Introduction to Quantum Field Theory

(student handout version)

Philip G. Ratcliffe
(philip.ratcliffe@uninsubria.it)

Dipartimento di Scienza e Alta Tecnologia
Università degli Studi dell'Insubria in Como
via Valleggio 11, 22100 Como (CO), Italy

(last revised 7th March 2016)

Preface

The present is a written version of lecture notes for the course on *Quantum Field Theory*, held within the framework of a Master's equivalent degree course in Physics at Insubria University in Como. The lectures were first delivered in the academic year 2003/04 and have since undergone some evolution. The notes have been augmented and edited with the aim of being self-contained and therefore of more general utility. They are thus primarily intended for use by students with a basic knowledge of classical electromagnetism and special relativity in addition to a good grounding in quantum mechanics (including the relativistic formulation due to Dirac). However, this volume should also hopefully represent a useful reference text and study aid.

Contents

Chapter 1

Introduction

1.1 Suggested supplementary reading

A (very short) list of suggested supplementary reading material follows. The book by Bailin and Love provides a very clear introduction to the Feynman path-integral formalism and covers much of the material to be presented, as does the book, by Ryder, which also contains a good deal of useful supplementary material. The remaining two volumes (Itzykson and Zuber; Bjorken and Drell) are intended as complementary reading in general. Finally, the Physics Report by Abers and Lee represents one of the first reviews of the path-integral formalism as applied to field quantisation in the standard model of elementary particle physics. Indeed, although intended for a more advanced readership, it covers such aspects as spontaneous symmetry breaking, a fundamental ingredient to the standard model, and remains a thorough and clear treatment of many of the topics covered in this course.

Reading list

Abers, E.S. and Lee, B.W. (1973), *Phys. Rep.* **9**, 1.

Bailin, D. and Love, A. (1993), *Introduction to Gauge Field Theory* (IOP Pub.), revised edition.

Bjorken, J.D. and Drell, S.D. (1965), *Relativistic Quantum Fields* (McGraw–Hill).

Itzykson, C. and Zuber, J.-B. (1980), *Quantum Field Theory* (McGraw–Hill).

Ryder, L.H. (1996), *Quantum Field Theory* (Cambridge U. Press), 2nd. edition.

1.2 Aims and philosophy

The main aim of this course is to introduce the student to the subject of *second* or *field* quantisation. As such, it is a natural continuation of the third and final part of the course on quantum mechanics. Indeed, it continues the studies started in that course and, in particular, the question of a relativistic formulation.

The path-integral formalism developed by Feynman (1950) is one of the most powerful tools in quantum theory; indeed, is has now become indispensable to almost every theoretical physicist dealing with the quantum world. While this approach has already been presented in the previous course on quantum mechanics and thus some familiarity is assumed, the method is derived here in detail from first principles. It is then applied to the quantisation of field theory, where the simplifications and gains in insight with respect to the canonical approach are priceless.

An important and pervasive "theme" in this course will be the implications and importance of symmetries in quantum field theory generally, as embodied in Noether's theorem (1918).

In the final section, having covered the main topics of field quantisation, such as the derivation of the Feynman rules and the question of renormalisation, we turn briefly to the question of spontaneous symmetry breaking and the generation of mass in the so-called standard model of elementary particle physics.

1.3 Conventions and notation

We shall indicate defining equations with the notation ":=", whereby the right-hand side defines the left-hand side, while the notation "$\widehat{=}$" will be used to indicate equivalence though not necessarily strict equality (such as between operators).

The momentum-space measure naturally conjugate to $\mathrm{d}x$ is $\mathrm{d}p/(2\pi)$ and in many cases the Dirac δ-function for momenta is also accompanied by a 2π factor; thus $2\pi\,\delta(p)$. Since most of such 2π factors generally simplify in the final expressions, for clarity of the intermediate steps by avoiding such inessential factors, we shall adopt the following "bar" notation (in analogy with the standard $\hbar$ and common $\lambda\!\!\!^{-}$):

$$đ := \frac{\mathrm{d}}{2\pi} \text{ and } \bar{\delta} := 2\pi\,\delta(p). \tag{1.3.1}$$

In particular, one then has

$$\int đp'\,\bar{\delta}(p'-p) = 1. \tag{1.3.2}$$

For clarity of notation, we shall generally adopt the universal "natural" units

of the high-energy physicist, in which $\hbar$ and c are set to unity and therefore disappear from all expressions. However, it will occasionally be useful, in order to fully appreciate the quantum or relativistic nature of a given situation to render the dependence on these two parameters explicit; in such cases $\hbar$ and c will be temporarily reinstated. Indeed, this restoration may always be achieved by noting the dimensions of the object under study: with the natural system all quantities have dimensions of powers of energy (say) and multiplication by suitable powers of $\hbar$ and c will regain the true physical dimensions.

As a final question of notation, to aid with identification of indexed objects, we shall attempt to adhere as strictly as possible to the following convention: indices from the middle of the Greek alphabet μ, ν, ρ, σ, λ will be used to indicate Lorentz four-vector indices (and thus run from 0 to 3); Latin letters i, j, k, l will usually stand for purely spatial indices (running from 1 to 3); letters from the beginning of the Greek alphabet α, β, γ will be used as Dirac-space indices (*i.e.* in Dirac spinors or gamma-matrix elements); and from the beginning of the Latin alphabet a, b, c will typically be used for internal spaces of various types. Unindexed four vectors will be printed in a normal math-italic font while unindexed three vectors will appear in bold face.

1.4 Bibliography

Feynman, R.P. (1950), *Phys. Rev.* **80**, 440.

Noether, E. (1918), *Nachr. v. d. Ges. d. Wiss. zu Göttingen*, 235; *transl.*, *Transport Theory and Statistical Mechanics* 183.

Chapter 2

Classical Field Theory

As a preliminary to the study of quantised fields, in this chapter we shall briefly review the principal formalism and results of classical field theory. It will be assumed, however, that the reader has some prior knowledge of the subject. We shall naturally immediately resort to a relativistic formulation, with which too a certain familiarity is assumed.

The concept of a field is the natural extension of the description of a single (point-like) particle, through many particles, to a continuum. It is, for example, necessary for the description of sound waves in a continuous medium. Classical electromagnetism is a further example: the electric and magnetic fields are described via a four-vector potential $A^\mu(t,\boldsymbol{x})$, which is defined at all space–time points $(t,\boldsymbol{x})$. A field is thus necessary for the description of any physical quantity that is defined continuously over space–time.

2.1 The Lagrangian and the action

The most natural starting point for field quantisation is the Lagrangian formulation. The Lagrangian *function* L is then defined as the continuum limit of the discrete case: for a system of n particles we would have

$$L(t) := \sum_{i=1}^{n} L_i(t), \tag{2.1.1}$$

where L_i is the Lagrangian corresponding to the i-th particle. For a field, we can imagine dividing all of three-dimensional space into n sub-volumes and then taking the limit $n \to \infty$. We are thus naturally led to

$$L(t) := \int \mathrm{d}^3\boldsymbol{x}\ \mathcal{L}(t,\boldsymbol{x}), \tag{2.1.2}$$

where $\mathcal{L}$ is the Lagrangian density.* For a single degree of freedom, the Lagrangian is a function of the generalised position and velocity, $q(t,\boldsymbol{x})$ and $\dot{q}(t,\boldsymbol{x}) := \partial q(t,\boldsymbol{x})/\partial t$. It is important to bear in mind the difference between the variable q (which might be, *e.g.*, a displacement with respect to an equilibrium position, a pressure, or a component of the electric field at a given point) and the three-dimensional space indicated here by $\boldsymbol{x}$. On moving over to the continuum, we should also consider the local gradients $\boldsymbol{\nabla} q_i$ of the generalised position variables. In other words, introducing the field φ (which may again be thought of as the limit of the mean value in a small sub-volume centred around the point $\boldsymbol{x}_i$), we expect to have

$$\mathcal{L} \;=\; \mathcal{L}(\varphi, \partial_\mu \varphi). \tag{2.1.3}$$

Note that we have already anticipated a relativistic formulation by combining the time derivative and gradient into the covariant form ∂_μ. We also recall that†

$$\partial^\mu \;=\; (\partial/\partial t, -\boldsymbol{\nabla}). \tag{2.1.4}$$

Finally, we introduce the action:

$$S[\varphi] \;:=\; \int_{t_i}^{t_f} \mathrm{d}t\, L(t) \;=\; \int_{t_i}^{t_f} \mathrm{d}^4x\; \mathcal{L}(\varphi, \partial_\mu \varphi). \tag{2.1.5}$$

This is what we shall call a *functional* of the field φ. It may always be thought of in terms of the continuum limit of a discrete but infinite set of degrees of freedom. This will be, in fact, the working definition of a functional. Initial states are normally assumed prepared in the remote past while measurements on final states are made in the remote future. Consequently, the upper and lower time limits may be set to $\pm\infty$, whence the previous expression becomes explicitly Lorentz covariant. It is perhaps worth remarking that the requirement of Lorentz covariance induces quite stringent relations between the temporal and spatial dependence, in terms of the derivatives appearing in Eq. (2.1.5). Note finally that the action $S[\varphi]$ is dimensionless; to remain so in the physical theory (in the natural system of units) it should be divided by $\hbar$.

2.2 Stationarity and the Euler–Lagrange equations

The Lagrange equations of motion are obtained by requiring that the action be *extremal* or *stationary* under variations of the trajectory $q(t,\boldsymbol{x})$. Similarly, we may

* We shall usually use a script letter to indicate a density and distinguish it from the corresponding integrated quantity.

† The reader who might wish to reinstate c should note that $x_0 \equiv ct$.

study variations of the field $\varphi(t,\boldsymbol{x})$. Consider thus an infinitesimal variation:

$$\varphi \;\to\; \varphi + \delta\varphi, \tag{2.2.1a}$$

under which

$$S \;\to\; S + \delta S. \tag{2.2.1b}$$

We then have

$$\begin{aligned}\delta S \;&=\; \int_{t_i}^{t_f} \mathrm{d}^4x\, \left[\frac{\partial\mathcal{L}}{\partial\varphi}\,\delta\varphi + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi)}\,\delta(\partial_\mu\varphi)\right] \\ &=\; \int_{t_i}^{t_f} \mathrm{d}^4x\, \left[\frac{\partial\mathcal{L}}{\partial\varphi} - \partial_\mu\!\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi)}\right)\right]\delta\varphi + \int_{t_i}^{t_f} \mathrm{d}^4x\,\partial_\mu\!\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi)}\,\delta\varphi\right),\end{aligned} \tag{2.2.2}$$

where we have exploited the commutativity of δ and ∂_μ (which is easily demonstrated starting from the discrete formulation) and have performed an integration by parts on the second term in square brackets of the first line. The last term on the right-hand side of the second line may be integrated immediately and its contribution is given by the expression in brackets evaluated at the three-surface of the four-volume of integration. If we now take as boundary conditions the vanishing of the variation $\delta\varphi$ on this surface (just as we take δq to vanish for $t = t_{i,f}$ in deriving the standard Lagrange equations), then this last term vanishes. We thus obtain the functional derivative:

$$\frac{\delta S}{\delta\varphi(x)} \;=\; \frac{\partial\mathcal{L}}{\partial\varphi} - \partial_\mu\!\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi)}\right). \tag{2.2.3}$$

The stationarity requirement

$$\frac{\delta S}{\delta\varphi(x)} \;=\; 0 \tag{2.2.4}$$

then leads to the Euler–Lagrange equations:

$$\frac{\partial\mathcal{L}}{\partial\varphi} \;=\; \partial_\mu\!\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi)}\right). \tag{2.2.5}$$

The extension to a multi-valued field φ^a is straightforward, there being one Euler–Lagrange equation associated with each field.

Now, while it is not a trivial matter to construct the Lagrangian density required to derive arbitrary given field equations, with a little "trial and error" the process is not overly difficult. Thus, for example, the standard scalar wave equation

$$\Box\,\varphi \;:=\; \partial^\mu\partial_\mu\varphi \;=\; 0 \tag{2.2.6}$$

may be obtained from the following Lagrangian density:

$$\mathcal{L}(\varphi, \partial_\mu \varphi) \;=\; \tfrac{1}{2}\rho\,(\partial^\mu \varphi)(\partial_\mu \varphi\,), \tag{2.2.7}$$

where ρ is a mass density (*e.g.* of a vibrating medium).

Exercise 2.2.1. *Derive the above scalar wave equation from the preceding Lagrangian density.*

Exercise 2.2.2. *Derive the Maxwell equations for an electromagnetic field,*

$$\partial_\mu F^{\mu\nu} \;=\; j^\nu, \tag{2.2.8}$$

from the following Lagrangian density:

$$\mathcal{L}(A^\mu, \partial_\nu A^\mu) \;=\; -\tfrac{1}{4}\,F^{\mu\nu}\,F_{\mu\nu} - j^\mu A_\mu, \tag{2.2.9}$$

with

$$F^{\mu\nu} \;:=\; \partial^\mu A^\nu - \partial^\nu A^\mu. \tag{2.2.10}$$

2.3 Noether's theorem

Besides providing a highly intuitive approach to second (or field) quantisation, we shall see that a particulary important aspect of the path integral is its very direct relation to the Lagrangian, which renders the symmetries of the physical situation under study most transparent. This in turn allows conserved quantities to be readily identified. The connection between the concepts of symmetry and conservation is supplied by Noether's theorem (1918), which states:

Theorem 2.3.1. *Corresponding to every continuous symmetry of the physical world, there exists a conserved quantity; and in correspondence to every conserved quantity, there exists a continuous symmetry.*

Consider, for example, a general infinitesimal coordinate transformation

defined by

$$x^\mu \;\to\; x'^\mu \;=\; x^\mu + \delta x^\mu, \tag{2.3.1a}$$

$$\delta x^\mu \;=\; X_i^\mu(x)\,\delta\omega^i, \tag{2.3.1b}$$

where the $\delta\omega^i$ form a set of infinitesimal parameters and the X_i^μ characterise the given type of transformation. In general, such a transformation will also be associated with transformations of the fields: first of all those *induced* by the variations in their position arguments and secondly by possible direct, *intrinsic*, variations in the fields themselves. We shall thus have

$$\varphi^a(x) \;\to\; \varphi'^a(x') \;=\; \varphi^a(x) + \delta\varphi^a(x), \tag{2.3.2}$$

where the overall field variations will be parametrised by the same infinitesimal $\delta\omega^i$ but characterised by some new quantities $\Phi_i^a(x)$:

$$\delta\varphi^a(x) \;=\; \Phi_i^a(x)\,\delta\omega^i. \tag{2.3.3}$$

Let us then examine the case in which the action

$$S \;=\; \int_R \mathrm{d}^4x\; \mathcal{L}\big(\varphi^a(x), \partial_\mu\varphi^a(x), x\big)\,, \tag{2.3.4}$$

is invariant under such a transformation, for an integration region R *any* compact sub-space of Minkowski space–time. Note that we also explicitly allow for other possible space–time dependence in $\mathcal{L}$ since, for example, there may be other fields present.

