2.$$

Write

$$\mu^\varepsilon \mathcal{I}_D(k, 0) = \frac{a}{\varepsilon} + b \ln \frac{k^2}{\mu^2} + c,$$

where a , b and c are finite for $\varepsilon \rightarrow 0$. Their values can be read from the calculations already made (in particular, $a = 1/(8\pi^2)$), but for what we are going to say we do not need to do that. Each counterterm of (2.53) equals $-a\lambda^2\mu^\varepsilon/(2\varepsilon)$, so

$$\frac{1}{3}(e) = \frac{1}{3}(f) = -\frac{1}{2\varepsilon}a\lambda^2\mu^\varepsilon \left(\frac{-\lambda\mu^\varepsilon}{2} \right) \mathcal{I}_D(k, 0).$$

Finally,

$$s_1 = -\frac{\lambda^3\mu^\varepsilon}{4} \left[\left(\frac{a}{\varepsilon} + b \ln \frac{k^2}{\mu^2} + c \right)^2 - 2\frac{a}{\varepsilon} \left(\frac{a}{\varepsilon} + b \ln \frac{k^2}{\mu^2} + c \right) \right].$$

In the difference, the (nonlocal) subdivergences

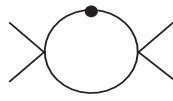
$$2\frac{ab}{\varepsilon} \ln \frac{k^2}{\mu^2} \quad (2.56)$$

do subtract away and the surviving (overall) divergences are purely local. We find

$$s_1 = -\frac{\lambda^3\mu^\varepsilon}{4} \left(-\frac{a^2}{\varepsilon^2} + \text{finite part} \right), \quad (2.57)$$

as expected. This example, although very simple, is sufficient to illustrate the most general facts. The subdivergences are in general nonlocal, because they are “products” of a divergent part, originated by some subdiagram, times a finite (thus nonlocal) part, due to the rest of the diagram. Subtracting something like (2.56) would really require to alter the original theory completely, turning it into a nonlocal theory. Fortunately, the subdivergences are automatically subtracted by the counterterms associated with the subdiagrams.

It remains to consider the diagram (d) of (2.49). Its subdivergence is subtracted by



(g)

(2.58)

where the dot denotes the counterterm of (2.41). The sum $s_4 = (d) + (g)$ is clearly free of subdivergences.

Consider now the two-loop corrections to the two-point function, which are given by the diagrams

$$\frac{1}{4} \text{ --- } \begin{array}{c} \bigcirc \\ \bigcirc \end{array} \quad \frac{1}{3!} \text{ --- } \bigcirc \quad (2.59)$$

The counterterms that subtract the subdivergences are

$$\frac{1}{2} \text{ --- } \frac{1}{2} \quad (c) \quad \frac{1}{2} \text{ --- } \frac{1}{2} \quad (d) \quad (2.60)$$

which vanish at vanishing masses, since they are tadpoles. Diagram (*h*) has two types of diverging subdiagrams, corresponding to both types of contributions (*i*) and (*j*). Instead, diagram (*k*) has a single type of subdiagram, but it appears three times, since freezing any internal line gives the same result. The total $t' = (h) + (k) + (i) + (j)$ can be arranged as $s_5 + s_6$, where

$$s_5 = (h) + \frac{1}{3}(i) + (j), \quad s_6 = (k) + \frac{2}{3}(i), \quad (2.61)$$

are both free of subdivergences. Explicitly, the (h) subdivergence due to the bottom loop is subtracted by $(i)/3$, while the (h) subdivergence due to the top loop is subtracted by (j) . This $(i)/3$ is obtained using the middle vertex of (2.53). Similarly, the (k) subdivergences obtained freezing any internal line (which gives a factor 3) are subtracted by $2(i)/3$. These two $(i)/3$'s are obtained using both the first and third vertices of (2.53).

Again, this proves that the overall divergences of the sums s_5 and s_6 are polynomial in the masses and the external momenta.

Exercise 5 Calculate Z_φ at two loops in the massless φ_4^4 theory.

Solution. The two-loop contribution to the self-energy is

$$\frac{\lambda^2 \mu^{2\varepsilon}}{3!} \int \frac{d^D p}{(2\pi)^D} \frac{d^D q}{(2\pi)^D} \frac{1}{p^2 q^2 (p+q+k)^2}, \quad (2.62)$$

where k is the external momentum. We first integrate over p by means of formula (2.36). We get

$$\frac{\lambda^2 \mu^{2\varepsilon}}{3!} \frac{\Gamma(2 - \frac{D}{2}) [\Gamma(\frac{D}{2} - 1)]^2}{(4\pi)^{D/2} \Gamma(D - 2)} \int \frac{d^D q}{(2\pi)^D} \frac{1}{q^2 [(q + k)^2]^{2-D/2}}.$$

Now we use the Feynman parameters again to calculate the integral over q . We obtain

$$\frac{\lambda^2 \mu^{2\varepsilon}}{3!} \frac{\Gamma(3 - D) [\Gamma(\frac{D}{2} - 1)]^3}{(4\pi)^D \Gamma(\frac{3D}{2} - 3)} (k^2)^{D-3}. \quad (2.63)$$

Extracting the divergent part, we finally obtain

$$Z_\varphi = 1 - \frac{\lambda^2}{12\varepsilon(4\pi)^4} + \mathcal{O}(\lambda^3). \quad (2.64)$$

Exercise 6 Calculate Z_λ at two loops in the massless φ_4^4 theory.

Solution. The diagrams we have to study are those of formula (2.49), plus the counterterms (2.50) and (2.58), plus two permutations of each. The diagram (d) and its counterterm (g) vanish in at $m = 0$. Since Z_λ is independent of the external momenta k , we can simplify the calculation by setting k to zero and working with nonvanishing artificial masses δ (to avoid infrared problems). Alternatively, we can keep the masses equal to zero and choose convenient configurations of the external momenta. We adopt the second option. The divergent contributions s_3^{div} and s_2^{div} of s_3 and s_2 coincide, so the total divergent part can be written as $3(s_1^{\text{div}} + 2s_2^{\text{div}})$, where s_1^{div} can be read from (2.57) and the overall factor 3 takes the permutations into account.

Now we evaluate the diagram (b). Let k denote the total incoming momentum of the two external legs on the left-hand side. We can simplify the calculation by setting the momentum k' of the top-right external leg to zero. Indeed, it is easy to see that the integral becomes fully convergent after one derivative with respect to k' , which means that the divergent part, although nonlocal, cannot depend on k' . The same trick does not work for k , so we cannot set $k = 0$. The subdiagram can be replaced with its exact expression (2.36). We get

$$(b) = -\frac{\lambda^3 \mu^{3\varepsilon} \Gamma(2 - \frac{D}{2}) [\Gamma(\frac{D}{2} - 1)]^2}{2(4\pi)^{D/2} \Gamma(D - 2)} \int \frac{d^D p}{(2\pi)^D} \frac{1}{(p^2)^{3-D/2} (p - k)^2}.$$

Now the calculation proceeds as usual. We get

$$\begin{aligned} (b) &= \frac{\lambda^3 \mu^{3\varepsilon} \left[\Gamma\left(\frac{D}{2} - 1\right) \right]^3 \Gamma(3 - D)}{(4\pi)^D (4 - D) \Gamma\left(\frac{3D}{2} - 4\right)} (k^2)^{D-4} \\ &= -\frac{\lambda^3 \mu^\varepsilon}{(4\pi)^4 \varepsilon^2} - \frac{\lambda^3 \mu^\varepsilon}{(4\pi)^4 \varepsilon} \left(\frac{5}{2} - \gamma_E - \ln \frac{k^2}{(4\pi)\mu^2} \right) + \text{finite}. \end{aligned}$$

On the other hand, it is easy to evaluate (e), which gives

$$\begin{aligned} (e) &= \frac{3\lambda^3 \mu^{2\varepsilon} \Gamma\left(2 - \frac{D}{2}\right) \left[\Gamma\left(\frac{D}{2} - 1\right) \right]^2}{2(4\pi)^2 \varepsilon (4\pi)^{D/2} \Gamma(D - 2)} (k^2)^{D/2-2} \\ &= \frac{3\lambda^3 \mu^\varepsilon}{(4\pi)^4 \varepsilon^2} + \frac{3\lambda^3 \mu^\varepsilon}{2(4\pi)^4 \varepsilon} \left(2 - \gamma_E - \ln \frac{k^2}{(4\pi)\mu^2} \right) + \text{finite}. \end{aligned}$$

The total gives

$$s_2 = (b) + \frac{2}{3}(e) = \frac{\lambda^3 \mu^\varepsilon}{(4\pi)^4 \varepsilon} \left(\frac{1}{\varepsilon} - \frac{1}{2} \right) + \text{finite}.$$

Note that the nonlocal subdivergences cancel out, as it must be. Finally, collecting the contributions of s_1 and s_2 , we get

$$3(s_1^{\text{div}} + 2s_2^{\text{div}}) = \frac{3\lambda^3 \mu^\varepsilon}{(4\pi)^4 \varepsilon} \left(\frac{3}{\varepsilon} - 1 \right).$$

Using (2.64) we obtain

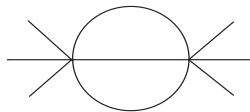
$$Z_\lambda(\lambda, \varepsilon) = 1 + \frac{3\lambda}{(4\pi)^2 \varepsilon} + \frac{9\lambda^2}{(4\pi)^4 \varepsilon^2} - \frac{17\lambda^2}{6(4\pi)^4 \varepsilon} + \mathcal{O}(\lambda^3). \quad (2.65)$$

Exercise 7 Compute the two-loop renormalization of the massless φ_3^6 theory.

Solution. The renormalized action is

$$S(\varphi) = \int d^D x \left(\frac{Z_\varphi}{2} (\partial_\mu \varphi)^2 + \lambda \mu^{2\varepsilon} Z_\lambda Z_\varphi^3 \frac{\varphi^6}{6!} \right),$$

where $\varepsilon = 3 - D$. It can be easily checked that there is no one-loop divergence, so we just have to consider the two-loop diagrams. Moreover, there is no two-loop contribution to the wave-function renormalization constant. Instead, the vertex gets a counterterm from the diagram



The combinatorial factor is $1/6$. The divergent part does not depend on permutations of external legs, which gives an extra factor of 10. We just have to calculate (2.62), multiplied by $10\mu^{2\varepsilon}$. Using (2.63) and expanding around $D = 3$ we get

$$Z_\lambda = 1 + \frac{5\lambda}{6\varepsilon(4\pi)^2} + \mathcal{O}(\lambda^2). \quad (2.66)$$

Note that using the dimensional regularization no other counterterm can be generated. For example, a counterterm of type φ^4 is in principle allowed by power counting, but it would have to be multiplied by a dimensionful parameter, which is absent in the massless case. On the other hand, the dimensional regularization kills the powerlike divergences, because it has a sort of dimensionless cutoff.

Exercise 8 *Compute the first contribution to the self-energy counterterm of the massless φ_3^6 theory.*

Solution. The first correction to the self-energy is of order λ^2 and has four loops. It can be computed with the method used in exercise 5 to go from (2.62) to (2.63). The difference is that now we have to iterate the integration four times instead of two. The result is

$$\frac{\lambda^2 \mu^{4\varepsilon}}{5!} \frac{\Gamma(5-2D)}{(4\pi)^{2D}} \frac{[\Gamma(\frac{D}{2}-1)]^5}{\Gamma(\frac{5D}{2}-5)} (k^2)^{2D-5}. \quad (2.67)$$

Extracting the divergent part, we obtain

$$Z_\varphi = 1 - \frac{4\lambda^2}{45\varepsilon(16\pi)^4} + \mathcal{O}(\lambda^3). \quad (2.68)$$

2.3 Renormalization to all orders

In the renormalizable theories, which we classify in the next sections, formulas like (2.45) generalize to all orders. Now we describe what happens, and later prove the theorems that justify our claims. Let φ , λ and m collectively denote the fields, the couplings and the masses, respectively. Start from the classical action, and interpret it as the bare action $S_B(\varphi_B, \lambda_B, m_B)$ of the quantum field theory, which depends on the bare fields and parameters, denoted by the subscript B. Then, there exist renormalization constants Z_φ ,

Z_m and Z_λ , which depend on λ and ε , and renormalized quantities φ , λ and m , defined by the map

$$\varphi_B = Z_\varphi^{1/2} \varphi, \quad m_B^2 = m^2 Z_m, \quad \lambda_B = \lambda \mu^{p\varepsilon} Z_\lambda, \quad (2.69)$$

such that all the renormalized generating functionals and the renormalized correlation functions are convergent in the limit $\varepsilon \rightarrow 0$. The renormalized generating functionals coincide with the bare generating functionals written in terms of the renormalized fields and parameters. The renormalized correlation functions are equal to the bare correlation functions, written in terms of the renormalized fields and parameters, apart from a multiplying factor, which is specified below. In formula (2.69) p denotes the difference between the continued and the physical dimensions of λ , the physical dimension being the one at $\varepsilon = 0$.

Precisely, define the renormalized action S_R , such that

$$S_B(\varphi_B, \lambda_B, m_B) = S_R(\varphi, \lambda, m, \mu). \quad (2.70)$$

Then, define the bare and renormalized generating functionals Z and W by means of the formulas

$$\begin{aligned} Z_B(J_B, \lambda_B, m_B) &= \int [d\varphi_B] e^{-S_B(\varphi_B, \lambda_B, m_B) + \int \varphi_B J_B} = e^{W_B(J_B, \lambda_B, m_B)} \\ &= \int [d\varphi] e^{-S_R(\varphi, \lambda, m, \mu) + \int \varphi J} = Z_R(J, \lambda, m, \mu) = e^{W_R(J, \lambda, m, \mu)}, \end{aligned} \quad (2.71)$$

with

$$J_B = Z_\varphi^{-1/2} J.$$

Define also bare and renormalized correlation functions, possibly connected and/or irreducible, as

$$\begin{aligned} G_B(x_1, \dots, x_n, \lambda_B, m_B) &= \langle \varphi_B(x_1) \cdots \varphi_B(x_n) \rangle \\ &= Z_\varphi^{n/2} \langle \varphi(x_1) \cdots \varphi(x_n) \rangle = Z_\varphi^{n/2} G_R(x_1, \dots, x_n, \lambda, m, \mu). \end{aligned}$$

Next, using

$$\begin{aligned} \Phi_B(J_B)_x &= \frac{\delta W_B(J_B)}{\delta J_B(x)} = \langle \varphi_B(x) \rangle_J \\ &= Z_\varphi^{1/2} \langle \varphi(x) \rangle_J = Z_\varphi^{1/2} \frac{\delta W_R(J)}{\delta J(x)} = Z_\varphi^{1/2} \Phi(x), \end{aligned}$$

perform the Legendre transforms, and construct the bare and renormalized generating functionals Γ . We find

$$\begin{aligned}\Gamma_B(\Phi_B, \lambda_B, m_B) &= -W_B(J_B(\Phi_B)) + \int J_B(\Phi_B)\Phi_B \\ &= -W_R(J(\Phi)) + \int J(\Phi)\Phi = \Gamma_R(\Phi, \lambda, m, \mu).\end{aligned}$$

By definition, the map (2.69) must be such that

$$\Gamma_R(\Phi, \lambda, m, \mu) < \infty,$$

in the limit $\varepsilon \rightarrow 0$, that is to say all the irreducible diagrams are convergent once expressed in terms of the renormalized quantities. This fact also implies

$$Z_R(J, \lambda, m, \mu) < \infty, \quad W_R(J, \lambda, m, \mu) < \infty, \quad G_R(x_1, \dots, x_n, \lambda, m, \mu) < \infty.$$

Observe that the renormalized action $S_R(\varphi, \lambda, m, \mu)$, instead, is not convergent for $\varepsilon \rightarrow 0$. Check for example, the one-loop renormalized action (2.42). However, the classical action does not have a direct physical meaning. It is just a tool that allows us to implement what remains of the correspondence principle in quantum field theory. As promised, renormalization amounts to a change of field variables, combined with a reparametrization, that is able to move the divergences away from all the physical quantities. It does not care if the nonphysical quantities, such as the classical action, remain or become meaningless.

Note that the renormalized sides of (2.69), (2.70), etc., depend on one quantity more than the bare sides, that is to say the *dynamical scale* μ . The nontrivial μ dependence of the renormalized correlation functions is the root of the renormalization-group flow, which will be studied later.

To prove the renormalizability to all orders, we need to prove two properties, to all orders, namely that the counterterms are local, and that they have the form of the terms already contained in the bare action S_B . We begin with the locality of the counterterms.

2.4 Locality of counterterms

Now we are ready to prove the theorem that ensures the locality of the counterterms.

Theorem 2 *Once subdivergences are subtracted, the overall divergences of a diagram are local and polynomial in the masses.*

Proof. Let V_0 denote a vertex of the starting Lagrangian and V_L an L -loop counterterm, $L \geq 1$. Let \mathcal{V}_L denote the set of V_L s, $L \geq 0$. Moreover, let G_L denote an L -loop diagram constructed with the vertices of \mathcal{V}_0 , and \mathcal{G}_L the set of such diagrams. Let C_L denote an L -loop diagram constructed with at least one counterterm V_N , $0 < N < L$, and \mathcal{C}_L the set of such diagrams. Note that a C_L cannot be a V_L . A counterterm V_L subtracts an overall divergence of some diagram G_L . Instead, a diagram C_L subtracts some subdivergence of a G_L .

Proceeding inductively, assume that the theorem is true up to the n th loop order included, which means that all V_m 's with $m \leq n$ are local and polynomial in the masses. Then, consider a diagram G_{n+1} . It corresponds to an integral over $n+1$ momenta p_i , $i = 1, \dots, n+1$. The momenta of the internal legs are linear combinations Δp_i of the p_i with coefficients ± 1 . The potentially diverging contributions to the integral can only come from the integration regions where the momenta p_i are sent to infinity. The overall divergence correspond to sending the p_i s to infinity with the same velocity. The subdivergences also correspond to sending them to infinity with the same velocity, but with the constraint that some Δp_i are kept fixed. Once we cure the behaviors in such integration regions, the integral is convergent, because any other integration regions, such as the ones corresponding to sending some Δp_i to infinity with different velocities, are automatically covered.

From the diagrammatic point of view, keeping the momenta of some internal lines fixed while all other integrated momenta are sent to infinity amounts to cut those internal lines and single out a proper subdiagram. Such a subdiagram is not necessarily connected, nor irreducible. We do not need to worry about that, since the inductive assumption ensures that all the diagrams of orders $\leq n$ are appropriately subtracted. Indeed, once the irreducible ones are cured to some order, the connected and disconnected diagrams are also cured to the same order.

Observe that the subdiagrams themselves have overall divergences and subdivergences. Nevertheless, again, the inductive assumption ensures that the necessary counterterms are right there. This is actually a nontrivial fact, since we must convince ourselves that the diagrams \mathcal{C}_L , which are built with at least one counterterms, appear in the right place and with the right

coefficient. Now we show that this property follows from Wick's theorem.

The examples studied before suggest that there exists a direct match among the coefficients of “terms and counterterms”, the G s and the C s. Formula (2.55) and (2.61) tell us that we need to multiply the subtractions C by appropriate coefficients in order to match the G s. This is the nontrivial part of the game: to distribute every C among various G s, and check that the total coefficient in front of C is still equal to one. For example, the subtraction (e) had to be split as follows: one third for (a) and two thirds for (b). In the end, everything worked perfectly, but what is not obvious is how to promote those examples to a general proof to all orders. Fortunately, we are making the problem harder than it actually is. It is sufficient to change the viewpoint, to realize that all the coefficients match in a rather natural way.

To see this, we rearrange the perturbative expansion not as a sum over diagrams G , but as a sum over sets of Wick contractions that lead to the diagrams. We anticipated that this trick was going to be useful for some theoretical proof (although it is definitely not convenient at the practical level). Let us denote the Wick contractions with \tilde{G} . We know that each Wick contraction has a simple combinatorial factor (the numerator s of (1.51) is always equal to one for bosons, and ± 1 if fermions are present), while diagrams have complicated combinatorial factors. It is convenient to apply a similar trick for counterterms. Rather than collecting the identical contributions altogether into a single counterterm, it is convenient to “mark” each counterterm, to keep track of the \tilde{G} it comes from. To make this operation clearer, we can refer to (2.33) and its splitting (2.51). In (2.51) we marked each contribution to keep track of the diagram it came from, rather than the Wick contraction. We stress that here we want to do even more, that is to say mark each contribution so as to remember the Wick contraction \tilde{G} it comes from. Clearly, all such marked counterterms have $s = \pm 1$. Moreover, after the decomposition, the combinatorial factors follow again the usual rules, as shown in (2.52).

Now, express the correlation function as a sum over the \tilde{G} s. Each \tilde{G} has subdivergences, which are just subsets of Wick contractions. Each such subset is certainly equipped with its own counterterm, and the coefficient is certainly correct, because in this expansion all the factors s are equal to plus or minus one.

Therefore, the sum

$$t = \sum_{\bar{\mathcal{G}}_{n+1}} \bar{G}_{n+1} + \sum_{\bar{\mathcal{C}}_{n+1}} \bar{C}_{n+1} = \sum_{\mathcal{G}_{n+1}} G_{n+1} + \sum_{\mathcal{C}_{n+1}} C_{n+1}$$

is free of subdivergences. Here the bars refer to the rearrangement of the sums over sets of Wick contractions and the associated counterterms.

The argument just given teaches us that, in practice, we do not need to worry about the combinatorial factors: they always turn out to be right, simply because there exists a rearrangement where the combinatorial factors of “terms and counterterms” naturally match. We emphasize that, ultimately, this fact is due to Wick’s formula, stating that each \bar{G} appears once and only once.

Once we know that t is free of subdivergences, we also know that a sufficient number of derivatives with respect to the masses and the external momenta kills the overall divergence and produces a fully convergent integral. Thus, the divergent part of t is polynomial in the external momenta and the masses, which ensures that the V_{n+1} s are local and polynomial in the masses. The inductive assumption is then promoted to the order $n + 1$ and so to all orders. \square

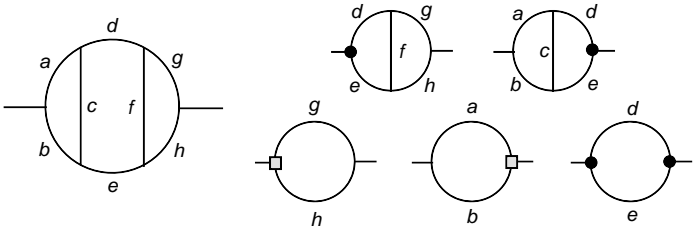
The one we have just considered is the “collective version” of the theorem about the locality of counterterms, which states that the sum of all the $(n + 1)$ -loop diagrams behaves correctly. A refined version of the theorem holds diagram by diagram, which states that every time the subdivergences of a diagram are subtracted away its overall divergence is polynomial in the external momenta and the masses. Precisely, there exist convex linear combinations

$$G_{n+1}^R = G_{n+1} + \sum_{\mathcal{C}_{n+1}} a_{n+1}^{C,G} C_{n+1}, \quad \sum_{\mathcal{G}_{n+1}} a_{n+1}^{C,G} = 1, \quad (2.72)$$

that are separately free of subdivergences. The $a_{n+1}^{C,G}$ are appropriate constants that can be worked out with the method described below. The divergent part of each G_{n+1}^R is local and polynomial in the masses. Examples of (2.72) are (2.55) and (2.61). The collective version of the theorem also follows from its single-diagram version.

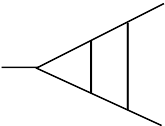
We illustrate the single-diagram version of the theorem by considering a φ_6^3 two-loop diagram together with the counterterms that subtract its sub-


divergences,



(2.73)

The dot denotes the one-loop vertex counterterm. Instead, the square denotes just one contribution to the two-loop vertex counterterm, the one that subtracts the overall divergence of the subdiagram (sd) of the following picture:


(sd)


(c)

(2.74)

Note that (c) does not include the contributions associated with the permutations of the (sd) external legs. As above, we mark each counterterm to remember which Wick contraction it comes from. Then, we sum the Wick contractions that give the same diagram. Thus, (c) is already equipped with the appropriate coefficient to fit into the linear combination (2.72).

Consider (2.73). Let $R(u_1, \dots, u_k)$ denote the region where the momenta u_1, \dots, u_k are sent to infinity and the other ones are kept fixed. Observe that the dot and the square “hide” certain momenta. Now, a counterterm subtracts the overall divergence of a subdiagram, which means that it corresponds to the integration region where the hidden momenta are sent to infinity. For example, the dot of the second diagram contributes to the regions $R(a, b, c)$ and $R(a, b, c, d, e, f)$, but does not contribute to the region $R(a, b, d, e, g, h)$.

Now we study the subdivergences region by region. Region $R(a, b, c)$: the diagram is corrected by the first counterterm. The other counterterms are not concerned. Region $R(a, b, c, d, e, f)$: the diagram is corrected by the first and third counterterms. Region $R(a, b, c, f, g, h)$: the diagram is corrected by the first, second and fifth counterterms. The regions $R(f, g, h)$ and $R(d, e, f, g, h)$ are symmetric to the first two already considered. All

other regions are trivial. We conclude that the sum (2.73) has only overall divergences, which are local. The argument we have illustrated in the case (2.73) generalizes to the most general diagram.

The locality of counterterms is a very general property. It does not depend on the theory, i.e. the types of fields, the forms of the propagators and the structures of the vertices, as long as a sufficient number of derivatives with respect to the external momenta, or the masses, kills the overall divergences. Any local, Lorentz invariant theory satisfies this property and produces local counterterms. For example, if the vertices contain derivatives, then the integrand of (2.21) are multiplied by certain polynomials of the momenta. Yet, it is true that: (i) every derivative with respect to the external momenta or the masses lowers the overall degree of divergence; and (ii) sufficient numbers of such derivatives kill the overall divergences. The subtraction of subdivergences described above is a matter of mere diagrammatics. In particular, it does not require to satisfy any conditions of renormalizability. Even more, the locality of counterterms is so general that it holds in several types of theories not considered so far, including Lorentz violating and nonlocal ones.

We mentioned before that a few tricks can simplify the calculation of the divergent part of a diagram. Now we can upgrade one of those tricks. In general, an L -loop diagram G corresponds to an integral of the form

$$\mathcal{I}_G(k, m) = \int \prod_{i=1}^L \frac{d^D p_i}{(2\pi)^D} \frac{P(p, k, m)}{Q(p, k, m)}, \quad (2.75)$$

instead of (2.21), where P and Q are polynomials of p , k and m . Nontrivial numerators P appear when the vertices contain derivatives. We know that the axioms satisfied by the analytic integral do not allow us to expand the integrand in powers of k and m , and integrate term by term. Nevertheless, it would be very convenient to do so, because it would allow us to efficiently isolate the overall divergent parts from the rest. We can make these operations legitimate by introducing artificial masses $\delta > 0$ in the denominators of propagators. Specifically, let $\mathcal{I}_G^R(k, m)$ denote the subtracted integral, that is to say the integral associated with the sum G_L^R of formula (2.72). Write

$$\mathcal{I}_G^R(k, m) = \lim_{\delta \rightarrow 0} \mathcal{I}_G^R(k, m + \delta).$$

Since \mathcal{I}_G^R is equipped with the counterterms that subtract its own subdivergences, the locality of counterterms ensures that $\mathcal{I}_G(k, m + \delta)$ only has

overall divergences, and those depend polynomially on k , m and δ . Separate $\mathcal{I}_G^R(k, m + \delta)$ into the sum of its divergent part $\mathcal{I}_{G_{\text{div}}}^R(k, m + \delta)$ and its convergent part $\mathcal{I}_{G_{\text{conv}}}^R(k, m + \delta)$. Since $\mathcal{I}_{G_{\text{div}}}^R(k, m + \delta)$ is a polynomial in δ , it admits a smooth limit $\mathcal{I}_{G_{\text{div}}}^R(k, m)$ for $\delta \rightarrow 0$. Then, $\mathcal{I}_{G_{\text{conv}}}^R$ also admits a smooth limit for $\delta \rightarrow 0$, because the sum $\mathcal{I}_{G_{\text{div}}}^R + \mathcal{I}_{G_{\text{conv}}}^R$ must tend to $\mathcal{I}_G^R(k, m)$. Thus, we can write

$$\mathcal{I}_G^R(k, m) = \mathcal{I}_{G_{\text{div}}}^R(k, m) + \lim_{\delta \rightarrow 0} \mathcal{I}_{G_{\text{conv}}}^R(k, m + \delta).$$

The second term on the right-hand side has no poles for $\varepsilon \rightarrow 0$, so it is convergent even after taking the limit $\delta \rightarrow 0$. Finally, the divergent part of the subtracted diagram G_L^R can be calculated as

$$\mathcal{I}_{G_{\text{div}}}^R(k, m) = \lim_{\delta \rightarrow 0} \mathcal{I}_{G_{\text{div}}}^R(k, m + \delta). \quad (2.76)$$

We stress again that it is not legitimate to expand the integrand of $\mathcal{I}_G(k, m)$ in powers of both k and m and then integrate term by term. However, these operations are legitimate on $\mathcal{I}_G^R(k, m + \delta)$, as long as δ is nonzero. Formula (2.76) tells us that when we set δ back to zero, we recover the full divergent part of $\mathcal{I}_G^R(k, m)$. The upgraded trick is particularly useful in massless theories, because the artificial mass allows us to compute the divergent parts by expanding in powers of the external momenta.

Exercise 9 *Prove that, in dimensional regularization, an odd-dimensional local quantum field theory has no nontrivial L -loop divergences, if L is odd.*

Solution. The integrals have the form (2.75). We insert artificial masses δ in the denominators, then expand in powers of the true masses m and the external momenta k . In the end, all the overall divergences are given by expressions of the form

$$\int \prod_{i=1}^L \frac{d^D p_i}{(2\pi)^D} \frac{p_{\mu_1} \cdots p_{\mu_n}}{Q'(p^2)}, \quad (2.77)$$

where the denominator is a polynomial in the squared momenta p_i^2 and Δp_j^2 . If both d and L are odd, then n must be odd, otherwise the integral is either convergent or powerlike divergent. We recall that powerlike divergences are fake divergences in dimensional regularization. If n is odd, the integral (2.77) is odd under the transformation $p_i \rightarrow -p_i$, so its overall divergent part vanishes.

2.5 Power counting

The renormalizability of a theory can be established with a simple dimensional analysis, called *power counting*.

Consider a d -dimensional theory of interacting bosonic fields φ and fermionic fields ψ . We assume that the bosonic fields have propagators $P_B(k)$ that behave like $\sim 1/k^2$ for large momenta k . By this we mean that also behaviors such as

$$P_B(k) \sim \sum_n \frac{k_{\mu_1} \cdots k_{\mu_{2n}}}{(k^2)^{n+1}} \quad (2.78)$$

are allowed. Similarly, we assume that the fermionic fields have propagators $P_F(k)$ that behave like $\sim k_\mu/k^2$, or more generally

$$P_F(k) \sim \sum_n \frac{k_{\mu_1} \cdots k_{\mu_{2n+1}}}{(k^2)^{n+1}}, \quad (2.79)$$

for large momenta. Such behaviors tell us that the dimensions of the bosonic and fermionic fields are $(d-2)/2$ and $(d-1)/2$, respectively.

More generally, consider fields χ_a of dimensions $d/2 - a$ with propagators that behave like

$$P_a(k) \sim \sum_n \frac{k_{\mu_1} \cdots k_{\mu_n}}{(k^2)^{n/2+a}}$$

at large momenta, where a is integer or half-integer, and n is even or odd, respectively. We are not making assumptions about the sign of a , nor the statistics of χ_a .

Let m collectively denote the masses of the fields. Let n_{iB} , n_{iF} , n_{ia} denote the numbers of legs of the types B , F and a of the i th vertex. Assume that the vertex is a polynomial $V_i(k)$ in the momenta, and that its dimension in units of mass is δ_i .

Consider a diagram G with E_B , E_F , E_a external legs and I_B , I_F , I_a internal legs of the various types, v_i vertices of the i th type and L loops. We have, from (2.18)

$$L - I_B - I_F - \sum_a I_a + V = 1, \quad V = \sum_i v_i. \quad (2.80)$$

We denote the external momenta with k and the loop momenta with p . The

integral associated with G has the form

$$\begin{aligned} \mathcal{I}_G(k, m) = & \int \frac{d^{Ld}p}{(2\pi)^{Ld}} \prod_{j=1}^{I_B} P_{Bj}(p, k, m) \prod_{l=1}^{I_F} P_{Fl}(p, k, m) \times \\ & \times \prod_a \prod_{j_a=1}^{I_a} P_{aj_a}(p, k, m) \prod_i \prod_{l_i=1}^{v_i} V_{il_i}(p, k, m), \end{aligned} \quad (2.81)$$

where the indices j, l, j_a and l_i of P_{Bj}, P_{Fl}, P_{aj_a} and V_{il_i} label the propagators and the vertices. Now, rescale k and m to λk and λm . It is convenient to rescale also the loop momenta, which is just a change of variables in the integral. Then $\mathcal{I}_G(k, m)$ rescales with a factor equal to its dimension in units of mass, which is

$$[\mathcal{I}_G] = Ld - 2I_B - I_F - 2 \sum_a a I_a + \sum_i v_i \delta_i. \quad (2.82)$$

Since the overall divergences are local, once the subdivergences have been subtracted away, we infer that they are a polynomial of degree $\omega_G \leq [\mathcal{I}_G]$ in the external momenta and the masses. Now, count the bosonic legs attached to the vertices: they can exit the diagram or be connected to other bosonic internal legs, so

$$E_B + 2I_B = \sum_i v_i n_{iB}.$$

Similarly, the countings of fermionic legs and the legs of the type χ_a , we obtain

$$E_F + 2I_F = \sum_i v_i n_{iF}, \quad E_a + 2I_a = \sum_i v_i n_{ia}.$$

Using (2.80)-(2.82) and $\omega_G \leq [\mathcal{I}_G]$ we get

$$\omega_G \leq d(E_B, E_F, E_a) + \sum_i v_i [\delta_i - d(n_{iB}, n_{iF}, n_{ia})],$$

where

$$d(x, y, z_a) \equiv d - \frac{d-2}{2}x - \frac{d-1}{2}y - \sum_a \frac{d-2a}{2}z_a.$$

We see that if all the vertices satisfy

$$\delta_i \leq d(n_{iB}, n_{iF}, n_{ia}) \quad (2.83)$$

then all the counterterms satisfy the same inequality, namely

$$\omega_G \leq d(E_B, E_F, E_a). \quad (2.84)$$

In other words, if the classical Lagrangian includes all the vertices that satisfy (2.83), then the divergent parts of all the diagrams can be subtracted by renormalizing the couplings, the fields and the masses. The condition (2.83) defines a theory that is *renormalizable by power counting*.

Instead, when the Lagrangian contains some vertex \bar{v} that does not satisfy (2.83), then the diagrams containing \bar{v} s can have an arbitrarily large degree of divergence. In general, in that case, it is necessary to add infinitely many new vertices and couplings to the Lagrangian, if we want to subtract the divergences by means of redefinitions of the fields and the parameters. This kind of theory is called *nonrenormalizable*.

The theories with all the δ_i s equal to $d(n_{iB}, n_{iF}, n_{ia})$ are called strictly renormalizable, those with all the δ_i s smaller than $d(n_{iB}, n_{iF}, n_{ia})$ super-renormalizable, and those with some δ_i greater than $d(n_{iB}, n_{iF}, n_{ia})$ nonrenormalizable.

It is easy to check that the requirement (2.83) is equivalent to demand that all the Lagrangian terms have coefficients of nonnegative dimensions in units of mass. Indeed, the dimension of the coupling λ_i that multiplies the i th vertex is

$$[\lambda_i] = d - \frac{d-2}{2}n_{iB} - \frac{d-1}{2}n_{iF} - \sum_a \frac{d-2a}{2}n_{ia} - \delta_i = d(n_{iB}, n_{iF}, n_{ia}) - \delta_i \geq 0.$$

Thus, a theory is renormalizable by power counting if it contains no parameters of negative dimension (and the propagators are well behaved). This conclusion can be derived more quickly as follows. At the level of the Lagrangian, a counterterm, being local, must have the structure

$$\left(\prod \lambda\right) \partial^p \varphi^{n_B} \psi^{n_F} \prod_a \chi_a^{n_{ia}}. \quad (2.85)$$

The coefficient is a certain product of couplings and masses. We do not need to specify where the derivatives act in (2.85), since it is not important for our discussion. Now, the dimension of (2.85) must be equal to d . If the theory contains no parameters of negative dimensions, we must have

$$p + n_{iB} \frac{d-2}{2} + n_{iF} \frac{d-1}{2} + \sum_a \frac{d-2a}{2} n_{ia} \leq d,$$

which is equivalent to (2.83). On the other hand, if the theory contains a parameter λ_- of negative dimension, then an arbitrarily large power h of λ_- can multiply the counterterm, and we just have an inequality of the form

$$p + n_{iB} \frac{d-2}{2} + n_{iF} \frac{d-1}{2} + \sum_a \frac{d-2a}{2} n_{ia} \leq d - h[\lambda_-], \quad (2.86)$$

which violates (2.83).

Now, it should be kept in mind that in general, in renormalization theory, the following “no-miracle” principle applies:

all the counterterms that are not a priori forbidden are generated by renormalization.

A counterterm can be forbidden by power counting, gauge symmetries, external symmetries. If it is not forbidden, there is practically no hope that it will not be generated as the divergent part of a diagram with an appropriate set of external legs. In other words, no miraculous cancellations should be expected. Then, the inequality (2.86) implies that infinitely many new types of counterterms will be effectively generated, so the theory is nonrenormalizable.

It is important to stress that the propagators must have the right behavior for large momenta. For example, the Proca vectors of formula (1.87) are in general not renormalizable, when interactions are present. Indeed, the propagator (1.89) contains a term $\sim p_\mu p_\nu / (m^2 p^2)$ that prevails over $\delta_{\mu\nu} / p^2$ at large momenta. This forces us to treat the field as a χ_a -field with $a = 0$, which means that its dimension, from the viewpoint of the power counting, is equal to $d/2$. The fields of such a dimension can appear at most quadratically in a local field theory, so they cannot have renormalizable self-interactions. Thus, in general the Proca vectors cannot be included in a renormalizable theory.

Gauge fields can instead be included consistently, although their propagators are naïvely not well behaved. For this reason the gauge theories deserve a special treatment, and we devote chapters of this book to prove their renormalizability.