First of all, let us separate out the *intrinsic* field variation $\hat{\delta}\varphi^a$ from the *induced* variation, we may thus write (to first order in the variations)

$$\varphi'^a(x') \;=\; \varphi^a(x) + \hat{\delta}\varphi^a(x) + \partial_\mu\varphi^a(x)\,\delta x^\mu. \tag{2.3.5}$$

The intrinsic field variation is then

$$\begin{aligned}\hat{\delta}\varphi^a(x) \;&=\; \delta\varphi^a(x) - \partial_\mu\varphi^a(x)\,\delta x^\mu \\ &=\; \big[\Phi_i^a(x) - X_i^\mu(x)\,\partial_\mu\varphi^a(x)\big]\,\delta\omega^i.\end{aligned} \tag{2.3.6}$$

Similarly, the variation of the action S receives different contributions: firstly from the variations considered above and secondly from the variation of the four-space integration measure. We thus have

$$\delta S \;=\; \int_R \mathrm{d}^4x\,\delta\,\mathcal{L}(\varphi^a, \partial_\mu\varphi^a, x) + \int_R \delta(\mathrm{d}^4x)\;\mathcal{L}(\varphi^a, \partial_\mu\varphi^a, x). \tag{2.3.7}$$

The second contribution is determined by the Jacobian of the transformation:

$$\begin{aligned}J \;&=\; \det\left[\frac{\partial x'^\mu}{\partial x^\nu}\right] \\ &=\; \det\big[g^\mu{}_\nu + \partial_\nu X_i^\mu(x)\delta\omega^i\big] \\ &=\; \big[1 + \partial_\mu X_i^\mu(x)\delta\omega^i\big]\,.\end{aligned} \tag{2.3.8}$$

Thus,

$$\delta(\mathrm{d}^4x) \;=\; \partial_\mu X_i^\mu\delta\omega^i\,\mathrm{d}^4x\,. \tag{2.3.9}$$

The first contribution, however, varies due to both the coordinate and field trans-

formations, thus

$$\delta\,\mathcal{L}\big(\varphi^a,\partial_\mu\varphi^a,x\big) \;=\; \frac{\partial\mathcal{L}}{\partial\varphi^a}\,\hat{\delta}\varphi^a + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi^a)}\,\hat{\delta}(\partial_\mu\varphi^a) + (\partial_\mu\,\mathcal{L})\,\delta x^\mu. \tag{2.3.10a}$$

If the fields φ^a represent a stationary solution, then they must satisfy the Euler–Lagrange equations (2.2.5) and we thus have (also substituting for δx^μ)

$$\begin{aligned}\delta\,\mathcal{L}\big(\varphi^a,\partial_\mu\varphi^a,x\big) \;&=\; \partial_\mu\!\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi^a)}\right)\hat{\delta}\varphi^a + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi^a)}\,\hat{\delta}(\partial_\mu\varphi^a) + (\partial_\mu\,\mathcal{L})\,X_i^\mu(x)\,\delta\omega^i \\ &=\; \partial_\mu\!\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi^a)}\,\hat{\delta}\varphi^a\right) + (\partial_\mu\,\mathcal{L})\,X_i^\mu(x)\,\delta\omega^i. \end{aligned}\tag{2.3.10b}$$

Inserting this expression into Eq. (2.3.7) and combining the second term here with the second there, via Eq. (2.3.9), we obtain

$$\delta S \;=\; \int_R \mathrm{d}^4x\;\partial_\mu\!\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi^a)}\,\hat{\delta}\varphi^a + X_i^\mu\delta\omega^i\;\mathcal{L}\right), \tag{2.3.11}$$

Substitution for $\hat{\delta}\varphi^a$ finally leads to

$$\delta S \;=\; \int_R \mathrm{d}^4x\;\partial_\mu\left(j_i^\mu\delta\omega^i\right), \tag{2.3.12}$$

where we have thus defined the following four-current:

$$j_i^\mu \;:=\; \left(g^\mu{}_\nu\,\mathcal{L} - \frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi^a)}\,\partial_\nu\varphi^a\right)X_i^\nu + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi^a)}\,\Phi_i^a. \tag{2.3.13}$$

Now, since our premise is that Eq. (2.3.12) should hold for any (compact) four-space volume, the vanishing of δS implies that the integrand vanishes. Finally, since $\delta\omega^i$ is an arbitrary set of infinitesimal parameters, we deduce that

$$\partial_\mu j_i^\mu \;=\; 0. \tag{2.3.14}$$

In other words, the current (2.3.13) is conserved. This is just Noether's theorem, as originally formulated within the context of classical field theory.

Exercise 2.3.1. *Show that the total charge*

$$Q_i(t) \;:=\; \int \mathrm{d}^3\boldsymbol{x}\,j_0(t,\boldsymbol{x}) \tag{2.3.15}$$

is constant in time.
Hint: *A necessary assumption is, of course, that the current vanish everywhere at the surface of the volume considered.*

Note in conclusion that the current so-defined is not unique. Its form depends on the use of the Euler–Lagrange equations; we may therefore add any divergence-less quantity, such as $\partial_\mu \Lambda^{\mu\nu}(x)$ with $\Lambda^{\mu\nu}(x)$ antisymmetric in the indices $\mu\nu$.

2.3.1 Internal symmetries

A natural and important extension is to the case of a continuous internal symmetry, to which Noether's theorem also applies. Now, the transformation

$$\delta\varphi \;=\; -\mathrm{i}\,\varepsilon q \varphi \tag{2.3.16}$$

relates to a simple U(1) symmetry group while a more general example is that for which the transformation is

$$\delta\varphi_i \;=\; -\mathrm{i}\,\varepsilon q_i \varphi_i, \tag{2.3.17}$$

with $[q_i, q_j] \neq 0$ for $i \neq j$. This is known as a *non-Abelian* group.

Consider, for example, SU(2) or SO(3), which are locally isomorphic. We then have

$$\varphi \;\to\; \mathrm{e}^{-\mathrm{i}\boldsymbol{L}\cdot\boldsymbol{\theta}}\;\varphi, \tag{2.3.18}$$

where now φ is a multi-component object (*e.g.* a spinor) and $\boldsymbol{L}$ is a matrix in a representation of, say, SU(2). For example, $\boldsymbol{L} \equiv \frac{1}{2}\boldsymbol{\sigma}$, the Pauli spin (or isospin) matrices. The algebra

$$\Big[L^a, L^b\Big] \;=\; \mathrm{i}\,\varepsilon^{abc}\,L^c \tag{2.3.19}$$

is that of the generators of SU(2).

The *global* symmetry of isospin (or rotations in $\mathbb{R}^3$) requires then that

$$\delta\,\mathcal{L} \;=\; 0 \quad \text{for} \quad \delta\varphi_i \;=\; -\mathrm{i}\,\delta\omega^a L^a_{ij}\,\varphi_j, \tag{2.3.20}$$

with $\delta\omega^a =$ constant. More in general, for the generators of the transformation T^a we shall have a Lie algebra

$$\Big[T^a, T^b\Big] \;=\; \mathrm{i}\,f^{abc}\,T^c. \tag{2.3.21}$$

2.3.2 Scalar electrodynamics

Electrodynamics possesses a larger symmetry, a *local* gauge symmetry. The local infinitesimal gauge transformation is

$$\delta\varphi_i(x) \;=\; -\mathrm{i}\,q_i\,\delta\omega(x)\,\varphi_i(x). \tag{2.3.22}$$

That is, the transformation may now depend on the space–time point. We then also have,

$$\delta\partial^\mu\varphi_i(x) \;=\; -\mathrm{i}\,q_i\,\delta\omega(x)\,\big(\partial^\mu\varphi_i(x)\big) - \mathrm{i}\,q_i\,\big(\partial^\mu\delta\omega(x)\big)\,\varphi_i(x). \tag{2.3.23}$$

The second, inhomogeneous, term would break the original invariance. We thus introduce the covariant derivative

$$D^\mu \;:=\; \partial^\mu - \mathrm{i}eq_iA^\mu(x), \tag{2.3.24}$$

where $A^\mu(x)$ is a spin-one, massless field (the so-called gauge field). This field should transform in such a way as to cancel the inhomogeneous term:

$$\delta\big(-\mathrm{i}eq_iA^\mu(x)\,\varphi_i(x)\big) \;=\; -\mathrm{i}eq_iA^\mu(x)\,\delta\varphi_i(x) - \mathrm{i}eq_i\delta A^\mu(x)\,\varphi_i(x), \tag{2.3.25}$$

with then

$$\delta A^\mu(x) \;=\; -\frac{1}{e}\,\partial^\mu\big(\delta\omega(x)\big)\,. \tag{2.3.26}$$

We may construct the corresponding field-strength tensor

$$F^{\mu\nu} \;:=\; \partial_\mu A_\nu - \partial_\nu A_\mu, \tag{2.3.27a}$$

with

$$\delta F^{\mu\nu} \;=\; 0. \tag{2.3.27b}$$

And thus the Lagrangian density

$$\mathcal{L} \;=\; -\tfrac{1}{4}F^{\mu\nu}F_{\mu\nu} \tag{2.3.28}$$

is naturally invariant. Note that an explicit mass term $-\frac{1}{2}m^2A^\mu A_\mu$ would necessarily break the gauge invariance. In other words, the masslessness of the photon and gauge invariance are intimately related.

In the non-Abelian case we have

$$\varphi(x) \;\to\; \varphi'(x) \;=\; \mathrm{e}^{-\mathrm{i}\boldsymbol{L}\cdot\boldsymbol{\theta}(x)}\;\varphi(x) \;=:\; U(\theta)\,\varphi(x), \tag{2.3.29}$$

with

$$\partial^\mu\varphi(x) \;\to\; U(\theta)\,\big(\partial^\mu\varphi(x)\big) + \big(\partial^\mu U(\theta)\big)\,\varphi(x). \tag{2.3.30}$$

As before, we introduce a covariant derivative, such that

$$D^\mu\varphi(x) \;\to\; U(\theta)\, D^\mu\varphi(x) \tag{2.3.31}$$

and $D^\mu\varphi^\dagger(x)D_\mu\varphi(x)$ therefore remains invariant. Here

$$D^\mu\varphi(x) \;:=\; (\partial^\mu - \mathrm{i}\, g\boldsymbol{L}\cdot\boldsymbol{A}^\mu(x))\,\varphi(x), \tag{2.3.32}$$

where $\boldsymbol{L}$ and $\boldsymbol{A}^\mu$ are vectors in the internal symmetry space and we write the gauge field as, *e.g.* $A^{a\mu}$.

Exercise 2.3.2. *Show that the requirement*

$$D^\mu\varphi(x) \;\to\; U(\theta)\, D^\mu\varphi(x) \tag{2.3.33}$$

leads to the transformation law

$$\boldsymbol{L}.\boldsymbol{A}'^\mu \;=\; U(\theta)\,\boldsymbol{L}\cdot\boldsymbol{A}^\mu\, U^{-1}(\theta) - \mathrm{i}\, g^{-1}\left(\partial^\mu U(\theta)\right) U^{-1}(\theta). \tag{2.3.34}$$

Exercise 2.3.3. *Show that the transformations so defined form a group. That is, that there exists a well-defined product with an identity element and an inverse.*

It is important to appreciate that the transformations themselves do *not* depend on the representation of the group. Consider the following infinitesimal transformation:

$$\begin{aligned} L^a\,\delta A^{a\mu} \;&=\; -g^{-1}L^a\,\partial^\mu\theta^a + \mathrm{i}\,L^b A^{b\mu}\,L^a\theta^a - \mathrm{i}\,L^a\theta^a\,L^b A^{b\mu} \\ &=\; -g^{-1}L^a\,\partial^\mu\theta^a + \mathrm{i}\,\theta^a A^{b\mu}\left[L^b, L^a\right] \\ &=\; -g^{-1}L^a\,\partial^\mu\theta^a - \theta^a A^{b\mu}\, f^{bac}L^c, \end{aligned} \tag{2.3.35}$$

where, recall, the coefficients f^{abc} are the *structure constants* defining the group. Now, since the L^a are linearly independent, we have

$$\delta A^{a\mu} \;=\; -g^{-1}\partial^\mu\theta^a + f^{abc}\theta^b A^{c\mu}, \tag{2.3.36}$$

which evidently only depends on the f^{abc}.

Exercise 2.3.4. *Show that* $L^a F^{a\mu\nu} := [L^b D^{b\mu}, L^c D^{c\nu}]$ *leads to the gauge invariant Lagrangian density*

$$\mathcal{L} \;=\; -\tfrac{1}{4}F^a_{\mu\nu}F^{a\mu\nu}. \tag{2.3.37}$$

Derive the explicit form of $F^a_{\mu\nu}$.

Finally, we see that the minimal coupling rule of quantum electrodynamics (QED) generalises to

$$\partial^\mu \;\to\; D^\mu \;:=\; \partial^\mu - \mathrm{i}gL^a A^{a\mu} \tag{2.3.38}$$

and, providing we define $F^{a\mu\nu}$ through the covariant derivative as above, the ensuing construction of the gauge Lagrangian density follows the same path. Note again that a mass term of the form $\frac{1}{2}m^2 A^{a\mu}A^a_\mu$ would break the gauge invariance.

2.4 Classical fields

We now briefly review the different types of fields that we shall later encounter: scalar, spinor and vector (both massive and massless).

2.4.1 Free massive scalar field

A single, free, massive, scalar field is governed by the wave equation

$$\left[\Box + m^2\right]\varphi(x) \;=\; 0. \tag{2.4.1}$$

where m is a constant corresponding to the particle mass. Indeed, if we consider a particle of energy E and momentum $\boldsymbol{p}$ then the solution of this wave equation requires

$$-E^2 + |\boldsymbol{p}|^2 + m^2 \;=\; 0. \tag{2.4.2}$$

The Lagrangian leading to such an equation of motion is just

$$\mathcal{L} \;=\; \tfrac{1}{2}\left[(\partial^\mu\varphi)(\partial_\mu\varphi) - m^2\varphi^2\right]. \tag{2.4.3}$$

We can add interactions to this description by adding further terms into the Lagrangian density; these will generate source (self-interaction) terms in the equation of motion. A term $V(\varphi) = \frac{1}{n!}\lambda\varphi^n$ in the Lagrangian leads to

$$\left[\Box + m^2\right]\varphi(x) \;=\; -\tfrac{1}{(n-1)!}\,\lambda\,\varphi^{n-1}(x). \tag{2.4.4}$$

There are two common choices for n: $n=3$, which leads to a three-body self-interaction, similar to the photon–fermion interaction in QED but which also evidently corresponds to a potential that is unbounded from below and which therefore corresponds to an unstable situation; and $n=4$, which leads to a four-body self-interaction but which instead provides a stable potential.

In the presence of interactions the definition of the mass of a particle requires

more care. In general, for

$$\mathcal{L} \;=\; \tfrac{1}{2}\big[(\partial^\mu\varphi)(\partial_\mu\varphi) - m^2\varphi^2\big] - V(\varphi), \tag{2.4.5}$$

the minimum of $V(\varphi)$, corresponding to the state of lowest energy, may no longer lie at the origin, but at some point $\varphi=\varphi_0\neq 0$, say, and so we must take

$$m^2 \;:=\; \left.\frac{\partial^2 V(\varphi)}{\partial\varphi^2}\right|_{\varphi=\varphi_0}, \quad \text{with} \quad \left.\frac{\partial V(\varphi)}{\partial\varphi}\right|_{\varphi=\varphi_0} = 0. \tag{2.4.6}$$

The descriptions "scalar", "spinor", "vector" *etc.* reflect the behaviour of the fields under general Poincaré transformations. In particular, a scalar field is invariant. Consider then an infinitesimal space–time translation:

$$x^\mu \;\to\; x'^\mu \;=\; x^\mu + \varepsilon^\mu. \tag{2.4.7}$$

This corresponds to Eq. (2.3.1b) with the transformation characterised by the tensor $X^\mu_\nu = g^\mu{}_\nu$, $\Phi_i = 0$ and $\delta\omega^\mu = \varepsilon^\mu$.

Exercise 2.4.1. *Show that the conserved current in this case is the so-called* energy–momentum tensor

$$\begin{aligned}\Theta^{\mu\nu} \;&=\; \frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi)}\,(\partial^\nu\varphi) - \mathcal{L}\,g^{\mu\nu} \\ &=\; (\partial^\mu\varphi)(\partial^\nu\varphi) - \big[\tfrac{1}{2}(\partial^\sigma\varphi)(\partial_\sigma\varphi) - \tfrac{1}{2}m^2\varphi^2 - V(\varphi)\big]\,g^{\mu\nu}.\end{aligned} \tag{2.4.8}$$

Show also that in the free-field case, $V(\varphi)=0$, the conserved "charge" or energy-momentum vector is

$$\begin{aligned}P^\mu \;&:=\; \int \mathrm{d}^3\boldsymbol{x}\,\Theta^{0\mu} \\ &=\; \int \mathrm{d}^3\boldsymbol{x}\,\big[(\partial^0\varphi)(\partial^\mu\varphi) - \tfrac{1}{2}g^{\mu 0}(\partial_0\varphi\,\partial^0\varphi + \varphi\,\partial_0\partial^0\varphi)\big]\end{aligned} \tag{2.4.9}$$

Hint: *Use the equations of motion to eliminate the term $-\frac{1}{2}g^{\mu 0}m^2\varphi^2$ and subtract the total spatial divergence $\frac{1}{2}g^{\mu 0}\nabla_r[\varphi\nabla^r\varphi]$.*[*]

Since φ is a real (classical) field, it may be expressed in terms of Fourier com-

[*] Note that in the book by Bailin and Love there is a mistake in both the derivation and the final expression here.

ponents (recall that both k and x are four-vectors here):

$$\varphi(x) \;=\; \int \frac{\mathrm{đ}^3\boldsymbol{k}}{2k_0}\left[b(k)\;\mathrm{e}^{-\mathrm{i}k\cdot x}+b^*(k)\;\mathrm{e}^{+\mathrm{i}k\cdot x}\right], \tag{2.4.10}$$

with $k_0 = +\left(\boldsymbol{k}^2+m^2\right)^{1/2}$. The factor $(2k_0)^{-1}$ guarantees Lorentz covariance of the measure. Inserting this into the expression for P^μ in Eq. (2.4.9), we find

$$P^\mu \;=\; \int \frac{\mathrm{đ}^3\boldsymbol{k}}{2k_0}\, b(k)\, b^*(k)\, k^\mu, \tag{2.4.11}$$

which is explicitly time-independent. In this decomposition each mode k^μ has probability amplitude $b(k)$ and the number of modes in the interval $\mathrm{đ}^3\boldsymbol{k}$ around the point $\boldsymbol{k}$ is

$$\frac{\mathrm{đ}^3\boldsymbol{k}}{2k_0}\,|b(k)|^2. \tag{2.4.12}$$

The energy is then

$$\begin{aligned} P^0 \;&=\; \int \mathrm{d}^3\boldsymbol{x}\,\left[\frac{\partial\mathcal{L}}{\partial(\partial^0\varphi)}\,(\partial^0\varphi)-\mathcal{L}\right] \\ &=\; \int \mathrm{d}^3\boldsymbol{x}\,\left[\tfrac{1}{2}\pi^2(x)+\tfrac{1}{2}(\boldsymbol{\nabla}\varphi)^2+V(\varphi)\right] \\ &=:\; \int \mathrm{d}^3\boldsymbol{x}\;\mathcal{H}(x) \;=:\; H, \end{aligned} \tag{2.4.13}$$

where $\mathcal{H}$, the Hamiltonian density, is now a function of $\varphi(x)$ and $\pi(x)$, with

$$\pi \;:=\; \frac{\partial\mathcal{L}}{\partial\dot\varphi} \qquad (\dot\varphi \;:=\; \partial^0\varphi). \tag{2.4.14}$$

It is thus possible to translate into Hamilton's formulation, with the equations

$$\pi(x) \;=\; \dot\varphi(x) \;=\; \frac{\partial\mathcal{H}}{\partial\pi} \tag{2.4.15}$$

and

$$\frac{\partial\mathcal{H}}{\partial\varphi} \;=\; -\boldsymbol{\nabla}^2\varphi+\frac{\partial V}{\partial\varphi}. \tag{2.4.16}$$

The Euler–Lagrange equation is then

$$\Box\,\varphi \;=\; -\frac{\partial V}{\partial\varphi} \tag{2.4.17}$$

and one thus has

$$\dot{\pi}(x) \;=\; -\frac{\partial \mathcal{H}}{\partial \varphi}. \tag{2.4.18}$$

Exercise 2.4.2. *Show that invariance with respect to the Lorentz transformation (a rotation) defined by* $x^\mu \to x^\mu + \varepsilon^\mu{}_\nu x^\nu$, *where* $\varepsilon^{\mu\nu} = -\varepsilon^{\nu\mu}$, *leads to the conserved "current"*

$$M^\mu_{\rho\sigma} \;=\; x_\rho\, \Theta^\mu_\sigma - x_\sigma\, \Theta^\mu_\rho \tag{2.4.19}$$

and "charge"

$$M_{\rho\sigma} \;:=\; \int \mathrm{d}^3\boldsymbol{x}\, M^0_{\rho\sigma}, \tag{2.4.20}$$

which is just the total angular momentum of the field.

As already noted, Noether's theorem also applies to an internal symmetry. A simple but instructive and important example is a single *complex* scalar field:

$$\varphi(x) \;=\; \tfrac{1}{\sqrt{2}} \left[\varphi_1(x) + \mathrm{i}\varphi_2(x)\right], \tag{2.4.21}$$

with $\varphi_{1,2}$ real. This turns out to be a representation of a single *charged* particle. Instead of $\varphi_{1,2}$ we shall adopt φ and φ^* as independent variables. The Lagrangian density (with a self interaction) then takes on the following form (note the different numerical coefficients chosen with respect to the real case):

$$\mathcal{L} \;=\; (\partial^\mu \varphi^*)(\partial_\mu \varphi) - m^2 \varphi^* \varphi - \tfrac{1}{2}\lambda (\varphi^* \varphi)^2. \tag{2.4.22}$$

Variation with respect to φ^* leads to the equation of motion for φ:

$$\left[\Box + m^2 + \lambda \varphi^* \varphi\right] \varphi(x) \;=\; 0. \tag{2.4.23}$$

Likewise, variation with respect to φ leads to the analogous equation of motion for φ^*. Note that, in the absence of interactions, we simply have two independent Klein–Gordon equations. Even in the presence of the interaction term, this system possesses an internal *global gauge* symmetry, that is, under the following transformation:

$$\varphi(x) \;\to\; \varphi'(x) \;=\; \mathrm{e}^{-\mathrm{i}q\theta}\; \varphi(x) \tag{2.4.24a}$$

and

$$\varphi^*(x) \;\to\; \varphi^{*\prime}(x) \;=\; \mathrm{e}^{+\mathrm{i}q\theta}\; \varphi^*(x), \tag{2.4.24b}$$

where q is the charge and θ is the transformation parameter (both are constant).

Consider now the infinitesimal form, for which

$$\delta\varphi(x) \;=\; -\mathrm{i}q\,\delta\theta\,\varphi(x) \quad \text{and} \quad \delta\varphi^*(x) \;=\; +\mathrm{i}q\,\delta\theta\,\varphi^*(x). \tag{2.4.25}$$

Exercise 2.4.3. *Show that invariance with respect to this variation leads to the following conserved Noether current*

$$j^\mu \;=\; (\partial^\mu\varphi^*)\mathrm{i}q\varphi(x) - (\partial^\mu\varphi)\mathrm{i}q\varphi^*(x) \;=\; -\mathrm{i}q\varphi^*\overset{\leftrightarrow}{\partial^\mu}\varphi, \tag{2.4.26}$$

where we have introduced $\overset{\leftrightarrow}{\partial} := \overset{\rightarrow}{\partial} - \overset{\leftarrow}{\partial}$*; the single arrows indicate the direction of action and the sign change is directly related to Gauss' theorem.*

2.4.2 Spinor fields

At the classical level it not obvious how to construct a theory for spin one-half objects since intrinsic spin is a purely quantum concept. It will therefore be useful to remind the reader of the basic notions of the Dirac equation (1928). We shall thus consider the *classical* spinor wave equation

$$\left[\mathrm{i}\gamma^\mu\partial_\mu - m\mathbb{1}\right]\psi(x) \;=\; 0, \tag{2.4.27}$$

where, in order that the operator version $\gamma^\mu p_\mu = m$ should agree with the Einstein relation $p^2 = m^2$ (recall that $c=1$ here), the γ-matrices should obey the Clifford algebra $\{\gamma^\mu,\gamma^\nu\} = 2g^{\mu\nu}$. Such matrices have (minimal) rank-four representations, which may be constructed block-wise with the help of the Pauli σ matrices. One explicit form (the Dirac representation) is

$$\gamma^0 \;=\; \begin{pmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{pmatrix} \quad \text{and} \quad \boldsymbol{\gamma} \;=\; \begin{pmatrix} 0 & \boldsymbol{\sigma} \\ -\boldsymbol{\sigma} & 0 \end{pmatrix}, \tag{2.4.28}$$

where the sub-matrices are 2×2. The wave-function is thus represented by a *four-component spinor.*[*] The (suppressed) indices on such a spinor, as too those on the matrices γ^μ are often referred to as Dirac indices and the space over which they run, Dirac space. We recall here too some of the (representation *independent*) properties of the γ-matrices:

$$\gamma^{0\dagger} \;=\; \gamma^0, \qquad \boldsymbol{\gamma}^\dagger \;=\; -\boldsymbol{\gamma}, \qquad \gamma^{\mu\dagger} \;=\; \gamma_\mu \;=\; \gamma^0\gamma^\mu\gamma^0. \tag{2.4.29}$$

Exploiting these properties, we may define the conjugate spinor, $\bar\psi(x) = \psi^\dagger\gamma^0$, obeying the conjugate equation:

$$\bar\psi(x)\left[-\mathrm{i}\gamma^\mu\overset{\leftarrow}{\partial_\mu} - m\mathbb{1}\right] \;=\; 0, \tag{2.4.30}$$

[*] This is not to be confused with a Lorentz four-vector: a spinor cannot be thought of as any sort of Lorentz vector since, for example, a spatial rotation through 2π reverses its sign.

Again, the negative sign in front of $\overleftarrow{\partial}_\mu$ reflects the use of integration by parts and Gauss' theorem. The Lagrangian density that generates the previous equations of motion is

$$\mathcal{L}(\psi,\bar{\psi}) \;=\; \bar{\psi}(x)\left[\mathrm{i}\gamma^\mu\partial_\mu - m\mathbb{1}\right]\psi(x), \tag{2.4.31}$$

where the derivative acts to the right. By considering variation with respect to $\bar{\psi}$, one obtains the equation of motion for ψ and *vice versa*.