Particular (scalar) fields of dimension $d/2$ can be useful as auxiliary fields. For example, in the massless φ_4^4 theory we can introduce an auxiliary field σ of dimension 2 and replace the φ^4 -vertex by

$$\mathcal{L}'_I = \frac{1}{2} \sigma^2 + i \mu^{\varepsilon/2} \alpha \sigma \varphi^2, \quad (2.87)$$

where $\alpha = \sqrt{\lambda/12}$. The integral over σ can be performed exactly, by means of a translation $\sigma' = \sigma + i\mu^{\varepsilon/2}\alpha\varphi^2$, which brings \mathcal{L}'_I to the form

$$\mathcal{L}'_I = \frac{1}{2}\sigma'^2 + \frac{\lambda\mu^\varepsilon}{4!}\varphi^4.$$

The field σ' decouples and can be dropped, so the modified theory is equivalent to the φ^4 theory. However, sometimes it can be useful to work out the Feynman rules and the diagrams from (2.87). In that case, σ has a propagator equal to 1, so it is a χ_a field with $a = 0$. The renormalizability by power counting still works. We just need to add an extra vertex φ^4 to the Lagrangian, because it is allowed by power counting. We multiply it by an independent coupling λ' and treat α as an independent coupling as well. In total, the renormalized Lagrangian \mathcal{L}'_I reads

$$\mathcal{L}'_{\text{IR}} = \frac{Z_\sigma}{2}\sigma^2 + i\mu^{\varepsilon/2}\alpha Z_\alpha Z_\sigma^{1/2} Z_\varphi \sigma \varphi^2 + \frac{\lambda'\mu^\varepsilon Z'_\lambda Z_\varphi^2}{4!}\varphi^4 = \frac{1}{2}\sigma'^2 + \frac{\lambda\mu^\varepsilon Z_\lambda}{4!}Z_\varphi^2\varphi^4$$

where $\sigma' = Z_\sigma^{1/2}\sigma + i\mu^{\varepsilon/2}\alpha Z_\alpha Z_\varphi \varphi^2$ and $\lambda Z_\lambda = \lambda' Z'_\lambda + 12\alpha^2 Z_\alpha$. The theory is equivalent to the ordinary massless φ^4 theory with the coupling $\lambda = \lambda' + 12\alpha^2$.

The no-miracle principle also implies that a renormalizable theory must contain *all* the Lagrangian terms that are not a priori forbidden. Indeed, assume that for some reason we start with a Lagrangian with some missing vertex \bar{v} . A divergent contribution \bar{c} of the same form will be generated by renormalization. To subtract it, it is necessary to go back to the classical Lagrangian and add \bar{v} , multiplied by a new coupling $\bar{\lambda}$. Once that is done, it is possible to remove \bar{c} by making a redefinition of $\bar{\lambda}$. We see that, because of renormalization, we are not free to choose the theory we like. Most classical theories make no sense at the quantum level, either because they do not contain enough vertices, the renormalizable ones, or because they contain nonrenormalizable vertices. Renormalization either guides us towards the right theory or blows up to (2.39). In this sense, it provides a way to *select* the theories.

Sometimes, the parameters of zero dimension are called *marginal*, those of positive dimensions *relevant* and those of negative dimensions *irrelevant*. This terminology refers to the low-energy behavior of the theory. For example, the parameters of negative dimensions multiply Lagrangian terms of dimensions larger than d , which are indeed negligible in the low-energy limit.

Instead, the parameters of positive dimensions multiply the terms that are more “relevant” at low energies. This terminology will be rarely used in this book.

2.6 Renormalizable theories

The list of the renormalizable theories depends on the spacetime dimension d . We start from four dimensions, where

$$d(n_B, n_F) = 4 - n_B - \frac{3}{2}n_F.$$

By locality, $d(n_B, n_F)$ must be non-negative, so n_B can be at most 4 and n_F can be at most 2. We have the following possibilities

(n_B, n_F)	(1, 0)	(2, 0)	(3, 0)
$d(n_B, n_F)$	3	2	1
Lagrangian terms	φ	$\varphi^2, \varphi \partial \varphi, (\partial \varphi)^2$	$\varphi^3, \varphi^2 \partial \varphi$
(n_B, n_F)	(4, 0)	(0, 2)	(1, 2)
$d(n_B, n_F)$	0	1	0
Lagrangian terms	φ^4	$\bar{\psi} \psi, \bar{\psi} \not{\partial} \psi$	$\varphi \bar{\psi} \psi$

The notation is symbolic, in the sense that we do not pay attention to where derivatives act and how indices are contracted. The most complicated bosonic interaction is φ^4 and the most complicated scalar-fermion interaction is the Yukawa vertex $\varphi \bar{\psi} \psi$. No fermion self-interaction is allowed.

The most general four-dimensional Lorentz invariant Lagrangian of scalar fields φ , vectors A and fermions ψ has the form

$$\begin{aligned}
\mathcal{L}_4 = & \frac{1}{2}(\partial_\mu A_\nu)^2 - \frac{\xi}{2}(\partial_\mu A_\mu)^2 + \frac{m_A^2}{2}A_\mu^2 + \frac{1}{2}(\partial_\mu \varphi)^2 + \frac{m_s^2}{2}\varphi^2 \\
& + \bar{\psi} \not{\partial} \psi + m_f \bar{\psi} \psi + \lambda_{1s} \varphi + \frac{\lambda_{3s}}{3!} \varphi^3 + \frac{\lambda_{4s}}{4!} \varphi^4 + \frac{g_3}{3!} A_\mu^2 (\partial_\nu A_\nu) \\
& + \frac{g'_3}{3!} A_\mu A_\nu \partial_\nu A_\mu + \frac{g_4}{4!} (A_\mu^2)^2 + Y_s \varphi \bar{\psi} \psi + Y_A A_\mu \bar{\psi} \gamma_\mu \psi + \frac{g_{2s}}{2} \varphi (\partial_\mu A_\mu) \\
& + \frac{g_{3s}}{3!} \varphi A_\mu^2 + \frac{g'_{3s}}{3!} \varphi^2 \partial_\nu A_\nu + \frac{g''_{3v}}{3!} A_\mu \varphi \partial_\mu \varphi + \frac{\lambda_{4sv}}{4!} \varphi^2 A_\mu^2,
\end{aligned}$$

where $\xi \neq 1$, plus fermionic terms equal to the listed ones with $\psi \rightarrow \gamma_5 \psi$, where γ_5 is the product of all the γ matrices. At $\xi \neq 1$ the vector propagator

behaves correctly at large momenta, even when the mass m_A does not vanish. However, we anticipate that at $\xi \neq 1$ the theory is not unitary, that is to say it propagates unphysical degrees of freedom. At $\xi = 1$, $m_A \neq 0$ the vector A_μ is of the Proca type, which has a bad behavior for large momenta. Instead, at $\xi = 1$, $m_A = 0$, the propagator does not exist. This is the case of the gauge theories, which will be treated in the next chapters. After a suitable “gauge-fixing” the propagator becomes well behaved and the theory can be proved to be renormalizable by power counting and unitarity.

Simple examples of renormalizable theories in four dimensions are the φ_4^4 theory (2.46), which is renormalized by (2.45) in the form (2.43), and, more generally, the Yukawa theory (2.17). Its bare action

$$S_B(\varphi_B, \psi_B) \equiv \int d^D x \left(\frac{1}{2} (\partial_\mu \varphi_B)^2 + \lambda_B \frac{\varphi_B^4}{4!} + \bar{\psi}_B \not{\partial} \psi_B + g_B \varphi_B \bar{\psi}_B \psi_B \right) \quad (2.88)$$

is renormalized by the map

$$\varphi_B = Z_\varphi^{1/2} \varphi, \quad \lambda_B = \lambda \mu^\varepsilon Z_\lambda, \quad \psi_B = Z_\psi^{1/2} \psi, \quad g_B = g \mu^{\varepsilon/2} Z_g,$$

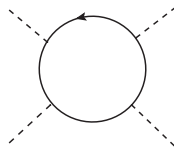
which gives the renormalized action

$$S_R(\varphi, \psi) \equiv \int d^D x \left(\frac{Z_\varphi}{2} (\partial_\mu \varphi)^2 + \lambda \mu^\varepsilon Z_\lambda Z_\varphi^2 \frac{\varphi^4}{4!} + Z_\psi \bar{\psi} \not{\partial} \psi + g \mu^{\varepsilon/2} Z_g Z_\varphi^{1/2} Z_\psi \varphi \bar{\psi} \psi \right). \quad (2.89)$$

We can use this example to illustrate what happens when we start from a theory with some missing vertices. Assume that we “forget” the φ^4 vertex and start with the Lagrangian

$$\mathcal{L} = \frac{1}{2} (\partial_\mu \varphi)^2 + \bar{\psi} \not{\partial} \psi + g \varphi \bar{\psi} \psi. \quad (2.90)$$

Then, consider the one-loop diagram



(2.91)

and its permutations, where the dashed lines refer to the scalars and the continuous ones to the fermions. It is easy to check that the divergent part

of (2.91) is nonvanishing. Thus, the theory (2.90) is not renormalizable as it stands. The missing φ^4 vertex must be added to the classical Lagrangian, and (2.90) becomes (2.17). Then, the redefinition of the φ^4 coupling can remove the divergent part of diagram (2.91). Thus, the theory (2.90) makes no sense at the quantum level, although it is a perfectly meaningful classical theory. Only (2.17) makes sense.

In three spacetime dimensions, we have

$$d(n_B, n_F) = 3 - \frac{1}{2}n_B - n_F - \frac{3-2a}{2}n_a,$$

so $n_B \leq 6$ and $n_F \leq 2$. The most complicated bosonic interaction is φ^6 and, again, no fermionic self-interaction is admitted. We have included a field χ_a , because in three dimensions there exist interesting bosonic vector fields with $a = 1/2$ and propagators $\sim k_\mu/k^2$ (Chern-Simons vectors). Their kinetic term reads

$$\mathcal{L}_{CS} = \frac{i}{2} A_\mu \partial_\rho A_\nu \varepsilon_{\mu\nu\rho}.$$

In this case, we have $n_{1/2} \leq 3$, and there can be vertices $A_\mu A'_\rho A''_\nu \varepsilon_{\mu\nu\rho}$ (with vectors of several types). Two boson-two fermion interactions $\varphi^2 \bar{\psi}\psi$, $A_\mu^2 \bar{\psi}\psi$, etc., are allowed. Apart from constraints coming from the statistics, the Chern-Simons fields behave like the fermions ψ . Summarizing,

(n_B, n_F)	(2, 0)	(4, 0)	(6, 0)	(2, 1)
$d(n_B, n_F)$	2	1	0	1
Lagrangian terms	$(\partial\varphi)^2$	$\partial\varphi^4$	φ^6	$\partial\varphi^2\psi$
(n_B, n_F)	(4, 1)	(0, 2)	(2, 2)	(0, 3)
$d(n_B, n_F)$	0	1	0	0
Lagrangian terms	$\varphi^4\psi$	$\psi\partial\psi$	$\varphi^2\psi^2$	ψ^3

where φ can stand for scalar fields and ordinary vector fields, while ψ can stand for fermions and Chern-Simons vectors. We have listed only the Lagrangian terms that have the largest powers of the fields and the largest numbers of derivatives. The missing terms are obtained from the listed ones by suppressing some powers of the fields and/or some derivatives.

In six dimensions

$$d(n_B, n_F) = 6 - 2n_B - \frac{5}{2}n_F,$$

which implies $n_B \leq 3$, $n_F \leq 2$. Moreover, for $n_F = 2$, n_B must vanish, so the fermions are free. It is said that they “decouple”, and so can be ignored for our present purposes. Hence, the only allowed interaction is φ^3 . However, the theory φ_6^3 is not physically interesting, since the potential φ^3 is not bounded from below. Vectors decouple also, by Lorentz invariance. In five dimensions the situation is the same as in six. In dimensions greater than six all the fields are free.

We see that only in dimensions smaller than or equal to four there exist physically acceptable interacting renormalizable theories. Unfortunately, gravity is not renormalizable in four dimensions. It is described by a spin-2 field, a symmetric tensor $h_{\mu\nu}$, which has derivative interactions of the form

$$\sim \kappa^n h^n \partial h \partial h,$$

which involve a coupling κ , the Newton constant, of dimension -1 in units of mass.

Newton’s constant is dimensionless in two dimensions, which suggests that gravity is power counting renormalizable there. However, it can be shown that general relativity in less than four dimensions contains no propagating graviton. In less than three dimensions gauge vectors have no propagating degree of freedom either. We discover that the renormalizable interactions are very few, which means that renormalizability is an extremely powerful criterion to select the theories. It is so restrictive that it almost selects the right dimension of spacetime: we have learned that four is the dimension with the largest variety of renormalizable theories. However, the fact that gravity is not renormalizable by power counting in $d > 2$ suggests that power counting renormalizability is not the final answer. A more profound renormalization principle must exist.

The renormalizable theories are those where the subtraction algorithm achieves its goal of removing all the divergences by redefining the fields and a finite number of independent parameters. Sometimes those theories are just called “renormalizable” in the literature. However, it should be kept in mind that there exist theories that are renormalizable by criteria different from power counting. Those theories will be studied the final chapters of this book.

Exercise 10 *Compute the one-loop renormalization of the massless scalar-fermion theory (2.17).*

Solution. The nontrivial divergent diagrams are

(2.92)

The calculation can be simplified by means of the tricks explained in this chapter. Note that the last diagram gives 6 identical contributions: a factor 2 comes from the orientation of the loop and a factor 3 comes from the permutations of the external legs. We find

$$Z_\varphi = 1 - \frac{4g^2}{(4\pi)^2\varepsilon}, \quad Z_\psi = 1 - \frac{g^2}{(4\pi)^2\varepsilon}, \quad Z_g = 1 + \frac{5g^2}{(4\pi)^2\varepsilon},$$

$$\lambda Z_\lambda = \lambda + \frac{1}{(4\pi)^2\varepsilon}(3\lambda^2 + 8g^2\lambda - 48g^4).$$

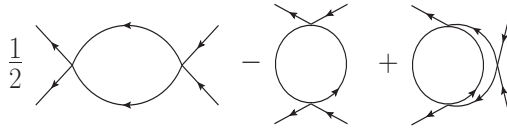
There is no diagram of order $g^2\lambda$, so $\lambda Z_\lambda Z_\varphi^2$ does not contain such type of contribution. Note that in general when the theory contains more couplings it may be not convenient to define vertex renormalization constants, such as Z_λ in this case, because they may contain negative powers of the couplings. Since λZ_λ is certainly polynomial, it is better to rewrite λZ_λ as $\lambda + \Delta_\lambda$, where Δ_λ collects the counterterms and is also polynomial.

Exercise 11 *Compute the one-loop renormalization of the four-fermion theory (1.102) in two dimensions, where ψ is a multiplet made of N copies of the basic spinor doublet.*

Solution. This theory is, in some respects, similar to the φ^4 theory in four dimensions. The Feynman rules are (1.103) with $\lambda \rightarrow \lambda\mu^\varepsilon$, where $\varepsilon = 2 - D$. There is no wave-function renormalization at one loop. The mass renormalization is given by a tadpole diagram, which turns out to be equal to $-\lambda(2N - 1)m\mu^\varepsilon/2$ times (2.6), where $2N - 1$ comes from evaluating the fermion loop. Expanding the left-hand side of (2.7) around two dimensions, we get

$$Z_m = 1 - \frac{(2N - 1)\lambda}{4\pi\varepsilon}.$$

The vertex renormalization is given by the diagrams



where the combinatorial factors and the signs due to the fermion exchanges are written explicitly. Observe that the first diagram is not multiplied by (-1) , since it does not contain a closed fermion loop. Instead, the third diagram has a plus sign, since the factor (-1) due to the closed fermion loop is compensated by another factor (-1) due to the permutation of two external identical fermions. Using the two-dimensional identity

$$(\gamma_\mu)^{\alpha\beta}(\gamma_\mu)^{\gamma\delta} - (\gamma_\mu)^{\gamma\beta}(\gamma_\mu)^{\alpha\delta} = -2(\delta^{\alpha\beta}\delta^{\gamma\delta} - \delta^{\alpha\delta}\delta^{\gamma\beta}),$$

where γ_μ are the first two Pauli matrices, we obtain

$$Z_\lambda = 1 - \frac{(N-1)\lambda}{2\pi\varepsilon}. \quad (2.93)$$

Exercise 12 Write the Lagrangian of the previous exercise in the equivalent form

$$\mathcal{L} = \bar{\psi}(\not{\partial} + m)\psi + \sqrt{\frac{\lambda}{2}}\mu^{\varepsilon/2}\sigma\bar{\psi}\psi + \frac{1}{2}\sigma^2,$$

having introduced an auxiliary field σ . Renormalize the theory in this form at one loop, and check the results already found.

Solution. The divergent one-loop diagrams are the first three of the list (2.92), plus a tadpole (the fermion loop with one external leg σ). The calculation is straightforward and gives the renormalized Lagrangian

$$\begin{aligned} \mathcal{L}_R = & \bar{\psi}\not{\partial}\psi + \left(1 + \frac{\lambda}{4\pi\varepsilon}\right) \left(m + \sqrt{\frac{\lambda}{2}}\mu^{\varepsilon/2}\sigma\right) \bar{\psi}\psi \\ & + \frac{1}{2}\sigma^2 \left(1 + \frac{\lambda N}{2\pi\varepsilon}\right) + \sqrt{\frac{\lambda}{2}}\frac{mN}{\pi\varepsilon}\mu^{-\varepsilon/2}\sigma. \end{aligned} \quad (2.94)$$

Integrating σ away, we retrieve the results of the previous exercise. Note that to have (2.94) real, the coupling λ must be positive. This is the reason why we have put a minus sign in front of the four fermion vertex of the Lagrangian

(1.102). If the fields were bosonic, that minus sign would be wrong. Instead, for the reason just explained it is the right sign for fermionic fields. \square

The last two exercises teach us that, if we make a change of field variables, the theory remains renormalizable, but the renormalization organizes itself in a different way. For example, the Lagrangian (1.102) has nontrivial renormalization constants for m and λ , at one loop, while (2.94) also has a renormalization constant for σ , and contains a σ linear term. The two renormalized Lagrangians are mapped into each other by a renormalized change of field variables. For the moment we content ourselves with these observations. We will say more later, where we investigate general changes of field variables in quantum field theory.

Exercise 13 Find the renormalized change of field variables that maps (2.94) into the renormalized version of (1.102).

Solution. It is

$$\sigma = \left(1 - \frac{\lambda N}{4\pi\varepsilon}\right) \sigma' - \left(1 - \frac{(2N-1)\lambda}{4\pi\varepsilon}\right) \sqrt{\frac{\lambda}{2}} \mu^{\varepsilon/2} \bar{\psi}\psi \\ - \sqrt{\frac{\lambda}{2}} \mu^{-\varepsilon/2} \left(1 - \frac{\lambda N}{2\pi\varepsilon}\right) \frac{mN}{\pi\varepsilon},$$

plus higher orders. Indeed, (2.94) becomes

$$\bar{\psi}\phi\psi + mZ_m\bar{\psi}\psi - \frac{\lambda}{4}\mu^\varepsilon Z_\lambda(\bar{\psi}\psi)^2 + \frac{1}{2}\sigma'^2,$$

plus a constant, plus higher orders. This is the renormalized version of (1.102), plus a quadratic term that decouples (and is not renormalized).

2.7 Composite fields

Composite fields are defined as products of elementary fields, and their derivatives, in the same spacetime point. Sometimes they are also called “operators”, or “composite operators”, although strictly speaking no operator is involved in the functional-integral approach. Being just nontrivial monomials of the fields and their derivatives, composite fields are local. Sometimes it is useful to consider also local functionals, that is to say the integrals of

composite fields over spacetime. Local functionals are also called integrated operators.

The renormalization of a composite field is in general not related in an obvious way to the renormalization of its component fields, and has to be calculated anew. The simplest example of a composite field is $\varphi^2(x)$ in the φ^4 theory. The renormalization of φ^2 is unrelated to the renormalization of φ . Specifically, the correlation functions that contain insertions of φ^2 can be made convergent with a new renormalization constant, Z_{φ^2} , which has no relation with Z_φ .

Let us recall that Z_φ renormalizes the divergences of the correlation functions where scalar fields are inserted at different spacetime points, e.g.

$$G_4(x, y, z, w) = \langle \varphi(x) \varphi(y) \varphi(z) \varphi(w) \rangle, \quad (2.95)$$

with $x \neq y, z, w$, etc. Instead, consider the φ^2 two-point function

$$\langle \varphi^2(x) \varphi^2(y) \rangle. \quad (2.96)$$

Recall that the correlation functions have to be meant as distributions. In a distribution it often makes no sense to take the limit of coinciding points. Therefore, (2.96) is *not* the $z \rightarrow x, w \rightarrow y$ limit of (2.95).

For a while, we write formulas assuming that we are dealing with a single composite field. Later we generalize our arguments by eliminating this restriction.

We must distinguish bare and renormalized composite fields. The bare composite fields are denoted by \mathcal{O}_B and are just the products of the bare factors. For example, the bare operator $\varphi^2(x)$ is just the product of two bare scalar fields in x , i.e. $\varphi_B^2(x)$. The renormalized composite fields are denoted by \mathcal{O}_R , or by writing the composite field between square brackets, such as $[\varphi^2(x)]$, to distinguish it from $\varphi^2(x)$.

Bare and renormalized operators are related by new renormalization constants $Z_{\mathcal{O}}$,

$$\mathcal{O}_B = Z_{\mathcal{O}} \mathcal{O}_R. \quad (2.97)$$

For example, we have $\varphi_B^2 = Z_{\varphi^2} [\varphi^2]$. On the other hand, we know that $\varphi_B = Z_\varphi^{1/2} \varphi$, hence

$$[\varphi^2] = Z_{\varphi^2}^{-1} Z_\varphi \varphi^2. \quad (2.98)$$

This formula emphasizes that the renormalized operator φ^2 does not coincide with the square of the renormalized field φ , unless $Z_{\varphi^2} = Z_\varphi$, which is in general not true.

Thus, the renormalized two-point function (2.96) reads

$$G_R(x, y) = \langle [\varphi^2(x)] [\varphi^2(y)] \rangle = Z_{\varphi^2}^{-2} Z_\varphi^2 \langle \varphi^2(x) \varphi^2(y) \rangle, \quad (2.99)$$

at $x \neq y$. Here $Z_{\varphi^2}^{-2}$ cancels the extra divergences due to the pairs of coinciding points.

A composite field can be described as a vertex. Its structure is visible when it is written in terms of renormalized (elementary) fields. Formula (2.98) tells us that $[\varphi^2]$ is a vertex with two φ legs. To exhibit the vertex associated with a composite field, we couple the composite fields to appropriate sources, which we denote by L , and add them to the action. At the bare level, we just need to add

$$-L_B \mathcal{O}(\varphi_B) \quad (2.100)$$

to the bare Lagrangian. At the renormalized level, we have to add

$$-L \mathcal{O}(\varphi) + f(\varphi, L), \quad (2.101)$$

where $f(\varphi, L)$ denotes counterterms that are at least quadratic in L . They renormalize the divergences of the correlation functions that contain more than one insertion of $\mathcal{O}(\varphi)$, such as those of (2.99) at $y \rightarrow x$ (see below).

The generating functionals Z , W and Γ are defined as usual. Now, they depend on the sources L , besides J or Φ . The correlation functions that carry \mathcal{O} insertions can be obtained by differentiating the generating functionals with respect to L .

Since the bare and renormalized actions are the same quantity, written in terms of different variables, we also have

$$L_B \mathcal{O}_B = L \mathcal{O}_R. \quad (2.102)$$

Let Z_L denote the renormalization constant of L ($L_B = Z_L L$). We clearly have

$$Z_L = Z_{\mathcal{O}}^{-1}.$$

In the case of φ^2 we have the new vertex $-L\varphi^2$, with one leg L and two legs φ . Once $L\mathcal{O}$ is written in terms of renormalized fields, the new

vertex can be treated as any other vertex, L being considered as an external, nonpropagating field. The Feynman rules are supplemented with the vertices generated by (2.100) or (2.101). We add

$$-\frac{1}{2} \int d^D x L_B(x) \varphi_B^2(x) = -\frac{Z_L Z_\varphi}{2} \int d^D x L(x) \varphi^2(x) \quad (2.103)$$

to the bare action (2.46) and

$$\begin{array}{c} \diagup \\ \text{---} L \text{---} \\ \diagdown \end{array} = 1 \quad (2.104)$$

to the Feynman rules. Observe that the source L has dimension two in units of mass, so by power counting the action must be completed with a term that is quadratic in L . We write it as

$$-\frac{1}{2a_B} \int L_B^2 = -\frac{\mu^{-\varepsilon}}{2} \left(\frac{1}{a} + \delta_a \right) \int L^2. \quad (2.105)$$

We have written

$$a_B = \frac{a\mu^\varepsilon Z_L^2}{1 + a\delta_a} \equiv \mu^\varepsilon a Z_a.$$

The terms quadratic in L are important when we consider multiple insertions of composite fields, as in (2.99).

The generating functional becomes

$$Z(J, L) = e^{W(J, L)} = \int [d\varphi] \exp \left(-S(\varphi, L) + \int J\varphi \right), \quad (2.106)$$

where in our example

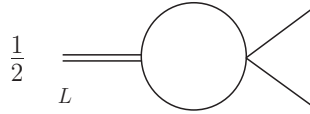
$$S(\varphi, L) = S(\varphi) - \frac{Z_L Z_\varphi}{2} \int L \varphi^2 - \frac{\mu^{-\varepsilon}}{2} \left(\frac{1}{a} + \delta_a \right) \int L^2$$

and $S(\varphi)$ is given by (2.16). Since L is an external field, the Feynman diagrams can have external legs L , but no internal legs L . There are only two overall divergent correlation functions that contain φ^2 insertions, namely the “vertex”

$$\langle \varphi^2(x) \varphi(y) \varphi(z) \rangle \quad (2.107)$$

and the two-point function (2.96). The counterterms associated with them give Z_L and Z_a , which we now calculate at one loop.

In the case (2.107) we have

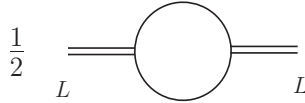


$$(2.108)$$

It gives $(-\lambda\mu^\varepsilon/2)\mathcal{I}_D(k, m)$, where \mathcal{I}_D is given by (2.28). We find

$$Z_L = Z_{\varphi^2}^{-1} = 1 + \frac{\lambda}{16\pi^2\varepsilon} + \mathcal{O}(\lambda^2).$$

The φ^2 two-point function (2.96) is given at one loop by the diagram



$$(2.109)$$

which gives $(1/2)\mathcal{I}_D(k, m)$, so

$$\delta_a = -\frac{1}{16\pi^2\varepsilon} + \mathcal{O}(\lambda). \quad (2.110)$$

Exercise 14 Calculate the functionals $W(J, L)$ and its Legendre transform $\Gamma(\Phi, L)$ with respect to J for a free massless scalar field in the presence of the composite field φ^2 .

Solution. We have the renormalized generating functional

$$e^{W(J, L)} = \int [d\varphi] \exp \left(-\frac{1}{2} \int \left\{ (\partial_\mu \varphi)^2 - L\varphi^2 - \frac{\mu^{-\varepsilon}}{a} (1 + a\delta_a) L^2 \right\} + \int J\varphi \right) \quad (2.111)$$

where δ_a is given by (2.110) with $\lambda = 0$. The functional integral is easy to work out, since it is Gaussian. The source L plays the role of a spacetime dependent mass. We obtain

$$\begin{aligned} W(J, L) &= \frac{1}{2} \int \left[J \frac{1}{-\square - L} J + \mu^{-\varepsilon} \left(\frac{1}{a} + \delta_a \right) L^2 \right] - \frac{1}{2} \text{tr} \ln(-\square - L), \\ \Gamma(\Phi, L) &= \frac{1}{2} \int \left[(\partial_\mu \Phi)^2 - L\Phi^2 - \mu^{-\varepsilon} \left(\frac{1}{a} + \delta_a \right) L^2 \right] + \frac{1}{2} \text{tr} \ln(-\square - L), \end{aligned} \quad (2.112)$$

where

$$\Phi = \int \frac{1}{-\square - L} J. \quad (2.113)$$

□

More generally, the renormalization of a composite field is not just multiplicative, but involves other composite fields. It is then said that composite fields “mix” under renormalization. Then, formulas (2.97), (2.100), (2.101) and (2.102) must be interpreted in a vector-matrix form.

Call $\mathcal{O}_B^I = \mathcal{O}^I(\varphi_B)$ the bare composite fields, \mathcal{O}_R^I the renormalized ones, L_B^I and L^I their sources, and Z^{IJ} the matrix of the renormalization constants, such that $\mathcal{O}_B^I = Z^{IJ} \mathcal{O}_R^J$. The linear terms in L that must be added to the Lagrangian read

$$L_B^I \mathcal{O}_B^I = L^I \mathcal{O}_R^I = L^I (Z^{-1})^{IJ} \mathcal{O}_B^J (Z_\varphi^{1/2} \varphi). \quad (2.114)$$

These terms are sufficient to describe the correlation functions that contain single insertions of the composite fields and arbitrary insertions of the elementary fields. Multiple insertions of composite fields, such as (2.96), are renormalized by terms that contain higher powers of the sources L :

$$S(\varphi, L) = S(\varphi) - L^I (Z^{-1})^{IJ} \mathcal{O}_B^J (Z_\varphi^{1/2} \varphi) - L^I \Delta_{IJ}(\varphi) L^J + \cdots$$

Organize the \mathcal{O}^I s in a row such that the composite fields of equal dimensions are close to one another and the composite fields of higher dimensions follow those of lower dimensions. Since the theory, by the renormalizability assumption, contains only parameters of non-negative dimensions, a composite field can only mix with composite fields of equal or smaller dimensions. For this reason, the matrix Z_{IJ} is block lower triangular. Each diagonal block encodes the renormalization mixing of the composite fields of equal dimensions. The off-diagonal blocks encode the mixing among composite fields of different dimensions.

Let us comment on the multiple insertions of composite fields, i.e. the terms of S_L that contain quadratic or higher powers of the sources L_I . In general, the renormalized action S_L is not polynomial in L_I . Indeed, if the dimension of \mathcal{O}_B^I is large, the dimension of L_{IB} is negative. Then, infinitely many local counterterms with high powers of the sources L_{IB} and their derivatives can be constructed. By the no-miracle principle of renormalization, S_L must contain all of them. This means, in particular, that,

strictly speaking, S_L is not even local, since it contains terms with arbitrarily many derivatives. However, it is perturbatively local, since each order of the perturbative expansion is local. At any rate, we do not need to worry, because we are not required to resum the L_I powers. Every correlation function contains a given, finite number of composite-field insertions, so it can be calculated by truncating S_L to appropriate finite powers of L_I . Every truncation is local and polynomial. Thus, we can still call S_L a local functional, according to the extended definition of local functionals introduced before.

Exercise 15 Calculate the one-loop renormalization of the composite field $\mathcal{O}(\varphi) = \varphi^4$ and the composite fields that mix with it, in the massless φ^4 theory.

2.8 Maximum poles of diagrams

An L -loop diagram has at most poles $1/\varepsilon^L$ of order L . However, sometimes the order of its maximum pole can be considerably smaller than L . For example, exercise (5) shows that the diagram (k) of figure (2.59) has two loops, but it has only a simple pole at $m = 0$. Here we prove a general theorem bounding the maximum pole of a diagram.

We are interested only in the UV divergences of the quantum theory, and their renormalization. Then it is consistent to treat the mass terms, if present, as vertices of two legs, the propagator being just the massless one. Any other dimensionful parameter that multiplies a quadratic term must be treated in a similar way. To avoid IR problems in the intermediate steps, it is convenient to calculate the UV divergences of the Feynman diagrams by means of deformed propagators that are equipped with an artificial mass δ , and let δ tend to zero at the end, as explained in formula (2.76). The tadpoles are loops with a single vertex, and vanish identically. Instead, the loops with at least two vertices are not tadpoles (even if one of the vertices is a two-leg “mass” vertex) and may give nontrivial divergent contributions.

Theorem 3 *The maximum pole of a diagram with V vertices and L loops is at most*

$$\frac{1}{\varepsilon^{m(V,L)}},$$

where

$$m(V, L) = \min(V - 1, L).$$

Proof. We prove the statement inductively in V and, for fixed V , inductively in L . The diagrams with $V = 1$ and arbitrary L are tadpoles, which vanish identically and trivially satisfy the theorem. Suppose that the statement is true for $V < \bar{V}$, $\bar{V} > 1$, and arbitrary L . Consider the diagrams that have \bar{V} vertices. For $L = 1$ the maximal divergence is $1/\varepsilon$, so the theorem is satisfied. Proceed inductively in L , i.e. suppose that the theorem is also satisfied by the diagrams that have \bar{V} vertices and $L < \bar{L}$ loops, and consider the diagrams $G_{\bar{V}, \bar{L}}$ that have \bar{V} vertices and \bar{L} loops. If $G_{\bar{V}, \bar{L}}$ has no subdivergence, its divergence is at most a simple pole, which satisfies the theorem. Higher-order poles are related to the subdivergences of $G_{\bar{V}, \bar{L}}$ and can be classified by replacing the subdiagrams by their counterterms. Consider the subdiagrams $\gamma_{v,l}$ of $G_{\bar{V}, \bar{L}}$ that have l loops and v vertices. Clearly, $1 \leq l < \bar{L}$ and $1 \leq v \leq \bar{V}$. By the inductive hypothesis, the maximal divergence of $\gamma_{v,l}$ is a pole of order $m(v, l)$. Contract the subdiagram $\gamma_{v,l}$ to a point and multiply by $1/\varepsilon^{m(v,l)}$. A diagram with $\bar{V} - v + 1 \leq \bar{V}$ vertices and $\bar{L} - l < \bar{L}$ loops is obtained, whose maximal divergence is at most a pole of order $m(v, l) + m(\bar{V} - v + 1, \bar{L} - l)$, if we take the factor $1/\varepsilon^{m(v,l)}$ into account. The inequality

$$m(v, l) + m(\bar{V} - v + 1, \bar{L} - l) \leq m(\bar{V}, \bar{L}),$$

which can be derived case by case, proves that the maximal divergence of $G_{\bar{V}, \bar{L}}$ associated with $\gamma_{v,l}$ satisfies the theorem. Since this is true for every subdiagram $\gamma_{v,l}$, the theorem follows for $G_{\bar{V}, \bar{L}}$. By induction, the theorem follows for every diagram. \square

Recall that this theorem holds after expanding in powers of the dimensional parameters that are contained in the propagators. The diagram (k) of figure (2.59) has $V = 2$ and $L = 2$, so $m(V, L) = 1$: indeed, its maximum pole in the massless limit is a simple pole instead of a double pole. It can be easily checked that at $m \neq 0$ the diagram has a double pole proportional to the squared mass. If we view the mass term as a two-leg vertex, that pole arises from the diagram obtained from (k) by attaching the two-leg vertex to one internal line. In that case, we have $V = 3$ and $L = 2$, so $m(V, L) = 2$, in agreement with the theorem.

2.9 Subtraction prescription

When we subtract a simple pole $1/\varepsilon$, we can equivalently subtract an arbitrary finite constant together with it, as shown in formula (2.32). Similarly, when we subtract a multiple pole $1/\varepsilon^n$, we can affect the less singular poles:

$$\frac{1}{\varepsilon^n} \rightarrow \frac{1}{\varepsilon^n} + \sum_{i=1}^n \frac{c_i}{\varepsilon^{n-i}}.$$

Sometimes, a prescription, called *subtraction scheme*, is adopted to associate finite constants c_i to the subtractions of the poles, according to a convenient rule. The *minimal* subtraction (MS) scheme is the convention according to which the poles are subtracted with no finite constants attached.

By locality of counterterms, the scheme arbitrariness can only affect the local terms. This means that it amounts to a finite redefinition of the constants that multiply the vertices and the kinetic terms contained in the Lagrangian. Since those constants, including the field normalizations, are arbitrary anyway, the arbitrariness amounts to a finite reparametrization of the theory. In any case, it does not affect the physical quantities.

In other words, renormalization is an infinite reparametrization of the theory, while a change of subtraction scheme is a finite reparametrization.

To be more explicit, consider the vertex φ^4 and its one-loop counterterm (2.32):

$$\lambda \mu^\varepsilon \frac{\varphi^4}{4!} + 3\lambda^2 \mu^\varepsilon \left(\frac{1}{16\pi^2 \varepsilon} + c_1 \right) \frac{\varphi^4}{4!}. \quad (2.115)$$

Now, move the arbitrary constant c_1 from the counterterm to the vertex φ^4 and define

$$\lambda'(\lambda) = \lambda + 3\lambda^2 c_1 + \mathcal{O}(\lambda^3). \quad (2.116)$$

We can rewrite (2.115) as

$$\mu^\varepsilon \lambda' \frac{\varphi^4}{4!} + \mu^\varepsilon \frac{3\lambda'^2}{16\pi^2 \varepsilon} \frac{\varphi^4}{4!} + \mathcal{O}(\lambda'^3). \quad (2.117)$$

We see that the finite reparametrization (2.116) converts the arbitrary subtraction (2.115) to the minimal form (2.117). It is always possible to make a similar rearrangement.

From the experimental point of view, the arbitrariness disappears when enough physical quantities are measured, and the theory is uniquely determined. Specifically, in the massive φ^4 theory three independent quantities

need to be measured. From them, the values of m and λ and the φ normalization can be derived, after which the theory is uniquely determined. Observe that the parameter m needs not be identified with the physical mass, sometimes denoted with m_{ph} . Since m_{ph} can only be a finite function of m and λ , it is determined once m and λ are.

In the minimally subtracted λ' parametrization (2.117), the theory does not depend on c_1 , so it is uniquely determined once λ' is measured (together m and the normalization of the field). On the other hand, in the nonminimally subtracted λ parametrization (2.115) there appears to be an additional arbitrary constant c_1 , so it seems that an extra measurement is necessary. This is just a blunder, because after the three measurements mentioned above, c_1 disappears from all the physical quantities.

The matter can be better explained as follows. Consider some physical quantity. Write it as a function $f(\lambda)$ of λ in the first scheme and a function $f'(\lambda')$ of λ' in the second scheme. When we change the scheme, we do not change just λ , but also the form of the function f of λ . The two changes compensate each other, so that the physical results remain the same, that is to say

$$f(\lambda) = f'(\lambda').$$

Check for example (2.115) and (2.117): the coupling changes, but also the function multiplying φ^4 changes, so that (2.115) and (2.117) coincide. So, if an experimental measurement gives $\lambda' = \ell$ in the second scheme, where ℓ is some number that we assume to be small, a measurement in the first scheme must give the number

$$\lambda = \ell - 3\ell^2 c_1 + \mathcal{O}(\ell^3),$$

whatever the value of c_1 is.

2.10 Regularization prescription

So far we have mostly worked using the dimensional regularization, but equivalent results can be obtained using any regularization technique we like. Now we prove that changing the regularization technique is equivalent to change the subtraction scheme, so it has no physical consequence. To this purpose, it is helpful to clarify the definition of regularization technique.

Definition 3 *We call naïve (or formal) limit, the limit in which the regularization parameters are removed by keeping the bare fields and parameters fixed.*

We emphasize that, in spite of its name, the naïve limit is a rigorous notion. The naïve limit of the action is the classical action. The naïve limit of the correlation functions is in general ill defined, because of the divergences.

Definition 4 *We call physical limit the limit in which the regularization parameters are removed by keeping the renormalized fields and parameters fixed.*

The physical limit of the action is ill defined, but the physical limit of the correlation functions exists.

Consider a quantum field theory \mathcal{T} , defined by an action $S(\varphi)$ and a functional measure $[d\varphi]$.