Consider next the general Poincaré transformation

with

$$x^\mu \;\to\; x'^\mu \;=\; \Lambda^\mu{}_\nu\, x^\nu + a^\mu, \tag{2.4.32a}$$

$$\psi(x) \;\to\; \psi'(x') \;=\; S(\Lambda)\,\psi(x). \tag{2.4.32b}$$

If the Lagrangian is thus invariant under an infinitesimal translation, we find, via the usual procedure, the following energy–momentum tensor:

$$\Theta^{\mu\nu} \;=\; \frac{\partial\mathcal{L}}{\partial(\partial_\mu\psi)}\,\partial^\nu\psi + \partial^\nu\bar{\psi}\,\frac{\partial\mathcal{L}}{\partial(\partial_\mu\bar{\psi})} - g^{\mu\nu}\,\mathcal{L}\,. \tag{2.4.33}$$

In this case, as a consequence of the Dirac equation (the equation of motion), we also have $\mathcal{L}=0$.* It therefore follows that

$$\Theta^{\mu\nu} \;=\; \bar{\psi}\,\mathrm{i}\gamma^\mu\partial^\nu\,\psi. \tag{2.4.34}$$

We may now again perform the Fourier decomposition of ψ:

$$\psi(x) \;=\; \int\frac{\mathrm{đ}^3\boldsymbol{k}}{2k_0}\sum_{\pm s}\Big[b(k,s)\,u(k,s)\;\mathrm{e}^{-\mathrm{i}k\cdot x} + d^*(k,s)\,v(k,s)\;\mathrm{e}^{+\mathrm{i}k\cdot x}\Big]\,, \tag{2.4.35}$$

where the positive- and negative-energy spinors satisfy

and

$$\big(\not{k} - m\big)\,u(k,\pm s) \;=\; 0 \tag{2.4.36a}$$

$$\big(\not{k} + m\big)\,v(k,\pm s) \;=\; 0. \tag{2.4.36b}$$

Recall that the spin four-vector s^μ is defined such that $s^2=-1$ and $k\cdot s=0$. And for $\bar{\psi}$, from its definition, we have

$$\bar{\psi}(x) \;=\; \int\frac{\mathrm{đ}^3\boldsymbol{k}}{2k_0}\sum_{\pm s}\Big[d(k,s)\,\bar{v}(k,s)\;\mathrm{e}^{-\mathrm{i}k\cdot x} + b^*(k,s)\,\bar{u}(k,s)\;\mathrm{e}^{+\mathrm{i}k\cdot x}\Big]\,. \tag{2.4.37}$$

* That is, $\mathcal{L}=0$ for the particular field configuration satisfying the equation of motion but is not, of course, necessarily constant with respect to its variation.

In terms of the Fourier components, the energy–momentum vector is

$$\begin{aligned} P^\mu &:= \int \mathrm{d}^3\boldsymbol{x}\, \Theta^{0\mu} \\ &= \int \frac{\mathrm{đ}^3\boldsymbol{k}}{2k_0} \sum_{\pm s} \left[b^*(k,s) b(k,s) k^\mu - d(k,s) d^*(k,s) k^\mu \right]. \end{aligned} \tag{2.4.38}$$

Note the ordering of the b's and d's.

A well-known problem now emerges: the second term in Eq. (2.4.38) would contribute a negative energy. For classical fields, this is usually resolved by *interpreting* such contributions as due to (negative-energy) *holes* or (positive-energy) *anti*particles. Here, however, anticipating the fermionic nature and, in particular, the anticommuting properties of spinor fields, we shall introduce the notion of Grassmann or anticommuting variables. We shall thus endow the b's and d's with anticommutation properties, as follows:

$$\{b, b'\} \,=\, \{b, d'\} \,=\, \{b, b'^*\} \,=\, \cdots \,=\, 0, \tag{2.4.39}$$

where $b = b(k,s)$, $b' = b'(k,s)$ *etc.* Here we are forced to come to terms with the intrinsically quantum nature of half-integer spin objects. In this manner we have effectively rendered the energy a Grassmann variable and, as such, its positivity becomes undetermined. Of course, when we eventually quantise this field theory, the corresponding physical quantity, or expectation value of the energy, will indeed always be found to be positive definite.

2.4.3 Vector fields

We now turn to the problem of vector fields. Here it is necessary to distinguish immediately between the *massive* and *massless* cases. The first is essentially little different to a set of four scalar fields while the second possesses a particular symmetry, known as a local gauge symmetry, precisely by virtue of the absence of a mass term in the Lagrangian and/or equations of motion. This symmetry, which requires $m = 0$ exactly, implies that the massless case *cannot* be considered as the zero-mass limit of a massive theory; it is a singular point.

Massive vector fields

In the absence of external sources (interactions), the Lagrangian is

$$\mathcal{L}(A^\mu, \partial_\nu A^\mu) \,=\, -\tfrac{1}{4}\, F^{\mu\nu}\, F_{\mu\nu} + \tfrac{1}{2} m_A^2 A^\mu A_\mu, \tag{2.4.40}$$

where we have simply added a mass term to Lagrangian (2.2.10). The Euler–Lagrange equations now yield

$$-\partial_\mu(\partial^\mu A^\nu - \partial^\nu A^\mu) \;=\; m_A^2 A^\nu. \tag{2.4.41}$$

From this (taking the divergence ∂_ν), we obtain

$$\partial_\nu A^\nu \;=\; 0. \tag{2.4.42}$$

Note that this is only true because $m_A \neq 0$. Substituting into the equations of motion, we finally obtain the equation that might have been anticipated:

$$\left(\Box + m_A^2\right) A^\mu \;=\; 0. \tag{2.4.43}$$

Note that, although the field apparently has four degrees of freedom, the derived condition (2.4.42) reduces these to three: the three spatial degrees of freedom for the polarisation vector.

We may now perform the usual Fourier transform and write

$$A^\mu(x) \;=\; \int \frac{\text{đ}^3\boldsymbol{k}}{2k_0}\left[\xi^\mu(k)\;\mathrm{e}^{-\mathrm{i}k\cdot x} + \xi^{*\mu}(k)\;\mathrm{e}^{+\mathrm{i}k\cdot x}\right], \tag{2.4.44}$$

with the condition $k\cdot\xi(k) = 0$ and where $k_0^2 = \boldsymbol{k}^2 + m^2$. In the rest frame $k^\mu = (m, \boldsymbol{0})$ and therefore $\xi^\mu = (0, \boldsymbol{\xi})$. It is thus natural to normalise ξ^μ by requiring $\xi\cdot\xi^* = -1$, the negative sign merely reflecting the space-like nature of the polarisation vector.

Exercise 2.4.4. *Derive the energy–momentum vector P^μ in terms of the above-defined Fourier components.*

Massless vector fields

A massless vector field in the presence of an external current j^μ obeys the Maxwell equations

$$\partial_\nu\left(\partial^\nu A^\mu - \partial^\mu A^\nu\right) \;=\; j^\mu, \tag{2.4.45a}$$

or

$$\partial_\nu\, F^{\nu\mu} \;=\; j^\mu. \tag{2.4.45b}$$

This does not determine A^μ unambiguously since for any solution A^μ, $A^\mu + \partial^\mu\Lambda(x)$ is also a solution and the resulting electromagnetic fields are identical. This is just the observation that the equations possess a *local* gauge invariance.

In the case of the Feynman integral, for example, this infinite degeneracy (an infinity of physically equivalent gauge-field configurations) will cause spurious infinities. In any case, we know that we need to choose or *fix* the gauge. A typical

choice is the Lorentz condition:

$$\partial^\mu A_\mu \;=\; 0. \tag{2.4.46}$$

However, this still does not entirely eliminate the gauge degree of freedom since the previous transformation still leaves the physical fields invariant provided that

$$\partial^\mu \partial_\mu \Lambda(x) \;=\; 0. \tag{2.4.47}$$

The fixing of a further condition will eliminate yet another degree of freedom, leaving just the two transverse polarisation states of a massless vector.

Now, in the Lagrange formalism, a condition such as that above may be applied via the so-called *Lagrange-multiplier* technique:

$$\mathcal{L}(A^\mu, \partial_\nu A^\mu) \;\to\; -\tfrac{1}{4}\, F^{\mu\nu}\, F_{\mu\nu} - j^\mu A_\mu - \tfrac{1}{2}\xi^{-1}(\partial_\mu A^\mu)^2, \tag{2.4.48}$$

where ξ is the so-called gauge parameter. With this expedient, the Euler–Lagrange equations become

$$\partial_\mu\, F^{\mu\nu} - \frac{1}{\xi}\,\partial^\nu(\partial_\mu A^\mu) \;=\; j^\nu. \tag{2.4.49}$$

If the external current is conserved, then

$$\Box(\partial_\mu A^\mu) \;=\; 0. \tag{2.4.50}$$

Thus, if both $\partial_\mu A^\mu$ and $\partial_0(\partial_\mu A^\mu)$ vanish at any given instant t, then $\partial_\mu A^\mu$ vanishes at all times. The Maxwell equations (from the Euler–Lagrange equations above) then reduce to

$$\Box\, A^\mu \;=\; j^\mu. \tag{2.4.51}$$

Again, however, for any solution A^μ, this equation is also satisfied by $A^\mu + \partial^\mu \Lambda$, with $\Box\Lambda = 0$. We therefore have

$$\frac{\partial \mathcal{L}}{\partial(\partial_\mu A_\nu)} \;=\; -F^{\mu\nu} - \frac{1}{\xi}\, g^{\mu\nu}\,(\partial_\sigma A^\sigma) \;=\; -F^{\mu\nu}. \tag{2.4.52}$$

Consider then a general Poincaré transformation, for the field A^μ this means

$$A^\mu(x) \;\to\; A'^\mu(x') \;=\; \Lambda^\mu{}_\nu A^\mu(x). \tag{2.4.53}$$

And, by imposing translation invariance, we derive the energy–momentum tensor:

$$\begin{aligned}\Theta^{\mu\nu} &= \frac{\partial\mathcal{L}}{\partial(\partial_\mu A_\sigma)}\,\partial^\nu A_\sigma - \mathcal{L}\; g^{\mu\nu} \\ &= -F^{\mu\sigma}\,\partial^\nu A_\sigma + g^{\mu\nu}\,\tfrac{1}{4}F_{\rho\sigma}F^{\rho\sigma}\end{aligned} \tag{2.4.54}$$

We perform the usual Fourier transform and write

$$A^\mu(x) \;=\; \int \frac{\text{đ}^3\boldsymbol{k}}{2k_0}\left[\xi^\mu(k)\;\mathrm{e}^{-\mathrm{i}k\cdot x} + \xi^{*\mu}(k)\;\mathrm{e}^{+\mathrm{i}k\cdot x}\right], \tag{2.4.55}$$

with the condition $k\cdot\xi(k)=0$ in order to satisfy $\partial^\mu A_\mu = 0$, just as with the massive vector field, but where now $k_0 = |\boldsymbol{k}|$ since $m=0$. Here it is not possible to determine the normalisation via reference to the rest frame. However, the Lorentz condition leads to

$$\xi_0(k) \;=\; \hat{\boldsymbol{k}}\cdot\boldsymbol{\xi}(k), \qquad (\hat{\boldsymbol{k}} \;=\; \boldsymbol{k}/|\boldsymbol{k}| \;=\; \boldsymbol{k}/k_0). \tag{2.4.56}$$

Let us decompose $\boldsymbol{\xi}$ into longitudinal and transverse components with respect to $\boldsymbol{k}$:

$$\boldsymbol{\xi} \;=\; (\boldsymbol{\xi}\cdot\hat{\boldsymbol{k}})\,\hat{\boldsymbol{k}} + \boldsymbol{\xi}_\perp. \tag{2.4.57}$$

We therefore have

$$\xi\cdot\xi^* \;=\; \xi_0\xi_0^* - \boldsymbol{\xi}\cdot\boldsymbol{\xi}^* \;=\; -\boldsymbol{\xi}_\perp\cdot\boldsymbol{\xi}_\perp^*. \tag{2.4.58}$$

And again we may normalise by requiring, *e.g.*, $\xi\cdot\xi^* = -1$.

Exercise 2.4.5. *Express the energy–momentum vector P^μ in terms of the above-defined Fourier components.*

The conserved "charge" (the total energy and momentum) is then

$$\begin{aligned}P^\mu &:= \int \mathrm{d}^3\boldsymbol{x}\,\Theta^{0\mu} \\ &= \int \frac{\text{đ}^3\boldsymbol{k}}{2k_0}[-\xi^*(k)\cdot\xi(k)\,k^\mu].\end{aligned} \tag{2.4.59}$$

The energy is thus positive definite owing to the space-like nature of $\xi(k)$.

It is worth examining this last result a little more in depth. In the massive-vector case the Lorentz condition merely removed one degree of freedom, leaving three possible independent space-like polarisation states. In contrast, in the massless case, where k^μ is then light-like, the very same Lorentz condition actually guarantees a cancellation between the *would-be* "time-like" and the "longitudinal" photon modes, leaving only the "transverse" modes as true degrees of freedom.

Indeed, returning to the gauge-choice freedom, for Λ such that $\Box\Lambda = 0$, we may write the following Fourier decomposition:

$$\Lambda(x) \;=\; \int \frac{\text{đ}^3\boldsymbol{k}}{2k_0}\left[\lambda(k)\;\mathrm{e}^{-\mathrm{i}k\cdot x} + \lambda^*(k)\;\mathrm{e}^{+\mathrm{i}k\cdot x}\right]. \tag{2.4.60}$$

Therefore, the transformation $A^\mu \to A'^\mu = A^\mu + \partial^\mu\Lambda$ results in

$$\xi^\mu(k) \;\to\; \xi'^\mu(k) \;=\; \xi^\mu(k) - \mathrm{i}k^\mu\lambda(k). \tag{2.4.61}$$

This means that there exists a choice of Λ, for which

$$\xi_0 \;=\; \hat{\boldsymbol{k}}\cdot\boldsymbol{\xi}(k) \;=\; 0, \tag{2.4.62}$$

thus, leaving only the transverse modes.

With this gauge choice, we can write

$$\xi^\mu(k) \;=\; \sum_{\lambda=1,2} a(k,\lambda)\,\varepsilon^\mu(k,\lambda), \tag{2.4.63}$$

with $\varepsilon^\mu(k,\lambda)$ for $\lambda = 1,2$ chosen such that:

$$\varepsilon^0(k,\lambda) \;=\; 0, \tag{2.4.64a}$$
$$\boldsymbol{\varepsilon}^*(k,\lambda)\cdot\boldsymbol{\varepsilon}(k,\lambda') \;=\; \delta_{\lambda\lambda'}, \tag{2.4.64b}$$
$$\boldsymbol{\varepsilon}(k,1)_\wedge\boldsymbol{\varepsilon}(k,2) \;=\; \hat{\boldsymbol{k}}, \tag{2.4.64c}$$
$$\boldsymbol{k}\cdot\boldsymbol{\varepsilon}(k,\lambda) \;=\; 0. \tag{2.4.64d}$$

And we thus finally obtain

$$P^\mu \;=\; \int \frac{\text{đ}^3\boldsymbol{k}}{2k_0}\sum_{\lambda=1,2} |a(k,\lambda)|^2\,k^\mu. \tag{2.4.65}$$

2.5 Bibliography

Bailin, D. and Love, A. (1993), *Introduction to Gauge Field Theory* (IOP Pub.), revised edition.

Dirac, P.A.M. (1928), *Proc. Royal Soc. (London)* **A117**, 610.

Gordon, W. (1926), *Z. Phys.* **40**, 117.

Klein, O. (1927), *Z. Phys.* **41**, 407.

Noether, E. (1918), *Nachr. v. d. Ges. d. Wiss. zu Göttingen*, 235; *transl.*, *Transport Theory and Statistical Mechanics* 183.

Chapter 3

Quantum Field Theory

3.1 The path integral

The approach to be adopted here for field or second quantisation will be that of the *path integral* (Feynman and Hibbs, 1965). While the technique pioneered by Feynman has already been introduced in the preceding course on advanced quantum mechanics, for the sake of completeness and clarity, we shall here provide a self-contained account. The basic definition of the path or *functional* integral starts with a discretised version of the field theory; we shall present the approach for a simple real scalar field and then cover the complications introduced in connection with spinor and gauge fields.

3.1.1 A real scalar field

The starting point is the simple Gaussian integral in one dimension,

$$\int_{-\infty}^{\infty} \mathrm{d}q \; \mathrm{e}^{-\frac{1}{2}aq^2} \;=\; \sqrt{2\pi}\, a^{-1/2}, \tag{3.1.1}$$

where q is intended to represent a generalised real coordinate. The natural extension to a real n-dimensional space leads to

$$\int_{-\infty}^{\infty} \mathrm{d}q_1 \, \mathrm{d}q_2 \, \ldots \mathrm{d}q_n \; \mathrm{e}^{-\frac{1}{2}\boldsymbol{q}^{\mathsf{T}}\mathsf{A}\,\boldsymbol{q}} \;=\; \left(\sqrt{2\pi}\right)^n \, (\det \mathsf{A})^{-1/2}, \tag{3.1.2}$$

where, for convergence of the integral, the $n\times n$-matrix A is necessarily taken to be real, symmetric and positive-definite while $\boldsymbol{q}$ is an n-component vector: $\boldsymbol{q} = (q_1, \ldots, q_n)$.

Exercise 3.1.1. *Prove the last equation above.*

Hint: *This is most easily performed by diagonalising the matrix* A.

It is convenient, for the path-integral quantisation procedure, to reformulate such expressions as exponentials and thus we rewrite this as

$$\int_{-\infty}^{\infty} \mathrm{d}q_1\, \mathrm{d}q_2\, \ldots \mathrm{d}q_n\, \mathrm{e}^{-\frac{1}{2}\boldsymbol{q}^{\mathsf{T}}\mathsf{A}\,\boldsymbol{q}} \;=\; \left(\sqrt{2\pi}\right)^n \mathrm{e}^{-\frac{1}{2}\,\mathrm{Tr}\ln \mathsf{A}}, \tag{3.1.3}$$

where we have exploited the following identity:

$$\ln\det \mathsf{A} \;\equiv\; \mathrm{Tr}\ln \mathsf{A}. \tag{3.1.4}$$

Exercise 3.1.2. *Prove this last equality.*
Hint: *The simplest proof is again obtained by diagonalising* A.

The application to field theory now proceeds by taking the limit $n \to \infty$, whence the vector $q_i \to q(x)$ and the matrix $A_{ij} \to A(x,y)$, with x and y representing space–time coordinates. One may think of the discretised quantities as local averages over infinitesimal space–time hypercubes centred around the points x, y *etc.* We shall thus take the continuum limit of (3.1.3) to define a *functional integral*:

$$\int_{-\infty}^{\infty} \mathcal{D}q\; \mathrm{e}^{-\frac{1}{2}\int \mathrm{d}x\,\mathrm{d}y\, q(x)\, A(x,y)\, q(y)} \;\equiv\; \mathrm{e}^{-\frac{1}{2}\,\mathrm{Tr}\ln A}. \tag{3.1.5}$$

In this definition the numerical $\sqrt{2\pi}$ factors have been absorbed into an overall (infinite) normalisation constant (which, as we shall soon see, is generally irrelevant). For clarity, we have limited the x integrals to one dimension; however, the extension to three-dimensional space or four-dimensional space–time is trivial. We must now learn how to evaluate the expression on the right-hand side.

To do this, let us first define a little more carefully what is meant by a function of an operator, such as $A(x,y)$. By considering the passage from the discrete case to the continuum, it is evident that multiplication is now defined by

$$(A{\cdot}B)(x,y) \;:=\; \int \mathrm{d}z\, A(x,z)\, B(z,y), \tag{3.1.6}$$

where the notation should be obvious and the integration is understood to be over the entire x-space. Moreover, the trace operation thus becomes

$$\mathrm{Tr}\, A \;:=\; \int \mathrm{d}x\, A(x,x). \tag{3.1.7}$$

Now, in all cases of interest here we shall be dealing with *local* interactions and

thus the most general form of operator to occur will be

$$A(x,y) \;=\; \delta(x-y)\,\mathcal{A}(x,y), \tag{3.1.8}$$

where $\mathcal{A}(x,y)$ is an analytic function of x, y and their derivatives. It is useful then to define the Dirac δ-function via a Fourier-transform representation

$$\delta(x-y) \;:=\; \int_{-\infty}^{\infty} \text{đ}k\; \mathrm{e}^{\mathrm{i}k(x-y)}\,. \tag{3.1.9}$$

We may thus write

$$A(x,y) \;=\; \int \text{đ}k\; \mathrm{e}^{\mathrm{i}k(x-y)}\,\tilde{\mathcal{A}}(k). \tag{3.1.10}$$

Consider, for example, the following operator (useful for later chapters):

$$A(x,y) \;=\; \left[\frac{\partial}{\partial x}\frac{\partial}{\partial y} + c\right]\delta(x-y), \tag{3.1.11a}$$

with c a constant, whence we have

$$A(x,y) \;=\; \int \text{đ}k\; \mathrm{e}^{\mathrm{i}k(x-y)}\,[k^2+c]. \tag{3.1.11b}$$

It is left as an exercise to show that an operator raised to a power is given by the following expression:*

$$\begin{aligned} A^n(x,y) \;&\equiv\; \int \mathrm{d}z_1\,\ldots\mathrm{d}z_{n-1}\,A(x,z_1)\,A(z_1,z_2)\ldots A(z_{n-1},y) \\ &=\; \int \text{đ}k\; \mathrm{e}^{\mathrm{i}k(x-y)}\;\mathcal{A}^n. \end{aligned} \tag{3.1.12}$$

In other words, inasmuch as a power-series representation exists, a function, or rather *functional*, of an operator may be represented as (note the use of square brackets)

$$F[A(x,y)] \;=\; \int \text{đ}k\; \mathrm{e}^{\mathrm{i}k(x-y)}\;F\big(\mathcal{A}(k)\big)\,. \tag{3.1.13}$$

In particular, for the inverse (which certainly exists if, as required, A is positive definite) we have

$$A^{-1}(x,y) \;=\; \int \text{đ}k\; \mathrm{e}^{\mathrm{i}k(x-y)}\;\mathcal{A}^{-1}(k). \tag{3.1.14}$$

* From now on we shall omit the tilde in Fourier transforms, as the argument suffices to indicate the space over which the object is defined.

And finally, for example, for the operator A defined in (3.1.11a), we then have

$$\ln A(x,y) \;=\; \int \text{đ}k\; \mathrm{e}^{\mathrm{i}k(x-y)}\; \ln[k^2+c]. \tag{3.1.15}$$

3.1.2 External source terms

A further useful generalisation of the previous formulæ is obtained by adding to the exponent a term linear in q; thus,

$$\int_{-\infty}^{\infty} \mathrm{d}q_1\, \mathrm{d}q_2\, \ldots \mathrm{d}q_n\; \mathrm{e}^{-\frac{1}{2}\boldsymbol{q}^{\mathsf{T}}\mathsf{A}\,\boldsymbol{q}+\boldsymbol{j}\cdot\boldsymbol{q}} \;=\; \left(\sqrt{2\pi}\right)^n\; \mathrm{e}^{-\frac{1}{2}\,\mathrm{Tr}\ln\mathsf{A}}\; \mathrm{e}^{\frac{1}{2}\boldsymbol{j}^{\mathsf{T}}\mathsf{A}^{-1}\boldsymbol{j}}, \tag{3.1.16}$$

where $\boldsymbol{j}=(j_1,\ldots,j_n)$ is some arbitrary external (*source* or *current*) vector.

Exercise 3.1.3. *Prove this last result.*
Hint: *Complete the square in the exponent and then shift the integration variables (such an operation is always possible since the integrals converge).*

In path-integral or *functional* language this becomes

$$\begin{aligned}\int_{-\infty}^{\infty} \mathcal{D}q\; \exp\!\left[-\tfrac{1}{2}\int \mathrm{d}x\,\mathrm{d}y\, q(x)\, A(x,y)\, q(y) + \int \mathrm{d}y\, j(y)\, q(y)\right] \\ = \;\exp\!\left[-\tfrac{1}{2}\,\mathrm{Tr}\ln A + \tfrac{1}{2}\int \mathrm{d}x\,\mathrm{d}y\, j(x)\, A^{-1}(x,y)\, j(y)\right].\end{aligned} \tag{3.1.17}$$

With this addition, we can now evaluate rather more complicated integrals. Differentiating Eq. (3.1.16) with respect to the discrete sources j_{i_1}, j_{i_2}, $\ldots j_{i_m}$ and then setting all of the $j_1=\cdots=j_n=0$, we obtain

$$\begin{aligned}\int_{-\infty}^{\infty} \mathrm{d}q_1\, \mathrm{d}q_2\, \ldots \mathrm{d}q_n\, q_{i_1}\, q_{i_2} \ldots q_{i_m}\; \mathrm{e}^{-\frac{1}{2}\boldsymbol{q}^{\mathsf{T}}\mathsf{A}\,\boldsymbol{q}+\boldsymbol{j}\cdot\boldsymbol{q}} \\ = \;\left[A^{-1}_{i_1 i_2}\cdots A^{-1}_{i_{n-1} i_n} + \text{perms}\right]\left(\sqrt{2\pi}\right)^n\; \mathrm{e}^{-\frac{1}{2}\,\mathrm{Tr}\ln\mathsf{A}},\end{aligned} \tag{3.1.18}$$

The *functional derivative*, the natural extension to the continuum of the previous derivatives, is intuitively defined via

$$\frac{\delta}{\delta j(x)} \int \mathrm{d}y\, j(y)\, q(y) \;:=\; q(x). \tag{3.1.19}$$

The functional-integral version is then simply (for m even)

$$\int_{-\infty}^{\infty} \mathcal{D}q\, q(x_1)\, q(x_2) \ldots q(x_m) \exp\left[-\tfrac{1}{2}\int \mathrm{d}y\, \mathrm{d}z\, q(y)\, A(y,z)\, q(z)\right]$$
$$= \left[A^{-1}(x_1,x_2)\cdots A^{-1}(x_{m-1},x_m) + \text{perms}\right] \exp\left[-\tfrac{1}{2}\,\mathrm{Tr}\ln A\right]. \quad (3.1.20)$$

3.1.3 A complex (charged) scalar field

The only immediately useful extension remaining is to complex variables $q_i \in \mathbb{C}$ and thus to complex fields $q(x) \in \mathbb{C}$.* In order that the integrals remain well-defined and convergent, we now require that the complex-valued matrix A be Hermitian (or self-adjoint) and positive definite. The natural complex measure is

$$\mathrm{d}q\, \mathrm{d}q^* \;\equiv\; 2\,\mathrm{d}(\mathrm{Re}\, q)\, \mathrm{d}(\mathrm{Im}\, q)\,. \quad (3.1.21)$$

And, starting again with the simple Gaussian integral (a real and positive)

$$\int \mathrm{d}q\, \mathrm{d}q^*\; \mathrm{e}^{-\frac{1}{2}aq^*q} \;=\; \sqrt{2\pi}\, a^{-1}, \quad (3.1.22)$$

through to the multi-dimensional version

$$\int \mathrm{d}q_1\, \mathrm{d}q_1^* \ldots \mathrm{d}q_n\, \mathrm{d}q_n^*\; \mathrm{e}^{-\frac{1}{2}\boldsymbol{q}^\dagger \mathsf{A}\, \boldsymbol{q}} \;=\; \left(\sqrt{2\pi}\right)^n \mathrm{e}^{-\,\mathrm{Tr}\ln \mathsf{A}}, \quad (3.1.23)$$

we obtain,

$$\int \mathcal{D}q\, \mathcal{D}q^*\, \exp\left\{-\tfrac{1}{2}\int \mathrm{d}x\, \mathrm{d}y\, q^*(x)\, A(x,y)\, q(y) + \int \mathrm{d}x\, \left[j^*(x)\, q(x) + j(x)\, q^*(x)\right]\right\}$$
$$= \exp\left\{-\,\mathrm{Tr}\ln A + \tfrac{1}{2}\int \mathrm{d}x\, \mathrm{d}y\, j^*(x)\, A^{-1}(x,y)\, j(y)\right\}. \quad (3.1.24)$$

Exercise 3.1.4. *Adopting the methods illustrated, prove the final result given here.*

3.2 The path-integral formulation of NRQM

3.2.1 The transition amplitude

In this section we shall briefly recall the path-integral formulation (Feynman, 1948) of non-relativistic quantum mechanics, born of a suggestion by Dirac (1933). The

* As already noted, such a case naturally represents a charged field.

basic approach we shall follow is close to that found in Abers and Lee (1973).