Definition 5 *A regularized theory for \mathcal{T} is a deformed theory $\mathcal{T}_{\mathcal{R}}$, defined by a deformed action $S_{\mathcal{R}}(\varphi)$ and a deformed functional measure $[d_{\mathcal{R}}\varphi]$, such that: (i) all the regularized diagrams are convergent; (ii) the propagators and the vertices tend to the ones of \mathcal{T} in the naïve limit; and (iii) all the diagrams, or derivatives of diagrams, that are convergent at the unregularized level are recovered by taking the naïve limits of their regularized versions.*

Now, consider an integral $\int f$ and define two regularized versions of it,

$$\int f_1(\Lambda_1) < \infty, \quad \int f_2(\Lambda_2) < \infty,$$

Λ_1 and Λ_2 denoting some regularization parameters. We just call them cut-offs and assume that they are removed by sending them to infinity. By definition, we must have

$$\lim_{\Lambda_1 \rightarrow \infty} f_1(\Lambda_1) = \lim_{\Lambda_2 \rightarrow \infty} f_2(\Lambda_2) = f. \quad (2.118)$$

Indeed, the integrands contain vertices and propagators, namely ingredients inherited from the classical action, so must tend to f in the naïve limit. However, we cannot extend the naïve limit to the integrals, because they might be divergent (this is the reason why the limit is called naïve or formal).

Expanding for large $\Lambda_{1,2}$, we can write

$$\int f_i(\Lambda_i) = I_{\text{div}}(\Lambda_i) + I_{\text{finite}} + I_{\text{ev}}(\Lambda_i), \quad (2.119)$$

where $i = 1, 2$, while I_{div} collects the terms that diverge, I_{ev} those that tend to zero and I_{finite} those that have finite limits.

We know that (assuming that the subdivergences have been subtracted with the usual algorithmic procedure) if we take a sufficient number of derivatives with respect to the external momenta k , the integrals become convergent. This property is independent of the regularization technique, to the extent that it also holds for the unregularized integral, namely there exists an n such that

$$\int \frac{\partial^n f}{\partial k^n} < \infty.$$

Now, because of (2.118), we also have

$$\lim_{\Lambda_1 \rightarrow \infty} \frac{\partial^n}{\partial k^n} f_1(\Lambda_1) = \lim_{\Lambda_2 \rightarrow \infty} \frac{\partial^n}{\partial k^n} f_2(\Lambda_2) = \frac{\partial^n f}{\partial k^n}.$$

Integrate each side of this equation. The first two integrals can be interchanged with the limits, which gives

$$\lim_{\Lambda_1 \rightarrow \infty} \int \frac{\partial^n}{\partial k^n} f_1(\Lambda_1) = \lim_{\Lambda_2 \rightarrow \infty} \int \frac{\partial^n}{\partial k^n} f_2(\Lambda_2) = \int \frac{\partial^n f}{\partial k^n} < \infty.$$

In the first two expressions we can also interchange the integrals and the derivatives. So doing, we obtain

$$\lim_{\Lambda_1 \rightarrow \infty} \frac{\partial^n}{\partial k^n} \int f_1(\Lambda_1) - \lim_{\Lambda_2 \rightarrow \infty} \frac{\partial^n}{\partial k^n} \int f_2(\Lambda_2) = 0.$$

Using (2.119), we get

$$\frac{\partial^n}{\partial k^n} I_{1\text{div}}(\Lambda_1) = \frac{\partial^n}{\partial k^n} I_{2\text{div}}(\Lambda_2) = 0, \quad \frac{\partial^n}{\partial k^n} (I_{1\text{finite}} - I_{2\text{finite}}) = 0.$$

The first formula is just the statement that counterterms are local with any regularization technique. The second formula, instead, states that the finite parts, calculated using two different regularizations, can differ at most by local terms:

$$I_{1\text{finite}} = I_{2\text{finite}} + \text{local}.$$

If the theory is renormalizable, such local terms are of the types already present in the Lagrangian, so they amount to a scheme change, but the physical quantities are unaffected. This concludes the proof.

Sometimes it is useful to regularize different classes of diagrams in different ways, or can be convenient to introduce multiple cutoffs Λ_i . Divergences expressed in terms of different cutoffs can be identified, up to local terms. The cutoffs Λ_i can be removed in different orders, e.g. $\Lambda_1 \rightarrow \infty$ followed by $\Lambda_2 \rightarrow \infty$, or $\Lambda_2 \rightarrow \infty$ followed by $\Lambda_1 \rightarrow \infty$. When the limits are interchanged, the results can differ at most by local terms, i.e. again a scheme change, but the physical quantities are always the same.

Ultimately, we have an enormous freedom. We can regularize a theory as a whole, or diagram by diagram. We can use one cutoff or many cutoffs, and we can remove the cutoffs in the order we like. We can even use a different regularization technique and a different subtraction scheme for each diagram. No matter how we regularize the theory, the physical results always come out right. The core of quantum field theory is finite and regularization independent: the divergences are confined to the “superficial” parts of the integrals, so to speak, since they are killed by a finite number of derivatives.

Different regularization techniques can demand very different amounts of effort. If we want to better keep track of what we do, it is convenient to use a simple regularization technique, with one or two cutoffs, defined on the theory as a whole.

2.11 Comments about the dimensional regularization

Some people use to say that the dimensional regularization “misses something” or “has problems of internal consistency”, because integrals such as (2.12) are set to zero and the powerlike divergences disappear, or because of other caveats that we will mention later.

The truth is that the dimensional regularization does not miss anything and has no problems of internal consistency. Actually, it is the most powerful regularization technique developed so far. It is very convenient both to make calculations (to the extent that the renormalization of QCD has been worked out to four loops and the one of the standard model to three loops) and to

prove theorems to all orders. Under both respects, no other regularization technique is even comparable with the dimensional one.

One of its virtues is that it smartly chooses a subtraction scheme where the powerlike divergences are automatically absent. Actually, it allows us to prove that, no matter what regularization technique we use, the powerlike divergences can always be subtracted away just as they come, without leaving any remnants. Later we will better understand what this means, studying the renormalization group. For the moment, it is sufficient to say that the powerlike divergences are completely scheme dependent, and can be washed away with a smart scheme choice, while the logarithmic divergences are only partially scheme dependent. The logarithmic divergences do contain physical information, while the powerlike divergences are devoid of any physical meaning.

Two main kinds of assumptions inspire some people to take unreasonable attitudes towards the dimensional regularization. One assumption is that behind empty space there should be a sort of lattice, or aether. The analytic way to regularize the integrals is not intuitive, they say, while a lattice spacing is supposed to be more “physical”. We leave to them to explain why a regularization should be intuitive, or “physical”, and what it is supposed to mean. We just observe that sometimes hypothetical links with condensed matter physics may be inspiring, but other times they may put us on the wrong track. More generally, there is no reason to assume that the human intuition (which is always the product of our interaction with a classical environment) should guide us. It may be helpful in some cases, misleading in others. Once we have given up the correspondence principle almost completely, we can live without intuition.

Another assumption is that the ultimate theory should be finite, that is to say a theory with no divergences. In that case, the powerlike divergences are not really divergences, but physical quantities that depend on a large energy scale and grow polynomially with it. The assumption that the final theory be finite turns out to be appealing to some people (for quite subjective and “human” reasons), but rather restrictive. Having learned that we can renormalize the divergences away, we no longer need to require that they are absent from the start. If one insists that the final theory must be finite, he/she should explain why we can make sense of theories that are not finite, and why we should privilege a small subset of the theories we can

work with, and ignore the other ones. Somebody explicitly advocates aesthetic criteria to answer these questions. We do not feel necessary to stress that such arguments are completely meaningless in physics. Other people try to disguise their arbitrary requirements under suspicious conditions of “simplicity”. Certainly, simplicity can be advocated for practical purposes. It cannot, however, be advocated to discriminate what is physical from what is not: that part pertains to nature.

A more serious point about the dimensional regularization is that it is just perturbative. However, at present we do not know how to define the functional integral nonperturbatively, so this problem goes beyond the dimensional regularization itself.

2.12 About the series resummation

We have stressed several times that our task is to define the functional integral as a perturbative expansion. We have converted the functional integral of the interacting theory into an infinite sum of functional integrals of the free-field theory, since those are the only ones we can deal with. Each property we use must be understood in the same spirit. For example, when we use that the integral of a functional total-derivative vanishes, we mean that each term of the perturbative expansion vanishes. At this level, the perturbative expansion must be regarded as a sequence, a list of terms, not as a series that should be summed. Indeed, our primary objective is to define the terms of the sequence and check that they are consistent with the key physical and mathematical requirements. As we have seen, this task already raises nontrivial problems. Several other difficulties will appear in gauge theories and dealing with anomalies. It makes no sense to investigate the summation properties, before defining the terms of the sequence.

In various cases, the sum of the perturbative series might not exist, at least naïvely. This, however, is not necessarily a limitation. It might just mean that different ways to organize the sum can give different results. Then, we must classify the resummation prescriptions that make physical sense. There might be more meaningful resummation prescriptions, each of them leading to a different physical theory, with the same perturbative expansion. Recalling that, so far, we have not been scared away by the divergences (and now we appreciate what we would have missed if we had), there is no point

in worrying about a problem that is not even there, yet.

We will actually see that whenever we have control on the perturbative expansion to arbitrarily high orders (such as in the cases of the anomalies, the renormalization-group flow and the conformal fixed points), the series does make sense, or the theory itself provides a natural resummation prescription. For example, there are anomalies that can be calculated exactly, since they receive no corrections beyond one loop (if we are careful enough, in a sense that will be specified, which includes choosing an appropriate subtraction scheme).

In this book, we make no attempt to define the functional integral beyond its perturbative expansion, unless that means searching for the physical prescriptions that allow us to resum the perturbative expansion when possible.

Chapter 2 — References

- [1] The analytic regularization was introduced in 1964 by three Argentinian physicists, Bollini, Giambiagi and Gonzáles Domínguez, with the paper C.G. Bollini, J.J. Giambiagi and A. Gonzáles Domínguez, *Analytic regularization and the divergences of quantum field theories*, Nuovo Cim. 31 (1964) 550.
- [2] Elaborating on those ideas, Bollini and Giambiagi also invented the dimensional regularization in 1971:
C.G. Bollini and J.J. Giambiagi, *The number of dimensions as a regularizing parameter*, Nuovo Cim. 12B (1972) 20;
C.G. Bollini and J.J. Giambiagi, *Lowest order divergent graphs in ν -dimensional space*, Phys. Lett. B40, 566 (1972).
- [3] Right away, the dimensional regularization was elaborated in great detail by 't Hooft and Veltman:
G. 't Hooft and M. Veltman, *Regularization and renormalization of gauge fields*, Nucl. Phys. B 44 (1972) 189.
- [4] A classic paper that contains mathematical aspects of the dimensional regularization is
P. Breitenlohner and D. Maison, *Analytic renormalization and the action principle*, Commun. Math. Phys. 52 (1977) 11
- [5] The dimensional regularization of chiral theories presents various annoying features. Several modifications have been studied in the literature. A recent deformation that has the virtue of simplifying the proofs of properties to all orders can be found in

D. Anselmi, *Weighted power counting and chiral dimensional regularization*, Phys. Rev. D 89 (2014) 125024, 14A2 Renormalization.com and arXiv:1405.3110 [hep-th]

- [6] The theorem of section 2.8 first appeared in D. Anselmi, *Renormalization of a class of nonrenormalizable theories*, JHEP 0507 (2005) 077, 05A1 Renormalization.com and arXiv:hep-th/0502237 [hep-th].

Chapter 3

Renormalization group

In this chapter we begin the exploration of the physical consequences of renormalization. Our considerations are very general, although we often illustrate them using specific models. We start by comparing the bare and renormalized actions

$$S_B(\varphi_B, \lambda_B, m_B^2, L_B) = S_R(\varphi, \lambda, m^2, L, \mu) \quad (3.1)$$

of a theory of fields, where φ denote the fields, λ and m the dimensionless and the dimensionful parameters, respectively, and L are the sources for the composite fields. We do not really need to keep λ and m distinct, but for the moment it is convenient to do so. Similarly, the relation between the bare and the renormalized Γ functionals is

$$\Gamma_B(\Phi_B, \lambda_B, m_B^2, L_B) = \Gamma_R(\Phi, \lambda, m^2, L, \mu). \quad (3.2)$$

We recall that in a theory with a single coupling λ , such as the φ_4^4 theory, we have relations of the form

$$\begin{aligned} \varphi_B &= Z_\varphi^{1/2}(\lambda, \varepsilon) \varphi, & \lambda_B &= \mu^\varepsilon \lambda Z_\lambda(\lambda, \varepsilon), \\ m_B^2 &= m^2 Z_{m^2}(\lambda, \varepsilon), & L_B &= Z_L(\lambda, \varepsilon) L. \end{aligned} \quad (3.3)$$

Theories with more couplings and fields will have more complicated relations, but these details do not really concern our discussion here. The key point is that the renormalized sides of (3.1) and (3.2) depend on one quantity more than the bare sides. Precisely, the renormalized sides depend on λ and μ separately, while the bare sides contains only λ_B , which depends on λ and μ .

Therefore, it must be possible to solve the μ dependence exactly in terms of the renormalization constants. The solution can be obtained by comparing the bare version and the renormalized version of any equation. The μ dependence of the physical correlation functions is called “renormalization-group flow”.

Let us see what originates the dependence on μ . At the tree level, the action depends on a unique combination of λ and μ , such as $\lambda\mu^\varepsilon$ in the case (3.3). However, that combination cannot survive beyond the tree level, because the subtraction of divergences is an operation that separates μ from λ : the counterterms are multiplied by higher powers of λ , which is dimensionless, by the same power of μ , which is fixed by dimensional analysis (see for example (2.33)). This produces the final relation $\lambda_B = \mu^\varepsilon \lambda Z_\lambda(\lambda, \varepsilon)$.

The key quantities that are used to describe the renormalization-group flow are the beta function and the anomalous dimensions. To introduce them, we define the total derivative $\mu d/d\mu$, which is the derivative calculated by keeping the bare quantities fixed, and the partial derivative $\mu\partial/\partial\mu$, which is the derivative calculated by keeping the renormalized quantities fixed. When we apply the total derivative to the functional Γ we obtain, by the Leibniz rule,

$$\begin{aligned} \mu \frac{d}{d\mu} = & \mu \frac{\partial}{\partial\mu} + \mu \frac{d\lambda}{d\mu} \frac{\partial}{\partial\lambda} + \mu \frac{dm^2}{d\mu} \frac{\partial}{\partial m^2} + \int d^D x \mu \frac{d\Phi(x)}{d\mu} \frac{\delta}{\delta\Phi(x)} \\ & + \int d^D x \mu \frac{dL^I(x)}{d\mu} \frac{\delta}{\delta L^I(x)}. \end{aligned} \quad (3.4)$$

If we apply the total derivative to the action, we obtain the same formula with Φ replaced by φ . If we apply the total derivative to the functional W , we obtain the same formula with Φ replaced by J .

Beta function

Define the “hat beta function” as

$$\hat{\beta}_\lambda = \mu \frac{d\lambda}{d\mu}.$$

At the tree level $\lambda_B = \mu^{p\varepsilon} \lambda$, so $\hat{\beta}_\lambda = -p\varepsilon\lambda + \mathcal{O}(\lambda^2)$. It is convenient to define the *beta function* β_λ such that

$$\hat{\beta}_\lambda(\lambda, \varepsilon) = \beta_\lambda(\lambda, \varepsilon) - p\varepsilon\lambda. \quad (3.5)$$

Clearly, $\beta_\lambda = \mathcal{O}(\lambda^2)$. Now, we apply the identity (3.4) to λ_B and recall that Z_λ depends only on λ and ε . We find

$$0 = \mu \frac{d\lambda_B}{d\mu} = \left(\mu \frac{\partial}{\partial \mu} + \hat{\beta}_\lambda \frac{\partial}{\partial \lambda} \right) (\mu^{p\varepsilon} \lambda Z_\lambda) = p\varepsilon \mu^\varepsilon \lambda Z_\lambda + \mu^\varepsilon \hat{\beta}_\lambda \frac{d(\lambda Z_\lambda)}{d\lambda},$$

whence

$$\frac{d \ln Z_\lambda}{d\lambda} = -\frac{\beta_\lambda}{\lambda \hat{\beta}_\lambda}. \quad (3.6)$$

Using (3.5) we have

$$\beta_\lambda = \frac{p\varepsilon \lambda^2 \frac{d \ln Z_\lambda}{d\lambda}}{1 + \lambda \frac{d \ln Z_\lambda}{d\lambda}}. \quad (3.7)$$

We also find the inverse formula

$$Z_\lambda(\lambda, \varepsilon) = \exp \left(- \int_0^\lambda \frac{d\lambda'}{\lambda'} \frac{\beta_\lambda(\lambda', \varepsilon)}{\beta_\lambda(\lambda', \varepsilon) - p\varepsilon \lambda'} \right). \quad (3.8)$$

The lower integration limit is fixed by demanding $Z_\lambda(0, \varepsilon) = 1$, since in the free-field limit the renormalization constants are equal to one.

Anomalous dimension

Let us study the total derivative of φ_B . Using (3.4) with $\Phi \rightarrow \varphi$ we find

$$0 = \mu \frac{d\varphi_B}{d\mu} = \mu \frac{d}{d\mu} \left(Z_\varphi^{1/2} \varphi \right) = \mu \frac{dZ_\varphi^{1/2}}{d\mu} \varphi + Z_\varphi^{1/2} \mu \frac{d\varphi}{d\mu},$$

that is to say

$$\mu \frac{d\varphi}{d\mu} = -\gamma_\varphi \varphi,$$

where

$$\gamma_\varphi \equiv \frac{1}{2} \mu \frac{d \ln Z_\varphi}{d\mu} = \frac{1}{2} \mu \frac{d\lambda}{d\mu} \frac{d \ln Z_\varphi}{d\lambda} = \frac{1}{2} \hat{\beta}_\lambda \frac{d \ln Z_\varphi}{d\lambda}. \quad (3.9)$$

The quantity γ_φ is called *anomalous dimension* of the field φ , and depends on λ and ε . Since $\Phi = \langle \varphi \rangle$ we also have

$$\mu \frac{d\Phi}{d\mu} = -\gamma_\varphi \Phi.$$

From (3.9) we find the inverse formula

$$Z_\varphi(\lambda, \varepsilon) = \exp \left(2 \int_0^\lambda d\lambda' \frac{\gamma_\varphi(\lambda', \varepsilon)}{\beta_\lambda(\lambda', \varepsilon) - p\varepsilon \lambda'} \right).$$

When the renormalization is multiplicative, we have

$$\int J_B \varphi_B = \int J \varphi,$$

so the renormalization constants of φ and J are the reciprocals of each other. Then we also find

$$\mu \frac{dJ}{d\mu} = \gamma_\varphi J. \quad (3.10)$$

Exercise 16 Calculate the beta functions and the anomalous dimensions of the φ_4^4 theory and the φ_6^3 theory at one loop in the minimal subtraction scheme.

Solution. In the minimal subtraction scheme, the constants c_1 and c_2 of formula (2.44) are equal to zero. Applying the formulas (3.7) and (3.9) to (2.44), (2.47) and (2.48), we get

$$\varphi_4^4: \quad \beta_\lambda = \frac{3\lambda^2}{16\pi^2} + \mathcal{O}(\lambda^3), \quad \gamma_\varphi = \mathcal{O}(\lambda^2), \quad (3.11)$$

$$\varphi_6^3: \quad \beta_\lambda = -\frac{3\lambda^3}{4(4\pi)^3} + \mathcal{O}(\lambda^5), \quad \gamma_\varphi = \frac{\lambda^2}{12(4\pi)^3} + \mathcal{O}(\lambda^4), \quad (3.12)$$

Exercise 17 Calculate the beta function and the anomalous dimension of the φ_4^4 theory at two loops.

Solution. Applying the formulas (3.7) and (3.9) to (2.64) and (2.65) we get

$$\beta_\lambda = \frac{3\lambda^2}{(4\pi)^2} - \frac{17\lambda^3}{3(4\pi)^4} + \mathcal{O}(\lambda^4), \quad \gamma_\varphi = \frac{\lambda^2}{12(4\pi)^4} + \mathcal{O}(\lambda^3).$$

Note that the divergences have canceled out. Later we will prove that this is a general fact.

Exercise 18 Calculate first nonvanishing contributions to the beta function and the anomalous dimension of the massless φ_3^6 theory.

Solution. Applying the formulas (3.7) and (3.9) to (2.66) and (2.68), and recalling that here $p = 2$, we get

$$\beta_\lambda = \frac{5\lambda^2}{3(4\pi)^2} + \mathcal{O}(\lambda^3), \quad \gamma_\varphi = \frac{\lambda^2}{90(8\pi)^4} + \mathcal{O}(\lambda^3).$$

Composite fields

Similarly, taking the total derivative $\mu d/d\mu$ of L_B we get

$$0 = \mu \frac{dL_B}{d\mu} = \mu \frac{dZ_L}{d\mu} L + Z_L \mu \frac{dL}{d\mu},$$

whence

$$\mu \frac{dL}{d\mu} = -\mu \frac{d \ln Z_L}{d\mu} L = -\hat{\beta}_\lambda \frac{d \ln Z_L}{d\lambda} L.$$

Recalling that $L_B \mathcal{O}_B = L \mathcal{O}$, and $Z_L = Z_{\mathcal{O}}^{-1}$, we have

$$\mu \frac{d\mathcal{O}}{d\mu} = -\gamma_{\mathcal{O}} \mathcal{O}, \quad \gamma_{\mathcal{O}} = -\gamma_L = \mu \frac{d \ln Z_{\mathcal{O}}}{d\mu} = -\mu \frac{d \ln Z_L}{d\mu}.$$

The inverse formula reads

$$Z_{\mathcal{O}}(\lambda, \varepsilon) = \exp \left(\int_0^\lambda d\lambda' \frac{\gamma_{\mathcal{O}}(\lambda', \varepsilon)}{\beta_\lambda(\lambda', \varepsilon) - p\lambda'\varepsilon} \right). \quad (3.13)$$

For example, consider the mass operator φ^2 . Its renormalization coincides with the renormalization of the integrated mass term

$$\int d^D x \varphi^2. \quad (3.14)$$

Indeed, the integral determines the integrand up to local total derivatives, which in this case must also have dimension 2 (for $D = 4$). Since there exist no local Lorentz invariant objects with these features, φ^2 and (3.14) renormalize in exactly the same way. Correspondingly, the source L_{φ^2} coupled to φ^2 renormalizes exactly as m^2 , and $Z_{\varphi^2} = Z_{m^2}^{-1}$. By dimensional analysis, $\mu dm^2/d\mu$ must be equal to m^2 times a function of λ and ε . We have

$$\mu \frac{dm^2}{d\mu} = m^2 \eta(\lambda, \varepsilon), \quad \eta = -\mu \frac{d \ln Z_{m^2}}{d\mu} = \mu \frac{d \ln Z_{\varphi^2}}{d\mu} = \gamma_{\varphi^2}.$$

More generally, the composite fields mix with one another. As explained in Chapter 2, it is convenient to collect all of them into a huge vector \mathcal{O}^I , where the composite fields of the same dimension are close to one another and the composite fields of higher dimensions follow those of lower dimensions. Then, we have

$$\mathcal{O}_B^I = Z^{IJ} [\mathcal{O}^J],$$

where the matrix Z^{IJ} of renormalization constants is block lower triangular. We find

$$\mu \frac{d[\mathcal{O}^I]}{d\mu} = -\gamma_{IJ}[\mathcal{O}^J], \quad \gamma_{IJ} \equiv Z_{IK}^{-1} \mu \frac{dZ^{KJ}}{d\mu}. \quad (3.15)$$

Due to (2.102), the sources L^I coupled to the \mathcal{O}^I s satisfy

$$\mu \frac{dL^I}{d\mu} = L^J \gamma_{JI}. \quad (3.16)$$

3.1 The Callan-Symanzik equation

Let us apply (3.4) to $\Gamma(\Phi)$. We obtain

$$0 = \mu \frac{d\Gamma_B}{d\mu} = \mu \frac{\partial \Gamma}{\partial \mu} + \hat{\beta}_\lambda \frac{\partial \Gamma}{\partial \lambda} + \eta m^2 \frac{\partial \Gamma}{\partial m^2} - \gamma_\varphi \int d^D x \Phi(x) \frac{\delta \Gamma}{\delta \Phi(x)}. \quad (3.17)$$

On W we have, instead,

$$0 = \mu \frac{dW_B}{d\mu} = \mu \frac{\partial W}{\partial \mu} + \hat{\beta}_\lambda \frac{\partial W}{\partial \lambda} + \eta m^2 \frac{\partial W}{\partial m^2} + \gamma_\varphi \int d^D x J(x) \frac{\delta W}{\delta J(x)}. \quad (3.18)$$

Let us take two functional derivatives of (3.17) with respect to Φ and set $\Phi = 0$ afterwards, or, equivalently, two derivatives of (3.18) with respect to J and then set $J = 0$. We obtain the Callan-Symanzik equations for the connected-irreducible two-point functions Γ_2 and the connected two-point function $W_2 = \langle \varphi(x) \varphi(y) \rangle_c$:

$$0 = \left(\mu \frac{\partial}{\partial \mu} + \hat{\beta}_\lambda \frac{\partial}{\partial \lambda} + \eta m^2 \frac{\partial}{\partial m^2} - 2\gamma_\varphi \right) \Gamma_2, \quad (3.19)$$

$$0 = \left(\mu \frac{\partial}{\partial \mu} + \hat{\beta}_\lambda \frac{\partial}{\partial \lambda} + \eta m^2 \frac{\partial}{\partial m^2} + 2\gamma_\varphi \right) W_2. \quad (3.20)$$

The two equations are indeed equivalent, because $\Gamma_2 W_2 = 1$.

For the moment let us work in the massless theory. We do not make the ε dependence explicit, because it is not important for the present discussion. Since W_2 has dimension $D - 2$ it is convenient to write

$$W_2(|x - y|, \lambda, \mu) = \frac{G_{2r}(t, \lambda)}{|x - y|^{D-2}}, \quad t \equiv -\ln(|x - y|\mu). \quad (3.21)$$

Then (3.20) at $m = 0$ becomes

$$0 = \left(-\frac{\partial}{\partial t} + \hat{\beta}_\lambda \frac{\partial}{\partial \lambda} + 2\gamma_\varphi \right) G_{2r}(t, \lambda). \quad (3.22)$$

We want to solve this equation. To this purpose, we define the “running coupling” $\tilde{\lambda}(t, \lambda)$, which is the solution of the first-order differential equation

$$\frac{d\tilde{\lambda}}{dt} = \hat{\beta}_\lambda(\tilde{\lambda}), \quad \tilde{\lambda}(0, \lambda) = \lambda. \quad (3.23)$$

We have $dt = d\tilde{\lambda}/\hat{\beta}_\lambda(\tilde{\lambda})$, hence

$$t = \int_\lambda^{\tilde{\lambda}(t, \lambda)} \frac{d\tilde{\lambda}}{\hat{\beta}_\lambda(\tilde{\lambda})}. \quad (3.24)$$

It is convenient to consider $\tilde{\lambda}$ as a function of both t and the initial condition λ . If so, the t derivative appearing in (3.23) must be written as a partial derivative $\partial\tilde{\lambda}(t, \lambda)/\partial t$. Differentiating each side of (3.24) with respect to λ , we can work out the derivative of the solution with respect to its initial condition, which is

$$\frac{\partial\tilde{\lambda}(t, \lambda)}{\partial\lambda} = \frac{\hat{\beta}_\lambda(\tilde{\lambda}(t, \lambda))}{\hat{\beta}_\lambda(\lambda)}. \quad (3.25)$$

The solution of the Callan-Symanzik equation (3.22) reads

$$G_{2r}(t, \lambda) = z^{-1}(\lambda, t) G_{2r}(0, \tilde{\lambda}(t, \lambda)), \quad (3.26)$$

with

$$z(\lambda, t) = \exp\left(-2 \int_0^t \gamma_\varphi(\tilde{\lambda}(s, \lambda)) ds\right). \quad (3.27)$$

We prove this statement by checking that (3.26) satisfies the equation and the initial condition. Given a function f of many variables, we write $f^{(n_1, n_2, \dots)}$ to denote its n_1 th partial derivative with respect to its first variable, n_2 th partial derivative with respect to its second variable, and so on.

The initial condition is certainly satisfied, since at $t = 0$ we have the identity $G_{2r}(0, \lambda) = G_{2r}(0, \lambda)$. Moreover, we can easily calculate the partial derivatives of G_{2r} with respect to t and λ . We find

$$\begin{aligned} G_{2r}^{(1,0)}(t, \lambda) &= 2\gamma_\varphi(\tilde{\lambda}(t, \lambda))G_{2r}(t, \lambda) + z^{-1}(\lambda, t)\hat{\beta}_\lambda(\tilde{\lambda}(t, \lambda))G_{2r}^{(0,1)}(0, \tilde{\lambda}(t, \lambda)), \\ G_{2r}^{(0,1)}(t, \lambda) &= 2G_{2r}(t, \lambda) \int_0^t \frac{\partial\tilde{\lambda}(s, \lambda)}{\partial\lambda} \gamma'_\varphi(\tilde{\lambda}(s, \lambda)) ds \\ &\quad + z^{-1}(\lambda, t) \frac{\partial\tilde{\lambda}(t, \lambda)}{\partial\lambda} G_{2r}^{(0,1)}(0, \tilde{\lambda}(t, \lambda)). \end{aligned}$$

Now, using (3.25) we also have

$$\begin{aligned}\hat{\beta}_\lambda(\lambda) \int_0^t \frac{\partial \tilde{\lambda}(s, \lambda)}{\partial \lambda} \gamma'_\varphi(\tilde{\lambda}(s, \lambda)) ds &= \int_0^t \hat{\beta}_\lambda(\tilde{\lambda}(s, \lambda)) \gamma'_\varphi(\tilde{\lambda}(s, \lambda)) ds \\ &= \int_0^t \frac{\partial \tilde{\lambda}(s, \lambda)}{\partial s} \gamma'_\varphi(\tilde{\lambda}(s, \lambda)) ds = \int_0^t \frac{\partial \gamma_\varphi(\tilde{\lambda}(s, \lambda))}{\partial s} ds = \gamma_\varphi(\tilde{\lambda}(t, \lambda)) - \gamma_\varphi(\lambda).\end{aligned}\quad (3.28)$$

Summing, we find immediately that (3.22) is satisfied.

When the theory contains more parameters λ_i (which can include also the masses), equation (3.22) becomes

$$0 = \left(-\frac{\partial}{\partial t} + \hat{\beta}_\lambda^i \frac{\partial}{\partial \lambda_i} + 2\gamma_\varphi \right) G_{2r}(t, \lambda). \quad (3.29)$$

Define the running parameters $\tilde{\lambda}_i(t, \lambda)$ as the solutions of the following system of first-order differential equations:

$$\frac{d\tilde{\lambda}_i}{dt} = \hat{\beta}_\lambda^i(\tilde{\lambda}), \quad \tilde{\lambda}_i(0, \lambda) = \lambda_i. \quad (3.30)$$

The solution (3.26) and formula (3.27) remain the same. However, formulas (3.24) and (3.25) do not hold. Define

$$f_i(t, \lambda) \equiv \hat{\beta}_\lambda^j(\lambda) \frac{\partial \tilde{\lambda}_i(t, \lambda)}{\partial \lambda_j},$$

where the sum over j is understood. We have $f_i(0, \lambda) = \hat{\beta}_\lambda^i(\lambda)$. Moreover, if $\hat{\beta}_{\lambda,k}^i(\lambda) \equiv \partial \hat{\beta}_\lambda^i(\lambda) / \partial \lambda_k$, we get

$$\frac{\partial f_i}{\partial t} = \hat{\beta}_\lambda^j(\lambda) \frac{\partial \hat{\beta}_\lambda^i(\tilde{\lambda}(t, \lambda))}{\partial \lambda_j} = \hat{\beta}_\lambda^j(\lambda) \frac{\partial \tilde{\lambda}_k(t, \lambda)}{\partial \lambda_j} \hat{\beta}_{\lambda,k}^i(\tilde{\lambda}(t, \lambda)) = f_k \hat{\beta}_{\lambda,k}^i(\tilde{\lambda}(t, \lambda)).$$

We obtain the system of first-order equations and initial conditions

$$\frac{\partial f_i(t, \lambda)}{\partial t} = f_k(t, \lambda) \hat{\beta}_{\lambda,k}^i(\tilde{\lambda}(t, \lambda)), \quad f_i(0, \lambda) = \hat{\beta}_\lambda^i(\lambda).$$

It is easy to check that

$$F_i(t, \lambda) \equiv \hat{\beta}_\lambda^i(\tilde{\lambda}(t, \lambda))$$

satisfies the equations and the initial conditions. In particular,

$$\frac{\partial F_i(t, \lambda)}{\partial t} = \hat{\beta}_\lambda^k(\tilde{\lambda}(t, \lambda)) \hat{\beta}_{\lambda,k}^i(\tilde{\lambda}(t, \lambda)) = F_k(t, \lambda) \hat{\beta}_{\lambda,k}^i(\tilde{\lambda}(t, \lambda)).$$

Therefore, we conclude that $F_i(t, \lambda) = f_i(t, \lambda)$, that is to say

$$\hat{\beta}_\lambda^j(\lambda) \frac{\partial \tilde{\lambda}_i(t, \lambda)}{\partial \lambda_j} = \hat{\beta}_\lambda^i(\tilde{\lambda}(t, \lambda)). \quad (3.31)$$

This formula is a generalization of (3.25). Going through the proof of (3.26) we realize that (3.25) was necessary only to derive (3.28). Extending the proof of (3.26) to the theories that contain more parameters, we see that (3.31) is just sufficient to derive the desired generalization of (3.28).

In the end, we find that (3.26) satisfies (3.29), as wanted.

General solution of the Callan-Symanzik equation

So far, we have studied the two-point function. However, the results can be extended to a generic correlation function

$$W_{\alpha_1 \dots \alpha_n}^{I_1 \dots I_m}(x, y, \lambda, \mu) = \langle \varphi_{\alpha_1}(x_1) \cdots \varphi_{\alpha_n}(x_n) \mathcal{O}^{I_1}(y_1) \cdots \mathcal{O}^{I_m}(y_m) \rangle_c \quad (3.32)$$

that contains both insertions of elementary and composite fields. The subscript α in φ_α is used to distinguish the different types of elementary fields, including the ghosts and the Lagrange multipliers. We denote the φ_α anomalous dimensions with $\gamma_\alpha(\lambda)$. Finally, λ_i collects the couplings and any other parameters, including the masses and the gauge-fixing parameters.

The Callan-Symanzik equation for (3.32) can be derived by applying (3.4) (with $\Phi \rightarrow J$) to $W(J, L)$ and using (3.10) and (3.16). We find

$$0 = \left(\mu \frac{\partial}{\partial \mu} + \sum_i \hat{\beta}_\lambda^i \frac{\partial}{\partial \lambda_i} + 2 \sum_{i=1}^n \gamma_{\alpha_i} \right) W_{\alpha_1 \dots \alpha_n}^{I_1 \dots I_m} + \sum_{j=1}^m \gamma_{I_j K_j} W_{\alpha_1 \dots \alpha_n}^{I_1 \dots I_{j-1} K_j I_{j+1} \dots I_m}. \quad (3.33)$$

Repeating the proof of the previous section it is easy to show that the solution reads

$$W_{\alpha_1 \dots \alpha_n}^{I_1 \dots I_m}(x, y, \lambda, \mu) = \prod_{i=1}^n z_{\alpha_i}^{-1/2}(t) \prod_{j=1}^m Z_{I_j K_j}^{-1}(t) W_{\alpha_1 \dots \alpha_n}^{K_1 \dots K_m}(x, y, \tilde{\lambda}(t), \tilde{\mu}), \quad (3.34)$$

where now $t = \ln(\tilde{\mu}/\mu)$, $\tilde{\lambda}(t)$ are the solutions of (3.30) and

$$z_{\alpha_i}(t) = \exp \left(-2 \int_0^t ds \gamma_{\alpha_i}(s) \right), \quad Z(t) = T \exp \left(- \int_0^t ds \gamma(s) \right), \quad (3.35)$$

where Z and γ are the matrices with entries Z_{IJ} and γ_{IJ} , $\gamma(t)$ stands for $\gamma(\tilde{\lambda}(t))$ and T denotes the T-ordered product. Precisely,

$$Z(t) = 1 + \sum_{k=1}^{\infty} (-1)^k \int_0^t dt_1 \int_0^{t_1} dt_2 \cdots \int_0^{t_{k-1}} dt_k \gamma(t_1) \cdots \gamma(t_{k-1}) \gamma(t_k), \quad (3.36)$$

Thus, formula (3.34) tells us how the correlation function depends on the scale μ .

3.2 Finiteness of the beta function and the anomalous dimensions

Formulas (3.11) and (3.49) show that the poles in ε cancel out in the one-loop beta functions and the anomalous dimensions. This is a very general fact: the beta functions and the anomalous dimensions are finite quantities. Consider the Callan-Symanzik equation (3.17) for $\Gamma(\Phi, \lambda, m^2, \mu)$. Restore \hbar and expand each quantity perturbatively,

$$\Gamma = \sum_{i=0}^{\infty} \hbar^i \Gamma_i, \quad \hat{\beta}_\lambda = \sum_{i=0}^{\infty} \hbar^i \hat{\beta}_{\lambda i}, \quad \eta = \sum_{i=0}^{\infty} \hbar^i \eta_i, \quad \gamma_\varphi = \sum_{i=0}^{\infty} \hbar^i \gamma_{\varphi i}. \quad (3.37)$$

Assume inductively that $\hat{\beta}_\lambda$, η and γ_φ are finite up to and including the order $n-1$, that is to say

$$\hat{\beta}_{\lambda i}, \eta_i, \gamma_{\varphi i} < \infty \quad \text{for } i \leq n-1. \quad (3.38)$$

The assumption is obviously true for $i=0$, since $\hat{\beta}_{\lambda 0} = -\varepsilon \lambda$, $\eta_0 = \gamma_{\varphi 0} = 0$. Consider the contribution to (3.17) of order n . We have

$$0 = \mu \frac{\partial \Gamma_n}{\partial \mu} + \sum_{i=0}^n \left(\hat{\beta}_{\lambda i} \frac{\partial \Gamma_{n-i}}{\partial \lambda} + \eta_i m^2 \frac{\partial \Gamma_{n-i}}{\partial m^2} - \gamma_{\varphi i} \int \Phi \frac{\delta \Gamma_{n-i}}{\delta \Phi} \right).$$

Recall that every Γ_i is convergent, and so are its derivatives with respect to the renormalized parameters. Using (3.38) we conclude

$$\hat{\beta}_{\lambda n} \frac{\partial \Gamma_0}{\partial \lambda} + \eta_n m^2 \frac{\partial \Gamma_0}{\partial m^2} - \gamma_{\varphi n} \int \Phi \frac{\delta \Gamma_0}{\delta \Phi} = \text{finite}. \quad (3.39)$$

Now, Γ_0 is just the classical action, and $\partial \Gamma_0 / \partial \lambda$, $\partial \Gamma_0 / \partial m^2$ and $\Phi(\delta \Gamma_0 / \delta \Phi)$ are independent terms, because they are the vertex, the mass term and the

field equation (which contains the vertex, the mass term and the kinetic term), respectively. Therefore, each coefficient of the linear combination (3.39) must be finite, which proves

$$\hat{\beta}_{\lambda n} < \infty, \quad \eta_n < \infty, \quad \gamma_{\varphi n} < \infty.$$

The inductive assumption (3.38) is thus promoted to $n = \infty$.

We have set $L^I = 0$, but clearly the argument can be repeated in the presence of sources for composite fields \mathcal{O}^I , which proves that the γ_{IJS} are also finite.

3.3 Fixed points of the RG flow

Consider the correlation function $W_{\alpha_1 \dots \alpha_n}^{I_1 \dots I_m}$ of formula (3.32) and rescale the coordinates, the momenta and the parameters λ by powers of ζ equal to their dimensions in units of mass. For example, rescale the coordinates x^μ to $\zeta^{-1}x^\mu$, the momenta p^μ to ζp^μ , the masses m to ζm , and so on. If we also rescale μ to $\zeta\mu$, we get

$$W_{\alpha_1 \dots \alpha_n}^{I_1 \dots I_m}(\zeta^{-1}x, \zeta^{-1}y, \zeta^{d_\lambda}\lambda, \zeta\mu) = \zeta^{d_W} W_{\alpha_1 \dots \alpha_n}^{I_1 \dots I_m}(x, y, \lambda, \mu),$$

where d_W and d_λ are the dimensions of $W_{\alpha_1 \dots \alpha_n}^{I_1 \dots I_m}$ and λ , respectively. Replacing μ by $\zeta^{-1}\mu$, we obtain

$$W_{\alpha_1 \dots \alpha_n}^{I_1 \dots I_m}(x, y, \lambda, \zeta^{-1}\mu) = \zeta^{-d_W} W_{\alpha_1 \dots \alpha_n}^{I_1 \dots I_m}(\zeta^{-1}x, \zeta^{-1}y, \zeta^{d_\lambda}\lambda, \mu).$$

The left-hand side of this equation tells us that the limit $\mu \rightarrow \infty$ in the correlation function is equivalent to let ζ tend to zero. The right-hand side tells us that this operation is equivalent to let the distances tend to infinity (and rescale the parameters of the theory appropriately). Thus, the limit $\mu \rightarrow \infty$ gives information about the infrared, or large-distance, limit of the correlation function. Similarly, the limit $\mu \rightarrow 0$ is equivalent to take ζ (in particular the distances) to infinity, so it gives information about the ultraviolet limit.

The solution (3.34) of the renormalization-group equations gives

$$W_{\alpha_1 \dots \alpha_n}^{I_1 \dots I_m}(x, y, \lambda, \zeta^{-1}\mu) = \prod_{i=1}^n z_{\alpha_i}^{-1/2}(t) \prod_{j=1}^m Z_{I_j K_j}^{-1}(t) W_{\alpha_1 \dots \alpha_n}^{K_1 \dots K_m}(x, y, \tilde{\lambda}(t), \mu), \quad (3.40)$$

where now $t = \ln \zeta$. Thus, to understand the infrared and the ultraviolet behaviors of the correlation functions, it is useful to work out the infrared and the ultraviolet behaviors of the beta functions and the anomalous dimensions.

For simplicity, we assume that the theory has a unique dimensionless coupling, and keep calling it λ . We also assume that λ is defined to be non-negative. Typically, as in the case of the theory φ^4 , this requirement is necessary to ensure that the potential is bounded from below.

An alternative way of defining the running coupling is by viewing it as a function $\lambda(\mu)$ of the energy scale μ . Start from formula (3.24), and set $\varepsilon = 0$, $t = \ln(\tilde{\mu}/\mu)$, $\tilde{\lambda} = \lambda(\tilde{\mu})$ and $\lambda = \lambda(\mu)$. In a generic subtraction scheme, define $\beta_\lambda(\lambda) \equiv \tilde{\beta}_\lambda(\lambda, 0)$. Exponentiating (3.24), relabeling the integration variable and splitting the integral into two symmetric parts with the help of an arbitrary constant $\bar{\lambda}$, we can write

$$\tilde{\mu} \exp \left(- \int_{\bar{\lambda}}^{\lambda(\tilde{\mu})} \frac{d\lambda}{\beta_\lambda(\lambda)} \right) = \mu \exp \left(- \int_{\bar{\lambda}}^{\lambda(\mu)} \frac{d\lambda}{\beta_\lambda(\lambda)} \right) = \text{constant} \equiv \Lambda_T.$$

The scale Λ_T (called Λ_{QCD} if the theory T is QCD and $\bar{\lambda}$ is chosen appropriately) is RG invariant, i.e. independent of μ . We also have

$$\int_{\bar{\lambda}}^{\lambda(\mu)} \frac{d\lambda}{\beta_\lambda(\lambda)} = \ln \frac{\mu}{\Lambda_T} \quad (3.41)$$

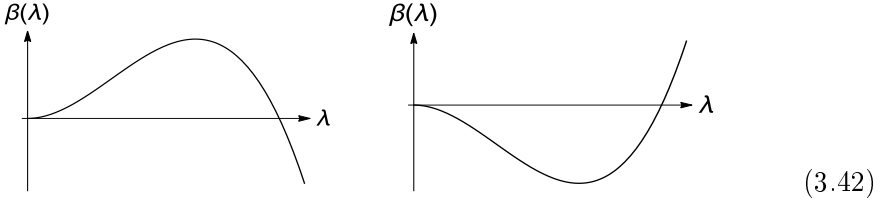
Now, the infrared (ultraviolet) behavior of $\tilde{\lambda}$ is studied for $t \rightarrow -\infty$ ($t \rightarrow \infty$), which is equivalent to take the limit $\tilde{\mu} \rightarrow 0$ ($\tilde{\mu} \rightarrow \infty$) of the function $\tilde{\lambda} = \lambda(\tilde{\mu})$. Then, it is also the limit $\mu \rightarrow 0$ ($\mu \rightarrow \infty$) of $\lambda(\mu)$. We see that in both the infrared and ultraviolet limits, the right-hand side of (3.41) diverges. On the other hand, the left-hand side can diverge in the following two cases: (i) the running coupling tends to a zero of the beta function, i.e.

$$\lim_{\mu \rightarrow 0} \lambda(\mu) = \lambda_{\text{IR}}, \quad \text{and/or} \quad \lim_{\mu \rightarrow \infty} \lambda(\mu) = \lambda_{\text{UV}},$$

where $\beta_\lambda(\lambda_{\text{IR}}) = \beta_\lambda(\lambda_{\text{UV}}) = 0$; or (ii) the running coupling tends to $\pm\infty$ and the infrared and/or ultraviolet limits. In all the cases the integral of (3.41) must diverge in the correct way.

The values of the couplings for which the beta functions vanish at $\varepsilon = 0$, i.e. the solutions λ_* of $\beta_\lambda(\lambda_*) = 0$, define a particular class of quantum field theories, which are called fixed points of the RG flow. Clearly, $\lambda = 0$ is

a trivial example of a fixed point, and corresponds to the free field theory we have been expanding around. However, there may exist interacting fixed points, that is to say solutions with $\lambda_* \neq 0$. This happens, for example, when the beta functions have the forms



In some cases the nontrivial fixed point can be reached perturbatively from the free fixed point. Then, the perturbative expansion allows us to follow the entire renormalization group flow in between the fixed points.

At any fixed point, the coupling does not run (at $\varepsilon = 0$), since $\beta_\lambda \equiv 0$ implies $\tilde{\lambda} \equiv \lambda_*$. However, a theory behaves rather differently around the free and the interacting fixed points. Now we study the typical behaviors, starting from the trivial fixed point.

Expand β_λ perturbatively around $\lambda = 0$:

$$\beta_\lambda(\lambda) = \beta_1 \lambda^2 + \beta_2 \lambda^3 + \beta_3 \lambda^4 + \mathcal{O}(\lambda^5). \quad (3.43)$$

If the running coupling $\tilde{\lambda}$ is small, we can keep the first nontrivial contribution to $\hat{\beta}_\lambda(\tilde{\lambda})$ in the RG equation (3.23) and neglect any higher orders. We assume here that $\beta_1 \neq 0$, so the first nontrivial contribution is the one-loop one. The running coupling reads

$$\tilde{\lambda}(t, \lambda) = \frac{\lambda}{1 - \beta_1 t \lambda} \quad (3.44)$$

up to two-loop corrections. Setting $t = \ln(\tilde{\mu}/\mu)$, $\tilde{\lambda} = \lambda(\tilde{\mu})$ and $\lambda = \lambda(\mu)$, we can also write the running in the form (3.41), or

$$\frac{1}{\lambda(\mu)} + \beta_1 \ln \mu = \text{constant}.$$

However, this result is just one loop, and can be trusted only if the running coupling is small. This happens when $\mu \rightarrow 0$ (which is the IR limit) for $\beta_1 > 0^1$, and when $\mu \rightarrow \infty$ (which is the UV limit) for $\beta_1 < 0$. Specifically,

¹Recall that λ is assumed to be non-negative.

formula (3.44) gives

$$\tilde{\lambda}(t, \lambda) \sim -\frac{1}{\beta_1 t} \ll 1 \quad \text{for } |t| \gg 1, \quad (3.45)$$

The running coupling tends to zero, so the theory tends to the free fixed point. Observe that the behavior (3.45) is λ independent.

The theory is said to be infrared free if $\beta_1 > 0$, and asymptotically free (or ultraviolet free) if $\beta_1 < 0$. Nontrivial examples of asymptotically free theories are provided by non-Abelian gauge field theories, as well as the two-dimensional four-fermion model (1.102). In the latter case, formula (2.93) allows us to work out the one-loop beta function, which reads

$$\beta_\lambda = -\frac{(N-1)\lambda^2}{2\pi} + \mathcal{O}(\lambda^3).$$

Now we study the behavior of the theory around an interacting fixed point. We expand by writing $\lambda = \lambda_* + \eta$ and taking $\eta \ll 1$. The Taylor expansion of the beta function gives, to the lowest order,

$$\beta_\lambda(\lambda) = \beta'_\lambda(\lambda_*)\eta + \mathcal{O}(\eta^2).$$

We assume that the slope $\beta'_\lambda(\lambda_*)$ of the beta function at the fixed point is nonvanishing. If not, we would have to go to the first nontrivial order of the Taylor expansion and modify the analysis accordingly. The RG equation (3.23) becomes

$$\frac{d\tilde{\eta}}{dt} = \beta'_\lambda(\lambda_*)\tilde{\eta} + \mathcal{O}(\tilde{\eta}^2), \quad \tilde{\eta}(0) = \eta,$$

The running coupling reads

$$\tilde{\eta}(\tilde{\mu}) = \eta e^{\beta'_\lambda(\lambda_*)t}.$$

Writing $t = \ln(\tilde{\mu}/\mu)$ to switch to the form (3.41), we obtain

$$\eta(\mu)\mu^{-\beta'_\lambda(\lambda_*)} = \text{constant}. \quad (3.46)$$

With the help of this formula, we can reach the fixed point. There, $\eta(\mu)$ must tend to zero. This happens when $\mu \rightarrow \infty$ for $\beta'_\lambda(\lambda_*) < 0$ and when $\mu \rightarrow 0$ for $\beta'_\lambda(\lambda_*) > 0$. We learn that when the slope of the beta function is

negative (positive) at the fixed point, the interacting fixed point is reached in the ultraviolet (infrared) limit.

Recapitulating, in the first (second) example of (3.42) the theory is free in the infrared (ultraviolet) limit, and tends to the interacting fixed point in the ultraviolet (infrared) limit.

At an interacting fixed point, the anomalous dimensions $\gamma_{\varphi^2}(\lambda_*) \equiv \gamma_{\varphi^2}^*$ (and $\gamma_{\alpha_i}(\lambda_*) \equiv \gamma_{\alpha_i}^*$, $\gamma_{IJ}(\lambda_*) \equiv \gamma_{IJ}^*$, in general) are just constants. Then, formulas (3.35) with $t = \ln(\tilde{\mu}/\mu)$ give

$$z_{\alpha_i}(t) = \left(\frac{\mu}{\tilde{\mu}}\right)^{2\gamma_{\alpha_i}^*}, \quad Z(t) = \left(\frac{\mu}{\tilde{\mu}}\right)^{\gamma^*}.$$

Finally, formula (3.34) gives

$$W_{\alpha_1 \dots \alpha_n}^{I_1 \dots I_m}(x, y, \lambda, \mu) = \left(\frac{\tilde{\mu}}{\mu}\right)^{\sum_{i=1}^n \gamma_{\alpha_i}^*} \prod_{j=1}^m \left(\frac{\tilde{\mu}}{\mu}\right)^{\gamma_{I_j K_j}^*} W_{\alpha_1 \dots \alpha_n}^{K_1 \dots K_m}(x, y, \lambda_*, \tilde{\mu}). \quad (3.47)$$

In the particular case of the two-point function $W_2 = \langle \varphi(x) \varphi(y) \rangle_c$ in the massless four-dimensional φ^4 theory, formulas (3.21), (3.26) and (3.27) give, in $D = 4$,

$$\langle \varphi(x) \varphi(y) \rangle_c = \frac{\mu^{-2\gamma_{\varphi}^*} C_{\varphi}}{|x - y|^{2(1+\gamma_{\varphi}^*)}},$$

where C_{φ} is a constant. If we compare this result with the two-point function of the free-fixed point, which is

$$\langle \varphi(x) \varphi(y) \rangle_c = \frac{1}{4\pi^2 |x - y|^2},$$

we see that the exponent is modified to twice the “critical exponent”

$$1 + \gamma_{\varphi}^*.$$

In turn, this is the sum of the naïve φ dimension, which is equal to one, plus γ_{φ}^* . Similarly, the two-point function of a composite field \mathcal{O} of naïve dimension $d_{\mathcal{O}}$ is

$$\langle \mathcal{O}(x) \mathcal{O}(y) \rangle_c = \frac{\mu^{-2\gamma_{\mathcal{O}}^*} C_{\mathcal{O}}}{|x - y|^{2(d_{\mathcal{O}} + \gamma_{\mathcal{O}}^*)}}.$$

These remarks justify the name “anomalous dimensions” for the quantities γ .

The solution (3.47) is simple at the fixed points, because the Callan-Symanzik equation (3.33) becomes simpler there. Dropping the terms proportional to the beta functions, we get

$$0 = \left(\mu \frac{\partial}{\partial \mu} + 2 \sum_{i=1}^n \gamma_{\alpha_i} \right) W_{\alpha_1 \dots \alpha_n}^{I_1 \dots I_m} + \sum_{j=1}^m \gamma_{I_j} K_j W_{\alpha_1 \dots \alpha_n}^{I_1 \dots I_{j-1} K_j I_{j+1} \dots I_m},$$

so the entire μ dependence of a correlation function is encoded in an appropriate multiplicative factor.

φ_4^4 at one loop The RG flow of the theory φ^4 in four dimensions can be worked out at one loop by means of the beta function (3.11). Formula (3.11) is correct also at $\varepsilon \neq 0$ in the minimal subtraction scheme, while it contains corrections of the form $\mathcal{O}(\varepsilon \lambda^2)$ in a generic scheme, where the constants c_1 and c_2 of formula (2.44) can be nonzero. In either case, the solution of (3.23) at $\varepsilon = 0$ reads

$$\tilde{\lambda}(t, \lambda) = \frac{\lambda}{1 - \frac{3t\lambda}{16\pi^2}}, \quad (3.48)$$

which is the running coupling in the one-loop approximation. Since $\beta_1 > 0$, the theory is infrared free.

The anomalous dimension of the composite operator φ^2 is

$$\eta = \gamma_{\varphi^2} = -\mu \frac{d \ln Z_m}{d\mu} = -\hat{\beta}_\lambda \frac{d \ln Z_m}{d\lambda} = \frac{\lambda}{16\pi^2} + \mathcal{O}(\lambda^2). \quad (3.49)$$

Let us study the φ^2 two-point function in the massless case. From (3.35) we have

$$z_{\varphi^2}(\lambda, t) = \exp \left(- \int_0^t \gamma_{\varphi^2}(\tilde{\lambda}(s, \lambda)) ds \right) = \left(1 - \frac{3t\lambda}{16\pi^2} \right)^{1/3}.$$

Applying the RG solution (3.34), we get (again at $\varepsilon = 0$)

$$\langle \varphi^2(x) \varphi^2(y) \rangle_c \equiv \frac{G_{\varphi^2 r}^{(2)}(t, \lambda)}{|x - y|^4} \sim \frac{\lambda_f^{2/3}(x - y, \mu)}{\lambda^{2/3}|x - y|^4} G_{\varphi^2 r}^{(2)}(0, \lambda_f(x - y, \mu)) \quad (3.50)$$

at large distances, where

$$\lambda_f(x - y, \mu) \equiv \frac{16\pi^2}{3 \ln(\mu|x - y|)}.$$

We cannot define a critical exponent here, since λ_f has a logarithmic behavior. The reason is that at the free fixed point the slope β'_λ of the beta function vanishes.

3.4 Scheme (in)dependence

Now we work out other useful properties of the beta function. Observe that in the minimal subtraction scheme the λ renormalization constant (which we denote with a bar) has the form $\bar{Z}_\lambda = 1 + \text{poles}$. Thus, formula (3.7) gives

$$\bar{\beta}_\lambda = \varepsilon \times \text{poles} = \text{finite} + \text{poles}, \quad (3.51)$$

with no orders ε^n , $n > 0$. However, we just proved that β_λ is finite, so the poles that appear on the right-hand side of (3.51) must cancel out. Thus, $\bar{\beta}_\lambda$ depends only on λ and not on ε . In the minimal subtraction scheme we can write

$$\widehat{\beta}_\lambda(\lambda, \varepsilon) = \bar{\beta}_\lambda(\lambda) - \varepsilon\lambda.$$

We know that the coefficients of the poles $1/\varepsilon$ are scheme independent at one loop. For this reason, the one-loop coefficients of beta functions and the anomalous dimensions are always scheme independent at $\varepsilon = 0$. Moreover, if the theory contains a unique dimensionless coupling λ , we can easily show that both the one-loop and the two-loop coefficients of the beta function are scheme independent at $\varepsilon = 0$. Expand $\beta_\lambda(\lambda)$ as in (3.43). A scheme change amounts to a perturbative reparametrization of λ . Write

$$\lambda = \lambda' + a_2\lambda'^2 + a_3\lambda'^3 + \mathcal{O}(\lambda'^4). \quad (3.52)$$

We have

$$\begin{aligned} \beta'_\lambda(\lambda') &= \mu \frac{d\lambda'}{d\mu} = \mu \frac{d\lambda}{d\mu} \frac{d\lambda'}{d\lambda} = \beta_\lambda(\lambda(\lambda')) \left(\frac{d\lambda}{d\lambda'} \right)^{-1} = \\ &= \beta_1\lambda'^2 + \beta_2\lambda'^3 + (\beta_3 + a_2\beta_2 + (a_2^2 - a_3)\beta_1)\lambda'^4 + \mathcal{O}(\lambda'^5). \end{aligned}$$

We see that the first two coefficients, and only those, are scheme independent. The result does not extend to $\varepsilon \neq 0$, since then we have to include reparametrizations of the form $\lambda = c(\varepsilon)\lambda' + \mathcal{O}(\lambda'^2)$, with $c(\varepsilon) = 1 + \mathcal{O}(\varepsilon)$.

With a suitable choice of a_2 and a_3 the third coefficient can be set to zero, for example

$$a_2 = 0, \quad a_3 = \frac{\beta_3}{\beta_1}. \quad (3.53)$$

It is easy to prove that with a suitable choice of the function $\lambda(\lambda')$ all the coefficients but the first two can be set to zero. However, this is just a curiosity. For example, the two-loop beta function cannot be trusted as an

exact formula, not even within the perturbative expansion. A warning that there is a problem here is the β_1 in the denominator of (3.53). Typically, β_1 is proportional to the number N of fields circulating in loops. Nowhere the perturbative expansion can generate inverse powers of N . In Yang-Mills theory, for example, β_1 is equal to a numerical factor times $11C(G) - 4N_f C(r)$, where $C(G)$ and $C(r)$ are the Casimir elements of the representations. It is obvious that, at the perturbative level, the Casimir elements cannot appear in the denominators.

Moreover, the reparametrization (3.52) can introduce spurious singularities at finite values of λ . For example, factors such as

$$\frac{\lambda^2}{\beta_\lambda(\lambda)} = \frac{1}{\beta_1} + \mathcal{O}(\lambda) \quad (3.54)$$

can easily be generated. If we ignore the awkward β_1 in the denominator (maybe because we are working with a given number of fields and are not aware of the importance of this point), such functions appear to have a perfectly good perturbative expansion around the free-fixed point of the RG flow. Nevertheless, they do not have a good perturbative behavior around an interacting fixed point, because they are singular there. If we make reparametrizations that involve expressions such as (3.54), we may lose the possibility of smoothly interpolating between two fixed points of the RG flow.

Finally, the “curiosity” mentioned above does not extend to theories that contain more than one coupling. When we generalize the argument given above, both λ and β_λ become vectors, while the coefficients β_i become tensors with $i + 2$ indices, and the coefficients a_i become tensors with $i + 1$ indices. The coefficients of the transformed beta function can be set to zero by solving linear recursive equations that have $i + 2$ indices, but their unknowns just have $i + 1$ indices. The solution does not exist, in general.

3.5 A deeper look into the renormalization group

If we insert the one-loop values (2.44) of the φ^4 renormalization constants into the inverse formulas (3.8) and (3.13), we can reconstruct the renormalization constants Z_λ and, for example, Z_{φ^2} . Then we find something interesting.

Working in the minimal subtraction scheme, we obtain

$$\bar{Z}_\lambda(\lambda, \varepsilon) = \frac{1}{1 - \frac{3\lambda}{16\pi^2\varepsilon}}, \quad \bar{Z}_{\varphi^2}(\lambda, \varepsilon) = \left(1 - \frac{3\lambda}{16\pi^2\varepsilon}\right)^{1/3}. \quad (3.55)$$

These results give the correct values (2.44) (at $c_1 = c_2 = 0$) at $\mathcal{O}(\lambda)$ and the two-loop double poles agree with formula (2.65). However, they tell us much more.

Now we include the higher-order corrections to the beta function. We insert them in (3.8) by writing $\bar{\beta}_\lambda(\lambda) = \lambda \sum_{i=1}^{\infty} \bar{\beta}_i \lambda^i$, and find

$$\begin{aligned} \bar{Z}_\lambda(\lambda, \varepsilon) &= \exp \left(\int_0^\lambda \frac{d\lambda'}{\lambda'} \frac{\sum_{i=1}^{\infty} \bar{\beta}_i \frac{\lambda'^i}{\varepsilon}}{1 - \sum_{j=1}^{\infty} \bar{\beta}_j \frac{\lambda'^j}{\varepsilon}} \right) \\ &= 1 + \bar{\beta}_1 \frac{\lambda}{\varepsilon} f_1 \left(\bar{\beta}_1 \frac{\lambda}{\varepsilon} \right) + \sum_{i=2}^{\infty} \bar{\beta}_i \frac{\lambda^i}{\varepsilon} f_i \left(\bar{\beta}_1 \frac{\lambda}{\varepsilon}, \bar{\beta}_2 \frac{\lambda^2}{\varepsilon}, \dots, \bar{\beta}_i \frac{\lambda^i}{\varepsilon} \right), \end{aligned} \quad (3.56)$$

where the functions f_i are power series of their arguments and receive contributions from the j th orders, $j \leq i$, of the beta function. We see that the maximum poles λ^n/ε^n , $n > 0$, even those that are due to diagrams with arbitrary many loops, are not affected by the higher-order corrections, but depend only on the one-loop coefficient of the beta function. Resumming them, we find

$$Z_\lambda(\lambda, \varepsilon) = \frac{1}{1 - \frac{\beta_1 \lambda}{\varepsilon}} + \sum_{i=2}^{\infty} \frac{\lambda^i}{\varepsilon} f_i.$$

The first two coefficients of the beta function contribute to the poles that have the form $(\lambda^2/\varepsilon)^n (\lambda/\varepsilon)^m$, with $n > 0$ and $m \geq 0$. However, they do not determine all of them, because the same powers of λ and ε can be obtained in different ways. For example, λ^4/ε^2 can be viewed as $(\lambda^2/\varepsilon)^2$, which is next-to-maximum, or $(\lambda^3/\varepsilon)(\lambda/\varepsilon)$. It is better to reorganize (3.56) as

$$\bar{Z}_\lambda(\lambda, \varepsilon) = 1 + \sum_{i=1}^{\infty} \frac{\lambda^i}{\varepsilon} g_i \left(\frac{\lambda}{\varepsilon} \right), \quad (3.57)$$

where the functions g_i are power series that depend only on the first i th coefficients of the beta function. The *next-to-maximum* poles are those of the form $(\lambda^2/\varepsilon)(\lambda/\varepsilon)^m$, $m \geq 0$, the next-to-next-to-maximum poles are those

of the form $(\lambda^3/\varepsilon)(\lambda/\varepsilon)^m$, $m \geq 0$, etc. Since the power of λ coincides with the power of \hbar , the poles are organized according to the general scheme

$$\begin{array}{ccccccc}
 L = 1 & \frac{\hbar}{\varepsilon} & & & & & \\
 L = 2 & \frac{\hbar^2}{\varepsilon^2} & \frac{\hbar^2}{\varepsilon} & & & & \\
 L = 3 & \frac{\hbar^3}{\varepsilon^3} & \frac{\hbar^3}{\varepsilon^2} & \frac{\hbar^3}{\varepsilon} & & & \\
 L = 4 & \frac{\hbar^4}{\varepsilon^4} & \frac{\hbar^4}{\varepsilon^3} & \frac{\hbar^4}{\varepsilon^2} & \frac{\hbar^4}{\varepsilon} & & \\
 \dots & & & & & &
 \end{array} \tag{3.58}$$

The one-loop coefficient of the beta function determine the first column (i.e. the maximum poles), the one- and two-loop coefficients determine the first two columns (i.e. the maximum and the next-to-maximum poles), and so on: the j -loop coefficients, $j \leq i$, of the beta function determine the first i columns. Some terms on the i th column receive contributions from the j -loop coefficients with $j < i$, but they can also receive contributions from the i -loop coefficient. Brand new information is contained only along the diagonal. Each nondiagonal element corresponds to some type of subdivergence.

For example, if we include the two-loop corrections to the beta function, $\beta_\lambda = \beta_1 \lambda^2 + \beta_2 \lambda^3 + \mathcal{O}(\lambda^4)$, we can determine Z_λ up to the next-to-maximum poles. We get

$$\bar{Z}_\lambda(\lambda, \varepsilon) = \frac{1 + \frac{\beta_2 \varepsilon}{\beta_1^2} \ln \left(1 - \frac{\beta_1 \lambda}{\varepsilon} \right)}{1 - \frac{\beta_1 \lambda}{\varepsilon}} + \frac{\beta_2 \lambda}{\beta_1 \left(1 - \frac{\beta_1 \lambda}{\varepsilon} \right)^2} = \frac{1}{1 - \frac{\beta_1 \lambda}{\varepsilon}} + \frac{\lambda^2}{\varepsilon} g_2 \left(\frac{\lambda}{\varepsilon} \right).$$

The poles of a generic correlation function G have a similar structure, where now the first i columns receive contributions from the j -loop coefficients, $j \leq i$, of the beta functions and the anomalous dimensions. We have, in the minimal subtraction scheme,

$$G(\hbar, \varepsilon) = G_c + \sum_{i=1}^{\infty} \frac{\hbar^i}{\varepsilon} G_i \left(\frac{\hbar}{\varepsilon} \right), \tag{3.59}$$

where G_c is the classical contribution and G_i are power series in \hbar/ε . The i th column of (3.58) is $\hbar^i G_i/\varepsilon$.

Now, assume that the first coefficient β_1 vanishes. Then (3.56) becomes

$$\bar{Z}_\lambda(\lambda, \varepsilon) = 1 + \sum_{i=2}^{\infty} \bar{\beta}_i \frac{\lambda^i}{\varepsilon} f_i \left(\bar{\beta}_2 \frac{\lambda^2}{\varepsilon}, \dots, \bar{\beta}_i \frac{\lambda^i}{\varepsilon} \right).$$

The first column of scheme (3.58) disappears, the second column is just made of its top element, and the other column are made by their top elements and sparse other elements. Again, the first i columns are determined by the first i coefficients of the beta function. Similar restrictions apply when the first two coefficients of the beta function vanish, and so on.

Another way to reach the conclusions derived above is to write

$$\ln \bar{Z}_\lambda = \sum_{i=1}^{\infty} \frac{c_i(\lambda)}{\varepsilon^i},$$

where $c_i(\lambda)$ are power series in λ that begin with $\mathcal{O}(\lambda^i)$. Next, we insert this expression into formula (3.7) (with $p = 1$). Equating each order in ε we get

$$\bar{\beta}_\lambda = \lambda^2 \frac{dc_1(\lambda)}{d\lambda}, \quad \frac{dc_i(\lambda)}{d\lambda} = \frac{\bar{\beta}_\lambda}{\lambda} \frac{dc_{i-1}(\lambda)}{d\lambda} \quad \text{for } i > 1.$$

Knowing that $c_i(\lambda) = \mathcal{O}(\lambda^i)$ we find

$$c_1(\lambda) = \int_0^\lambda \frac{\bar{\beta}_{\lambda'}(\lambda')}{\lambda'^2} d\lambda', \quad c_i(\lambda) = \int_0^\lambda \frac{dc_{i-1}(\lambda')}{d\lambda'} \frac{\bar{\beta}_{\lambda'}(\lambda')}{\lambda'} d\lambda' \quad \text{for } i > 1.$$

We see that the one-loop beta function determines the orders $\mathcal{O}(\lambda^i)$ of all the $c_i(\lambda)$ s. Similarly, the two-loop beta function determines all the $c_i(\lambda)$ s up to and including $\mathcal{O}(\lambda^{i+1})$ and the k -loop beta function determines them to and including $\mathcal{O}(\lambda^{i+k-1})$.

To summarize, the power of the renormalization group is that it relates infinitely many quantities, such as the entries of the columns of (3.58), and allows us to resum them. A consequence is that computing the entries of the same column involves more or less the same level of difficulty. As a check, we suggest the reader to compute the two-loop double poles of the φ_4^4 theory, which is part of exercise 6. It may be easily realized that if we just want the poles $1/\varepsilon^2$, we can considerably reduce the effort of the calculation. In the end, the two-loop double poles do not involve the typical difficulty of a two-loop calculation, but the one of a one-loop calculation. The same occurs with the three-loop triple poles, and so on.

To further appreciate the meaning of these facts, consider the formula (3.48) of the one-loop running coupling and compare it to the one-loop bare coupling:

$$\tilde{\lambda}(t, \lambda) = \frac{\lambda}{1 - \frac{3t\lambda}{16\pi^2}}, \quad \lambda_B \mu^{-\varepsilon} = \frac{\lambda}{1 - \frac{3\lambda}{16\pi^2\varepsilon}}.$$

We see that $\lambda_B \mu^{-\varepsilon}$ is nothing but the running coupling $\tilde{\lambda}(t, \lambda)$ at $t = 1/\varepsilon$. In a cutoff regularization framework, it would be the running coupling at an energy scale equal to the cutoff Λ ($1/\varepsilon \sim \ln(\Lambda/\mu)$). The resummation of the powers of (λ/ε) in $\lambda_B \mu^{-\varepsilon}$ is just like the resummation of the powers of $(t\lambda)$ in $\tilde{\lambda}(t, \lambda)$.

Actually, there is a much closer relation between the two resummations. For definiteness, assume that the theory contains a unique dimensionless coupling λ and consider the correlation functions

$$\begin{aligned} W_{\alpha_1 \dots \alpha_n}(x, y, \lambda, \mu) &= \langle \varphi_{\alpha_1}(x_1) \cdots \varphi_{\alpha_n}(x_n) \rangle_c \\ &= \prod_{i=1}^n \exp \left(\int_{\lambda}^{\tilde{\lambda}(t)} d\lambda' \frac{\gamma_{\alpha_i}(\lambda')}{\beta_{\lambda}(\lambda')} \right) W_{\alpha_1 \dots \alpha_n}(x, y, \tilde{\lambda}(t), \mu), \end{aligned} \quad (3.60)$$

where we have used the solution (3.34) of the Callan-Symanzik equation, and rewritten (3.35) by means of an integral on the coupling. Now, the limit of integration λ that appears on the right-hand side of (3.60) has no physical meaning, since it can be absorbed into the normalization of the fields. Obviously, the cross sections and the other physical quantities do not depend on such normalizations. Thus, ignoring that limit of integration, the right-hand side of (3.60) depends on the coupling just through the running coupling $\tilde{\lambda}(t)$. If $\tilde{\lambda}$ is small, the perturbative expansion of the right-hand side does in powers of $\tilde{\lambda}$ makes sense. Formula (3.44) shows that when $|t|$ is large, the running coupling can be small even if λ is of order one. The point is that if λ is of order one, the perturbative expansion of the left-hand side of (3.60) does not make sense. In other words, the renormalization group teaches us that, after resumming the powers of λ into $\tilde{\lambda}$, the perturbative expansion can become meaningful, if it is understood in the sense offered by the right-hand side of (3.60) (apart from an overall constant that is not physically important).

We have already remarked that little is known about the resummation of the perturbative expansion in quantum field theory, to the extent that different resummations may give different results and correspond to different physical theories. The renormalization group helps us on this, because it allows us to partially resum the perturbative expansion in powers of λ . Precisely, although we started with the aim of defining a perturbative expansion in powers of λ , and we ended by discovering that we may be able to define the perturbative theory even if λ is of order one, at least in a domain of

energies where the running coupling $\tilde{\lambda}$ is small. Indeed, the renormalization group equations tell us that, if we appropriately resum the powers of $t\lambda$, we can *reorganize* the expansion as an expansion in powers of $\tilde{\lambda}$. So doing, we are resumming the analogues of the columns of (3.58).

Chapter 4

Gauge symmetry

In this chapter we begin the study of Abelian and non-Abelian gauge symmetries in quantum field theory. After giving the basic notions and the main properties, we discuss the problems raised by their quantization, such as the gauge fixing and unitarity. In the next chapter we upgrade the formalism to make it suitable to prove the renormalizability of gauge theories to all orders. Then we proceed by proving the renormalizability of quantum electrodynamics (chapter 6) and the renormalizability of non-Abelian gauge field theories (chapter 7). In this part of the book, we assume that the theory is parity invariant, so no chiral fermions are present. The renormalization of parity violating quantum field theories raises bigger issues.

4.1 Abelian gauge symmetry

The propagation of free massless vector fields A_μ is described by the massless limit of the Proca action (1.87),

$$S_{\text{free}}(A) = \int d^D x \frac{1}{4} F_{\mu\nu}^2, \quad (4.1)$$

where $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ is the field strength. This action is invariant under the gauge symmetry

$$A'_\mu = A_\mu + \partial_\mu \Lambda, \quad (4.2)$$

where Λ is an arbitrary function of the position. In infinitesimal form, the symmetry transformation reads

$$\delta_\Lambda A_\mu = \partial_\mu \Lambda. \quad (4.3)$$

We have already written the action (4.1) in complex D dimensions, because one of the main virtues of the dimensional regularization is that it is manifestly gauge invariant, as long as the theory does not contain chiral fermions. Gauge invariance looks formally the same in all (integer) dimensions, so it is easy to generalize it to the formal objects A_μ , ∂_μ , x^μ , γ^μ , ψ , etc., that are used in the dimensional regularization. Instead, the notion of chirality is dimension dependent, so gauge invariance is not manifest in D dimensions when the Lagrangian explicitly contains γ_5 .

A direct consequence of the local gauge symmetry is that the quadratic part of the action (4.1) is not invertible. Indeed, it is proportional to $k^2\delta_{\mu\nu} - k_\mu k_\nu$ in momentum space, and has a null eigenvector k_ν . Therefore, the Green function $\langle A_\mu(k)A_\nu(-k) \rangle$ is not well defined. This fact is also evident by taking the massless limit of the Proca propagator (1.89), which is singular.

The free fermion action (1.100) is invariant under the global $U(1)$ transformation

$$\psi' = e^{-ie\Lambda}\psi, \quad (4.4)$$

where Λ is constant. The photon A_μ is the gauge field that promotes the global $U(1)$ invariance (4.4) to a local symmetry

$$A'_\mu = A_\mu + \partial_\mu\Lambda, \quad \psi' = e^{-ie\Lambda}\psi, \quad \bar{\psi}' = e^{ie\Lambda}\bar{\psi}, \quad (4.5)$$

where now Λ is a function of the spacetime point.

Replacing the simple derivative ∂_μ with the covariant derivative $\partial_\mu + ieA_\mu$ and adding (4.1), we obtain the Lagrangian of quantum electrodynamics (QED)

$$\mathcal{L}_0 = \frac{1}{4}F_{\mu\nu}^2 + \bar{\psi}(\not{\partial} + ie\not{A} + m)\psi, \quad (4.6)$$

which is invariant under the gauge transformation (4.5). In infinitesimal form, (4.5) becomes

$$\delta_\Lambda A_\mu = \partial_\mu\Lambda, \quad \delta_\Lambda\psi = -ie\Lambda\psi, \quad \delta_\Lambda\bar{\psi} = ie\Lambda\bar{\psi}. \quad (4.7)$$

4.2 Gauge fixing

To define the functional integral of a gauge theory, actually its perturbative expansion, we first need to choose a gauge, by imposing a condition of the form

$$\mathcal{G}(A) = 0,$$

where $\mathcal{G}(A)$ is a suitable local function. Later we must show that the physical quantities do not depend on the choice we make. Among the most popular gauge choices we mention the Lorenz gauge fixing

$$\mathcal{G}(A) = \partial_\mu A_\mu,$$

and the Coulomb gauge fixing

$$\mathcal{G}_C(A) = -\nabla \cdot \mathbf{A}. \quad (4.8)$$

Now we describe how to implement the gauge fixing at the quantum level. Start from the functional integral

$$Z(J) = \int [dA] \exp \left(-S(A) + \int JA \right)$$

in the absence of matter. The functional integration measure is certainly gauge invariant, since the gauge transformation is just a translation. For the moment, we assume that the current J is divergenceless, so the generating functional $Z(J)$ is formally gauge invariant. We know that $Z(J)$ is ill defined, because we are integrating also on the longitudinal mode (4.3) and the integrand is independent of it. For the moment, we ignore this fact and proceed.