Given a time-independent *operator* $Q_{\rm S}$ (in the Schrödinger picture) "measuring" the generalised coordinate q, as usual, the eigenstates are defined through

$$Q_{\rm S}|q\rangle_{\rm S} \equiv q|q\rangle_{\rm S}. \tag{3.2.1}$$

Moving over to the Heisenberg picture, we have

$$Q_{\rm H}(t) := {\rm e}^{{\rm i}Ht}\, Q_{\rm S}\, {\rm e}^{-{\rm i}Ht}, \tag{3.2.2}$$

where H is the Hamiltonian, and we may write

$$Q_{\rm H}|q,t\rangle_{\rm H} \equiv q|q,t\rangle_{\rm H}. \tag{3.2.3}$$

Recall that in the Heisenberg picture the physical states are *frozen* in time and all time dependence is absorbed into the operators, which consequently then have time-*dependent* eigenstates. The eigenstates (or basis states) in the two pictures are related via

$$|q,t\rangle_{\rm H} = {\rm e}^{{\rm i}Ht}\, |q\rangle_{\rm S}. \tag{3.2.4}$$

The transition matrix element

$$\begin{aligned} M(q_f,t_f;q_i,t_i) &= {}_{\rm H}\langle q_f;t_f|q_i;t_i\rangle_{\rm H} \\ &= {}_{\rm S}\langle q_f|{\rm e}^{-{\rm i}H(t_f-t_i)}|q_i\rangle_{\rm S} \end{aligned} \tag{3.2.5}$$

describes the probability amplitude for a state initially having eigenvalue q_i at some instant t_i to find itself with q_f at a later instant t_f. It encodes all important information (dynamics) of the quantum theory. The idea is then to express this object as a path integral. The first step is to divide the finite interval $[t_i,t_f]$ into $n+1$ subintervals of infinitesimal duration $\Delta t=(t_f-t_i)/(n+1)$ by inserting n complete sets of states ($t_i<t_1<\cdots<t_n<t_f$):

$$M(q_f,t_f;q_i,t_i) = \int {\rm d}q_1\, {\rm d}q_2\, \ldots {\rm d}q_n\, \langle q_f,t_f|q_n,t_n\rangle \cdots \langle q_2,t_2|q_1,t_1\rangle\langle q_1,t_1|q_i,t_i\rangle, \tag{3.2.6}$$

where (and henceforth) we have omitted the suffix S, on the understanding that such time-dependent eigenstates belong exclusively to the Schrodinger picture.

Consider a generic infinitesimal time slice:

$$\begin{aligned} \langle q',t+\Delta t|q,t\rangle &= \langle q'|{\rm e}^{-{\rm i}H\Delta t}|q\rangle \\ &= \delta(q'-q) - {\rm i}\Delta t\langle q'|H|q\rangle + O\big((\Delta t)^2\big)\,. \end{aligned} \tag{3.2.7}$$

The Hamiltonian $H = H(P,Q)$ is a function of the generalised position and conjugate momentum *operators* Q and P.* The simplest, but most useful and common, case is a massive particle subject to a potential depending only on position Q:

$$H \;=\; \frac{P^2}{2m} + V(Q). \tag{3.2.8}$$

Using the fact that $\langle q|p\rangle = \mathrm{e}^{\mathrm{i}pq}$ and $\int \mathrm{đ}p\, \mathrm{e}^{\mathrm{i}p(q'-q)} = \delta(q'-q)$, we have

$$\begin{aligned}\langle q'|H(P,Q)|q\rangle \;&=\; \int \mathrm{đ}p'\, \mathrm{đ}p \left[\frac{\langle q'|p'\rangle\langle p'|P^2|p\rangle\langle p|q\rangle}{2m}\right] + \langle q'|V(Q)|q\rangle \\ &=\; \int \mathrm{đ}p\; \mathrm{e}^{\mathrm{i}p(q'-q)} \left[\frac{p^2}{2m} + V(q)\right] \\ &=\; \int \mathrm{đ}p\; \mathrm{e}^{\mathrm{i}p(q'-q)}\; H(p,q),\end{aligned} \tag{3.2.9}$$

where $H(p,q)$ is now simply the *classical* Hamiltonian function.†

To first order in Δt, we may thus rewrite (3.2.7) as

$$\langle q',t+\Delta t|q,t\rangle \;\simeq\; \int \mathrm{đ}p\; \mathrm{e}^{\mathrm{i}\,[p(q'-q)-H(p,q)\Delta t]} \tag{3.2.10}$$

and substituting this expression into that above for the matrix element, we have

$$\begin{aligned}M(q_f,t_f;q_i,t_i) \;=\; \lim_{n\to\infty} \int \prod_{k=1}^{n} \mathrm{d}q_k \prod_{k=1}^{n+1} \mathrm{đ}p_k \\ \times \exp\left\{\mathrm{i}\sum_{j=1}^{n+1} \big[p_j(q_j - q_{j-1}) - H(p_j,q_j)\Delta t\big]\right\},\end{aligned} \tag{3.2.11}$$

with the identification $q_0 \equiv q_i$ and $q_{n+1} \equiv q_f$. In the limit $(q_j - q_{j-1}) \to \dot{q}_j \Delta t$, the sum becomes an integral and we may rewrite the (now infinite) products as‡

$$M(q_f,t_f;q_i,t_i) \;=\; \int \mathcal{D}q\, \mathcal{D}p\, \exp\left\{\mathrm{i}\int_{t_i}^{t_f} \mathrm{d}t\; \big[p\,\dot{q} - H(p,q)\big]\right\}. \tag{3.2.12}$$

* Capital letters will be used here for operators while lower case will indicate ordinary variables.

† In some texts one finds the more symmetric $\frac{1}{2}(q+q')$ in place of q as the argument of H, but this really is unnecessary in view of the limit to be taken.

‡ As already noted, in general, any numerical factors accompanying the measures $\mathrm{d}q_k$ and $\mathrm{đ}p_k$ may simply be absorbed into an irrelevant overall normalisation constant.

This passage then *defines* the Feynman path integral.

With the above choice of Hamiltonian, the complex integrals in p are Gaussian and may therefore be performed via the usual Wick rotation:

$$\int_{-\infty}^{\infty} \text{đ}p \; \mathrm{e}^{\mathrm{i}\Delta t\,(p\,\dot{q}-p^2/2m)} \;=\; \Big[\frac{m}{2\pi\,\mathrm{i}\,\Delta t}\Big]^{1/2} \; \mathrm{e}^{\frac{1}{2}\,\mathrm{i}\,\Delta t\,\dot{q}^2 m}\,. \tag{3.2.13}$$

For the transition matrix element, we thus finally obtain

$$\begin{aligned} M(q_f,t_f;q_i,t_i) \;&=\; \lim_{n\to\infty} \int \prod_{k=1}^{n} \frac{\mathrm{d}q_k}{\big[2\pi\,\mathrm{i}\,\Delta t/m\big]^{1/2}} \\ &\qquad\qquad \times \exp\!\left\{ \mathrm{i} \sum_{j=1}^{n+1} \left[\frac{m}{2} \left(\frac{q_j - q_{j-1}}{\Delta t} \right)^2 - V(q_j) \right] \Delta t \right\} \\ &=\; \lim_{n\to\infty} \int \prod_{k=1}^{n} \frac{\mathrm{d}q_k}{\big[2\pi\,\mathrm{i}\,\Delta t/m\big]^{1/2}} \exp\!\left\{ \mathrm{i} \int_{t_i}^{t_f} \mathrm{d}t\, L(q,\dot{q}) \right\}, \end{aligned} \tag{3.2.14}$$

where $L = \frac{1}{2} m\dot{q}^2 - V(q)$ is simply the classical Lagrangian. We thus see that the exponent is precisely the classical action

$$S \;:=\; \int_{t_i}^{t_f} \mathrm{d}t\, L(q,\dot{q}), \tag{3.2.15}$$

which then determines the temporal evolution in quantum mechanics. The important point is that while the expression for the transition amplitude presented here has been derived *starting* from the canonical formulation of quantum mechanics, we may now equally take the path integral as the starting point and from it *derive* the Schrödinger equation.

It is important to stress, however, that the final simple form shown above is a direct consequence of the particular choice of Hamiltonian, which is, in fact, generally sufficient for our purposes. If, on the other hand, the Hamiltonian were to involve products of P and Q, then the final form would not depend simply on the classical action, but on what might be called an *effective* action, which could then be calculated from Eq. (3.2.11).

Exercise 3.2.1. *Starting from the non-linear Lagrangian*

$$L(q,\dot{q}) \;=\; \tfrac{1}{2}\,\dot{q}\,f(q), \tag{3.2.16}$$

where $f(q)$ is some non-singular function of q, show that the transition amplitude

is as above but with an effective action (Lee and Yang, 1962)

$$S_{\mathit{eff}} \;=\; \int_{t_i}^{t_f} \mathrm{d}t\;\left[L(q,\dot{q}) - \tfrac{1}{2}\mathrm{i}\,\delta(0)\ln f(q)\right]. \tag{3.2.17}$$

One then finds that the scattering matrix calculated using such a transition amplitude generates another infinite contribution, which precisely cancels the $\delta(0)$ term.

In any case, the evident advantage of this approach is that we express *quantum-mechanical* quantities in terms of the *classical* Lagrangian. In particular, this permits a direct study of the importance of the known symmetries at the classical level of such quantities and their properties under various symmetry transformations.

3.2.2 The generalisation to N degrees of freedom

It is straightforward to generalise Eq. (3.2.12) to a system with N degrees of freedom:

$$M(\boldsymbol{q}_f,t_f;\boldsymbol{q}_i,t_i) \;=\; \int \mathcal{D}\boldsymbol{q}\,\mathcal{D}\boldsymbol{p}\,\exp\!\left\{\mathrm{i}\int_{t_i}^{t_f}\mathrm{d}t\;\left[\boldsymbol{p}\cdot\dot{\boldsymbol{q}} - H(\boldsymbol{p},\boldsymbol{q})\right]\right\}, \tag{3.2.18}$$

where $\boldsymbol{q}:=(q_1,q_2,\ldots q_N)$ and $\boldsymbol{p}:=(p_1,p_2,\ldots p_N)$. This will be the starting point for a quantum field theory. However, for the rest of this section a single degree of freedom will suffice.

3.2.3 Transition matrix elements for products of fields

Let us now try to calculate more complicated matrix elements: inserting the operator $Q(t)$ into the transition matrix element (3.2.5), for some instant t such that $t_i<t<t_f$, leads to

$$\begin{aligned}\langle q_f;t_f|Q(t)|q_i;t_i\rangle \;=\; &\int \mathrm{d}q_1\,\mathrm{d}q_2\,\ldots\mathrm{d}q_n\\ &\times\langle q_f,t_f|q_n,t_n\rangle\cdots\langle q_k,t_k|Q(t)|q_{k-1},t_{k-1}\rangle\cdots\langle q_2,t_2|q_1,t_1\rangle\langle q_1,t_1|q_i,t_i\rangle,\end{aligned} \tag{3.2.19}$$

where we have identified t_k and t_{k-1} such that $t_{k-1}<t<t_k$. To first order in Δt we can consider $|q_{k-1},t_{k-1}\rangle$ to be an approximate eigenstate of $Q(t)$ and thus replace $Q(t)$ with $q(t)$. It is then easy to see that we may proceed as before, with simply an extra factor $q(t)$ inside the integral, to obtain

$$\langle q_f;t_f|Q(t)|q_i;t_i\rangle \;=\; \int \mathcal{D}q\,\mathcal{D}p\,q(t)\,\exp\!\left\{\mathrm{i}\int_{t_i}^{t_f}\mathrm{d}t\;\left[p\dot{q} - H(p,q)\right]\right\}. \tag{3.2.20}$$

Consider now a product of two such operators $Q(t_a)Q(t_b)$: if $t_a > t_b$, we have

$$\begin{aligned}\langle q_f;t_f|Q(t_a)Q(t_b)|q_i;t_i\rangle \;=\; &\int \mathrm{d}q_1\,\mathrm{d}q_2\,\ldots\mathrm{d}q_n \\ &\times \langle q_f,t_f|q_n,t_n\rangle\cdots\langle q_{k_a},t_{k_a}|Q(t_a)|q_{k_a-1},t_{k_a-1}\rangle\cdots \\ &\cdots\langle q_{k_b},t_{k_b}|Q(t_b)|q_{k_b-1},t_{k_b-1}\rangle\cdots\langle q_1,t_1|q_i,t_i\rangle, \quad (3.2.21a)\end{aligned}$$

which eventually leads us to

$$\langle q_f;t_f|Q(t_a)Q(t_b)|q_i;t_i\rangle \;=\; \int \mathcal{D}q\,\mathcal{D}p\,q(t_a)\,q(t_b)\exp\biggl\{\mathrm{i}\int_{t_i}^{t_f}\mathrm{d}t\;\bigl[p\dot{q}-H(p,q)\bigr]\biggr\}. \quad (3.2.21b)$$