Let us insert “1”, written in the form

$$1 = \int [d\Lambda] (\det \square) \delta_F(-\partial_\mu A_\mu + Q + \square\Lambda), \quad (4.9)$$

where Q is an arbitrary function. Here $\delta_F(Y)$ denotes the “functional delta function”, which means the product of $\delta(Y(x))$ over all spacetime points x , where $Y(x)$ is a function of the point. Formula (4.9) is the functional version of the ordinary formula

$$\int \prod_{i=1}^n dx_i \mathcal{J}(x) \prod_{k=1}^n \delta(f_k(x)) = \sum_{\bar{x}} \text{sign} \mathcal{J}(\bar{x}). \quad (4.10)$$

where

$$\mathcal{J}(x) = \det \frac{\partial f_i(x)}{\partial x_j}$$

and \bar{x} are the points where the functions $f_k(x)$ simultaneously vanish. As long as it does not vanish, the right-hand side of (4.10) is just a normalization factor, which can be omitted. In our case it is precisely 1, so we get

$$Z(J) = \int [dA d\Lambda] (\det \square) \delta_F(-\partial_\mu A_\mu + Q + \square \Lambda) \exp \left(-S(A) + \int JA \right).$$

Now, perform a change of variables $A' = A - \partial \Lambda$. Recalling that J is divergenceless and the functional measure is invariant under translations, after dropping the primes we obtain

$$Z(J) = \int [d\Lambda] \int [dA] (\det \square) \delta_F(-\partial_\mu A_\mu + Q) \exp \left(-S(A) + \int JA \right).$$

We see that the integral over the longitudinal mode Λ factorizes. Since the normalization of Z is physically irrelevant, we can equivalently define

$$Z(J) = \int [dA] (\det \square) \delta_F(-\partial_\mu A_\mu + Q) \exp \left(-S(A) + \int JA \right),$$

which is gauge-fixed. At this point, introduce a “Lagrange multiplier” B^1 and write the functional delta function as

$$\delta_F(-\partial_\mu A_\mu + Q) = \int [dB] \exp \left(-i \int d^D x B (\partial_\mu A_\mu - Q) \right).$$

We prefer to work with Hermitian quantities in the exponent, so we replace B with $-iB$. This operation factorizes an irrelevant constant in front of the integral, which we omit henceforth. Then

$$Z(J) = \int [dA dB] (\det \square) \exp \left(- \int d^D x \left(\frac{1}{4} F_{\mu\nu}^2 + B \partial_\mu A_\mu - BQ - JA \right) \right). \quad (4.11)$$

We see that the function Q plays the role of an external source for the Lagrange multiplier B . We can easily work out the propagator of the pair (A_μ, B) , and find

$$\begin{pmatrix} \langle A_\mu(k) A_\nu(-k) \rangle & \langle A_\mu(k) B(-k) \rangle \\ \langle B(k) A_\nu(-k) \rangle & \langle B(k) B(-k) \rangle \end{pmatrix} = \frac{1}{k^2} \begin{pmatrix} \delta_{\mu\nu} - \frac{k_\mu k_\nu}{k^2} & -ik_\mu \\ ik_\nu & 0 \end{pmatrix}. \quad (4.12)$$

¹Also known as Nakanishi-Lautrup auxiliary field, in this context.

The functional integral can be easily evaluated and gives

$$Z(J) = (\det \square)^{-1} \exp \left(\frac{1}{2} \int d^D x d^D y J_\mu(x) G_A(x-y) J_\mu(y) \right),$$

where the Green function $G_A(x-y)$ coincides with $G_B(x-y)$ at $m=0$, and we have used that J is divergenceless. We see that the result is independent of Q , which was expected, since Q is arbitrary in formula (4.9).

Formula (4.11) contains a $\det \square$ in the numerator, for which it is not easy to write Feynman rules. We could ignore this factor, because it is a just constant in QED. However, in more general gauge theories the analogue of this factor depends on the fields. If we introduce suitable anticommuting fields C and \bar{C} , called Faddeev-Popov *ghosts* and *antighosts*, respectively, we can exponentiate the determinant. The complete gauge-fixed functional then reads

$$Z(J) = \int [d\mu] \exp \left(- \int d^D x \left(\frac{1}{4} F_{\mu\nu}^2 + B \partial_\mu A_\mu - \bar{C} \square C - BQ - JA \right) \right), \quad (4.13)$$

where $[d\mu] \equiv [dAdCd\bar{C}dB]$. The ghost propagator is simply

$$\langle C(k) \bar{C}(-k) \rangle = \frac{1}{k^2}. \quad (4.14)$$

We can also relax the assumption that J is divergenceless. Then the functional integral (4.13) does depend on Q , but is still well defined. If we average over Q with the Gaussian measure

$$\int [dQ] \exp \left(- \frac{1}{2\lambda} \int d^D x Q(x)^2 \right),$$

where λ is an arbitrary parameter, we obtain

$$Z'(J) = \int [d\mu] \exp \left(- \int d^D x \left(\frac{1}{4} F_{\mu\nu}^2 - \frac{\lambda}{2} B^2 + B \partial \cdot A - \bar{C} \square C - JA \right) \right). \quad (4.15)$$

Since B now appears quadratically, it can be easily integrated away, giving

$$Z'(J) = \int [d\mu] \exp \left(- \int d^D x \left(\frac{1}{4} F_{\mu\nu}^2 + \frac{1}{2\lambda} (\partial \cdot A)^2 - \bar{C} \square C - JA \right) \right), \quad (4.16)$$

where now $[d\mu] = [dAdCd\bar{C}]$.

In this framework, the ghost propagator (4.14) is unchanged, while the gauge-field propagator reads

$$\langle A_\mu(k) A_\nu(-k) \rangle = \frac{1}{k^2} \left(\delta_{\mu\nu} + (\lambda - 1) \frac{k_\mu k_\nu}{k^2} \right). \quad (4.17)$$

A simple gauge choice is the *Feynman gauge*, where $\lambda = 1$ and

$$\langle A_\mu(k) A_\nu(-k) \rangle = \frac{\delta_{\mu\nu}}{k^2}. \quad (4.18)$$

The choice $\lambda = 0$ is also known as the *Landau gauge*.

Everything we said so far can be repeated by replacing $\partial_\mu A_\mu$ in (4.9) with the Coulomb gauge fixing (4.8). Then we get

$$Z'(J) = \int [d\mu] \exp \left(- \int d^D x \left(\frac{1}{4} F_{\mu\nu}^2 - \frac{\lambda}{2} B^2 - B \nabla \cdot \mathbf{A} + \bar{C} \Delta C - JA \right) \right), \quad (4.19)$$

and, after integrating B away,

$$Z'(J) = \int [d\mu] \exp \left(- \int d^D x \left(\frac{1}{4} F_{\mu\nu}^2 + \frac{1}{2\lambda} (\nabla \cdot \mathbf{A})^2 + \bar{C} \Delta C - JA \right) \right). \quad (4.20)$$

For the moment we content ourselves with these two choices of gauge fixing. However, in principle the gauge-fixing function $\mathcal{G}(A)$ can be arbitrary, as long as it properly fixes the gauge. Later we will see how to define the functional integral with an arbitrary $\mathcal{G}(A)$.

The arguments given above are formal, however the final result is correct. We can take the final result as the definition of the functional integral for gauge fields and then prove that this definition satisfies the properties we need.

Physical degrees of freedom

The physical degrees of freedom are more clearly visible in the Coulomb gauge. Indeed, formula (4.20) gives us the gauge-field propagators, which in Minkowski spacetime read

$$\begin{aligned} \langle A_0(k) A_0(-k) \rangle_{\text{M}} &= \frac{1}{\mathbf{k}^2} - \frac{\lambda E^2}{(\mathbf{k}^2)^2}, & \langle A_0(k) A_i(-k) \rangle_{\text{M}} &= -\frac{\lambda E k_i}{(\mathbf{k}^2)^2}, \\ \langle A_i(k) A_j(-k) \rangle_{\text{M}} &= \frac{1}{E^2 - \mathbf{k}^2} \left(\delta_{ij} - \frac{k_i k_j}{\mathbf{k}^2} \right) - \frac{\lambda k_i k_j}{(\mathbf{k}^2)^2}. \end{aligned} \quad (4.21)$$

To switch from the Euclidean framework to the Minkowskian one we have written $A = (iA_0, \mathbf{A})$ and $k = (iE, \mathbf{k})$, and then recalled that the Fourier transforms of the fields get a further factor i . Studying the poles of (4.21), we see that only $\langle A_i A_j \rangle$ has any, precisely two. They have positive residues and their dispersion relations are $E = |\mathbf{k}|$. The ghost propagator is now

$$\langle C(k) \bar{C}(-k) \rangle_M = \frac{1}{\mathbf{k}^2} \quad (4.22)$$

and has no pole. In total, the physical degrees of freedom are 2, as expected.

In the Lorenz gauge the propagators have a more complicated pole structure. For example, the ghost propagator (4.14) has poles and the gauge-field propagators (4.17) and (4.18) have extra poles. We will show that the unphysical degrees of freedom that appear with an arbitrary choice of gauge fixing compensate one another. More precisely, we will prove that the physical quantities do not depend on the gauge fixing, and that property will allow to freely switch to the Coulomb gauge, where no unphysical poles are present.

When we add matter, the gauge-fixing procedure does not change. For example, in the Lorenz gauge the gauge-fixed Lagrangian of QED is

$$\mathcal{L}_{\text{gf}} = \frac{1}{4} F_{\mu\nu}^2 + \bar{\psi}(\not{\partial} + ie\not{A} + m)\psi - \frac{\lambda}{2} B^2 + B \partial \cdot A - \bar{C} \square C,$$

before integrating B out. For completeness, we report the propagator of the multiplet made of A_μ and B

$$\begin{pmatrix} \langle A_\mu(k) A_\nu(-k) \rangle & \langle A_\mu(k) B(-k) \rangle \\ \langle B(k) A_\nu(-k) \rangle & \langle B(k) B(-k) \rangle \end{pmatrix} = \frac{1}{k^2} \begin{pmatrix} \delta_{\mu\nu} + (\lambda - 1) \frac{k_\mu k_\nu}{k^2} & -ik_\mu \\ ik_\nu & 0 \end{pmatrix}, \quad (4.23)$$

which coincides with (4.12) at $\lambda = 0$.

The gauge-fixing procedure we have described breaks the local symmetry (4.3). Nevertheless, the symmetry is not truly lost, because the functional integral acquires new properties. Thanks to those, we will be able to prove that the physical quantities are gauge invariant and gauge independent, before and after renormalization. Such properties are elegantly combined in a very practical and compact canonical formalism. That formalism is actually more than we need for Abelian theories, but has the virtue of providing a

unified treatment that is also suitable to treat the non-Abelian gauge theories, as well as quantum gravity and every general gauge theory. Thus, before introducing the canonical formalism for the gauge symmetry, we introduce non-Abelian gauge symmetry.

4.3 Non-Abelian global symmetry

Consider a multiplet ψ^i of fermionic fields. The free Lagrangian

$$\sum_i \left(\bar{\psi}^i \not{\partial} \psi^i + m \bar{\psi}^i \psi^i \right)$$

is invariant under the global transformation

$$\psi^{i'} = U^{ij} \psi^j, \quad \bar{\psi}^{i'} = \bar{\psi}^j U^{\dagger j i}, \quad (4.24)$$

where U is a unitary matrix. More generally, given a non-Abelian group G , we can consider multiplets ψ^i of fermionic fields that transform according to some representation of G , and theories that are symmetric with respect to these global transformations.

We focus our attention on the case $G = SU(N)$, where U can be parametrized as

$$U = \exp(-g \Lambda^a T^a), \quad (4.25)$$

by using a basis of $N \times N$ traceless anti-Hermitian matrices T_{ij}^a , where g and Λ^a are real constants and a is an index ranging from 1 to $\dim G = N^2 - 1$.

Consider the commutator $[T^a, T^b]$: since it is traceless and anti-Hermitian, it can be expanded in the basis T^a . We have

$$[T^a, T^b] = f^{abc} T^c, \quad (4.26)$$

where f^{abc} are real numbers such that

$$f^{abc} = -f^{bac}, \quad (4.27)$$

$$f^{abc} f^{cde} + f^{dac} f^{cbe} + f^{bdc} f^{cae} = 0. \quad (4.28)$$

The second line follows from the Jacobi identity of the anticommutator.

The matrices T^a can be normalized so that

$$\text{tr}[T^a T^b] = -\frac{1}{2} \delta^{ab}, \quad (4.29)$$

where the sign is determined by the anti-Hermiticity, while the factor $1/2$ is conventional. In a basis where (4.29) holds, the constants f^{abc} are completely antisymmetric, which can be proved from

$$\text{tr}[T^c[T^a, T^b]] = -\frac{1}{2}f^{abc},$$

by using the cyclicity of the trace.

For example, in the case $G = SU(2)$ we have $T^a = -i\sigma^a/2$, where σ^a are the Pauli matrices in the standard basis, and $f^{abc} = \epsilon^{abc}$.

Any real constants f^{abc} that satisfy the properties (4.27) and (4.28) define a *Lie algebra*. The f^{abc} s are called *structure constants* of the algebra. We can introduce abstract *generators* \mathcal{T}^a that satisfy the formal commutation rules

$$[\mathcal{T}^a, \mathcal{T}^b] = f^{abc}\mathcal{T}^c. \quad (4.30)$$

When the generators \mathcal{T}^a are given an explicit form, as matrices of some kind, we have a *representation* of the Lie algebra.

The commutation rules (4.30) define the Lie algebra associated with the group G . The set of $N \times N$ traceless anti-Hermitian matrices T^a form the *fundamental* representation of the algebra, which is the one of minimal dimension > 1 . It is commonly denoted with its dimension, which is equal to N . Taking the complex conjugate of (4.26), we obtain a new representation with $\mathcal{T}^a = \bar{\mathcal{T}}^a$, called *antifundamental*, commonly denoted with \bar{N} . The trivial representation, which has dimension 1, is called *singlet*.

In a generic representation r , the matrices \mathcal{T}^a can be normalized so that

$$\text{tr}[\mathcal{T}^a \mathcal{T}^b] = -C(r)\delta^{ab}, \quad (4.31)$$

where $C(r)$ is a constant depending on r . We have $C(N) = C(\bar{N}) = 1/2$. Another important identity is

$$\mathcal{T}^a \mathcal{T}^a = -C_2(r)\mathbb{1}, \quad (4.32)$$

where $C_2(r)$ is called *quadratic Casimir operator* of the representation r . The property (4.32) is proved below.

A consequence of the Jacobi identity (4.28) is that the matrices

$$(\mathcal{T}^a)^{bc} = -f^{abc}$$

satisfy the commutation rules (4.30), so they form another representation of the Lie algebra, called *adjoint representation*, normally denoted with G . Contracting a and b in (4.31) and tracing the equation (4.32), we get

$$C(r)d(G) = C_2(r)d(r),$$

where $d(r)$ is the dimension of the representation r . With $r = N$ we find $C_2(N) = (N^2 - 1)/(2N)$. Choosing $r = G$ we obtain that the two Casimirs of the adjoint representation coincide: $C(G) = C_2(G)$. It can be shown that $C(G) = C_2(G) = N$. In particular, using (4.31) for $r = G$, we get

$$f^{acd}f^{bcd} = N\delta^{ab} \quad (4.33)$$

Observe that any $N \times N$ matrix can be written as a complex linear combination of the identity matrix and the matrices T^a . Consider the tensor $\delta_{ij}\delta_{kl}$ as a $N \times N$ matrix in the indices j and k . It can be expanded as

$$\delta_{ij}\delta_{kl} = \alpha_{il}\delta_{kj} + \alpha_{il}^a T_{jk}^a, \quad (4.34)$$

where α_{il} and α_{il}^a are complex numbers. Taking $j = k$ we get

$$\delta_{il} = N\alpha_{il}. \quad (4.35)$$

Moreover, we also have

$$T_{li}^a = T_{kj}^a \delta_{ij} \delta_{kl} = \alpha_{il}^b \text{tr}[T^a T^b] = -\frac{1}{2}\alpha_{il}^a. \quad (4.36)$$

Collecting (4.35) and (4.36), formula (4.34) can be rephrased as

$$T_{ij}^a T_{kl}^a = -\frac{1}{2} \left(\delta_{il} \delta_{kj} - \frac{1}{N} \delta_{ij} \delta_{kl} \right), \quad (4.37)$$

an identity that is often handy.

We have started this discussion with fields ψ^i and $\bar{\psi}^i$ that, according to (4.24), transform in the fundamental and antifundamental representations, respectively. It is convenient to distinguish these two types of indices. We introduce the following notation. We call

$$v_{j_1 \dots j_m}^{i_1 \dots i_n}$$

a tensor whose upper n indices transform in the fundamental representation and lower m indices transform in the antifundamental representation. Globally and infinitesimally, we have

$$v_{j_1 \dots j_m}^{i_1 \dots i_n} = U_{j_1}^{\dagger l_1} \dots U_{j_m}^{\dagger l_m} U_{k_1}^{i_1} \dots U_{k_n}^{i_n} v_{l_1 \dots l_m}^{k_1 \dots k_n},$$

and

$$\begin{aligned} \delta_\Lambda v_{j_1 \dots j_m}^{i_1 \dots i_n} = & -g \sum_{a=1}^{\dim G} \Lambda^a \left(T^{a i_1}_{l_1} v_{j_1 \dots j_m}^{l_1 i_2 \dots i_n} + \dots + T^{a i_n}_{l_n} v_{j_1 \dots j_m}^{i_1 \dots i_{n-1} l_n} \right) \\ & + g \sum_{a=1}^{\dim G} \Lambda^a \left(T^{a k_1}_{j_1} v_{k_1 j_2 \dots j_m}^{i_1 \dots i_n} + \dots + T^{a k_m}_{j_m} v_{j_1 \dots j_{m-1} k_m}^{i_1 \dots i_n} \right), \end{aligned} \quad (4.38)$$

respectively. We have written the matrices T_{ij}^a as T^a_{ij} to emphasize the roles of their indices.

The tensors

$$\delta_j^i, \quad \varepsilon^{i_1 \dots i_N}, \quad \varepsilon_{i_1 \dots i_N}, \quad (4.39)$$

are clearly invariant. Observe that δ_j^i can contract only different types of indices.

Let u^i, v^i, \dots and u_i, v_i, \dots denote vectors transforming in the fundamental and antifundamental representations, respectively. We can construct new representations by considering the products $u^i v_j w^k \dots$. Using the tensors (4.39), the products of fundamental and antifundamental representations can be decomposed into sums of *irreducible* representations. The decomposition is obtained by repeatedly subtracting contributions proportional to the invariant tensors, until what remains vanishes whenever it is contracted with invariant tensors.

For example, the product $u^i v_j$ of a fundamental and an antifundamental representation can be decomposed as the sum of two irreducible representations, as follows

$$u^i v_j = \left(P_{1jk}^{il} + P_{2jk}^{il} \right) u^k v_l, \quad (4.40)$$

by means of the projectors

$$P_{1jk}^{il} = \delta_k^i \delta_j^l - \frac{1}{N} \delta_j^i \delta_k^l, \quad P_{2jk}^{il} = \frac{1}{N} \delta_j^i \delta_k^l. \quad (4.41)$$

It can be shown that the traceless combination given by P_1 is equivalent to the adjoint representation. The term proportional to δ_j^i is obviously a singlet.

We symbolically write the decomposition (4.40) as

$$N \otimes \bar{N} = \text{Adg} \oplus \mathbb{1}.$$

The matrices \mathcal{T}^{ail}_{jk} of the representation acting on the product $u^k v_l$ are $T^{ai}_{\ k} \delta^l_j - T^{al}_{\ j} \delta^i_k$. They do not need to be projected with (4.41), since they act nontrivially only on the adjoint combination.

Another example is the product $u^i v^j$ of two fundamental representations. It can be decomposed as the sum of the symmetric and antisymmetric combinations,

$$u^i v^j = \frac{1}{2}(u^i v^j + u^j v^i) - \frac{1}{2(N-2)!} \varepsilon^{ijk_1 \dots k_{N-2}} \varepsilon_{k_1 \dots k_{N-2} mn} u^m v^n,$$

which are new representations of dimensions $N(N+1)/2$ and $N(N-1)/2$, respectively. We have

$$\mathcal{T}^{aij}_{\ \ kl} = \frac{1}{2} \left(T^{ai}_{\ k} \delta^j_l + \delta^i_k T^{aj}_{\ l} \pm T^{aj}_{\ k} \delta^i_l \pm \delta^j_k T^{ai}_{\ l} \right),$$

or, symbolically,

$$\mathcal{T}^a_{\pm} = P_{\pm} (T^a \otimes \mathbb{1} + \mathbb{1} \otimes T^a) P_{\pm},$$

where P_{\pm} are the projectors on the symmetric and antisymmetric combinations, respectively.

A theorem ensures that all the representations can be obtained by means of a similar procedure.

Theorem 4 *All the irreducible finite dimensional representations can be obtained from the products of fundamental and antifundamental representations, decomposed by means of the invariant tensors (4.39).*

Actually, even the antifundamental representation can be obtained from the fundamental one. Indeed,

$$u_i \equiv \frac{1}{(N-1)!} \varepsilon_{ik_2 \dots k_N} v^{k_2} \dots v^{k_N}$$

does transform in the antifundamental representation. Thus,

Theorem 5 *All the irreducible finite dimensional representations can be obtained by decomposing products of fundamental representations.*

The theorem just stated ensures that

Corollary 6 *the generators \mathcal{T}^a of every representation can be written using the matrices $T^{ai}{}_j$ and the invariant tensors (4.39).*

Symbolically, we can write

$$\mathcal{T}_r^a = P_r(T^a \otimes \mathbf{1} \cdots \otimes \mathbf{1} + \cdots + \mathbf{1} \otimes \cdots \otimes \mathbf{1} \otimes T^a)P_r, \quad (4.42)$$

where P_r is the projector on the representation r , constructed with the tensors (4.39).

The identity (4.32) can be proved as follows. Observe that $\mathcal{T}_r^a \mathcal{T}_r^a$ is an invariant tensor, since T^a , as well as any \mathcal{T}_r^a , transforms in the adjoint representation. Using (4.42) and (4.37) we know that it has the form $P_r(Q_r)P_r$, where Q_r is constructed by means of Kronecker symbols. Thus, $\mathcal{T}_r^a \mathcal{T}_r^a$ is a constant times P_r itself, as wanted.

Expanding by means of the Kronecker tensor, contracting in all possible ways and using the identities already proved, we also find

$$f^{abc}T^{ai}{}_j T^{bk}{}_l T^{cm}{}_n = \frac{1}{4}(\delta_l^i \delta_n^k \delta_j^m - \delta_n^i \delta_j^k \delta_l^m). \quad (4.43)$$

4.4 Non-Abelian gauge symmetry

Now we want to gauge the non-Abelian global symmetry. We promote the unitary matrix (4.25) to a family of spacetime-dependent unitary matrices

$$U(x) = e^{-g\Lambda^a(x)T^a}. \quad (4.44)$$

and introduce the non-Abelian gauge fields A_μ , as well as the covariant derivative

$$(D_\mu \psi)^i = \partial_\mu \psi^i + ig(A_\mu \psi)^i, \quad (4.45)$$

where g denotes the gauge coupling. Formula (4.45) shows that the A_μ s must be matrices with indices ij . The gauge field A_μ is often called *gauge connection*.

We determine the transformation A'_μ of the gauge field by requiring that $(D_\mu \psi)^i$ transform exactly as ψ^i . We have

$$\begin{aligned} (D_\mu \psi)'^i &= \partial_\mu \psi'^i + ig(A'_\mu \psi')^i \\ &= (\partial_\mu U^{ij})\psi^j + ig(A'_\mu U - U A_\mu)^{ij}\psi^j + U^{ij}D_\mu \psi^j = U^{ij}D_\mu \psi^j, \end{aligned}$$

hence

$$A'_\mu = \frac{i}{g}(\partial_\mu U)U^{-1} + UA_\mu U^{-1}, \quad (4.46)$$

The transformation rule for D_μ is

$$D'_\mu = UD_\mu U^{-1}. \quad (4.47)$$

Since the covariant differential operator iD_μ must be Hermitian, as is $i\partial_\mu$, the matrices A_μ are also Hermitian, so they can be parametrized as

$$A_\mu^{ij}(x) = -iA_\mu^a(x)T_{ij}^a, \quad (4.48)$$

where $A_\mu^a(x)$ are real functions. We can write, in matrix and index notations, $D_\mu = \partial_\mu + igA_\mu$ and $D_\mu^{ij} = \delta^{ij}\partial_\mu + igA_\mu^{ij}$, respectively. Distinguishing upper and lower indices, we have $A_\mu^{ij} = -iT^{ai}{}_j A_\mu^a$, which emphasizes that A_μ is the traceless product of a fundamental and an antifundamental representation.

Define the field strength

$$F_{\mu\nu} = \frac{1}{g}[D_\mu, D_\nu] \equiv F_{\mu\nu}^a T^a.$$

Clearly, (4.47) implies the transformation rule

$$F'_{\mu\nu} = UF_{\mu\nu}U^{-1}. \quad (4.49)$$

We find

$$F_{\mu\nu}^a = \partial_\mu A_\nu^a - \partial_\nu A_\mu^a + gf^{abc}A_\mu^b A_\nu^c.$$

So far, we have taken matter fields ψ^i in the fundamental representation. For fields χ^I in a generic representation r , described by matrices \mathcal{T}_r^a , we have $A_\mu^{IJ} = -iA_\mu^a \mathcal{T}_{rIJ}^a$ and still $D_\mu = \partial_\mu + igA_\mu$, but $D_\mu^{IJ} = \delta^{IJ}\partial_\mu + igA_\mu^{IJ}$.

The infinitesimal forms of the transformations (4.24), (4.49) and (4.46) read

$$\begin{aligned} \delta_\Lambda A_\mu^a &= \partial_\mu \Lambda^a + gf^{abc}A_\mu^b \Lambda^c \equiv D_\mu \Lambda^a, \\ \delta_\Lambda F_{\mu\nu}^a &= gf^{abc}F_{\mu\nu}^b \Lambda^c, \quad \delta_\Lambda \chi^I = -g\mathcal{T}_{rIJ}^a \Lambda^a \chi^J. \end{aligned} \quad (4.50)$$

We can state that the functions Λ^a belong to the adjoint representation of G , so $\delta_\Lambda A_\mu^a$ is just the covariant derivative of Λ^a .

The gauge-invariant action of fermionic fields χ^I can be constructed by means of the covariant derivative, and reads

$$S_\psi = \int d^D x \bar{\chi} (\not{D} + m) \chi = \int d^D x \bar{\chi}^I (\delta_{IJ} \not{D} + g A_\mu^a \mathcal{T}_{rIJ}^a + m \delta_{IJ}) \chi^J. \quad (4.51)$$

Similarly, if φ are (complex) scalar fields transforming according to some representation of the gauge group, the scalar action is

$$S_\varphi = \int d^D x (|D_\mu \varphi|^2 + m^2 \bar{\varphi} \varphi), \quad (4.52)$$

up to other interactions. The invariant action of the gauge fields is

$$S_A = \frac{1}{4} \int d^D x F_{\mu\nu}^a{}^2 = -\frac{1}{2} \int d^D x \text{tr}[F_{\mu\nu}^2], \quad (4.53)$$

where $F_{\mu\nu}$ is written using the matrices T^a of the fundamental representation. The gauge invariance of S_A is a consequence of (4.49) and the cyclicity of the trace. The theory (4.53) is called non-Abelian Yang-Mills theory. Note that (4.53) is an interacting theory.

The free-field limit $g \rightarrow 0$ of (4.53) describes a set of $N^2 - 1$ free photons. For this reason, the propagator of non-Abelian Yang-Mills theory has the same problems as the propagator of QED, and can be defined only after a gauge fixing. Now, the gauge-fixing procedure is more involved than in quantum electrodynamics. It becomes relatively simple if we endow the gauge symmetry with a suitable canonical formalism, which is introduced in the next chapter.

Chapter 5

Canonical gauge formalism

Gauge symmetry can be treated efficiently by means of a canonical formalism of new type, known as Batalin-Vilkovisky formalism. It is equipped with suitable notions of parentheses, canonical transformations, and a number of tools that allow us to make a number of crucial operation with a relatively small effort. The “time” evolution associated with this formalism is just the gauge transformation. We do not need to explicitly introduce a “time” coordinate θ for this kind of evolution, because this θ would be constant anticommuting parameter, so every function of θ has a Taylor expansion that stops at the first order.

Briefly, the Batalin-Vilkovisky formalism is a practical tool to (i) gather the key properties of the infinitesimal symmetry transformations and their algebra in a single equation, (ii) fix the gauge and have control on the gauge fixing with a straightforward procedure, (iii) prove that gauge theories can be renormalized by preserving gauge invariance to all orders, (iv) prove that the physical quantities are gauge independent, and (v) study the anomalies of the global and gauge symmetries to all orders. Combined with the dimensional regularization (or its modifications and upgrades, to be defined later on), the Batalin-Vilkovisky formalism allows us to derive these and several other properties in a systematic way.

We generically refer to the resulting formal apparatus by calling it “canonical gauge formalism” for quantum field theory. Its main virtue is that it allows us to prove old and new results to all orders with much less effort than is required by the other approaches.

Although we mainly work with gauge (that is to say local) symmetries,

everything we say also works for global symmetries. At the same time, we stress that the canonical formalism is well suited to study infinitesimal symmetries. At present, there exists no equally compact and elegant formalism for finite or discrete symmetries.

In the rest of this chapter we mostly work at the bare level, but drop the subscripts $_B$ that we normally use to denote bare quantities. The properties we are going to derive can be interpreted at the classical level, if they concern the action S , and at the bare and tree levels, if they concern the generating functionals. The renormalization program will be carried out in detail in the next sections. Among other things, we will have to prove that the properties derived here are preserved by the subtraction of divergences.

We collect the classical fields into a single row

$$\phi^i = (A_\mu^a, \bar{\psi}, \psi, \varphi)$$

and assume that a classical action $S_c(\phi)$ is given, which is invariant under some infinitesimal transformations

$$\delta_\Lambda \phi^i = R_c^i(\phi, \Lambda), \quad (5.1)$$

that is to say

$$\delta_\Lambda S_c = \int \delta_\Lambda \phi^i \frac{\delta_l S_c}{\delta \phi^i} = \int R_c^i(\phi, \Lambda) \frac{\delta_l S_c}{\delta \phi^i} = 0. \quad (5.2)$$

Here $\Lambda(x)$ denote the local parameters of the symmetry.

5.1 General idea behind the canonical gauge formalism

We first introduce the basic ideas behind the canonical formalism, without paying too much attention to the details, such as the statistics of the fields, the correct relative positions of fields and sources and some crucial minus signs that will be dealt with shortly after. It is useful to have a general idea of what we want to do, before plunging into the technical aspects. Later, we go through the systematics.

The functions $R_c^i(\phi, \Lambda)$ are local, and linear in Λ . Apart from this, they can be arbitrary functions, nonlinear in ϕ , and renormalize independently of

the elementary fields. In other words, they must be regarded as composite fields. We know that composite fields can be treated by adding them to the action, coupled to external sources K . This defines the extended action

$$S_c^\Lambda(\phi, K) = S_c(\phi) + \int R_c^i(\phi, \Lambda) K_i.$$

The identity (5.2) can be written in the form

$$\delta_\Lambda S_c = (S_c^\Lambda, S_c) = \int d^D x \left\{ \frac{\delta S_c}{\delta \phi^i(x)} \frac{\delta S_c^\Lambda}{\delta K_i(x)} - \frac{\delta S_c^\Lambda}{\delta \phi^i(x)} \frac{\delta S_c}{\delta K_i(x)} \right\} = 0. \quad (5.3)$$

This expression is appealing, because it reminds us of a canonical formalism, once the sources K are viewed as canonically conjugate to the elementary fields ϕ . The symmetry transformation of a functional $X(\phi)$ can be expressed as the parenthesis with S_c^Λ :

$$\delta_\Lambda X = (S_c^\Lambda, X) = \int \frac{\delta X}{\delta \phi^i} R_c^i(\phi, \Lambda).$$

The identity (5.3), however, just tells us about the symmetry transformations, but does not incorporate the algebra of the transformations, in particular their closure. Closure means that the commutator $[\delta_\Lambda, \delta_\Sigma]$ of two transformations δ_Λ and δ_Σ , with parameters Λ and Σ , is a symmetry transformation $\delta_{\Delta(\Lambda, \Sigma)}$ of the same algebra, with certain parameters $\Delta(\Lambda, \Sigma)$:

$$[\delta_\Lambda, \delta_\Sigma] = \delta_{\Delta(\Lambda, \Sigma)}. \quad (5.4)$$

A priori, renormalization may affect both the action S_c and the transformations R_c^i , as well as the closure relations (5.4). Thus, it is important to collect these three pieces of information into a unique extended functional. So doing, we will be able build a powerful formalism that allows us to easily understand how those pieces of information renormalize and what role they play inside the generating functionals.

In formula (5.3) we have two different functionals, S_c and S_c^Λ . Moreover, S_c , does not contain the sources K . As said, we would like to collect all pieces of information into a unique extended functional, and find an identity that involves only that functional. Now, the parenthesis $(S_c^\Lambda, S_c^\Lambda) = 0$ is trivial, because it is the subtraction of the term

$$\int \frac{\delta S_c^\Lambda}{\delta \phi^i} \frac{\delta S_c^\Lambda}{\delta K_i} \quad (5.5)$$

with itself. Nevertheless, the expression (5.5) goes into the right direction, as we can see if we split it into two pieces, the contributions that do not contain K and the contributions that are linear in K :

$$\int R_c^i \frac{\delta S_c}{\delta \phi^i}, \quad \int R_c^i \frac{\delta R_c^j}{\delta \phi^i} K_j.$$

The former give the transformation of the action S_c and the latter somehow point to the transformation of the transformation, that is to say the closure of the algebra.

A trick to make the terms (5.5) to sum up instead of cancel out, is to change the relative statistics of ϕ^i and K_i , and distinguish left- and right-derivatives. The resulting definition of parentheses and all other details are given below. The formalism satisfies all the usual properties, appropriately adapted, including a generalized Jacobi identity. For the moment, we just anticipate that we get something like

$$(S_c^\Lambda, S_c^\Lambda) = 2 \int d^D x \frac{\delta S_c^\Lambda}{\delta \phi^i} \frac{\delta S_c^\Lambda}{\delta K_i}. \quad (5.6)$$

Again, this cannot be the final answer, because the “double” Λ transformation contained in (5.6) is not really a commutator. However, it becomes the commutator once we also play with the statistics of Λ and provide suitable transformation rules for the Λ s themselves.

Now, closure demands two independent parameters, e.g. Λ and Σ , so S_c^Λ is certainly inadequate to contain the closure relations (5.4). On a functional X , we have

$$[\delta_\Lambda, \delta_\Sigma]X = (S_c^\Lambda, (S_c^\Sigma, X)) - (S_c^\Sigma, (S_c^\Lambda, X)) = \delta_{\Delta(\Lambda, \Sigma)}X = (S_c^{\Delta(\Lambda, \Sigma)}, X). \quad (5.7)$$

We may expect that the generalized Jacobi identity for the parentheses allows us to replace the sum of terms that follows the first equal sign in (5.7) with something like $((S_c^\Lambda, S_c^\Sigma), X)$. Thus, (5.7) gives

$$((S_c^\Lambda, S_c^\Sigma) - S_c^{\Delta(\Lambda, \Sigma)}, X) = 0, \quad (5.8)$$

for every X , so we can express closure by means of a relation of the form

$$(S_c^\Lambda, S_c^\Sigma) = S_c^{\Delta(\Lambda, \Sigma)}.$$

We have two functionals instead of one here. We can collect everything into a single functional, if we replace $\Lambda(x)$ and $\Sigma(x)$ by $\theta C(x)$ and $\theta' C(x)$, respectively, where θ and θ' are anticommuting parameters (that we drop after moving them to the right or left of each identity) and $C(x)$ is an anticommuting field, to be identified with the Fadeev-Popov ghosts. Their main virtue is that they can carry an anticommutator by themselves. For example, if θ^a are also anticommuting quantities, we have

$$(\theta^a C^a)^2 = -\frac{1}{2}[\theta^a, \theta^b] C^a C^b.$$

This trick allows us to work with a unique, but anticommuting, $C(x)$ and a unique extended action. Later we will show that, if done properly, the operations encoded into the replacements $\Lambda \rightarrow \theta C$, $\Sigma \rightarrow \theta' C$ are completely reversible, so they do not cause any loss of information. Finally, $\Delta(C, C)$ is identified with the transformation of C itself, apart from a proportionality factor.

The new extended action is something of the form

$$S'_c(\phi, C, K) = S_c(\phi) + \int d^D x \left(K_i R_c^i(\phi, C) - \frac{1}{2} K_C \Delta(C, C) \right),$$

where K_C are sources for the C transformations. Next, we have an identity of the form

$$(S'_c, S'_c) = 2 \int d^D x \frac{\delta S'_c}{\delta \phi^i} \frac{\delta S'_c}{\delta K_i} + 2 \int d^D x \frac{\delta S'_c}{\delta C} \frac{\delta S'_c}{\delta K_C} = 0. \quad (5.9)$$

The terms proportional to K_i in this expression do give the closure of the algebra. The terms proportional to K_C cancel out by themselves, because they are just the Jacobi identity of the Lie algebra.

Summarizing, once the crucial identity $(S'_c, S'_c) = 0$ is satisfied, the extended action S'_c incorporates the invariant action, the symmetry transformations of the fields, the closure of the algebra and its Jacobi identity.

5.2 Systematics of the canonical gauge formalism

Without further premises, we are now ready to present the systematics of the canonical gauge formalism. Make the substitution $\Lambda(x) \rightarrow \theta C(x)$ in the

identity (5.2), then move θ to the far left and drop it. Since R^i is linear in Λ , we get an identity of the form

$$\int R^i(\phi, C) \frac{\delta_l S_c}{\delta \phi^i} = 0. \quad (5.10)$$

The functions $R^i(\phi, C)$ are such that

$$R_c^i(\phi, \theta C) = \theta R^i(\phi, C) \quad (5.11)$$

and may differ from $R_c^i(\phi, C)$ by a sign, depending on the statistics of ϕ .

The fields C are called this way, because they coincide with the Faddeev-Popov ghosts already met. For the moment, we do not need to introduce antighosts \bar{C} and Lagrange multipliers B . They are useful to fix the gauge, but they are not basic ingredients of the canonical formalism. We include the fields and the ghosts into the extended row

$$\Phi^\alpha = (A_\mu^a, C^a, \bar{\psi}, \psi, \varphi).$$

The conjugate row made by the sources is

$$K_\alpha = (K_a^\mu, K_C^a, K_{\bar{\psi}}, K_\psi, K_\varphi).$$

We define the statistics ε_Φ , ε_K , ε_λ , ε_X of a field Φ , a source K , a parameter λ or a functional X to be zero if the field, source, parameter or functional is bosonic, one if it is fermionic. We define the statistics of the sources as opposite to the statistics of the fields that are conjugate to them:

$$\varepsilon_{K_\alpha} = \varepsilon_{\Phi^\alpha} + 1 \bmod 2. \quad (5.12)$$

Given two functionals $X(\Phi, K)$ and $Y(\Phi, K)$ of the fields and sources, we define their *antiparentheses* as the functional

$$(X, Y) \equiv \int d^D x \left\{ \frac{\delta_r X}{\delta \Phi^\alpha(x)} \frac{\delta_l Y}{\delta K_\alpha(x)} - \frac{\delta_r X}{\delta K_\alpha(x)} \frac{\delta_l Y}{\delta \Phi^\alpha(x)} \right\}, \quad (5.13)$$

where the sum over α is understood. Observe that if X and Y are local functionals, then (X, Y) is a local functional.