The derivation rests on the ordering $t_a > t_b$, *i.e.* the operator on the right acts temporally before that on the left. Indeed, it is not possible to derive such a formula with the opposite ordering. We thus see that, to avoid specific restrictions on t_a and t_b, we should write

$$\begin{aligned}&\langle q_f;t_f|\mathrm{T}\bigl[Q(t_a)Q(t_b)\bigr]|q_i;t_i\rangle \\ &\qquad = \int \mathcal{D}q\,\mathcal{D}p\,q(t_a)\,q(t_b)\;\exp\biggl\{\mathrm{i}\int_{t_i}^{t_f}\mathrm{d}t\;\bigl[p\dot{q}-H(p,q)\bigr]\biggr\}, \quad (3.2.22)\end{aligned}$$

where we have introduced the standard *time-ordered product* $\mathrm{T}\bigl[Q(t_a)Q(t_b)\bigr]$.

$$\mathrm{T}\bigl[Q(t_a)\,Q(t_b)\bigr] \;:=\; \begin{cases} Q(t_a)\,Q(t_b) & \text{for } t_a \,>\, t_b, \\ Q(t_b)\,Q(t_a) & \text{for } t_b \,>\, t_a. \end{cases} \quad (3.2.23)$$

The extension to a product of M such operators is obvious

$$\begin{aligned}&\langle q_f;t_f|\mathrm{T}\bigl[Q(t_1)Q(t_2)...Q(t_M)\bigr]|q_i;t_i\rangle \\ &\qquad = \int \mathcal{D}q\,\mathcal{D}p\;q(t_1)\,q(t_2)\ldots q(t_M)\;\exp\biggl\{\mathrm{i}\int_{t_i}^{t_f}\mathrm{d}t\;\bigl[p\dot{q}-H(p,q)\bigr]\biggr\}. \quad (3.2.24)\end{aligned}$$

3.2.4 External source terms

We next wish to consider the effect of adding external source terms (or driving forces) into the Hamiltonian, which, as always, we assume does not itself depend explicitly on time. We shall need the unperturbed energy eigenstates $|n\rangle$ and their

corresponding wave-functions:

$$\varphi_n(q,t) \;=\; \langle q,t|n\rangle \;=\; \mathrm{e}^{-\mathrm{i}E_n t}\,\langle q|n\rangle, \quad \text{with} \quad \varphi_n(q) \;=\; \langle q|n\rangle. \tag{3.2.25}$$

Let $\varphi_0(q,t)$ be the lowest-energy or ground state of the system. We aim now calculate the "transition" amplitude for a system in the ground state at some initial instant t_i in the distant past to be in the same state at some final instant t_f in the distant future. In particular, we consider a source term $J(t)Q(t)$, with $J(t)\neq 0$ only during the interval $t_a<t<t_b$, added to the Hamiltonian.* That is, the arbitrary source term is initially switched off and remains zero until $t=t_a$ and then remains active only until $t=t_b$, when it is again set to zero. We thus write

$$\langle q_f;t_f|q_i;t_i\rangle^J \;=\; \int \mathcal{D}q\,\mathcal{D}p\,\exp\!\left\{\mathrm{i}\int_{t_i}^{t_f}\mathrm{d}t\;\left[p\,\dot{q}-H(p,q)+J\,q\right]\right\}. \tag{3.2.26}$$

Inserting complete sets of states, we can now split the amplitude into a product of pieces running from t_i to t_a, to t_b, to t_f:

$$\langle q_f;t_f|q_i;t_i\rangle^J \;=\; \int \mathrm{d}q_a\,\mathrm{d}q_b\,\langle q_f;t_f|q_b;t_b\rangle\langle q_b;t_b|q_a;t_a\rangle^J\langle q_a;t_a|q_i;t_i\rangle. \tag{3.2.27}$$

The first and last elements on the right-hand side are given by expressions such as Eq. (3.2.26) *without* the source term. Consider, for example, the last:

$$\begin{aligned}\langle q_a;t_a|q_i;t_i\rangle \;&=\; \langle q_a|\mathrm{e}^{-\mathrm{i}H(t_a-t_i)}|q_i\rangle\\ &=\; \sum_n \langle q_a|\mathrm{e}^{-\mathrm{i}H(t_a-t_i)}|n\rangle\langle n|q_i\rangle\\ &=\; \sum_n \langle q_a|n\rangle\langle n|q_i\rangle\,\mathrm{e}^{-\mathrm{i}E_n(t_a-t_i)}\\ &=\; \sum_n \varphi_n(q_a)\,\varphi_n^*(q_i)\;\mathrm{e}^{-\mathrm{i}E_n(t_a-t_i)}\,.\end{aligned} \tag{3.2.28}$$

Since the dependence on t_i is explicit, we may make the analytic continuation $t_i\to\mathrm{i}\infty$ (or Wick rotation). The rapidly oscillating imaginary phases then turn into decaying exponentials, which kill all contributions for $n>0$, leaving

$$\lim_{t_i\to\mathrm{i}\infty}\langle q_a;t_a|q_i;t_i\rangle \;=\; \lim_{t_i\to\mathrm{i}\infty}\varphi_0(q_a,t_a)\,\varphi_0^*(q_i,t_i). \tag{3.2.29a}$$

* The external source $J(t)$ is not, of course, really an operator, but we use a capital letter for clarity of notation.

Likewise, the first term becomes

$$\lim_{t_f\to-\mathrm{i}\infty}\langle q_f;t_f|q_b;t_b\rangle \;=\; \lim_{t_f\to-\mathrm{i}\infty}\varphi_0(q_f,t_f)\,\varphi_0^*(q_b,t_b). \tag{3.2.29b}$$

Substituting these back into the original expression, we obtain

$$\lim_{\substack{t_i\to+\mathrm{i}\infty\\ t_f\to-\mathrm{i}\infty}}\frac{\langle q_f;t_f|q_i;t_i\rangle^J}{\exp[-\mathrm{i}\,E_0(t_f-t_i)]\,\varphi_0^*(q_i)\,\varphi_0(q_f)} = \int \mathrm{d}q_a\,\mathrm{d}q_b\,\varphi_0^*(q_b,t_b)\,\langle q_b;t_b|q_a;t_a\rangle^J\,\varphi_0(q_a,t_a). \tag{3.2.30}$$

On examining this equation, we find that the right-hand side is just the ground-state to ground-state amplitude—note that $t_{a,b}$ may be taken as large as we like. This object is what we shall call $Z[J]$.* The left-hand side above then simply provides the method to calculate it.

The generating functional $Z[J]$ also allows the derivation of ground-state to ground-state expectations values of products of fields, as in Eq. (3.2.24), simply by taking functional derivatives with respect to $J(t)$, which pull down factors $\mathrm{i}\varphi(t)$. Taking n such derivatives and then setting $J(t)=0$ leads to

$$\left.\frac{\delta^n Z[J]}{\delta J(t_1)\delta J(t_2)\ldots\delta J(t_n)}\right|_{J(t)=0} = \mathrm{i}^n\int \mathrm{d}q_a\,\mathrm{d}q_b\,\varphi_0^*(q_b,t_b)\,\varphi_0(q_a,t_a) \int \mathcal{D}q\,\mathcal{D}p\,q(t_n)\ldots q(t_2)\,q(t_1)\,\exp\left\{\mathrm{i}\int_{t_i}^{t_f}\mathrm{d}t\;\left[p\dot{q}-H(p,q)\right]\right\}, \tag{3.2.31}$$

where the time variables are ordered $t_i<t_1,t_2,\ldots,t_n<t_f$. This then is just the expectation value of the time-ordered product $\mathrm{T}\left[Q(t_1)Q(t_2)\ldots Q(t_n)\right]$ evaluated between ground states at instants t_i and t_f. These will be the Green functions of the quantised field theory.

Up to constant factors, which may, as always, be absorbed into the overall

* The reader should be warned that in some texts, the symbol $W[J]$ is used. However, in order to underline the close ties between the object defined here and the partition function $Z(\beta)$ in statistical mechanics, we shall use $Z[J]$.

normalisation, on performing the p integral we have

$$\begin{aligned} Z[J] \;&\sim\; \lim_{\substack{t_i\to+\mathrm{i}\infty\\ t_f\to-\mathrm{i}\infty}} \langle q_f;t_f|q_i;t_i\rangle^J \\ &\sim\; \lim_{\substack{t_i\to+\mathrm{i}\infty\\ t_f\to-\mathrm{i}\infty}} \int \mathcal{D}q\, \exp\!\left\{ \mathrm{i}\int_{t_i}^{t_f} \mathrm{d}t\, \left[L_{\text{eff}}(q,\dot{q}) + J(t)\,q(t)\right]\right\}, \end{aligned} \tag{3.2.32}$$

with boundary conditions

$$\lim_{t_i\to-\infty} q(t_i) \;=\; q_i \quad \text{and} \quad \lim_{t_f\to+\infty} q(t_f) \;=\; q_f, \tag{3.2.33}$$

where q_i and q_f are some fixed constants. We can then write

$$\begin{aligned} &\langle \mathrm{T}\big[Q(t_1)Q(t_2)...Q(t_n)\big]\rangle_0 \\ &\sim\; \lim_{\substack{t_i\to+\mathrm{i}\infty\\ t_f\to-\mathrm{i}\infty}} \int \mathrm{d}q_1\, \ldots \mathrm{d}q_n\, \langle q_f;t_f|q_n;t_n\rangle\, q_n \ldots q_2\, \langle q_2;t_2|q_1;t_1\rangle\, q_1\, \langle q_1;t_1|q_i;t_i\rangle, \end{aligned} \tag{3.2.34}$$

where $q_i := q(t_i)$, $t_1 < t_2 < \cdots < t_n$ and $\langle\cdots\rangle_0$ indicates a ground-state expectation value.

We shall again need the Wick rotation or analytic continuation to imaginary time $t_i \to +\mathrm{i}\infty$ and $t_f \to -\mathrm{i}\infty$. To perform the continuation, note that

$$\begin{aligned} \langle q_f;t_f|q_i;t_i\rangle \;=\; \lim_{n\to\infty} \int \prod_{k=1}^{n} \frac{\mathrm{d}q_k}{\left[2\pi\,\mathrm{i}\,\Delta t/m\right]^{1/2}} \\ \times \exp\!\left\{ \mathrm{i}\sum_{j=1}^{n+1} L_{\text{eff}}\!\left(\frac{q_j+q_{j-1}}{2}, \frac{q_j-q_{j-1}}{\Delta t}\right)\Delta t\right\}, \end{aligned} \tag{3.2.35}$$

only depends on time through Δt. We may therefore write

$$\begin{aligned} \langle q_f;t_f|q_i;t_i\rangle\Big|_{\substack{t_i=-\mathrm{i}\tau_i\\ t_f=-\mathrm{i}\tau_f}} \;=\; \lim_{n\to\infty} \int \prod_{k=1}^{n} \frac{\mathrm{d}q_k}{\left[2\pi\Delta\tau/m\right]^{1/2}} \\ \times \exp\!\left\{ \sum_{j=1}^{n+1} L_{\text{eff}}\!\left(\frac{q_j+q_{j-1}}{2}, \frac{q_j-q_{j-1}}{-\,\mathrm{i}\,\Delta\tau}\right)\Delta\tau\right\}, \end{aligned} \tag{3.2.36}$$

where $\Delta\tau := (\tau_f - \tau_i)/(n+1)$. The analytic continuation of the ground-state ex-

pectation value of the time-ordered product of operators thus becomes

$$\langle \mathrm{T}\big[Q(t_1)Q(t_2)...Q(t_n)\big]\rangle_0\Big|_{t_j=-\mathrm{i}\tau_j} \sim \lim_{\substack{\tau_i\to-\infty\\ \tau_f\to+\infty}} \int \mathcal{D}q \;\; q(\tau_n)\ldots q(\tau_2)\,q(\tau_1)\,\exp\left\{\int_{\tau_i}^{\tau_f} \mathrm{d}\tau\, L_{\mathrm{eff}}\left(q, \mathrm{i}\frac{\mathrm{d}q}{\mathrm{d}\tau}\right)\right\}. \quad (3.2.37)$$

Such a move to imaginary time is none other than a Euclidean formulation.* Thus, we shall also define a Euclidean version of the functional generator:

$$Z_{\mathrm{E}}[J] \sim \int \mathcal{D}q\, \exp\left\{\int_{\infty}^{\infty} \mathrm{d}\tau\, \left[L_{\mathrm{eff}}\left(q, \mathrm{i}\frac{\mathrm{d}q}{\mathrm{d}\tau}\right) + J(\tau)\,q(\tau)\right]\right\}, \quad (3.2.38)$$

with the boundary conditions that $q(\tau)$ approaches some constant values for $\tau \to \pm\infty$. Note that the problem of overall normalisation of these expressions is explicitly avoided when relating the Minkowski and Euclidean versions:

$$\frac{1}{Z[J]}\,\frac{\delta^n Z[J]}{\delta J(t_1)\delta J(t_2)\ldots\delta J(t_n)}\bigg|_{J=0} = \frac{\mathrm{i}^n}{Z_{\mathrm{E}}[J]}\,\frac{\delta^n Z_{\mathrm{E}}[J]}{\delta J(\tau_1)\delta J(\tau_2)\ldots\delta J(\tau_n)}\bigg|_{\substack{J=0\\ \tau_i=\mathrm{i}t_i}}. \quad (3.2.39)$$

For an example of the Euclidean formulation of the approach presented here, see the treatment of the harmonic oscillator presented by Feynman (1950) and also Feynman and Hibbs (1965).

It may be worth stressing that we can go back one step to the more general formula, before performing the functional p integration:

$$Z[J] \sim \lim_{\substack{t_i\to+\mathrm{i}\infty\\ t_f\to-\mathrm{i}\infty}} \int \mathcal{D}q\,\mathcal{D}p\, \exp\left\{\mathrm{i}\int_{t_i}^{t_f} \mathrm{d}t\, \left[p\dot{q} - H(q,p) + J(t)\,q(t)\right]\right\}. \quad (3.2.40)$$

In any case, we then finally have

$$\langle \mathrm{T}\big[Q(t_1)Q(t_2)...Q(t_n)\big]\rangle_0 = -\mathrm{i}^n\,\frac{\delta^n Z[J]}{\delta J(t_1)\delta J(t_2)\ldots\delta J(t_n)}\bigg|_{J(t)=0}. \quad (3.2.41)$$

* Note that for $t\to-\mathrm{i}\tau$, we have $x^\mu x_\mu \to -(\tau^2+\boldsymbol{x}^2)$. In other words, $g^{\mu\nu}\to-\mathrm{diag}[1,1,1,1]$.