The antiparentheses satisfy the properties

$$\begin{aligned} (Y, X) &= -(-1)^{(\varepsilon_X+1)(\varepsilon_Y+1)}(X, Y), \\ (-1)^{(\varepsilon_X+1)(\varepsilon_Z+1)}(X, (Y, Z)) &+ \text{cyclic permutations} = 0, \end{aligned} \quad (5.14)$$

and $\varepsilon_{(X,Y)} = \varepsilon_X + \varepsilon_Y + 1$, which can be verified straightforwardly. In particular, formula (5.14) is the Jacobi identity. Immediate consequences are

$$(F, F) = 0, \quad (B, B) = 2 \int \frac{\delta_r B}{\delta \Phi^\alpha} \frac{\delta_l B}{\delta K_\alpha} = -2 \int \frac{\delta_r B}{\delta K_\alpha} \frac{\delta_l B}{\delta \Phi^\alpha}, \quad (5.15)$$

if the functionals F and B have fermionic and bosonic statistics, respectively. In (5.15), as often below, we understand integrations over spacetime points associated with repeated indices α, β, \dots . Another important consequence is

$$(X, (X, X)) = 0 \quad (5.16)$$

for every functional X . This property follows from the Jacobi identity (5.14) and is useful to study the anomalies.

The action $S(\Phi, K)$ is defined as the solution of the *master equation*

$$(S, S) = 0, \quad (5.17)$$

with the *boundary conditions*

$$S(\Phi, 0) = S_c(\phi), \quad -\frac{\delta_r S(\Phi, K)}{\delta K_i} \Big|_{K=0} = R^i(\phi, C). \quad (5.18)$$

In the naïve derivation given above, the extended action S'_C was linear in the sources K . This is actually true in all the applications we have in mind, at least at the tree level. Thus, we write the solution of the master equation in the form

$$S(\Phi, K) = S_c(\phi) + S_K(\Phi, K), \quad (5.19)$$

where

$$S_K(\Phi, K) = - \int R^\alpha(\Phi) K_\alpha = S_c(\phi) - \int R^i(\phi, C) K_i - \int R_C^a(\Phi) K_C^a, \quad (5.20)$$

and $R_C^a(\Phi)$ are functions to be determined, related somehow to $\Delta(\Lambda, \Lambda)$. The signs have been adjusted to match the choices of statistics we have made. It can be shown that the linearity of $S(\Phi, K)$ in K means that the algebra closes off shell.

More explicitly, using the last expression of (5.15), we find the formula

$$0 = (S, S) = 2 \int R^\alpha(\Phi) \frac{\delta_l S}{\delta \Phi^\alpha} = 2 \int \left[R^i(\phi, C) \frac{\delta_l S}{\delta \phi^i} + R_C^a(\Phi) \frac{\delta_l S}{\delta C^a} \right].$$

The terms of order 0 in K are twice the identity (5.10), while the terms of order 1 in K give the formula

$$0 = -2 \int R^\alpha(\Phi) \frac{\delta_l}{\delta \Phi^\alpha} \int R^\beta(\Phi) K_\beta,$$

which implies

$$0 = \int R^\alpha(\Phi) \frac{\delta_l R^\beta(\Phi)}{\delta \Phi^\alpha} \quad (5.21)$$

for every β . Taking $\beta = i$, we find

$$0 = \int R^j(\phi, C) \frac{\delta_l R^i(\phi, C)}{\delta \phi^j} + \int R_C^a(\Phi) \frac{\delta_l R^i(\phi, C)}{\delta C^a}. \quad (5.22)$$

Since R^i are linear in C , the last term is equals to $R_c^i(\phi, R_C(\Phi))$. Setting $C^a = \theta \Lambda^a + \theta' \Sigma^a$ in (5.22), where θ and θ' are both anticommuting parameters, we obtain

$$0 = \int R^j(\phi, \theta \Lambda^a + \theta' \Sigma) \frac{\delta_l R^i(\phi, \theta \Lambda^a + \theta' \Sigma)}{\delta \phi^j} + R_c^i(\phi, R_C(\theta \Lambda^a + \theta' \Sigma)). \quad (5.23)$$

Note that setting $C^a = \theta \Lambda^a$ in formula (5.11), we also obtain

$$\theta R_c^i(\phi, \Lambda^a) = R^i(\phi, \theta \Lambda^a). \quad (5.24)$$

Using this formula and taking the terms proportional to $\theta \theta'$ (all other terms being zero), let us write

$$R_C^a(\theta \Lambda + \theta' \Sigma) = -\theta \theta' \Delta(\Lambda, \Sigma). \quad (5.25)$$

This formula can be taken as the definition of $R_C(\Phi)$, where $\Delta(\Lambda, \Sigma)$ is assumed to be known from the closure relation (5.4).

Then formula (5.23) gives

$$\int \theta R_c^j(\phi, \Lambda^a) \frac{\delta_l \theta' R_c^i(\phi, \Sigma)}{\delta \phi^j} + \int \theta' R_c^j(\phi, \Sigma) \frac{\delta_l \theta R_c^i(\phi, \Lambda^a)}{\delta \phi^j} = \theta \theta' R_c^i(\phi, \Delta(\Lambda, \Sigma)).$$

Moving θ and θ' to the left and using (5.1), we obtain

$$\theta \theta' \int \left(\delta_\Lambda \phi^j \frac{\delta_l (\delta_\Sigma \phi^i)}{\delta \phi^j} - \delta_\Sigma \phi^j \frac{\delta_l (\delta_\Lambda \phi^i)}{\delta \phi^j} \right) = \theta \theta' R_c^i(\phi, \Delta(\Lambda, \Sigma)),$$

or, finally,

$$\theta \theta' [\delta_\Lambda, \delta_\Sigma] \phi^i = \theta \theta' \delta_{\Delta(\Lambda, \Sigma)} \phi^i,$$

which is equivalent to (5.4).

Finally, taking $R^\beta \rightarrow R_C^a$ in (5.21), we get

$$0 = \int R_C^b(C) \frac{\delta_l R_C^a(C)}{\delta C^b}, \quad (5.26)$$

having observed that $R_C^a(\Phi)$ in the end depends just on C . Formula (5.26), which is the closure of the closure, in some sense, is just the Jacobi identity of the lie algebra.

For example, in non-Abelian Yang-Mills theories, we have (on fermions ψ , for definiteness) $\delta_\Lambda \psi^i = -g T_{ij}^a \Lambda^a \psi^j$, so

$$[\delta_\Lambda, \delta_\Sigma] \psi^i = g^2 [T^a, T^b]_{ij} \psi^j \Sigma^a \Lambda^b = -g^2 T_{ij}^c \psi^j f^{abc} \Lambda^a \Sigma^b = \delta_{\Delta(\Lambda, \Sigma)} \psi^i,$$

hence

$$\Delta^a(\Lambda, \Sigma) = g f^{abc} \Lambda^b \Sigma^c.$$

Using this expression in (5.25), we find

$$R_C^a(\theta \Lambda + \theta' \Sigma) = -\theta \theta' g f^{abc} \Lambda^b \Sigma^c = -\frac{g}{2} g f^{abc} (\theta \Lambda + \theta' \Sigma)^b (\theta \Lambda + \theta' \Sigma)^c,$$

whence

$$R_C^a(C) = -\frac{g}{2} f^{abc} C^b C^c.$$

Thus, the identity (5.26) gives

$$0 = f^{abc} f^{bde} C^c C^d C^e.$$

Since the C s are anticommuting, this equation is equivalent to (4.28).

In the Abelian case, $\Delta(\Lambda, \Sigma) = 0$, so $R_C(\Phi) = 0$.

Given a functional Y , we can view the antiparentheses (Y, X) as a map acting on the space of functionals X . Choosing $Y = S$, the map (S, X) is nilpotent, because of the Jacobi identity (5.14) and the master equation (5.17). Indeed,

$$(S, (S, X)) = \frac{1}{2}((S, S), X) = 0.$$

On the fields and sources, we have

$$(S, \Phi^\alpha) = R^\alpha(\Phi), \quad (S, K_\alpha) = \frac{\delta_r S}{\delta \Phi^\alpha}.$$

Basically, (S, K_α) is the Φ^α field equation, plus $\mathcal{O}(K)$.

The map (S, X) sends functionals $\mathcal{G}(\Phi)$ that depend only on the fields into functionals that depend only on the fields:

$$(S, \mathcal{G}(\Phi)) = \int R^\alpha(\Phi) \frac{\delta_l \mathcal{G}(\Phi)}{\delta \Phi^\alpha}.$$

On the functionals $\mathcal{G}(\phi)$ that depend only on the physical fields ϕ the map is precisely the gauge transformation:

$$(S, \mathcal{G}(\phi)) = \int R^i(\phi, C) \frac{\delta_l \mathcal{G}(\phi)}{\delta \phi^i}.$$

In particular, $(S, S_c(\phi)) = 0$, which is nothing but the gauge invariance of the classical action $S_c(\phi)$.

It is always possible to generate solutions $S(\Phi, K)$ of the master equation that are nonlinear in K by means of field and source redefinitions that preserve the master equation (or the antiparentheses, in which case they are canonical transformations, see below). However, not all the solutions that are nonlinear in K can be obtained this way. To understand this issue better, consider again the relation

$$0 = (S, S) = -2 \int d^D x \frac{\delta_r S}{\delta K_\alpha} \frac{\delta_l S}{\delta \Phi^\alpha}. \quad (5.27)$$

The K -independent contributions are always the identity (5.10), but when $S(\Phi, K)$ is not linear in K , the terms of (5.27) that are linear in K (which encode the closure of the algebra) contain extra contributions proportional to the field equations. If there exists no canonical transformation that absorbs the extra terms away, it means that the gauge algebra does not close off shell, but just on shell.

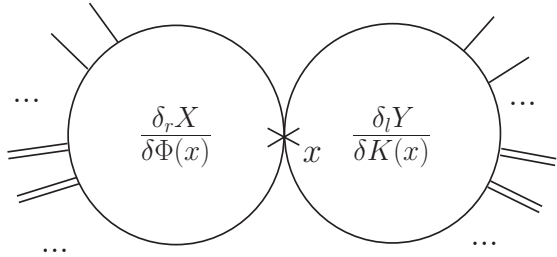
In most physical applications the symmetry algebras that close off shell play a major role. In this book we mainly focus on those. Nevertheless, it is important to know that more general structures exist. Still, we see that the solution (5.19) does depend on the field variables we choose.

The canonical formalism does not apply only to local functionals, such as the action S , but also to the generating functionals Z , W , Γ and Ω , which are nonlocal. For this reason, it is necessary to prove some general properties before proceeding. We have remarked above that the antiparentheses map local functionals X and Y into a local functional (X, Y) . We now prove that

they also map one-particle irreducible functionals X and Y into a one-particle irreducible functional (X, Y) . Define the operator

$$\mathcal{V} \equiv \int d^D x \frac{\overleftarrow{\delta}_r}{\delta \Phi^i(x)} \frac{\overrightarrow{\delta}_l}{\delta K_i(x)}.$$

We focus on the contribution $X\mathcal{V}Y$ to (X, Y) in (5.13), since the other contribution can be treated in an analogous way. Note that if X and Y are one-particle irreducible, a functional derivative with respect to $\Phi^\alpha(x)$ is an amputated Φ^α leg and a functional derivative with respect to $K_\alpha(x)$ is an insertion of $R^\alpha(\Phi(x))$. In particular, no propagators are attached to such legs. The operator \mathcal{V} produces a sort of new vertex, whose legs are the legs attached to $\Phi(x)$ in $\delta_r X / \delta \Phi(x)$ plus the legs attached to $K(x)$ in $\delta_l Y / \delta K(x)$. Since the diagrams of $\delta_r X / \delta \Phi(x)$ and $\delta_l Y / \delta K(x)$ are irreducible, the contribution $X\mathcal{V}Y$ to (X, Y) is also irreducible. Diagrammatically, we have


(5.28)

where the double lines are sources and the single lines are fields.

The solution (5.19) to the master equation is called *minimal*, because it contains the minimal set of fields. The minimal solution is not sufficient to gauge-fix the theory and define the propagators of the gauge fields, because it does not contain the antighosts and the Lagrange multipliers. We can include them by enlarging the sets of fields and sources to

$$\Phi^\alpha = (A_\mu^a, C^a, \bar{C}^a, B^a, \bar{\psi}, \psi, \varphi), \quad K_\alpha = (K_a^\mu, K_C^a, K_{\bar{C}}^a, K_B^a, K_{\bar{\psi}}, K_\psi, K_\varphi). \quad (5.29)$$

Again, the statistics of the sources are defined to be opposite of those of their conjugate fields.

It is easy to prove that if $S_{\min}(\Phi, K)$ is a minimal solution to the master equation the extended action

$$S_{\min}(\Phi, K) - \int B^a K_C^a$$

is also a solution. We call it *extended solution* to the master equation. This extension is sufficient for the purposes of gauge fixing. From now we understand that the sets of fields and sources are (5.29) and the general solution to the master equation is

$$S(\Phi, K) = S_c(\phi) - \int [R^i(\phi, C)K_i + R_C^a(\Phi)K_C^a + B^a K_C^a].$$

It is also useful to introduce the *ghost number*,

$$\text{gh}(A) = \text{gh}(\psi) = \text{gh}(\bar{\psi}) = \text{gh}(\varphi) = \text{gh}(B) = 0, \quad \text{gh}(C) = 1, \quad \text{gh}(\bar{C}) = -1.$$

Indeed, the global $U(1)$ transformation

$$\Phi \rightarrow \Phi e^{i\sigma \text{gh}(\Phi)}, \quad (5.30)$$

σ being a constant parameter, is a symmetry of the actions we are going to work with, as well as the functional integration measure. The ghost number is trivially preserved by the Feynman rules and the diagrammatics, so also by the radiative corrections and renormalization.

The fermionic number of a field or a source is equal to zero or one, depending on whether the field or source is a boson or a fermion. The statistics of a field or a source is equal to the sum of its fermionic number plus its ghost number, modulo 2. For example, the sources K_ψ and $K_{\bar{\psi}}$ associated with the Dirac fermions are commuting objects, since they are fermions, but they also have odd ghost numbers. Thus, K_ψ and $K_{\bar{\psi}}$ are “fermions with bosonic statistics”, while C and \bar{C} are “bosons with fermionic statistics”.

Now we are ready to derive the solutions of the master equation for Abelian and non-Abelian gauge theories. In quantum electrodynamics formulas (4.7) give the functions $R_c^i(\phi, \Lambda)$. Replacing Λ by θC and using $R_c^i(\phi, \theta C) = \theta R^i(\phi, C)$, we obtain the functions $R^i(\phi, C)$ for A_μ , ψ and $\bar{\psi}$, which read

$$\partial_\mu C, \quad -ieC\psi, \quad -ie\bar{\psi}C,$$

respectively. The functions R_C^a associated with the ghosts can be derived from the closure of the algebra. Since it is trivial in the Abelian case, we just have $R_C^a = 0$. Thus, the extended solution of the master equation reads

$$S(\Phi, K) = S_c(\phi) - \int d^D x \left(\partial_\mu C K_\mu - ie\bar{\psi} C K_{\bar{\psi}} - ieK_\psi C \psi + B K_{\bar{C}} \right), \quad (5.31)$$

where the classical action is

$$S_c(\phi) = \int d^D x \left[\frac{1}{4} F_{\mu\nu}^2 + \bar{\psi}(\not{\partial} + ie\not{A} + m)\psi \right].$$

It is easy to check that (5.17) is indeed satisfied.

In non-Abelian Yang-Mills theory we start from (4.50) to read the functions $R_c^i(\phi, \Lambda)$, replace Λ by θC and use $R_c^i(\phi, \theta C) = \theta R^i(\phi, C)$ again. We find that

$$\partial_\mu C^a + g f^{abc} A_\mu^b C^c, \quad -g T_{ij}^a C^a \psi^j, \quad -g \bar{\psi}^j T_{ij}^a C^a, \quad (5.32)$$

are the functions $R^i(\phi, C)$ for the gauge potential A_μ^a , the fermions ψ^i in the fundamental representation and their conjugates, respectively.

The solution reads

$$S(\Phi, K) = S_c(\phi) + g \int \left(\bar{\psi}^i T_{ij}^a C^a K_\psi^j + K_\psi^i T_{ij}^a C^a \psi^j \right) - \int \left[(\partial_\mu C^a + g f^{abc} A_\mu^b C^c) K_\mu^a - \frac{g}{2} f^{abc} C^b C^c K_C^a + B^a K_C^a \right], \quad (5.33)$$

where the classical action is

$$S_c(\phi) = \frac{1}{4} \int d^D x F_{\mu\nu}^2 + \int d^D x \bar{\psi}^i (\delta_{ij} \not{\partial} + g T_{ij}^a \not{A} + m \delta_{ij}) \psi^j. \quad (5.34)$$

In the next chapters we prove the renormalizability of both theories.

From (5.33) and (5.34) we can read the dimensions [] of the fields and the sources, as well as their statistics $\varepsilon_{\Phi A}$, $\varepsilon_{K A}$. The ghost numbers $\text{gh}(K)$ of the sources are obtained by demanding that (5.33) be invariant under (5.30) combined with $K \rightarrow K e^{i\sigma \text{gh}(K)}$. We have the tables

	A_μ^a	C^a	\bar{C}^a	B^a	$\bar{\psi}$	ψ	φ
[]	$\frac{D}{2} - 1$	$\frac{D}{2} - 1$	$\frac{D}{2} - 1$	$\frac{D}{2}$	$\frac{D-1}{2}$	$\frac{D-1}{2}$	$\frac{D}{2} - 1$
gh	0	1	-1	0	0	0	0

(5.35)

	K_a^μ	K_C^a	\bar{K}_C^a	K_B^a	$K_{\bar{\psi}}$	K_ψ	K_φ
[]	$\frac{D}{2}$	$\frac{D}{2}$	$\frac{D}{2}$	$\frac{D}{2} - 1$	$\frac{D-1}{2}$	$\frac{D-1}{2}$	$\frac{D}{2}$
gh	-1	-2	0	-1	-1	-1	-1

(5.36)

Now we stress a property that will be useful later, in the proof of the renormalizability of non-Abelian Yang-Mills theories. Theorem (5) allows us

to work without the matrices T^a or \mathcal{T}^a , just using the invariant tensors (4.39) and the gauge-field variables (4.48). Distinguishing upper and lower indices, we have $A_{\mu j}^i = -iT^{ai}{}_j A_\mu^a$. Using formula (4.29) the converse formula reads $A_\mu^a = -2iT^{aj}{}_i A_{\mu j}^i$. Similarly, the ghosts, antighosts and Lagrange multipliers can be written as $C_j^i = -iT^{ai}{}_j C^a$, $\bar{C}_j^i = -iT^{ai}{}_j \bar{C}^a$, and $B_j^i = -iT^{ai}{}_j B^a$.

When the fermions are in the fundamental representation the solution of the master equation is written as

$$\begin{aligned} S(\Phi, K) = S_c(\phi) + ig \int d^D x \left(\bar{\psi}_i C_j^i K_{\bar{\psi}}^j + K_{\psi i} C_j^i \psi^j \right) \\ - \int d^D x \left[\left(\partial_\mu C_j^i + ig(A_{\mu k}^i C_j^k - A_{\mu j}^k C_k^i) \right) K_{\mu i}^j \right. \\ \left. - 2ig C_k^i C_j^k K_{C i}^j + B_j^i K_{\bar{C} i}^j \right], \end{aligned} \quad (5.37)$$

where

$$S_c(\phi) = \int d^D x \left[\frac{1}{4} (F_{\mu\nu j}^i)^2 + \bar{\psi}_i (\delta_j^i \not{\partial} + g A_\mu^a T^{ai}{}_j + m \delta_j^i) \psi^j \right], \quad (5.38)$$

and

$$F_{\mu\nu j}^i = \partial_\mu A_{\nu j}^i - \partial_\nu A_{\mu j}^i + ig(A_{\mu k}^i A_{\nu j}^k - A_{\nu k}^i A_{\mu j}^k).$$

To derive these expressions we have also used formula (4.43).

A matter field ψ^I in an irreducible representation r can be denoted by $\psi_{j_1 \dots j_m}^{i_1 \dots i_n}$, if the indices have appropriate symmetry properties. From (4.38) we derive that its contribution to the solution $S(\Phi, K)$ proportional to the sources K becomes

$$\begin{aligned} ig \int d^D x \left(C_{l_1}^{i_1} \psi_{j_1 \dots j_m}^{l_1 i_2 \dots i_n} + \dots + C_{l_n}^{i_n} \psi_{j_1 \dots j_m}^{i_1 \dots i_{n-1} l_n} \right) K_{i_1 \dots i_n}^{j_1 \dots j_m} \\ - ig \int d^D x \left(C_{j_1}^{k_1} \psi_{k_1 j_2 \dots j_m}^{i_1 \dots i_n} + \dots + C_{j_m}^{k_m} \psi_{j_1 \dots j_{m-1} k_m}^{i_1 \dots i_n} \right) K_{i_1 \dots i_n}^{j_1 \dots j_m}, \end{aligned} \quad (5.39)$$

Finally, a generic vertex has the form

$$\phi_{j_1 \dots j_m}^{i_1 \dots i_n} \psi_{l_1 \dots l_q}^{k_1 \dots k_p} \chi_{v_1 \dots v_s}^{u_t \dots u_r} \dots \quad (5.40)$$

with indices contracted by means of the invariant tensors (4.39).

An important theorem states that Yang-Mills theory in practice exhausts the gauge theories of vector fields.

Theorem 7 *The most general local, power counting renormalizable quantum field theory of vector fields is a Yang-Mills theory based on a Lie algebra.*

Proof. To prove this theorem we can take advantage of the canonical formalism, because we know that it collects the properties of the Lie algebra in a compact form. Let A_μ^I denote the set of gauge vectors contained in the theory. In the free-field limit, the theory must obey the Abelian gauge symmetry $\delta A_\mu^I = \partial_\mu \Lambda^I$. Writing $\Lambda^I = \theta C^I$, as usual, the ghosts C^I can always be defined so that the derivative term in $R_\mu^I(\phi, C)$ is $\partial_\mu C^I$. By locality, ghost number conservation and power counting, the most general K sector of the minimal solution to the master equation must have the form

$$-(\partial_\mu C^I + A_\mu^J C^K \kappa^{JKI}) K^I + \frac{1}{2} C^J C^K h^{JKI} K_C^I,$$

where κ^{IJK} and h^{JKI} are numerical constants and h^{JKI} are antisymmetric in J and K . Now we study the constraints imposed by the master equation $(S, S) = 0$. It is easy to show that the terms proportional to K_C^I in $(S, S) = 0$ imply that the constants h^{JKI} satisfy the Jacobi identity (4.28)

$$h^{IJK} h^{KLM} + h^{LIK} h^{KJM} + h^{JLK} h^{KIM} = 0.$$

Since both assumptions (4.27) and (4.28) are satisfied, the constants h^{JKI} define a Lie algebra. It is also straightforward to check that the terms proportional to K^I in $(S, S) = 0$ give $\kappa^{IJK} = h^{IJK}$. Thus, the gauge transformations have the Yang-Mills form

$$\delta_\Lambda A_\mu^I = \partial_\mu \Lambda^I + A_\mu^J \Lambda^K h^{JKI},$$

which proves the theorem. \square

We stress again that we have not proved the renormalizability, yet, but this theorem anticipates that if Yang-Mills theory is renormalizable, it is unique.

5.3 Canonical transformations

A canonical transformation \mathcal{C} of the fields and the sources is a transformation

$$\Phi^{\alpha'}(\Phi, K), \quad K'_\alpha(\Phi, K),$$

that preserves the antiparentheses, that is to say such that

$$(X', Y')' = (X, Y)$$

for every pairs of functionals X and Y , where X' and Y' are defined as

$$\begin{aligned} X'(\Phi', K') &= X(\Phi(\Phi', K'), K(\Phi', K')), \\ Y'(\Phi', K') &= Y(\Phi(\Phi', K'), K(\Phi', K')), \end{aligned}$$

and $(.,.)'$ denotes the antiparentheses calculated with respect to Φ' and K' . Sometimes we simply write

$$X' = \mathcal{C}X, \quad Y' = \mathcal{C}Y.$$

By this we mean that X and Y are regarded as functions of the new variables Φ', K' , obtained by expressing their arguments Φ, K in terms of Φ', K' .

A canonical transformation is generated by a functional $\mathcal{F}(\Phi, K')$ and can be expressed as

$$\Phi^{\alpha'} = \frac{\delta \mathcal{F}}{\delta K'_\alpha}, \quad K_\alpha = \frac{\delta \mathcal{F}}{\delta \Phi^\alpha}. \quad (5.41)$$

Formula (5.12) implies that \mathcal{F} is a functional of fermionic statistics. For this reason, there is no need to specify whether the derivatives of (5.41) are left or right.

The generating functional of the identity transformation is

$$\mathcal{I}(\Phi, K') = \int d^D x \Phi^\alpha(x) K'_\alpha(x).$$

Observe that if X is such that $(S, X) = 0$, then X' is such that $(S', X')' = 0$.

Let us inspect the most general canonical transformation, to understand what it can be useful for. We can write the generating functional as a sum of a term independent of the sources plus the rest:

$$\mathcal{F}(\Phi, K') = \Psi(\Phi) + \int K'_\alpha U^\alpha(\Phi, K'). \quad (5.42)$$

Then (5.41) gives

$$\Phi^{\alpha'} = U^\alpha(\Phi, K') + \int K'_\beta \frac{\delta_r U^\beta}{\delta K'_\alpha}, \quad K_\alpha = \frac{\delta \Psi(\Phi)}{\delta \Phi^\alpha} + \int K'_\beta \frac{\delta_r U^\beta}{\delta \Phi^\alpha}. \quad (5.43)$$

Recall that at the end we must set the sources K to zero, since they are introduced just to have control on the gauge symmetry. To illustrate the meaning of (5.42), we set them to zero after the canonical transformation. If we drop the terms proportional to K' in (5.43) we obtain

$$\Phi^{\alpha'} = U^\alpha(\Phi, 0), \quad K_\alpha = \frac{\delta\Psi(\Phi)}{\delta\Phi^\alpha}. \quad (5.44)$$

The Φ transformation is a field redefinition. Instead, as we will explain later, the K transformation is a gauge fixing, or a change of gauge fixing. The K -dependent terms contained in (5.43) do not have a particular meaning. They are there to promote (5.44) to a canonical transformation, which is much easier to manipulate, because it preserves the antiparentheses. Thus,

Proposition 8 *the most general canonical transformation (5.42) is the combination of the most general field redefinition and the most general gauge-fixing.*

Of course, the physics should not depend on the field variables we use.

5.4 Gauge fixing

We have gauge-fixed quantum electrodynamics in chapter 4. Now we are ready to gauge-fix non-Abelian Yang-Mills theory. Call $\mathcal{G}^a(A)$ the gauge-fixing functions, i.e. $\mathcal{G}^a(A) = \partial_\mu A_\mu^a$ in the Lorenz gauge, and $\mathcal{G}^a(A) = -\nabla \cdot \mathbf{A}^a$ in the Coulomb gauge. Define the *gauge fermion*

$$\Psi(\Phi) = \int d^D x \bar{C}^a \left(-\frac{\lambda}{2} B^a + \mathcal{G}^a(A) \right). \quad (5.45)$$

The gauge fermion is a local functional of fermionic statistics that fixes the gauge in the way explained below. Its typical form is (5.45), but more general functionals can be chosen.

Working in the Lorenz gauge, for definiteness, define the gauge-fixed action

$$S_\Psi(\Phi, K) = S(\Phi, K) + (S, \Psi). \quad (5.46)$$

It is easy to prove that S_Ψ and S are related by the canonical transformation generated by the functional

$$\mathcal{F}(\Phi, K') = \int \Phi^\alpha K'_\alpha + \Psi(\Phi). \quad (5.47)$$

Indeed, (5.41) gives

$$\Phi^{\alpha'} = \Phi^\alpha, \quad K'_\alpha = K_\alpha - \frac{\delta\Psi(\Phi)}{\delta\Phi^\alpha}.$$

Recalling that the action is linear in the sources, we have from (5.19)

$$S(\Phi', K') = S(\Phi, K) + \int R^\alpha(\Phi) \frac{\delta\Psi(\Phi)}{\delta\Phi^\alpha} = S(\Phi, K) + (S, \Psi) = S_\Psi(\Phi, K).$$

Moreover,

Proposition 9 *If S satisfies the master equation, then every $S_\Psi = S + (S, \Psi(\Phi))$ satisfies the master equation.*

The reason is that the canonical transformations preserve the antiparentheses, so $(S, S) = 0$ if and only if $(S_\Psi, S_\Psi) = 0$. In particular, the gauge-fixing procedure preserves the master equation.

Working out S_Ψ explicitly, we find

$$S_\Psi(\Phi, K) = S_c(\phi) + S_{\text{gf}}(\Phi) - \int R^\alpha K_\alpha, \quad (5.48)$$

where

$$S_c(\phi) + S_{\text{gf}}(\Phi) = \int d^D x \left(\frac{1}{4} F_{\mu\nu}^a{}^2 - \frac{\lambda}{2} (B^a)^2 + B^a \partial \cdot A^a - \bar{C}^a \partial_\mu D_\mu C^a \right). \quad (5.49)$$

Observe that the ghosts do not decouple in non-Abelian Yang-Mills theory.

The gauge-field propagator can be worked out from the free subsector of (5.49), after integrating B^a out, which gives the equivalent gauge-fixed action

$$\mathcal{S} = \int d^D x \left(\frac{1}{4} F_{\mu\nu}^a{}^2 + \frac{1}{2\lambda} (\partial_\mu A_\mu^a)^2 - \bar{C}^a \partial_\mu D_\mu C^a \right). \quad (5.50)$$

The result is

$$\langle A_\mu^a(k) A_\nu^b(-k) \rangle = \frac{\delta^{ab}}{k^2} \left(\delta_{\mu\nu} + (\lambda - 1) \frac{k_\mu k_\nu}{k^2} \right), \quad (5.51)$$

The ghost propagator is

$$\langle C^a(k) \bar{C}^b(-k) \rangle = \frac{\delta^{ab}}{k^2}. \quad (5.52)$$

Repeating the argument that leads to (4.21) we can check, in the Coulomb gauge, that the physical degrees of freedom are $2 \dim G$, as it must be.

The argument just given does not change when we add matter fields, since they are not interested by the gauge-fixing procedure. Clearly, in QED we get back (4.15) and (4.16).

Exercise 19 *Show that the action $S_\Psi(\Phi, K)$ continues to satisfy the master equation after integrating the Lagrange multipliers B^a out.*

Solution. Integrating B out is equivalent to replace B with the solution of its own field equation, that is to say make the replacement

$$B^a \rightarrow \frac{1}{\lambda}(\partial \cdot A^a - K_C^a).$$

Then $S_\Psi(\Phi, K)$ becomes

$$\begin{aligned} \bar{S}_\Psi(\Phi, K) = & \frac{1}{4} \int F_{\mu\nu}^a{}^2 - \int \bar{C}^a \partial_\mu D_\mu C^a + \frac{1}{2\lambda} \int (K_C^a - \partial \cdot A^a)^2 \\ & - \int (D_\mu C^a) K_\mu^a + \frac{g}{2} \int f^{abc} C^b C^c K_C^a. \end{aligned}$$

At this point, it is straightforward to check that the master equation $(\bar{S}_\Psi, \bar{S}_\Psi) = 0$ holds. Note that $(\bar{S}_\Psi, \bar{C}) = (\partial \cdot A^a - K_C^a)/\lambda$.

Observe that the action $\bar{S}_\Psi(\Phi, K)$ is no longer linear in the sources K , but contains a term that is quadratic in K_C . This means that after integrating B out the Φ transformations do not close off shell anymore, which the reader can verify directly. Working with the canonical formalism this problem is cured by itself, since the master equation is satisfied both before and after the integration over B .

Exercise 20 *Derive the Feynman rules of (5.50) coupled to fermions.*

Solution. The propagators have been given above. The vertices are

$$\begin{aligned}
 & \begin{array}{c} a \text{---} q \text{---} c \\ | \\ \text{wavy line} \\ \mu b \end{array} = igf^{abc}q^\mu, & \begin{array}{c} i \text{---} \text{wavy line} \text{---} j \\ | \\ \mu a \end{array} = -g\gamma^\mu(T^a)_{ij}, \\
 & \begin{array}{c} \mu a \text{---} k_1 \text{---} \nu b \\ | \quad | \\ k_3 \text{---} \rho c \end{array} = igf^{abc}(\delta_{\mu\nu}(k_{1\rho} - k_{2\rho}) + \delta_{\mu\rho}(k_{3\nu} - k_{1\nu}) + \delta_{\nu\rho}(k_{2\mu} - k_{3\mu})), \\
 & \begin{array}{c} \mu a \text{---} \text{wavy line} \text{---} \nu b \\ | \quad | \\ \rho c \text{---} \text{wavy line} \text{---} \sigma d \end{array} = -g^2 [f^{eab}f^{ecd}(\delta_{\mu\rho}\delta_{\nu\sigma} - \delta_{\mu\sigma}\delta_{\nu\rho}) \\
 & \quad + f^{ead}f^{ebc}(\delta_{\mu\nu}\delta_{\rho\sigma} - \delta_{\mu\rho}\delta_{\nu\sigma}) \\
 & \quad + f^{eac}f^{ebd}(\delta_{\mu\nu}\delta_{\rho\sigma} - \delta_{\mu\sigma}\delta_{\nu\rho})].
 \end{aligned} \tag{5.53}$$

Exercise 21 Prove that the term $\int R^\alpha K_\alpha$ can be written as (S, χ) for a local functional χ .

Solution. Consider the canonical transformation generated by

$$\mathcal{F}(\Phi, K') = \int \Phi^\alpha K'_\alpha + (e^\zeta - 1) \int C^a K_C^a + (e^{-\zeta} - 1) \int \bar{C}^a K_{\bar{C}}^a. \tag{5.54}$$

Let $S_\zeta(\Phi, K)$ denote the rescaled action. Expanding in ζ we obtain

$$S_\zeta(\Phi, K) = S(\Phi, K) + \zeta \left(S, \int (C^a K_C^a - \bar{C}^a K_{\bar{C}}^a) \right) + \mathcal{O}(\zeta^2). \tag{5.55}$$

Now, the transformation rescales the ghosts by a factor e^ζ , the antighosts by the reciprocal factor $e^{-\zeta}$, and their sources K_C and $K_{\bar{C}}$ by $e^{-\zeta}$ and e^ζ , respectively. Applied to (5.31) and (5.33), even after including the gauge fixing (5.48), it is equivalent to rescale all the sources K by e^ζ , which gives $S_\zeta(\Phi, K) = S(\Phi, e^\zeta K)$. Differentiating this equation and (5.55) with respect to ζ and setting $\zeta = 0$ we get

$$\int R^\alpha K_\alpha = \left(S, \int (\bar{C}^a K_{\bar{C}}^a - C^a K_C^a) \right).$$

The reader is invited to check this formula explicitly in both QED and Yang-Mills theory. This result teaches us that $\int R^\alpha K_\alpha$ is exact in the cohomology defined by the application $X \rightarrow (S, X)$, acting on the local functionals X . \square

From now on, we drop the subscript Ψ in $S_\Psi(\Phi, K)$ and when we write $S(\Phi, K)$ we mean the gauge-fixed action (5.48).

5.5 Generating functionals

Define the generating functionals as

$$Z(J, K) = \int [d\Phi] \exp \left(-S(\Phi, K) + \int \Phi^\alpha J_\alpha \right) = \exp(W(J, K)), \quad (5.56)$$

and $\Gamma(\Phi, K)$ as the Legendre transform of $W(J, K)$ with respect to J , the sources K remaining inert:

$$J_\alpha = \frac{\delta_l \Gamma(\Phi, K)}{\delta \Phi^\alpha}, \quad \Gamma(\Phi, K) = -W(J, K) + \int \Phi^\alpha J_\alpha. \quad (5.57)$$

Observe that Γ is the generating functional of one-particle irreducible diagrams, including the diagrams that have ghosts, Lagrange multipliers and sources K on their external legs. We are tacitly assuming that the integral (5.56) makes sense, at least perturbatively. This means that the action $S(\Phi, K)$ should be gauge-fixed, so that the propagators are well defined. In the next subsection we show how the action can be gauge-fixed preserving the master equation. For the moment we study the canonical formalism for the traditional functionals Z , W and Γ . Later we introduce the master functional Ω for gauge theories and discuss the composite fields and the changes of field variables in detail.

It is apparent from (5.43) that the canonical transformations cannot be implemented as changes of field variables inside the functional integral. Indeed, in general they mix the fields Φ , over which we integrate, with the external sources K . While it is legitimate to make a change of field variables $\Phi \rightarrow \Phi'(\Phi, K)$ in the functional integral, it is not legitimate to redefine the external sources as functions of the integrated fields. Thus, when we use canonical transformations we means that we apply them to the action $S(\Phi, K)$, while the generating functionals Z , W and Γ are just replaced with the ones associated with the transformed action. We will not be able to describe our operations, including the renormalization, as true changes of field variables until we introduce the master functional for gauge theories.

To some extent, we can study the composite fields already at the level of Z , W and Γ . We can treat the correlation functions

$$\langle \mathcal{O}^{I_1}(x_1) \cdots \mathcal{O}^{I_n}(x_n) \rangle \quad (5.58)$$

of gauge-invariant composite fields $\mathcal{O}^I(\phi)$ by adding them to the action, multiplied by suitable sources L . Thus, we consider the generating functionals

$$Z(J, K, L) = \int [d\Phi] \exp \left(-S(\Phi, K) + \int L_I \mathcal{O}^I(\phi) + \int \Phi^\alpha J_\alpha \right) = e^{W(J, K, L)}, \quad (5.59)$$

and the Lagrange transform $\Gamma(\Phi, K, L)$ of $W(J, K, L)$ with respect to Φ . We have

$$J_\alpha = \frac{\delta_l \Gamma}{\delta \Phi^\alpha}, \quad \frac{\delta_r W}{\delta K_\alpha} = -\frac{\delta_r \Gamma}{\delta K_\alpha}, \quad \frac{\delta_l W}{\delta L_I} = -\frac{\delta_l \Gamma}{\delta L_I}. \quad (5.60)$$

Recall that in this chapter we are working at the bare level. Indeed, formula (5.59) is the correct bare form for the generating functionals, while the renormalized structure is considerably more involved. Among other things, the exponent of the integrand becomes nonpolynomial in the sources L and K , when higher-dimensional composite fields are present. For the moment, the correlation functions (5.58) that are gauge invariant and gauge independent (see below), but still divergent.

Consider the change of field variables

$$\Phi^{\alpha'} = \Phi^\alpha + \theta R^\alpha = \Phi^\alpha + \theta(S, \Phi^\alpha), \quad (5.61)$$

in the functional integral (5.56), where θ is a constant anticommuting parameter. In a sense that we now explain, (5.61) is equivalent to a canonical transformation generated by

$$\mathcal{F}(\Phi, K') = \int (\Phi^\alpha K'_\alpha + \theta R^\alpha K'_\alpha).$$

Indeed, formulas (5.41) give

$$\Phi^{\alpha'} = \Phi^\alpha + \theta R^\alpha, \quad K'_\alpha = K_\alpha - \int \frac{\delta_l R^\beta}{\delta \Phi^\alpha} K_\beta \theta. \quad (5.62)$$

We have inverted the second relation using $\theta^2 = 0$, which ensures that the Taylor expansions in θ stops after the first order in θ . The K transformation

appearing here does not affect the action, because S depends on K only via the combination $-\int R^\alpha(\Phi)K_\alpha$, which gets an extra contribution equal to

$$\int R^\alpha(\Phi) \frac{\delta_l R^\beta}{\delta \Phi^\alpha} K_\beta \theta = \int (S, R^\alpha) K_\alpha \theta = \int (S, (S, \Phi^\alpha)) K_\alpha \theta = 0.$$

Thus, (5.62) is equivalent to just (5.61).

We know that, using the dimensional regularization the functional integration measure is invariant under the local change of field variables (5.61), by theorem 1. There actually exists a stronger argument to prove the same result, which can be applied to a more general class of regularization techniques. Thanks to (1.98) we have

$$\mathcal{J} = \text{sdet} \frac{\delta \Phi^{\alpha'}(x)}{\delta \Phi^\beta(y)} = \text{sdet} \left(\delta^{\alpha\beta} \delta(x-y) + \frac{\delta[\theta R^\alpha(x)]}{\delta \Phi^\beta(y)} \right) = 1 + \text{str} \frac{\delta[\theta R^\alpha(x)]}{\delta \Phi^\beta(y)}. \quad (5.63)$$

We have again used $\theta^2 = 0$. In QED the matrix

$$\frac{\delta[\theta R^\alpha(x)]}{\delta \Phi^\beta(y)} = \frac{\delta[\theta(\partial_\mu C, 0, B, 0, -ie\bar{\psi}C, -ieC\psi)]}{\delta(A_\nu, C, \bar{C}, B, \bar{\psi}, \psi)}$$

has no diagonal elements except for the block

$$\frac{\delta(ie\bar{\psi}\theta C, -ie\theta C\psi)}{\delta(\bar{\psi}, \psi)} = ie \begin{pmatrix} \theta C & 0 \\ 0 & -\theta C \end{pmatrix}, \quad (5.64)$$

but the trace vanishes. Clearly, this is due to the fact that $\bar{\psi}$ and ψ have opposite charges. Using (1.97) we see that the supertrace of (5.63) vanishes, so $\mathcal{J} = 1$.

In non-Abelian gauge theories formulas (5.32) and (5.33) give

$$\begin{aligned} \frac{\delta(\theta R_\mu^a)}{\delta A_\nu^b} &= g\delta_{\mu\nu} f^{abc} \theta C^c, & \frac{\delta(\theta R_C^a)}{\delta C^b} &= g f^{abc} \theta C^c, \\ \frac{\delta_l(\theta R_\psi^i)}{\delta \psi^j} &= -g T_{ij}^a \theta C^a, & \frac{\delta_l(\theta R_{\bar{\psi}}^i)}{\delta \bar{\psi}^j} &= g T_{ij}^a \theta C^a. \end{aligned}$$

The scalar contribution is similar to the fermion one. When the representation is not the fundamental one it is sufficient to replace T^a by the appropriate matrices \mathcal{T}^a . The A and C contributions to (5.63) are zero, because f^{abc} is completely antisymmetric. If the gauge group has no Abelian factors,

then $\text{tr}[\mathcal{T}^a] = 0$, so the traces of the ψ and φ contributions are also zero. If the gauge group has Abelian factors, the traces $\text{tr}[\mathcal{T}^a]$ are given by the $U(1)$ charges. They cancel out summing the contributions of both ψ and $\bar{\psi}$, or both φ and $\bar{\varphi}$, as in (5.64). Finally, the contributions of \bar{C} and B are obviously zero.

Now we prove that

Theorem 10 *If the action S satisfies the master equation, the generating functionals Z and W are invariant under the transformation*

$$\tau K_\alpha = (-1)^{\varepsilon_\alpha+1} J_\alpha, \quad \tau J_\alpha = 0, \quad \tau L_I = 0.$$

Observe that the operator τ flips the statistics. Because of this, it produces a minus sign every time it crosses an object with fermionic statistics. This property can be proved by observing that $\delta_\tau = \theta\tau$ obeys the ordinary Leibniz rule, where θ is an anticommuting constant.

Proof. Apply the operator δ_τ to the Z functional (5.59). Using (5.19), we see that the exponent of the integrand is changed into itself plus

$$\theta \int R^\alpha(\Phi) J_\alpha. \quad (5.65)$$

Thus, we obtain the formula

$$\tau W = \left\langle \int R^\alpha(\Phi) J_\alpha \right\rangle. \quad (5.66)$$

We can prove that this average vanishes by performing the change of field variables (5.61) in (5.59). Indeed, recall that the functional measure is invariant, the action S satisfies the master equation and the composite fields \mathcal{O}^I are gauge invariant. Then, (5.61) affects only $\int \Phi^\alpha J_\alpha$, by an amount equal to

$$\theta \int (S, \Phi^\alpha) J_\alpha = \theta \int R^\alpha(\Phi) J_\alpha, \quad (5.67)$$

and W by an amount equal to the average of (5.67). Since a change of field variables cannot modify the result of the integral, we conclude that

$$\tau Z(J, K, L) = 0, \quad \tau W(J, K, L) = 0. \quad (5.68)$$

□

Using (5.60), we can write

$$\tau W = \int \tau K_\alpha \frac{\delta_l W}{\delta K_\alpha} = (-1)^{\varepsilon_\alpha} \int J_\alpha \frac{\delta_l \Gamma}{\delta K_\alpha} = \int \frac{\delta_r \Gamma}{\delta \Phi^\alpha} \frac{\delta_l \Gamma}{\delta K_\alpha}. \quad (5.69)$$

Using (5.15) and (5.68), we obtain

$$\tau W = \frac{1}{2}(\Gamma, \Gamma) = 0, \quad (5.70)$$

which is the master equation for Γ . Later we will show that it encodes the gauge invariance of physical correlation functions. We have thus proved that

Theorem 11 *If S satisfies the master equation, then Γ satisfies the master equation.*

When the action S is not assumed to satisfy the master equation, a more general result tells us that the violation of the Γ master equation $(\Gamma, \Gamma) = 0$ is given by the average of (S, S) . This gives a formula that, due to its importance, we call *master identity*. It will be crucial in the proofs of renormalizability and in the study of anomalies and gauge independence to all orders.

Theorem 12 *The generating functional Γ satisfies the master identity*

$$(\Gamma, \Gamma) = \langle (S, S) \rangle.$$

Proof. It can be proved by going through the argument that lead to (5.70), and making the necessary modifications. Formula (5.66) is unaffected. Instead, the change of variables (5.61) does not only affect $\int \Phi J$, by an amount equal to (5.67), but also $-S$, by an amount equal to $-\theta(S, S)/2$. Since W cannot change under a change of variables, we obtain

$$\tau W = \left\langle \int R^\alpha(\Phi) J_\alpha \right\rangle = \frac{1}{2} \langle (S, S) \rangle$$

Formula (5.69) is also unmodified, so in the end

$$\frac{1}{2} \langle (S, S) \rangle = \tau W = \frac{1}{2}(\Gamma, \Gamma).$$

5.6 Ward identities

Consider the change of variables (5.61) in the functional integral

$$\int [d\Phi] \mathcal{Q}(\Phi) \exp \left(-S(\Phi, K) + \int L_I \mathcal{O}^I(\phi) \right), \quad (5.71)$$

where now \mathcal{Q} denotes a completely arbitrary function of the fields. It can include any string of insertions of elementary and composite fields, including ghosts and Lagrange multipliers, as well as functionals, and does not need to be local. However, for the derivation that we give below \mathcal{Q} cannot depend on the sources K . The reason is that the functional integral is only over Φ , so the change of variables cannot transform K . Note that in (5.71) we have set the sources J for the elementary fields Φ to zero. The reason is that most sources J are not gauge invariant. By means of (5.71), we can study the correlation functions (5.58).

If S satisfies the master equation, then only $\mathcal{Q}(\Phi)$ is affected by (5.61), and we easily obtain

$$\left\langle \int R^\alpha \frac{\delta \mathcal{Q}}{\delta \Phi^\alpha} \right\rangle = \langle (S, \mathcal{Q}) \rangle_0 = 0, \quad (5.72)$$

where the subscript 0 reminds us that the sources J for the elementary fields are set to zero.

This identity is called *Ward identity*. Its meaning is that an object of the form (S, \mathcal{Q}) is zero for every physical purposes, that is to say a completely unobservable quantity. Observe that (S, \mathcal{Q}) is just a functional of the fields. Replacing \mathcal{Q} with $\mathcal{Q}(S, \mathcal{Q})^{n-1}$ in (5.72), it follows that

$$\langle (S, \mathcal{Q})^n \rangle_0 = 0$$

for every n . Then, if we specialize \mathcal{Q} to be a local functional Ψ of fermionic statistics, we also have the identity

$$\int [d\Phi] e^{-S_\Psi(\Phi, K) + \int L_I \mathcal{O}^I(\phi)} = \int [d\Phi] e^{-S(\Phi, K) + \int L_I \mathcal{O}^I(\phi)}, \quad (5.73)$$

where S_Ψ and S are related by formula (5.46), or, which is the same, the canonical transformation (5.47). Identity (5.73) tells us that we are free to add an arbitrary functional of the form (S, Ψ) to the action, and no correlation function (5.58) will depend on it. We have already seen that this freedom allows us to gauge-fix the theory, by choosing a Ψ of the form (5.45).

This proves that

Theorem 13 *The correlation functions (5.58) are invariant under the canonical transformations of the form (5.47), for an arbitrary local $\Psi(\Phi)$.*

Since the most general canonical transformation is a combination of a canonical transformation of type (5.47) and a change of variables for the fields Φ , we conclude that

Theorem 14 *The physical quantities are invariant under the most general canonical transformation.*

Among the freedom we have, we can replace $\partial_\mu A_\mu$ in (5.45) by another gauge-fixing function $\mathcal{G}(A)$. From the arbitrariness of $\Psi(\Phi)$ and theorem 13, we conclude that

Theorem 15 *The correlation functions (5.58) are gauge-independent,*

that is to say they are independent of the gauge fixing. Even if we stick to the same $\mathcal{G}(A)$, they are independent of the gauge-fixing parameter λ that appears in (5.45).

Note that the notion of gauge independence does not coincide with the notion of gauge invariance. A gauge invariant quantity is a quantity that does not change when a gauge transformation is applied to it. A gauge independent quantity is a quantity that does not change by modifying the gauge-fixing function $\mathcal{G}(A)$ that is used to define the functional integral.

Gauge independence ensures that the value of the physical correlation functions, such as (5.58), is the same with any gauge choice. In particular, it coincides with the value we would find, for example, in the Coulomb gauge (4.8), where only the physical degrees of freedom propagate. For this reason, gauge independence is crucial to prove unitarity.

We will have more to say about the independence of the physical quantities on canonical transformations later on. Moreover, we still have to prove that the theory is renormalizable. So far, we have been working with physical quantities that may be gauge invariant and gauge independent (see below), but still divergent. We must show that the subtraction of divergences can be organized so as to preserve the properties proved above.

Chapter 6

Quantum electrodynamics

In this chapter we study quantum electrodynamics and prove its renormalizability to all orders. Since the action does not contain chiral fermions the properties we have derived in the previous chapter, such as the master equation (5.17), hold in arbitrary complex D dimensions. In particular, the Lagrangian

$$\mathcal{L}_0 = \frac{1}{4}F_{\mu\nu}^2 + \bar{\psi}(\not{\partial} + ie\not{A} + m)\psi \quad (6.1)$$

is gauge invariant in D dimensions and the dimensionally regularized gauge-fixed extended action

$$S(\Phi, K) = \int \mathcal{L}_{\text{tot}} + \int (K_\mu \partial_\mu C + ie\bar{\psi}CK_{\bar{\psi}} + ieK_\psi C\psi - BK_{\bar{C}}), \quad (6.2)$$

where

$$\mathcal{L}_{\text{tot}} = \mathcal{L}_0 - \frac{\lambda}{2}B^2 + B\partial \cdot A - \bar{C}\square C, \quad (6.3)$$

satisfies $(S, S) = 0$ identically.

After integrating B out, the Feynman rules are

$$\begin{aligned} \mu \text{---} \text{wiggled line} \text{---} \nu &= \frac{1}{k^2} (\delta_{\mu\nu} + (\lambda - 1) \frac{k_\mu k_\nu}{k^2}) \\ \text{---} p \text{---} &= \frac{1}{i\not{p} + m} \end{aligned} \quad \begin{aligned} &\text{Feynman diagram: a vertex with two incoming lines and one outgoing wiggled line labeled } \mu \\ &= -ie\gamma^\mu \end{aligned} \quad (6.4)$$

where the wiggled line denotes the photon. We do not need rules for the ghosts, since they decouple.

The first thing to note is that (6.1) does not contain all the terms that are allowed by power counting. The missing ones, such as

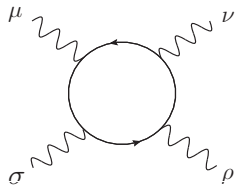
$$\frac{1}{2}m_\gamma^2 A_\mu^2, \quad \frac{1}{3!}A_\mu^2 \partial_\nu A_\nu, \quad \frac{1}{4!}(A_\mu^2)^2, \quad (6.5)$$

etc., are forbidden by gauge invariance. We know that such terms are absent at the tree level, because the tree-level Lagrangian is gauge invariant. In principle, renormalization might generate them at one loop or higher orders. More precisely, it might be necessary to introduce the vertices (6.5) as counterterms, to remove divergences proportional to them. However, if that happened, renormalization would ruin the gauge invariance of the theory. We need to prove that, instead, the divergent parts of the Feynman diagrams are gauge invariant, and can be removed by redefining the ingredients (fields, sources and parameters) of the tree-level action $S(\Phi, K)$. Fortunately, in most cases, which include QED, renormalization and gauge invariance are compatible with each other.

For the moment, we just assume that this compatibility holds and work out some consequences. The renormalizability of (6.3) is proven in section 6.2.

Exercise 22 *Using the dimensional regularization prove by explicit computation that the photon four-point function $\langle A_\mu(x)A_\nu(y)A_\rho(z)A_\sigma(w) \rangle$ is one-loop convergent.*

Solution. By power counting and locality, the divergent part is just a constant, so it can be calculated at vanishing external momenta. Although the divergent part is also independent of the mass, we keep m nonzero, because the limit when both the external momenta and the masses tend to zero cannot be taken inside the integral in dimensional regularization. We have the diagram


(6.6)

plus permutations of external legs, which means exchanges of ν , ρ and σ .

The integral corresponding to (6.6) is

$$-e^4 \int \frac{d^D p}{(2\pi)^D} \frac{\text{tr}[(-i\not{p} + m)\gamma_\mu(-i\not{p} + m)\gamma_\nu(-i\not{p} + m)\gamma_\rho(-i\not{p} + m)\gamma_\sigma]}{(p^2 + m^2)^4}.$$

The masses in the numerator can be dropped, since they contribute only to the finite part. We get

$$-e^4 \int \frac{d^D p}{(2\pi)^D} \frac{p_\alpha p_\beta p_\gamma p_\delta}{(p^2 + m^2)^4} \text{tr}[\gamma_\alpha \gamma_\mu \gamma_\beta \gamma_\nu \gamma_\gamma \gamma_\rho \gamma_\delta \gamma_\sigma]. \quad (6.7)$$

By Lorentz covariance, the integral can only be proportional to $\delta_{\alpha\beta}\delta_{\gamma\delta} + \delta_{\alpha\gamma}\delta_{\beta\delta} + \delta_{\alpha\delta}\delta_{\beta\gamma}$. The factor in front of this tensor can be calculated by contracting α with β and γ with δ . We can thus write

$$\int \frac{d^D p}{(2\pi)^D} \frac{p_\alpha p_\beta p_\gamma p_\delta}{(p^2 + m^2)^4} = \frac{\delta_{\alpha\beta}\delta_{\gamma\delta} + \delta_{\alpha\gamma}\delta_{\beta\delta} + \delta_{\alpha\delta}\delta_{\beta\gamma}}{D(D+2)} \int \frac{d^D p}{(2\pi)^D} \frac{(p^2)^2}{(p^2 + m^2)^4}.$$

Evaluating the integral with the help of formula (A.5) and using (2.14) and (2.15) to compute the trace, we can easily find that the divergent part of (6.7) is nontrivial, equal to

$$-\frac{8e^4}{3\varepsilon(4\pi)^2} (\delta_{\mu\nu}\delta_{\rho\sigma} - 2\delta_{\mu\rho}\delta_{\nu\sigma} + \delta_{\mu\sigma}\delta_{\nu\rho}),$$

where $\varepsilon = 4 - D$. However, the pole disappears by summing over the permutations of the external legs. Without this cancellation, there would be a divergent part proportional to $(A_\mu^2)^2$, which would violate gauge invariance. This exercise is an explicit check that the dimensional regularization is manifestly gauge invariant.

Exercise 23 *Show that the three-point function $\langle A_\mu(x)A_\nu(y)A_\rho(z) \rangle$ is also convergent at one loop.*

Solution. We leave the details to the reader. The fermion loop with three external photons has a nontrivial divergent part, which is linear in the external momenta. As before, the pole cancels when the permutations of the external legs are included. Note that exchanging two photon legs is equivalent to flip the arrow of the fermion loop.

There exist more powerful methods, based on the invariance under charge conjugation, to show that the n -photon correlation functions identically vanish when n is odd. However, it is not straightforward to use those arguments

together with the dimensional regularization. The reason is that the charge-conjugation matrix, like the matrix γ_5 , does not admit a simple extension to D dimensions, and the dimensionally regularized QED Lagrangian is not exactly invariant under charge conjugation. \square

As usual, we have bare and renormalized versions of \mathcal{L}_0 , which read

$$\begin{aligned}\mathcal{L}_{0B} &= \frac{1}{4}F_{\mu\nu B}^2 + \bar{\psi}_B(\not{\partial} + ie_B\not{A}_B + m_B)\psi_B = \\ \mathcal{L}_{0R} &= \frac{1}{4}Z_A F_{\mu\nu}^2 + Z_\psi \bar{\psi}(\not{\partial} + ie\mu^\varepsilon Z_e Z_A^{1/2}\not{A} + m Z_m)\psi,\end{aligned}\quad (6.8)$$

having defined

$$A_{\mu B} = Z_A^{1/2} A_\mu, \quad \psi_B = Z_\psi^{1/2} \psi, \quad e_B = e\mu^\varepsilon Z_e, \quad m_B = m Z_m.$$

We have replaced e by $e\mu^\varepsilon$ at the tree level, to have the renormalized electric charge e dimensionless.

The renormalization of the gauge-fixing sector is rather simple. Since C and \bar{C} decouple, they are not renormalized, so $C_B = C$, $\bar{C}_B = \bar{C}$. Moreover, since B appears only quadratically in (6.3), no one-particle irreducible diagram with external legs B can be constructed. Therefore, the Lagrangian terms involving B are not renormalized either. Writing

$$B_B = Z_B^{1/2} B, \quad \lambda_B = \lambda Z_\lambda,$$

we have

$$-\frac{\lambda}{2}B^2 + B\partial\cdot A = -\frac{\lambda_B}{2}B_B^2 + B_B\partial\cdot A_B = -\frac{\lambda Z_\lambda}{2}Z_B B^2 + Z_B^{1/2} Z_A^{1/2} B\partial\cdot A, \quad (6.9)$$

that is to say

$$Z_B = Z_A^{-1}, \quad Z_\lambda = Z_A. \quad (6.10)$$

We see that B can have a nontrivial renormalization constant.

Now, let us consider the terms proportional to the sources in (6.2). The term $BK_{\bar{C}}$ is not renormalized by the argument just given. Moreover, since the ghosts decouple, no irreducible diagrams with sources K_ψ , $K_{\bar{\psi}}$ and/or K_μ on the external legs can be constructed. This means that the entire K sector of the solution (6.2) to the master equation is nonrenormalized and

$$K_{\bar{\psi}B} = Z_e^{-1} Z_\psi^{-1/2} K_{\bar{\psi}}, \quad K_{\psi B} = Z_e^{-1} Z_\psi^{-1/2} K_\psi, \quad K_{\bar{C}B} = Z_B^{-1/2} K_{\bar{C}}.$$

The renormalized solution of the master equation reads

$$S_R(\Phi, K) = \int (\mathcal{L}_{0R} + \mathcal{L}_{gf} + \mathcal{L}_K) = S_B(\Phi_B, K_B),$$

where

$$\begin{aligned}\mathcal{L}_{gf} &= -\frac{\lambda}{2}B^2 + B \partial \cdot A - \bar{C}\square C, \\ \mathcal{L}_K &= K_\mu \partial_\mu C + ie\mu^\varepsilon \bar{\psi} C K_{\bar{\psi}} + ie\mu^\varepsilon K_\psi C \psi - BK_{\bar{C}}.\end{aligned}$$

6.1 Ward identities

The Ward identities (5.72) allow us to derive relations among the correlation functions and the renormalization constants. Before deriving the main formulas, let us mention two simple, but useful properties concerning the functional integral over the ghosts C , \bar{C} and the Lagrange multiplier B .

Since B does not propagate and appears quadratically in the action, integrating over B is equivalent to replace it with the solution

$$B = \frac{1}{\lambda} \partial \cdot A \quad (6.11)$$

of its own field equation. Precisely, let $X(B)$ be a local functional of B (and possibly other fields). Making a translation we find

$$\begin{aligned}\langle X \rangle_B &\equiv \int [dB] X(B) \exp \left(\frac{\lambda B^2}{2} - B \partial \cdot A \right) \\ &= \int [dB] X \left(B + \frac{\partial \cdot A}{\lambda} \right) \exp \left(\frac{\lambda B^2}{2} - \frac{(\partial \cdot A)^2}{2\lambda} \right).\end{aligned}$$

Now, expand $X(B)$ in powers of B . Observe that each odd power integrates to zero. On the other hand, nonvanishing even powers give $\delta(0)$ s or derivatives of $\delta(0)$ s, e.g.

$$\int [dB] B(x) \partial_\mu B(x) \exp \left(\frac{\lambda B^2}{2} \right) = \partial_\mu \delta(x-y)|_{y=x},$$

which, by formulas (2.12) and (2.13), vanish using the dimensional regularization. We conclude that

$$\langle X \rangle_B = X \left(\frac{\partial \cdot A}{\lambda} \right) \exp \left(-\frac{(\partial \cdot A)^2}{2\lambda} \right). \quad (6.12)$$

Another useful property is that, since the ghosts decouple, the correlation functions involving ghost insertions factorize, i.e.

$$\begin{aligned} & \langle C(x_1) \cdots C(x_m) \bar{C}(y_1) \cdots \bar{C}(y_n) \chi \rangle \\ &= \langle C(x_1) \cdots C(x_m) \bar{C}(y_1) \cdots \bar{C}(y_n) \rangle \langle \chi \rangle, \end{aligned} \quad (6.13)$$

where χ is any string of elementary fields other than the ghosts, e.g.

$$\chi = A_{\mu_1}(x_1) \cdots A_{\mu_n}(x_n) \bar{\psi}(y_1) \cdots \bar{\psi}(y_m) \psi(z_1) \cdots \psi(z_m).$$

Formula (6.13) can be easily proved by writing down the expressions of the averages as functional integrals.

We obtain the first Ward identity by choosing $\Psi = \bar{C}(x) \partial \cdot A_B(y)$ in formula (5.72), which gives

$$0 = \langle B_B(x) \partial \cdot A_B(y) \rangle_0 - \langle \bar{C}(x) \square C(y) \rangle_0.$$

We recall that the subscript 0 reminds us that the sources J are set to zero. Using (6.9) and (6.12) we can replace B_B with $(\partial \cdot A_B)/\lambda_B$. Next, using

$$\langle C(y) \bar{C}(x) \rangle_0 = G_{\text{free}}(y - x) \quad (6.14)$$

where $G_{\text{free}}(y - x)$ is the solution of $-\square G_{\text{free}}(y - x) = \delta(y - x)$, we find

$$\langle \partial \cdot A_B(x) \partial \cdot A_B(y) \rangle_0 = \lambda_B \delta(x - y).$$

In terms of renormalized quantities, this identity becomes

$$\langle \partial \cdot A(x) \partial \cdot A(y) \rangle_0 = \frac{\lambda Z_\lambda}{Z_A} \delta(x - y).$$

Since the left-hand side is convergent, by construction, the right-hand side must also be convergent, so we find

$$\frac{Z_\lambda}{Z_A} = \text{finite}. \quad (6.15)$$

In the minimal subtraction scheme every Z has the form $1 + \text{poles in } \varepsilon$, so

$$\bar{Z}_\lambda = \bar{Z}_A. \quad (6.16)$$

The bar over the Z s is to remind us that the renormalization constants are evaluated in the $\overline{\text{MS}}$ scheme. The result (6.16) agrees with (6.10). When we

derived (6.10), indeed, we implicitly used the minimal subtraction scheme, since we concentrated on the form of divergences. More generally, we know that we can always “subtract” arbitrary finite local counterterms. If we do this in the sector (6.9), we end up with (6.15).

As a second example, take $\Psi = \bar{C}_B(x)\bar{\psi}_B(y)\psi_B(z)$ in (5.72), which gives

$$0 = \langle B_B(x)\bar{\psi}_B(y)\psi_B(z) \rangle_0 + ie_B \langle \bar{C}(x)\bar{\psi}_B(y)C(y)\psi_B(z) \rangle_0 \\ - ie_B \langle \bar{C}(x)\bar{\psi}_B(y)C(z)\psi_B(z) \rangle_0.$$

Using (6.14), (6.12) and (6.13) we find

$$\frac{1}{\lambda_B} \langle \partial \cdot A_B(x)\bar{\psi}_B(y)\psi_B(z) \rangle_0 = -ie_B \langle \bar{\psi}_B(y)\psi_B(z) \rangle_0 [G_{\text{free}}(x-y) - G_{\text{free}}(x-z)].$$

In terms of the renormalized quantities, we have

$$\frac{Z_A^{1/2}}{\lambda Z_\lambda} \langle \partial \cdot A(x)\bar{\psi}(y)\psi(z) \rangle_0 = -ie\mu^\varepsilon Z_e \langle \bar{\psi}(y)\psi(z) \rangle_0 [G_{\text{free}}(x-y) - G_{\text{free}}(x-z)].$$

Since the correlation functions appearing in this equation are finite, we conclude

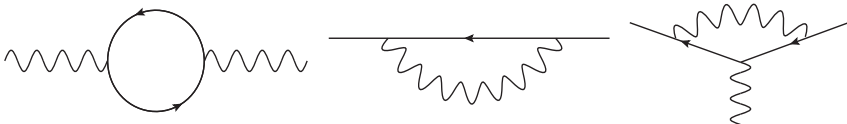
$$\frac{Z_A^{1/2}}{Z_\lambda Z_e} = \text{finite}. \quad (6.17)$$

Summarizing, in the minimal subtraction scheme

$$\bar{Z}_A = \bar{Z}_\lambda = \bar{Z}_e^{-2}. \quad (6.18)$$

Exercise 24 *Using the dimensional regularization, compute the renormalization of QED at one loop and check (6.18).*

Solution. We have already checked in exercises 22 and 23 that the photon four- and three-point functions are convergent. The surviving diagrams are



(6.19)

The first diagram is called “vacuum polarization”. Its divergent part is

$$-\frac{e^2}{6\pi^2\varepsilon}(k^2\delta_{\mu\nu} - k_\mu k_\nu), \quad (6.20)$$

where k is the external momentum, and gives

$$Z_A = 1 - \frac{e^2}{6\pi^2\varepsilon}. \quad (6.21)$$

Note that (6.20) is transverse, namely it vanishes if contracted with k_μ or k_ν . This means that the gauge-fixing term $(\partial \cdot A)^2/(2\lambda)$ is nonrenormalized, so

$$Z_\lambda = 1 - \frac{e^2}{6\pi^2\varepsilon} = Z_A,$$

in agreement with the first Ward identity (6.16).

The second diagram of (6.19) is the electron self-energy. Its divergent part is

$$\frac{-ie^2\lambda}{8\pi^2\varepsilon}\not{p} - \frac{me^2}{8\pi^2\varepsilon}(\lambda + 3), \quad (6.22)$$

where p is the external momentum, oriented according to the arrow. We find

$$Z_\psi = 1 - \frac{\lambda e^2}{8\pi^2\varepsilon}, \quad Z_m = 1 - \frac{3e^2}{8\pi^2\varepsilon}. \quad (6.23)$$

Finally, by locality and power counting the divergent part of the vertex-diagram can be calculated at vanishing external momenta. Moreover, masses in numerators can be dropped. We then easily find

$$- \frac{i\lambda e^3}{8\pi^2\varepsilon}\gamma_\mu, \quad (6.24)$$

whence

$$Z_e = 1 + \frac{e^2}{12\pi^2\varepsilon} = Z_A^{-1/2}, \quad (6.25)$$

in agreement with the second Ward identity (6.18). Observe that only Z_ψ is gauge dependent. Later we will appreciate why. We will be also able to characterize the gauge dependence more precisely. \square

Two interesting consequences of the Ward identities in the minimal subtraction scheme can be derived very easily.

i) The covariant derivative is not renormalized. Precisely,

$$D_\mu \equiv \partial_\mu + ie_B A_{B\mu} = \partial_\mu + ie\mu^{\varepsilon/2} \bar{Z}_e \bar{Z}_A^{1/2} A_\mu = \partial_\mu + ie\mu^{\varepsilon/2} A_\mu.$$

ii) The renormalization of the fields and the sources can be expressed in the form

$$\Phi_B^\alpha = (\bar{Z}_\Phi)^{1/2} \Phi^\alpha, \quad K_{\alpha B} = \bar{Z}_e^{-1} (\bar{Z}_\Phi)^{-1/2} K_\alpha, \quad (6.26)$$

where no sum over α is understood. Indeed, collecting all the pieces of information found so far, we have

$$\begin{aligned}
 A_{\mu\text{B}} &= \bar{Z}_e^{-1} A_\mu, & K_{\mu\text{B}} &= K_\mu, & \psi_{\text{B}} &= \bar{Z}_\psi^{1/2} \psi, \\
 K_{\psi\text{B}} &= \bar{Z}_e^{-1} \bar{Z}_\psi^{-1/2} K_\psi, & K_{C\text{B}} &= \bar{Z}_e^{-1} K_C, & C_{\text{B}} &= C, \\
 B_{\text{B}} &= \bar{Z}_e B, & K_{B\text{B}} &= \bar{Z}_e^{-2} K_B, & K_{\bar{C}\text{B}} &= \bar{Z}_e^{-1} K_{\bar{C}}, & \bar{C}_{\text{B}} &= \bar{C}, \\
 e_{\text{B}} &= e\mu^\varepsilon \bar{Z}_e, & m_{\text{B}} &= m\bar{Z}_m, & \lambda_{\text{B}} &= \lambda\bar{Z}_e^{-2}.
 \end{aligned} \tag{6.27}$$

The renormalizations of K_B and K_C are completely arbitrary, since the action does not depend on them. We have chosen them to enforce (6.26).

We see that only three renormalization constants are independent. The meaning of (6.26) is that the renormalization of the fields and the sources is \bar{Z}_e^{-1} times a canonical transformation. The complete renormalization is made of these two operations plus a redefinition of the electric charge e , the electron mass m and the gauge-fixing parameter λ .

Precisely, we have the canonical transformation $(\Phi_{\text{B}}, K_{\text{B}}) \rightarrow (\Phi', K')$ generated by

$$\begin{aligned}
 \mathcal{F}(\Phi', K_{\text{B}}) &= \int \left(\bar{Z}_e^{-1} A'_\mu K_{\text{B}}^\mu + \bar{Z}_\psi^{1/2} K_{\psi\text{B}} \psi' + \bar{Z}_\psi^{1/2} \bar{\psi}' K_{\bar{\psi}\text{B}} \right. \\
 &\quad \left. + K_{C\text{B}} C' + \bar{C}' K_{\bar{C}\text{B}} + \bar{Z}_e B' K_{B\text{B}} \right).
 \end{aligned}$$

composed with the source redefinition

$$\Phi' = \Phi, \quad K' = \bar{Z}_e^{-1} K,$$

and

$$e_{\text{B}} = e\mu^\varepsilon \bar{Z}_e, \quad m_{\text{B}} = m\bar{Z}_m, \quad \lambda_{\text{B}} = \lambda\bar{Z}_e^{-2}.$$

We can write the relation between the bare and the renormalized antiparentheses as

$$\bar{Z}_e^{-1} (X, Y)_{\text{B}} = (X, Y). \tag{6.28}$$

Details about nonminimal subtraction schemes are given in the next section.

6.2 Renormalizability of QED to all orders

Now we prove that quantum electrodynamics is renormalizable to all orders in a gauge invariant way. We first work out the proof in the minimal subtraction

scheme and at the end extend the proof to a generic gauge invariant scheme. Consider the bare generating functional

$$Z_B(J_B, K_B) = \int [d\Phi_B] \exp \left(-S_B(\Phi_B, K_B) + \sum_i \Phi_B^i J_{Bi} \right) = e^{W_B(J_B, K_B)}, \quad (6.29)$$

written in terms of bare fields and sources. The action S_B is the one of formula (5.31) once the subscript B is inserted everywhere.

We know that S_B satisfies the master equation, $(S_B, S_B)_B = 0$ and then theorem (11) ensures that the bare Γ functional Γ_B also satisfies $(\Gamma_B, \Gamma_B)_B = 0$. This identity implies

$$0 = - \int \frac{\delta_r \Gamma_B}{\delta K_{\alpha B}} \frac{\delta_l \Gamma_B}{\delta \Phi_B^\alpha} = \int \langle R_B^\alpha(\Phi) \rangle \frac{\delta_l \Gamma_B}{\delta \Phi_B^\alpha}. \quad (6.30)$$

Now, observe that $\langle R_B^A(\Phi_B) \rangle = R_B^A(\Phi_B)$. This is obvious for $\Phi_B = A_B, C_B, \bar{C}_B$ and B_B , because their functions $R_B^A(\Phi_B)$ vanish or are linear in the fields themselves. It is less obvious for $\Phi_B = \bar{\psi}_B, \psi_B$, yet true, because the ghosts decouple, so by (6.13) we have $\langle C_B \psi_B \rangle = \langle C_B \rangle \langle \psi_B \rangle$ and $\langle \bar{\psi}_B C_B \rangle = \langle \bar{\psi}_B \rangle \langle C_B \rangle$. We conclude that the bare functional Γ_B satisfies

$$\int R_B^\alpha(\Phi_B) \frac{\delta_l \Gamma_B}{\delta \Phi_B^\alpha} = 0. \quad (6.31)$$

More explicitly,

$$\int \left(\partial_\mu C_B \frac{\delta_l \Gamma_B}{\delta A_B^\mu} - ie_B \bar{\psi}_B C_B \frac{\delta_l \Gamma_B}{\delta \bar{\psi}_B} + ie_B \frac{\delta_r \Gamma_B}{\delta \psi_B} C_B \psi_B + B_B \frac{\delta_l \Gamma_B}{\delta \bar{C}_B} \right) = 0 \quad (6.32)$$

Now we proceed inductively. Assume that the theory can be renormalized up to and including the n th loop order by means of renormalization constants $\bar{Z}_{e,n}$, $\bar{Z}_{\psi,n}$ and $\bar{Z}_{m,n}$ and the renormalized action

$$S_n(\Phi, K) = \int \frac{1}{4} \bar{Z}_{e,n}^{-2} F_{\mu\nu}^2 + \int \bar{Z}_{\psi,n} \bar{\psi} (\not{\partial} + ie\mu^\varepsilon \not{A} + m \bar{Z}_{m,n}) \psi + \int (\mathcal{L}_{\text{gf}} + \mathcal{L}_K),$$

in the minimal subtraction scheme. The relations between the bare and the renormalized quantities are (6.27) with $\bar{Z}_e, \bar{Z}_\psi, \bar{Z}_m \rightarrow \bar{Z}_{e,n}, \bar{Z}_{\psi,n}, \bar{Z}_{m,n}$. Let $\Gamma_n(\Phi, K, e, \lambda) = \Gamma_B(\Phi_B, K_B, e_B, \lambda_B)$ denote the n -loop renormalized and bare generating functionals of one-particle irreducible diagrams.

We must prove that the inductive hypotheses are promoted to the $(n+1)$ -th loop order. Switching formula (6.32) to the renormalized quantities, we find that all the renormalization constants simplify apart from a common factor \bar{Z}_e , which we can drop. At the end, we have

$$\int \left(\partial_\mu C \frac{\delta_l \Gamma_n}{\delta A^\mu} - ie\mu^\varepsilon \bar{\psi} C \frac{\delta_l \Gamma_n}{\delta \bar{\psi}} + ie\mu^\varepsilon \frac{\delta_r \Gamma_n}{\delta \psi} C \psi + B \frac{\delta_l \Gamma_n}{\delta \bar{C}} \right) = 0. \quad (6.33)$$

We know that the gauge-fixing sector and the K sector do not renormalize. Actually they do not receive any radiative corrections, because no diagrams can be constructed with those sets of external legs. Thus we have

$$\Gamma_n(\Phi, K) = \tilde{\Gamma}_n(A, \bar{\psi}, \psi) + \int (\mathcal{L}_{\text{gf}} + \mathcal{L}).$$

Inserting this formula in (6.33), we get

$$\int \left(\partial_\mu C \frac{\delta_l \tilde{\Gamma}_n}{\delta A^\mu} - ie\mu^\varepsilon \bar{\psi} C \frac{\delta_l \tilde{\Gamma}_n}{\delta \bar{\psi}} + ie\mu^\varepsilon \frac{\delta_r \tilde{\Gamma}_n}{\delta \psi} C \psi \right) = 0.$$

Multiplying by a constant anticommuting parameter ξ to the left and identifying ξC with a commuting function Λ , we obtain that $\tilde{\Gamma}_n$ is gauge invariant, that is to say

$$\delta_\Lambda \tilde{\Gamma}_n = 0, \quad (6.34)$$

where δ_Λ is given by (4.7).

To keep track of the orders of the expansion, we reintroduce \hbar for a moment. Define

$$\tilde{\Gamma}_n = \sum_{k=0}^{\infty} \hbar^k \tilde{\Gamma}_n^{(k)}. \quad (6.35)$$

Observe that δ_Λ is independent of \hbar . Taking the $(n+1)$ -th order of (6.34), we obtain

$$\delta_\Lambda \tilde{\Gamma}_n^{(n+1)} = 0. \quad (6.36)$$

By the inductive assumption, Γ_n and $\tilde{\Gamma}_n$ are convergent up to and including the n th order. Instead, $\tilde{\Gamma}_n^{(n+1)}$ is the sum of a divergent part, which we denote by $\tilde{\Gamma}_{n\text{div}}^{(n+1)}$, and a finite part. Since all $\tilde{\Gamma}_n^{(k)}$, $k \leq n$, are convergent by the inductive assumption, all the subdivergences of the Feynman diagrams of order \hbar^{n+1} are subtracted by appropriate counterterms. By the theorem of

the locality of counterterms, $\tilde{\Gamma}_{n\text{div}}^{(n+1)}$ is a local functional. Taking the divergent part of (6.36) (i.e. its poles in ε), we obtain

$$\delta_{\Lambda} \tilde{\Gamma}_{n\text{div}}^{(n+1)} = 0. \quad (6.37)$$

Thus, we learn that $\tilde{\Gamma}_{n\text{div}}^{(n+1)}$ is gauge invariant. Summarizing, $\tilde{\Gamma}_{n\text{div}}^{(n+1)}(A, \bar{\psi}, \psi)$ is a local, gauge-invariant functional. Precisely, it is the integral of a local function $\Delta_{n+1}\mathcal{L}(A, \bar{\psi}, \psi)$ of dimension four.

Now we use power counting. With the Lorenz gauge-fixing, the photon propagator behaves correctly for large momenta. Moreover, the theory does not contain parameters of negative dimensions. These facts ensure that the function $\Delta_{n+1}\mathcal{L}(A, \psi, \bar{\psi})$ is a linear combination of the local terms of dimensions ≤ 4 that are not total derivatives and can be built with the fields A , ψ and $\bar{\psi}$ and their derivatives. Such terms are F^2 , $(\partial \cdot A)^2$, $\bar{\psi}\not{\partial}\psi$, $\bar{\psi}\not{A}\psi$ and $\bar{\psi}\psi$. We cannot use neither the tensor $\varepsilon^{\mu\nu\rho\sigma}$, nor the matrix γ_5 , since the Feynman rules do not contain them. Finally, (6.37) reduces the list to the gauge-invariant combinations F^2 , $\bar{\psi}\not{D}\psi$ and $\bar{\psi}\psi$, so we can write

$$\Delta_{n+1}\mathcal{L} = a_{n+1}F_{\mu\nu}^2 + b_{n+1}\bar{\psi}\not{D}\psi + c_{n+1}m\bar{\psi}\psi, \quad \tilde{\Gamma}_{n\text{div}}^{(n+1)} = \int d^Dx \Delta_{n+1}\mathcal{L}, \quad (6.38)$$

for suitable divergent coefficients a_{n+1} , b_{n+1} and c_{n+1} . These divergences can be subtracted by means of new renormalization constants

$$\begin{aligned} \bar{Z}_{e,n+1} &= (\bar{Z}_{e,n}^{-2} - a_{n+1})^{-1/2}, & \bar{Z}_{\psi,n+1} &= \bar{Z}_{\psi,n} - b_{n+1}, \\ \bar{Z}_{m,n+1} &= (\bar{Z}_{\psi,n} - b_{n+1})^{-1}(\bar{Z}_{\psi,n}\bar{Z}_{m,n} - c_{n+1}). \end{aligned}$$

The renormalized action

$$\begin{aligned} S_{n+1}(\Phi, K) &= \int \frac{1}{4} \bar{Z}_{e,n+1}^{-2} F_{\mu\nu}^2 + \int \bar{Z}_{\psi,n+1} \bar{\psi}(\not{\partial} + ie\mu^\varepsilon \not{A} + m\bar{Z}_{m,n+1})\psi \\ &\quad + \int (\mathcal{L}_{\text{gf}} + \mathcal{L}_K) = S_n(\Phi, K) - \tilde{\Gamma}_{n\text{div}}^{(n+1)}, \end{aligned}$$

produces a generating functional Γ_{n+1} that is convergent up to and including $n+1$ loops. Indeed, since the actions differ by $\mathcal{O}(\hbar^{n+1})$, the Feynman diagrams with n loops or less are exactly the same, which ensures $\Gamma_{n+1} = \Gamma_n + \mathcal{O}(\hbar^{n+1})$. Moreover, at $n+1$ loops we have exactly the same diagrams plus the vertices of $-\tilde{\Gamma}_{n\text{div}}^{(n+1)}$, which subtract the overall divergent

parts. In conclusion,

$$\Gamma_{n+1} \equiv \sum_{k=0}^{\infty} \hbar^k \tilde{\Gamma}_{n+1}^{(k)} = \Gamma_n - \tilde{\Gamma}_{n\text{div}}^{(n+1)} + \mathcal{O}(\hbar^{n+2}),$$

that is to say $\tilde{\Gamma}_{n+1}^{(k)} = \tilde{\Gamma}_n^{(k)} < \infty$ for $k \leq n$ and $\tilde{\Gamma}_{n+1}^{(n+1)} = \tilde{\Gamma}_n^{(n+1)} - \tilde{\Gamma}_{n\text{div}}^{(n+1)} < \infty$.

This result extends the inductive hypotheses to $n+1$ loops, as we wanted. Iterating the argument to $n = \infty$, the map relating the bare and renormalized quantities is (6.27) with $\bar{Z}_e = \bar{Z}_{e,\infty}$, $\bar{Z}_\psi = \bar{Z}_{\psi,\infty}$ and $\bar{Z}_m = \bar{Z}_{m,\infty}$. The renormalized action is

$$S_R = S_\infty = \int \frac{1}{4} \bar{Z}_e^{-2} F_{\mu\nu}^2 + \int \bar{Z}_\psi \bar{\psi} (\not{\partial} + ie\mu^\varepsilon \not{A} + m\bar{Z}_m) \psi + \int (\mathcal{L}_{\text{gf}} + \mathcal{L}_K) \quad (6.39)$$

and the renormalized generating functional of the one-particle irreducible correlation functions is

$$\Gamma_R(\Phi, K) = \Gamma_\infty(\Phi, K) = \tilde{\Gamma}_\infty(A, \bar{\psi}, \psi) + \int (\mathcal{L}_{\text{gf}} + \mathcal{L}_K). \quad (6.40)$$

Moreover, we have

- (i) $\bar{Z}_e^{-1}(X, Y)_B = (X, Y)$;
- (ii) $(S_R, S_R) = 0$;
- (iii) $(\Gamma_R, \Gamma_R) = 0$.

Point (i) follows from (6.27), as shown in (6.28). Point (ii) follows from $(S_B, S_B)_B = 0$ and point (i). It can also be verified immediately by using (6.39). Point (iii) follows from point (ii) and theorem 11.

So far, we have worked in the absence of composite fields, which is enough to derive the S matrix. When we include gauge-invariant composite fields, built with the physical fields A , ψ and $\bar{\psi}$, both the gauge-fixing sector and the K sector remain uncorrected, because no nontrivial diagrams affecting them can be constructed. The derivation given above is unmodified up to and including (6.37).

Let $\mathcal{O}^I(\Phi, e\mu^\varepsilon)$ denote a basis of gauge-invariant composite fields, which includes the identity. The \mathcal{O}^I s may depend on e by gauge invariance, but there is no need to assume that they depend on m . The bare action is extended to

$$S_B(\Phi_B, K_B, L_B) = S_B(\Phi_B, K_B) + \int L_B^I \mathcal{O}^I(\Phi_B, e_B).$$

Clearly, the master equation $(S_B, S_B)_B = 0$ is still satisfied. We write the n -loop renormalized action as

$$S_n(\Phi, K, L) = S_n(\Phi, K) + \int f_n^I(L) \mathcal{O}^I(\Phi, e\mu^\varepsilon), \quad (6.41)$$

where $f_n^I(L)$ are local functions to be determined that have the form $L_n^I + \text{poles}$ in ε , with $f_0^I(L) = L^I$. Obviously, $(S_n, S_n) = 0$.

The sources L that multiply the composite fields of dimensions > 4 have negative dimensions in units of mass. This means that the divergent part $\tilde{\Gamma}_{n\text{div}}^{(n+1)}$ is no longer restricted by power counting. Nevertheless, we can write

$$\tilde{\Gamma}_{n\text{div}}^{(n+1)}(\Phi, K, L) = \int \Delta_{n+1} \mathcal{L} + \int h_n^I(L) \mathcal{O}^I(\Phi, e\mu^\varepsilon),$$

where the divergent functions $h_n^I(L) = \mathcal{O}(L)$ are local. As before, the divergent terms of $\Delta_{n+1} \mathcal{L}$ can be reabsorbed in the constants \bar{Z}_e , \bar{Z}_ψ and \bar{Z}_m . Instead, the L -dependent divergent part can be reabsorbed by defining

$$f_{n+1}^I(L) = f_n^I(L) - h_n^I(L). \quad (6.42)$$

The relations among the bare sources L_B^I and the renormalized sources L^I , and the solutions of (6.42), are

$$L_B^I \equiv f_n^J(L) (d_n^{-1})^{JI}, \quad f_n^I(L) = L^I - \sum_{k=1}^n h_k^I(L),$$

where the matrices of constants $d_n^{IJ} = \delta^{IJ} + \text{poles}$ in ε are defined by

$$\mathcal{O}^I(\Phi_B, e_B) = \mathcal{O}^I(\bar{Z}_{\Phi,n}^{1/2} \Phi, e\mu^\varepsilon \bar{Z}_{e,n}) \equiv d_n^{IJ} \mathcal{O}^J(\Phi, e\mu^\varepsilon).$$

The $(n+1)$ -renormalized action $S_{n+1}(\Phi, K, L)$ has the form (6.41) with $n \rightarrow n+1$. We still have

$$S_{n+1}(\Phi, K, L) = S_n(\Phi, K, L) - \tilde{\Gamma}_{n\text{div}}^{(n+1)}(\Phi, K, L),$$

which ensures that the Γ functional $\Gamma_{n+1}(\Phi, K, L)$ is renormalized up to and including $n+1$ loops. Iterating the argument to $n = \infty$, we find the renormalized action $S_R(\Phi, K, L) = S_\infty(\Phi, K, L)$ and the renormalized Γ functional $\Gamma_R(\Phi, K, L) = \Gamma_\infty(\Phi, K, L)$, which still satisfy the properties (i), (ii) and (iii) listed above.

Chapter 7

Non-Abelian gauge field theories

In this chapter we use the Batalin-Vilkovisky formalism to prove the renormalizability of Yang-Mills theory to all orders in the perturbative expansion. We concentrate on gauge theories with a simple gauge group, since the generalization to product groups is straightforward. We also assume that the theories are parity invariant, which ensures that the classical Lagrangian does not contain the matrix γ_5 , the tensor $\varepsilon^{\mu\nu\rho\sigma}$, or their d -dimensional analogues, where d denotes the physical dimension of spacetime.

7.1 Renormalizability of non-Abelian gauge theories to all orders

Denote the bare fields and the bare sources with Φ_B and K_B , respectively. Denote the bare action and the bare Γ functional, defined according to (5.59), with $S_B(\Phi_B, K_B, L_B, \zeta_B, \xi_B)$ and $\Gamma_B(\Phi_B, K_B, \zeta_B, L_B, \xi_B)$, where ζ denote the physical parameters, ξ are the gauge-fixing parameters and L are sources for gauge-invariant composite fields. At $L = 0$ the bare action can be read from (5.34) and (5.33), or (5.38) and (5.37), if all the quantities that appear in those formulas are interpreted as bare quantities.

From (5.17) we have the master equation

$$(S_B, S_B)_B = 0, \tag{7.1}$$

which implies, according to theorem 11,

$$(\Gamma_B, \Gamma_B)_B = 0. \quad (7.2)$$

The subscript $_B$ attached to the antiparentheses means that they are calculated with respect to the bare fields and sources, the other bare quantities being kept fixed.

As usual, renormalizability is proved by proceeding inductively. We give two proofs: a raw one and a more detailed one.

Raw subtraction The simpler proof amounts to subtract the counterterms “as they come” in the minimal subtraction scheme. We will see in a moment what this means. We do not need to preserve the master equation at each step of the subtraction. Instead, higher-order violations are allowed.

Call S_n and Γ_n the action and the Γ functional renormalized up to and including n loops. Assume the inductive hypotheses

$$S_n = S_0 + \text{poles}, \quad (S_n, S_n) = \mathcal{O}(\hbar^{n+1}), \quad \Gamma_n^{(k)} < \infty \quad \forall k \leq n, \quad (7.3)$$

having used the expansion (6.35). The last requirement is just the statement that Γ_n is convergent up to and including n loops. Clearly, the inductive assumptions are trivially satisfied for $n = 0$. In particular, S_0 coincides with the bare action S_B , and formula (7.1) ensures $(S_0, S_0) = 0$. Using the master identity 12 we have

$$(\Gamma_n, \Gamma_n) = \langle (S_n, S_n) \rangle. \quad (7.4)$$

Since the antiparenthesis (S_n, S_n) is a local functional of order \hbar^{n+1} , the contributions to $\langle (S_n, S_n) \rangle$ of order \hbar^{n+1} are given by tree diagrams, so they coincide with the order- \hbar^{n+1} contributions to (S_n, S_n) , which we denote by $(S_n, S_n)|_{n+1}$. Such quantity is divergent, by $(S_0, S_0) = 0$ and the first assumption of (7.3).

Use the expansion (6.35) and think of (7.4) diagrammatically, as shown in (5.28). The order \hbar^{n+1} of (7.4) gives

$$\sum_{k=0}^{n+1} \left(\Gamma_n^{(k)}, \Gamma_n^{(n-k)} \right) = (S_n, S_n)|_{n+1}. \quad (7.5)$$

We know that $\Gamma_n^{(k)}$ are convergent for $k \leq n$, by the inductive assumption. Taking the divergent part of (7.5), we obtain

$$2 \left(\Gamma_n^{(0)}, \Gamma_n^{(n+1)} \right)_{\text{div}} = (S_n, S_n)|_{n+1}. \quad (7.6)$$

where $\Gamma_{n \text{ div}}^{(n+1)}$ is the order- \hbar^{n+1} divergent part of Γ_n . By the third inductive assumption (7.3), all the subdivergences are subtracted away, so $\Gamma_{n \text{ div}}^{(n+1)}$ is a local functional. Now, $\Gamma_n^{(0)}$ coincides with the classical action S_0 , so (7.6) becomes

$$(S_0, \Gamma_{n \text{ div}}^{(n+1)}) = \frac{1}{2} (S_n, S_n)|_{n+1}. \quad (7.7)$$

At this point, define

$$S_{n+1} = S_n - \Gamma_{n \text{ div}}^{(n+1)}. \quad (7.8)$$

The first inductive assumption of the list (7.3) is clearly promoted to S_{n+1} . Formulas (7.8) and (7.7) give

$$(S_{n+1}, S_{n+1}) = (S_n, S_n) - 2 \left(S_n, \Gamma_{n \text{ div}}^{(n+1)} \right) + \left(\Gamma_{n \text{ div}}^{(n+1)}, \Gamma_{n \text{ div}}^{(n+1)} \right) = \mathcal{O}(\hbar^{n+2}),$$

so the second of (7.3) is also promoted to S_{n+1} . Finally, the diagrams constructed with the vertices of S_{n+1} coincide with the diagrams of S_n , plus new diagrams containing the vertices of $-\Gamma_{n \text{ div}}^{(n+1)}$. However, the first contributions of the new diagrams have order \hbar^{n+1} , so

$$\Gamma_n^{(k)} = \Gamma_{n+1}^{(k)} \quad \forall k \leq n.$$

Moreover, at $n+1$ loops any vertex of $-\Gamma_{n \text{ div}}^{(n+1)}$ can be used only once and alone, since it is already of order \hbar^{n+1} . Thus,

$$\Gamma_{n+1}^{(n+1)} = \Gamma_n^{(n+1)} - \Gamma_{n \text{ div}}^{(n+1)} < \infty,$$

which promotes the third inductive assumption of (7.3) to Γ_{n+1} .

We conclude that formulas (7.3) and (7.4) also hold for the renormalized action $S_R \equiv S_\infty$ and the renormalized generating functional $\Gamma_R \equiv \Gamma_\infty$, i.e.

$$(S_R, S_R) = 0, \quad (\Gamma_R, \Gamma_R) = 0. \quad (7.9)$$

The subtraction algorithm just given is clearly compatible with propositions 16 and 17. In particular, formula (7.8) ensures that those propositions hold at every step of the subtraction procedure.

We now study the renormalized action.

Proposition 16 *The renormalized action is independent of K_B^a and depends on B^a , $K_{\bar{C}}^a$ only by means of the terms*

$$\int \left(-\frac{\lambda}{2} (B^a)^2 + B^a \partial \cdot A^a - B^a K_{\bar{C}}^a \right),$$

which are nonrenormalized.

Proof. Clearly the classical action (5.37) satisfies these properties. Then, no one-particle irreducible diagrams with external legs B^a , K_B^a and $K_{\bar{C}}^a$ can be constructed, so no counterterms can depend on B^a , K_B^a or $K_{\bar{C}}^a$. Note that, in particular, the absence of vertices with B legs is due to the linearity of the gauge fixing \mathcal{G}^a in A .

Proposition 17 *The renormalized action depends on \bar{C} and K_a^μ only by means of the combination*

$$K_a^\mu + \partial^\mu \bar{C}^a. \quad (7.10)$$

Proof. Again, this property is satisfied by the classical action (5.37). Then, the vertices that contain an antighost leg always have a derivative ∂ acting on \bar{C} . Moreover, the vertex containing $\partial \bar{C}$ has an identical vertex-partner with $\partial \bar{C}$ replaced by K_a^μ . Therefore, given a diagram G with a K_a^μ external leg, there exists an almost identical diagram G' , which differs from G only because the external K_a^μ leg is replaced by a $\partial^\mu \bar{C}^a$ leg, and vice versa. Thus all the counterterms satisfy the property, and so does the renormalized action.

Proposition 18 *The renormalized action is linear in K .*

Proof. Indeed, from (5.35) and (5.36) it follows that in the absence of composite fields any local terms that are quadratic in K have either dimension greater than four or ghost number different from zero. \square

Propositions 16 and 17 also hold in the presence of source for composite fields, because their proofs do not require arguments based on power counting. Instead, 18 does not hold at $L \neq 0$, in general, because the sources L for the composite fields can have arbitrarily large negative dimensions, as well as vanishing ghost number. Then, local Lagrangian terms with arbitrarily large powers of K can be constructed, provided we adjust their dimensions

by means of powers of L and their ghost numbers by means of powers of C . We know that, in renormalization, when a term cannot be excluded *a priori* by advocating power counting, symmetries or other properties, it is typically generated as the divergent part of some diagram. For this reason, we cannot guarantee that proposition 18 holds at $L \neq 0$. For a while we argue at $L = 0$, then generalize our arguments to $L \neq 0$.

Since the renormalized action S_R is linear in the sources K at $L = 0$, we can write

$$S_R(\Phi, K) = S'_R(\Phi) - \int d^D x R_\infty^\alpha(\Phi) K_\alpha.$$

The functions $R_\infty^\alpha(\Phi)$ that multiply the sources inside S_R are the renormalized field transformations. By proposition 16, ghost number conservation, locality and power counting, we must have, in the notation of formulas (5.37) and (5.39),

$$\begin{aligned} S_R(\Phi, K) = S'_R(\Phi) - \int (a \partial_\mu C_j^i + b A_{\mu k}^i C_j^k - c A_{\mu j}^k C_k^i) K_{\mu i}^j + \int h C_k^i C_j^k K_{C i}^j \\ - \int B_j^i K_{C i}^j + \int \left[\left(a_1 C_{l_1}^{i_1} \psi_{j_1 \dots j_m}^{l_1 i_2 \dots i_n} + \dots + a_n C_{l_n}^{i_n} \psi_{j_1 \dots j_m}^{i_1 \dots i_{n-1} l_n} \right) K_{i_1 \dots i_n}^{j_1 \dots j_m} + \text{h.c.} \right], \end{aligned}$$

where a, b, c, h, a_k and b_k are numerical constants. Note that, out of the three tensors of (4.39), we can only use δ_j^i . The ε tensors cannot appear, just because they are not present in the Feynman rules. It is easy to check that the terms proportional to $K_{\mu i}^j$ in the master equation $(S_R, S_R) = 0$ give $2b = 2c = h$. Moreover, the terms proportional to $K_{i_1 \dots i_n}^{j_1 \dots j_m}$ give $2a_k = h$ for every k, l . Writing $h = 2iga'$, we have

$$\begin{aligned} S_R(\Phi, K) = S'_R(\Phi) + ga' \int d^D x \left(\bar{\psi}^I \mathcal{T}_{IJ}^a C^a K_\psi^J + K_\psi^I \mathcal{T}_{IJ}^a C^a \psi^J \right) \\ - \int d^D x \left[\left(a \partial_\mu C^a + ga' f^{abc} A_\mu^b C^c \right) K_\mu^a - \frac{ga'}{2} f^{abc} C^b C^c K_C^a + B^a K_C^a \right], \end{aligned}$$

Propositions 16 and 17 ensure that the B -dependent terms are nonrenormalized. Then, by locality, power counting and ghost number conservation, $S'_R(\Phi)$ has the form

$$\begin{aligned} S'_R(\Phi) = S_{cR}(\phi) - \frac{\lambda}{2} \int B_j^i B_i^j + \int B_i^j \partial \cdot A_j^i \\ - \int \bar{C}_i^j \partial_\mu \left(\tilde{a} \partial_\mu C_j^i + \tilde{b} A_{\mu k}^i C_j^k - \tilde{c} A_{\mu j}^k C_k^i \right), \end{aligned}$$

where \tilde{a} , \tilde{b} and \tilde{c} are other constants. Imposing $(S_R, S_R) = 0$, we get $\tilde{a} = a$, $\tilde{b} = \tilde{c} = ga'$ and $(S_R, S_{cR}) = 0$. Then, we can write

$$\begin{aligned} S'_R(\Phi) &= S_{cR}(\phi) + (S_R, \Psi) = S_{cR}(\phi) + S_{\text{gf}}(\Phi) = \\ &= S_{cR}(\phi) + \int \left[-\frac{\lambda}{2}(B^a)^2 + B^a \partial \cdot A^a - \bar{C}^a \partial_\mu \left(a \partial_\mu C^a + ga' f^{abc} A_\mu^b C^c \right) \right]. \end{aligned}$$

Note that the gauge fermion

$$\Psi = \int \bar{C}^a \left(-\frac{\lambda}{2} B^a + \partial \cdot A^a \right)$$

is nonrenormalized.

Let us focus, for simplicity, on the pure gauge theory. Writing the most general local form of $S_{cR}(A)$, it is easy to check, by explicit computation, that the most general solution to $(S_R, S_{cR}) = 0$ is

$$S_{cR}(A) = \frac{a''}{4} \int d^D x \left(a \partial_\mu A_\nu^a - a \partial_\nu A_\mu^a + ga' f^{abc} A_\mu^b A_\nu^c \right)^2,$$

where a'' is another constant. Writing

$$a = Z_C, \quad a'' = Z_A Z_C^{-2}, \quad a' = \mu^{\varepsilon/2} Z_g Z_A^{1/2} Z_C,$$

we finally obtain

$$S_R(\Phi, K, g, \lambda) = S_B(\Phi_B, K_B, g_B, \lambda_B), \quad (7.11)$$

with

$$\begin{aligned} A_{\mu B}^a &= Z_A^{1/2} A_\mu^a, & C_B^a &= Z_C^{1/2} C^a, & g_B &= g \mu^{\varepsilon/2} Z_g, \\ B_B^a &= Z_A^{-1/2} B^a, & \bar{C}_B^a &= Z_C^{1/2} \bar{C}^a, & \lambda_B &= \lambda Z_A, \\ K_{aB}^\mu &= Z_C^{1/2} K_a^\mu, & K_{CB}^a &= Z_A^{1/2} K_C^a, & K_{\bar{C}B}^a &= Z_A^{1/2} K_{\bar{C}}^a. \end{aligned} \quad (7.12)$$

The inclusion of matter is straightforward: only $S_{cR}(\phi)$ changes, since it must include all the terms of dimensions ≤ 4 that are invariant with respect to the renormalized gauge transformations.

At $L \neq 0$ the renormalized action has a more involved structure, since higher-dimensional composite fields make it nonpolynomial in Φ , K and L . In the sector $L \neq 0$, we just subtract the counterterms as they come, according

to formula (7.8). We do not need to worry about rewriting the subtraction as a redefinition of the fields, the sources and the parameters. If we wanted to, we would have nonpolynomial field redefinitions. The gauge transformations would also be affected, so the L sector would have to include nonpolynomial corrections that are not invariant under the starting gauge transformations, but invariant under suitably corrected gauge transformations. For the moment, we do not need to go through this, because formula (7.8) is sufficient for most practical purposes.

Expressing renormalization as a redefinition of the fields, the sources and the parameters (which is the true meaning of the word “re-normalization”) is more useful in the $L = 0$ sector, which contains, among other things, the physical parameters. So doing, we can show that the renormalization program can be carried out to the end by keeping the number of independent physical parameters finite. This is a necessary requirement to ensure that predictivity is retained. The composite fields, on the other hand, do not add physical parameters to the theory, since the sources L are just tools to simplify the derivations of various properties. Thus, we do not lose much, if we renormalize the divergences belonging to the L -dependent sector by subtracting them away just as they come.

Appendix A

Notation and useful formulas

The flat space metric tensor reads

$$\eta_{\mu\nu} = \eta^{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}.$$

The Pauli matrices are

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

The γ matrices in four dimensions read

$$\gamma^\mu = \begin{pmatrix} 0 & \sigma^\mu \\ \bar{\sigma}^\mu & 0 \end{pmatrix}, \quad \gamma^5 = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix},$$

where $\sigma^\mu = (1, \sigma^1, \sigma^2, \sigma^3)$ and $\bar{\sigma}^\mu = (1, -\sigma^1, -\sigma^2, -\sigma^3)$. The ε tensor $\varepsilon^{\mu\nu\rho\sigma}$ is defined so that $\varepsilon^{0123} = 1$.

In Minkowski spacetime the Fourier transform is defined as

$$\varphi(x) = \int \frac{d^D p}{(2\pi)^D} e^{-ip \cdot x} \tilde{\varphi}(p), \quad (\text{A.1})$$

while in Euclidean space it is

$$\varphi(x) = \int \frac{d^D p}{(2\pi)^D} e^{ip \cdot x} \tilde{\varphi}(p).$$

To manipulate the denominators of Feynman diagrams it is useful to introduce Feynman parameters by means of the formula

$$\prod_i \frac{1}{A^{\alpha_i}} = \frac{\Gamma(\sum_i \alpha_i)}{\prod_j \Gamma(\alpha_j)} \int_0^1 \prod_i \left(dx_i x_i^{\alpha_i-1} \right) \frac{\delta(1 - \sum_k x_k)}{(\sum_m x_m A_m)^{\sum_n \alpha_n}}.$$

Particular cases are

$$\begin{aligned} \frac{1}{AB} &= \int_0^1 dx \frac{1}{[Ax + B(1-x)]^2}, \\ \frac{1}{A^\alpha B^\beta} &= \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} \int_0^1 dx \frac{x^{\alpha-1}(1-x)^{\beta-1}}{[Ax + B(1-x)]^{\alpha+\beta}}, \\ \frac{1}{ABC} &= 2 \int_0^1 dx \int_0^{1-x} dy \frac{1}{[Ax + By + C(1-x-y)]^3}. \end{aligned} \quad (\text{A.2})$$

The integration over Feynman parameters often reduces to the integral

$$\int_0^1 dx x^{\alpha-1}(1-x)^{\beta-1} = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha + \beta)}. \quad (\text{A.3})$$

The most frequently used D -dimensional integral is

$$\int \frac{d^D p}{(2\pi)^D} \frac{1}{(p^2 - m^2)^\alpha} = \frac{i(-1)^\alpha \Gamma(\alpha - \frac{D}{2})}{(4\pi)^{D/2} \Gamma(\alpha)} (m^2)^{\frac{D}{2} - \alpha}. \quad (\text{A.4})$$

More generally,

$$\int \frac{d^D p}{(2\pi)^D} \frac{(p^2)^\beta}{(p^2 - m^2)^\alpha} = \frac{i(-1)^{\alpha+\beta} \Gamma(\beta + \frac{D}{2}) \Gamma(\alpha - \beta - \frac{D}{2})}{(4\pi)^{D/2} \Gamma(\alpha) \Gamma(\frac{D}{2})} (m^2)^{\frac{D}{2} - \alpha + \beta},$$

In Euclidean space this result reads

$$\int \frac{d^D p}{(2\pi)^D} \frac{(p^2)^\beta}{(p^2 + m^2)^\alpha} = \frac{\Gamma(\beta + \frac{D}{2}) \Gamma(\alpha - \beta - \frac{D}{2})}{(4\pi)^{D/2} \Gamma(\alpha) \Gamma(\frac{D}{2})} (m^2)^{\frac{D}{2} - \alpha + \beta}. \quad (\text{A.5})$$

We also recall that

$$\int \frac{d^D p}{(2\pi)^D} (p^2)^\alpha = 0, \quad (\text{A.6})$$

for every α .

We have $\Gamma(x+1) = x\Gamma(x)$, $\Gamma(n+1) = n!$ and

$$\begin{aligned}\Gamma\left(\frac{n}{2}\right) &= \sqrt{\pi} \frac{(n-2)!!}{2^{(n-1)/2}}, \\ \Gamma(z) &= \frac{1}{z} - \gamma_E + \mathcal{O}(z), \\ \Gamma(z) &= \sqrt{\pi} \left[1 + \left(z - \frac{1}{2}\right) \psi^{(0)}(1/2) + \mathcal{O}\left(\left(z - \frac{1}{2}\right)^2\right) \right],\end{aligned}\tag{A.7}$$

where $\gamma_E = 0.5772\dots$ is the Euler-Mascheroni constant, while $\psi^{(0)}(1/2) = -1.96351\dots$ and $\psi^{(m)}(z)$ are the polygamma functions.

Popular textbooks on quantum field theory

- C. Itzykson and J.B. Zuber, *Quantum field theory*, McGraw-Hill Inc., 1980
- M.E. Peskin and D.V. Schroeder, *An introduction to quantum field theory*, Westview Press, 1995
- S. Weinberg, *The quantum theory of fields*, vols. I and II, Cambridge University Press, 1995
- L.S. Brown, *Quantum field theory*, Cambridge University Press, 1992
- J.C. Collins, *Renormalization*, Cambridge University Press, 1984.
- G. 't Hooft and M. Veltman, *Diagrammar*
- M. Veltman, *Diagrammatica. The path to Feynman diagrams*, Cambridge University Press 1984

Research papers

Dimensional regularization

- G.t Hooft and M.Veltman, Regularization and renormalization of gauge fields, Nucl. Phys. B 44 (1972) 189
- C.G. Bollini and J.J. Giambiagi, Lowest order divergent graphs in ν -dimensional space, Phys. Lett. B40, 566 (1972)
- G.M. Cicuta and E. Montaldi, Analytic renormalization via continuous space dimension, Lett. Nuovo Cimento 4 (1972) 329
- P. Breitenlohner and D. Maison, Analytic renormalization and the action principle, Commun. Math. Phys. 52 (1977) 11

— D. Anselmi, Weighted power counting and chiral dimensional regularization, Phys. Rev. D 89 (2014) 125024, 14A2 Renormalization.com and arXiv:1405.3110 [hep-th]

Batalin-Vilkovisky formalism

— I.A. Batalin and G.A. Vilkovisky, Gauge algebra and quantization, Phys. Lett. B 102 (1981) 27-31

— I.A. Batalin and G.A. Vilkovisky, Quantization of gauge theories with linearly dependent generators, Phys. Rev. D 28 (1983) 2567, Erratum-ibid. D 30 (1984) 508

— I.A. Batalin and G.A. Vilkovisky, Existence theorem for gauge algebra, J. Math. Phys. 26 (1985) 172

Renormalization with the Batalin-Vilkovisky formalism

— D. Anselmi, Removal of divergences with the Batalin-Vilkovisky formalism, Class. and Quantum Grav. 11 (1994) 2181-2204, 93A2 Renormalization.com and arXiv:hep-th/9309085

— D. Anselmi, More on the subtraction algorithm, Class. and Quantum Grav. 12 (1995) 319, 94A1 Renormalization.com and arXiv:hep-th/9407023

— see also S. Weinberg, *The quantum theory of fields*, vol. II, Cambridge University Press, 1995

Anomalies

— S.L. Adler, Axial-vector vertex in spinor electrodynamics, Phys. Rev. 177 (1969) 2426

— J.S. Bell and R. Jackiw, A PCAC puzzle: $\pi^0 \rightarrow \gamma\gamma$ in the σ -model, Nuovo Cimento A60 (1969) 47

— W.A. Bardeen, Anomalous Ward identities in spinor field theories, Phys. Rev. 184 (1969) 1848

— J. Wess and B. Zumino, Consequences of anomalous Ward identities, Phys. Lett. 37 B (1971) 95

— S.L. Adler, J.C. Collins and A. Duncan, Energy-momentum-tensor trace anomaly in spin 1/2 quantum electrodynamics, Phys. Rev. D15 (1977) 1712

— N.K. Nielsen, The energy-momentum tensor in a non-Abelian quark gluon theory, Nucl. Phys. B (1977) 212

- S.J. Hathrell, Trace anomalies and $\lambda\varphi^4$ theory in curved space, *Ann. Phys.* (NY) 139 (1982) 136
- S.J. Hathrell, Trace anomalies and QED in curved space, *Ann. Phys.* (NY) 142 (1982) 34
- M.D. Freeman, Renormalization of non-Abelian gauge theories in curved space-time, *Ann. Phys.* (NY) 153 (1984) 339

Ward identities and gauge independence

- J.C. Ward, An Identity in quantum electrodynamics, *Phys. Rev.* 78, (1950) 182
- Y. Takahashi, On the generalized Ward identity, *Nuovo Cimento*, 6 (1957) 371
- A.A. Slavnov, Ward identities in gauge theories, *Theor. Math. Phys.* 10 (1972) 99
- J.C. Taylor, Ward identities and charge renormalization of Yang-Mills field, *Nucl. Phys.* B33 (1971) 436
- D. Anselmi, Background field method, Batalin-Vilkovisky formalism and parametric completeness of renormalization, *Phys. Rev. D* 89 (2014) 045004, 13A3 *Renormalization.com* and arXiv:1311.2704 [hep-th]
- D. Anselmi, Ward identities and gauge independence in general chiral gauge theories, 15A1 *Renormalization.com* and arXiv:1501.06692 [hep-th]

Adler-Bardeen theorem

- S.L. Adler and W.A. Bardeen, Absence of higher order corrections in the anomalous axial vector divergence, *Phys. Rev.* 182 (1969) 1517
- S.L. Adler, Anomalies to all orders, in “*Fifty Years of Yang-Mills Theory*”, G. 't Hooft ed., World Scientific, Singapore, 2005, p. 187-228, and arXiv:hep-th/0405040
- A. Zee, Axial-vector anomalies and the scaling property of field theory, *Phys. Rev. Letters* 29 (1972) 1198
- J. Collins, *Renormalization*, Cambridge University Press, 1984, Chapter 13
- T. Marinucci and M. Tonin, Dimensional regularization and anomalies, *Nuovo Cimento A* 31 (1976) 381
- G. Costa, J. Julve, T. Marinucci and M. Tonin, Non-Abelian gauge theories and triangle anomalies, *Nuovo Cimento A* 38 (1977) 373

- C. Lucchesi, O. Piguet and K. Sibold, The Adler-Bardeen theorem for the axial $U(1)$ anomaly in a general non-Abelian gauge theory, *Int. J. Mod. Phys. A* 2 (1987) 385
- O. Piguet and S. Sorella, Adler-Bardeen theorem and vanishing of the gauge beta function, *Nucl.Phys. B* 395 (1993) 661 and [arXiv:hep-th/9302123](#)
- E. Witten, Global aspects of current algebra, *Nucl. Phys. B* 223 (1983) 422
- Wess, J. and B. Zumino, Consequences of anomalous Ward identities, *Phys. Lett. B* 37 (1971) 95
- D. Anselmi, Adler-Bardeen theorem and manifest anomaly cancellation to all orders in gauge theories, *Eur. Phys. J. C* 74 (2014) 3083, 14A1 Renormalization.com and [arXiv:1402.6453 \[hep-th\]](#)
- D. Anselmi, Adler-Bardeen theorem and cancellation of gauge anomalies to all orders in nonrenormalizable theories, 15A2 Renormalization.com and [arXiv:1501.07014 \[hep-th\]](#)

Index

- γ matrices, 68
- Analytic integral, 61
- Anomalous dimension, 131, 133, 134, 138, 144
- Bare action, 81, 90, 112
- bare action, 211
- Batalin-Vilkovisky formalism, 211
- Berezin integral, 47
- Beta function, 130, 138, 145, 149
- Combinatorial factor, 26
- Composite fields, 110
- Correlation functions, 8, 16, 22
- Coulomb gauge fixing, 155
- Counterterm, 69, 72, 75, 77, 80, 83, 87, 92, 95, 97, 113, 115, 118, 130, 198, 208, 214
- Cutoff, 57
- Dimensional regularization, 60, 68, 119, 154, 191, 201
- Domain of convergence, 62
- Feynman parameters, 75
- Feynman rules, 22, 23, 25, 52, 197
- Free field theory, 17
- Free particle, 14
- Free particle kernel, 15
- Functional Z , 15, 21
- Functional equation for Γ , 33
- Functional equation for W , 32
- Functional equation for Z , 31
- Functional integral, 7
- Functional integration measure, Theorem 1, 66
- Gauge fermion, 185
- Gauge independence, Theorem 15, 195
- Gaussian integral, 8
- Generators of representations, Corollary 6, 165
- Grassmann variables, 46
- Green function, 16
- Invariance of Z and W , Theorem 10, 192
- Invariance of correlation functions, Theorem 13, 195
- Invariance of physical quantities, Theorem 14, 195
- Irreducible representations, Theorem 5, 164
- Irreducible representations, Theorem 4, 164
- Locality of counterterms, 73, 74, 92, 118
- Locality of counterterms, Theorem 2, 93

- Logarithmic divergence, 63
- Lorenz gauge fixing, 155
- Master equation, 175, 193
- master equation, 211
- Master identity, 193
- Master identity, Theorem 12, 193
- Maximum poles of diagrams, Theorem 3, 117
- Minimal subtraction scheme, 81, 118, 145, 147, 202, 206
- Most general Yang-Mills theory, Theorem 7, 183
- Nonrenormalizability, 78
- Overall divergence, 71
- Path integral, 12
- Perturbative expansion, 19, 21
- Perturbative locality, 67
- Poles, 63
- Power counting, 99
- Powerlike divergence, 63
- Proposition 16, 214
- Proposition 17, 214
- Proposition 18, 214
- Regularization, 58, 59
- Renormalizability, 78
- Renormalization, 58
- Renormalization constants, 81, 83, 88–90
- Renormalized action, 81, 91, 105, 112, 115
- Scheme change, 122
- Schrödinger equation, 13
- Subdivergence, 71
- Subtraction scheme, 81, 118, 119
- Superdeterminant, 49
- Tadpole, 25, 59, 80
- Theorem 1, Functional integration measure, 66
- Theorem 11, 193
- Theorem 12, 193
- Theorem 15, Gauge independence, 195
- Theorem 13, Invariance of correlation functions, 195
- Theorem 7, Most general Yang-Mills theory, 183
- Theorem 2, Locality of counterterms, 93
- Theorem 12, Master identity, 193
- Theorem 14, Invariance of physical quantities, 195
- Theorem 3, Maximum poles of diagrams, 117
- Theorem 10, Invariance of Z and W , 192
- Theorem 5, Irreducible representations, 164
- Theorem 4, Irreducible representations, 164
- Ultralocality, 66
- Vertex, 22, 23
- Ward identities, 201, 203
- Wick's theorem, 17, 51