3.3 Scalar quantum field theory

We now apply the path-integral formulation of quantum theory to the simplest case of a scalar field theory. The idea is that the field $\varphi(t,\boldsymbol{x})$ will become an operator (in the Heisenberg picture). The operator $\hat{\varphi}(t,\boldsymbol{x})$ should act on the occupation-number space—in canonical quantisation the $b(k)$ and $b^*(k)$ create and destroy particles (seen as occupation levels), in analogy with, *e.g.* the ladder operators $a_\pm := p \pm \mathrm{i}q$ (or a and a^*) for the harmonic oscillator.

However, the important point to bear in mind is that the approach outlined so far is an *alternative* to the canonical quantisation procedure via substitution of operators: indeed, the procedure is far more direct and conceptually simpler. One merely needs to identify the Lagrangian governing the *classical* dynamics and use this in the path integral, which is then unambiguously defined.

The extension of the path integral to N degrees of freedom was given in Eq. (3.2.18). This can now be applied to a scalar field theory as follows. We divide the spatial volume into cubic sub-volumes of dimension $(\Delta x)^3$ and use the index a, the index to the vectors in Eq. (3.2.18), to label them: V_a. The a-th component of the vector $\boldsymbol{q}(t)$ thus becomes $\varphi_a(t)$, which we *define* by

$$\varphi_a(t) \;:=\; (\Delta x)^{-3} \int_{V_a} \mathrm{d}^3\boldsymbol{x}\, \varphi(t,\boldsymbol{x}). \tag{3.3.1}$$

It is thus the average of φ over the cell volume (we might equally have taken simply the value of φ at the cell centre). The Lagrangian then becomes

$$L(t) \;\equiv\; \int \mathrm{d}^3\boldsymbol{x}\; \mathcal{L}(\boldsymbol{x},t) \;\to\; \sum_a (\Delta x)^3\; \mathcal{L}\big(\dot{\varphi}_a(t), \varphi_a(t)\big)\,. \tag{3.3.2}$$

As usual, we may now define the canonical momenta

$$p_a(t) \;:=\; \frac{\partial L}{\partial \dot{\varphi}_a(t)} \;=\; (\Delta x)^3\, \frac{\partial \mathcal{L}}{\partial \dot{\varphi}_a(t)} \;=:\; (\Delta x)^3\, \pi_a(t). \tag{3.3.3}$$

The Hamiltonian is then

$$H \;=\; \sum_a p_a\, \dot{\varphi}_a(t) - L \;:=\; \sum_a (\Delta x)^3\; \mathcal{H}_a, \tag{3.3.4}$$

where, naturally, the Hamiltonian density is just

$$\mathcal{H}\big(\pi_a(t), \varphi_a(t)\big) \;=\; \pi_a(t)\, \dot{\varphi}_a(t) - \mathcal{L}\big(\dot{\varphi}_a(t), \varphi_a(t)\big)\,. \tag{3.3.5}$$

With these definitions, the path integral (3.2.18) becomes

$$\lim_{\substack{n\to\infty\\ \Delta x\to 0}} \int \prod_a \left(\prod_{k=1}^{n} \mathrm{d}\varphi_a(t_k) \prod_{k=1}^{n+1} (\Delta x)^3 \mathrm{d}\pi_a(t_k) \right) \times \exp\left\{ \mathrm{i} \sum_{j=1}^{n+1} \Delta t \sum_b (\Delta x)^3 \left[\pi_b(t_j) \frac{\varphi_b(t_j) - \varphi_b(t_{j-1})}{\Delta t} - \mathcal{H}\big(\pi_b(t_j), \varphi_b(t_j)\big) \right] \right\}$$
$$= \int \mathcal{D}\varphi\, \mathcal{D}\pi\, \exp\left\{ \mathrm{i} \int_{t_i}^{t_f} \mathrm{d}t\, \mathrm{d}^3\boldsymbol{x} \left[\pi(t,\boldsymbol{x})\, \dot{\varphi}(t,\boldsymbol{x}) - \mathcal{H}(t,\boldsymbol{x}) \right] \right\}, \quad (3.3.6)$$

where, following our earlier formulation, we have defined $\pi(t,\boldsymbol{x})$, the field conjugate to $\varphi(t,\boldsymbol{x})$, by

$$\pi(t,\boldsymbol{x}) \;:=\; \frac{\partial \mathcal{L}}{\partial \dot{\varphi}(t,\boldsymbol{x})}. \quad (3.3.7)$$

In analogy with the quantum-mechanical case, the path integral runs over all field configurations $\varphi(t,\boldsymbol{x})$ and $\pi(t,\boldsymbol{x})$, subject to boundary conditions of the form

$$\varphi(t_i,\boldsymbol{x}) \;=\; \varphi_i(\boldsymbol{x}) \quad \text{and} \quad \varphi(t_f,\boldsymbol{x}) \;=\; \varphi_f(\boldsymbol{x}). \quad (3.3.8)$$

Let us stress once more, the quantities involved here, $\varphi(t,\boldsymbol{x})$ and $\pi(t,\boldsymbol{x})$, remain *classical* field configurations and *not* operators.

3.3.1 The generating functional $Z[J]$

The natural extension of Eq. (3.2.26) to a field is simply

$$Z[J] \;\propto\; \langle \varphi_f(\boldsymbol{x}),t_f | \varphi_i(\boldsymbol{x}),t_i \rangle^J$$
$$= \int \mathcal{D}\varphi\, \mathcal{D}\pi\, \exp\left\{ \mathrm{i} \int_{t_i}^{t_f} \mathrm{d}^4x \left[\pi(x)\, \dot{\varphi}(x) - \mathcal{H}\big(\pi(x),\varphi(x)\big) + J(x)\,\varphi(x) \right] \right\}. \quad (3.3.9)$$

The normalisation of $Z[J]$ is chosen such that $Z[J]\big|_{J=0} = 1$ and, as before, the auxiliary, external, source term $J(x)\varphi(x)$ serves to study the generic ground-state transition amplitude. Since we are dealing here with the description of particle dynamics in a relativistic context and, in particular, we expect creation and destruction to occur, it is natural to take the *ground state* as coinciding with the *vacuum*; that is, the *empty* state.* Following the example from non-relativistic

* That the vacuum state should be the lowest-energy state is *not* guaranteed *a priori*; it is indeed possible that non-trivial interactions lead to configurations with non-zero expectation value for

quantum mechanics, *i.e.* Eq. (3.2.41), we shall thus define the n-point Green function:

$$\begin{aligned} \mathcal{G}^{(n)}(x_1,\ldots,x_n) &:= (-\mathrm{i})^n \frac{\delta^n Z[J]}{\delta J(x_1)\ldots\delta J(x_n)}\bigg|_{J=0} \\ &= \langle 0|\mathrm{T}\big[\varphi(x_1)\ldots\varphi(x_n)\big]|0\rangle, \end{aligned} \tag{3.3.10}$$

which is none other than the vacuum expectation value of the time-ordered product of n fields.

On a historical note, the relationship between $\mathcal{G}^{(n)}$ and the expectation value of the time-ordered product of fields was originally derived by Schwinger (1951) within the context of his source theory, but without the auxiliary of the path integral. The important point of the path-integral formulation, apart from the elegant proof it permits, is that it also provides an explicit (albeit formal) expression for calculation.

Recall that all quantities are classical and therefore commute; we may thus rewrite $Z[J]$ as an expansion in powers of J, for which the Green functions are then just the coefficients:

$$Z[J] = \sum_{n=0}^{\infty} \frac{\mathrm{i}^n}{n!} \int \mathrm{d}^4x_1 \ldots \mathrm{d}^4x_n \, \mathcal{G}^{(n)}(x_1,\ldots,x_n)\, J(x_1)\ldots J(x_n), \tag{3.3.11}$$

where the first term ($n=0$) is just 1, representing the trivial zero-point function.

We can again perform the analytic continuation to Euclidean space–time (*i.e.* Wick rotate to imaginary time). However, since the process should be well understood by now, we shall remain in Minkowski space and simply appeal, when necessary, to the possibility of a Euclidean formulation in order to justify the various operations. To render the transition from one to the other unambiguous, we may introduce an "$\mathrm{i}\varepsilon$" term (equivalent to a complex mass or finite lifetime):

$$\mathcal{H} \to \mathcal{H} - \tfrac{1}{2}\mathrm{i}\varepsilon\varphi^2 \qquad (\varepsilon = 0^+). \tag{3.3.12}$$

As before, the special case, in which the Lagrangian density has the simple separated form (*i.e.* the φ's and $\dot{\varphi}$'s are not mixed)

$$\mathcal{L}(\varphi, \partial_\mu\varphi) = \tfrac{1}{2}\dot{\varphi}^2 + \mathcal{F}(\varphi, \boldsymbol{\nabla}\varphi), \tag{3.3.13}$$

some fields and yet with lower energy than the *naïve* vacuum (as in the case of the phenomenon known as spontaneous symmetry breaking, which we shall study later). For the moment we shall simply ignore such possible complications.

allows us to perform the π integral explicitly. We start from, *cf.* (3.3.9),

$$Z[J] \;=\; \int \mathcal{D}\varphi\,\mathcal{D}\pi\,\exp\!\left\{\mathrm{i}\int_{t_i}^{t_f}\mathrm{d}^4x\,\left[\pi\,\dot{\varphi}-\tfrac{1}{2}\pi^2-\mathcal{F}(\varphi,\boldsymbol{\nabla}\varphi)+J\varphi\right]\right\}. \tag{3.3.14}$$

Completing the square in π and performing the Gaussian functional integral, leads directly to

$$\begin{aligned} Z[J] \;&\propto\; \int \mathcal{D}\varphi\,\exp\!\left\{\mathrm{i}\int_{t_i}^{t_f}\mathrm{d}^4x\,\left[\tfrac{1}{2}\dot{\varphi}^2+\mathcal{F}(\varphi,\boldsymbol{\nabla}\varphi)+J\varphi\right]\right\} \\ &=\; \int \mathcal{D}\varphi\,\exp\!\left\{\mathrm{i}\int_{t_i}^{t_f}\mathrm{d}^4x\,\left[\mathcal{L}(\varphi,\partial_\mu\varphi)+J\varphi\right]\right\}. \end{aligned} \tag{3.3.15}$$

As usual, the normalisation is fixed by requiring $Z[J]\big|_{J=0}=1$. Again, this simple result is obtained *only* by virtue of the particular form of $\mathcal{L}$. Otherwise, the integral in π must be performed explicitly first and the final result will then depend on some $\mathcal{L}_{\text{eff}}\neq\mathcal{L}$.

3.3.2 The free-field case

Let us first examine the simple free-field case in which $\mathcal{F}$ does not contain terms higher than quadratic in φ. That is, the Lagrangian takes the form

$$\mathcal{L} \;\to\; \mathcal{L}_0 \;=\; \tfrac{1}{2}[(\partial^\mu\varphi)(\partial_\mu\varphi)-m^2\varphi^2], \tag{3.3.16}$$

where the subscript "$_0$" stands for "free". For clarity, we return to the discrete formulation and for simplicity we shall choose $\Delta t=\Delta=\Delta x$. Thus,

$$\begin{aligned} Z_0[J] \;&\propto\; \int \mathcal{D}\varphi\,\exp\!\left\{\mathrm{i}\int_{t_i}^{t_f}\mathrm{d}^4x\,\left[\mathcal{L}_0(\varphi,\partial_\mu\varphi)+J\varphi\right]\right\} \\ &\to\; \lim_{\Delta\to 0}\int\prod_a \mathrm{d}\varphi_a\,\exp\!\left\{\mathrm{i}\,\left[\Delta^8\sum_{a,b}\tfrac{1}{2}\varphi_a\,\mathcal{K}_{ab}\,\varphi_b+\Delta^4\sum_a J_a\varphi_a\right]\right\}, \end{aligned} \tag{3.3.17}$$

The matrix operator $\mathcal{K}_{ab}$ is such that in the limit we have

$$\lim_{\Delta\to 0}\mathcal{K}_{ab} \;=\; -\left[\Box+m^2-\mathrm{i}\varepsilon\right]\delta^4(x-y) \;=:\; \mathcal{K}(x,y), \tag{3.3.18}$$

with $a \to x$ and $b \to y$ as $\Delta \to 0$; we shall call this the Klein–Gordon operator. The φ_a integrations can be performed by exploiting the formulæ derived earlier:

$$Z_0[J] \;=\; \lim_{\Delta\to 0} \frac{1}{\sqrt{\det \mathcal{K}_{ab}}} \prod_c \sqrt{\frac{2\pi}{i\Delta^8}} \exp\!\left\{ -\tfrac{1}{2}\,\mathrm{i} \sum_{a,b} J_a (\mathcal{K}^{-1})_{ab} J_b \right\}, \tag{3.3.19}$$

where the inverse of $\mathcal{K}_{ab}$ (the Green function) is naturally defined by

$$\sum_c (\mathcal{K}^{-1})_{ac}\, \mathcal{K}_{cb} \;=\; \delta_{ab}. \tag{3.3.20}$$

Note that the continuum limits for the delta and sum are:

$$\Delta^{-4}\delta_{ab} \;\to\; \delta^4(x-y) \quad \text{and} \quad \sum_a \Delta^4 \;\to\; \int \mathrm{d}^4 x\,. \tag{3.3.21}$$

In the continuum then (with the usual normalisation) we have just

$$Z_0[J] \;=\; \exp\!\left\{ -\tfrac{1}{2}\,\mathrm{i} \int \mathrm{d}^4x\, \mathrm{d}^4y\, J(x)\; \mathcal{K}^{-1}(x,y)\, J(y) \right\}. \tag{3.3.22}$$

We now rewrite the inverse of $\mathcal{K}(x,y)$ in the following suggestive form:

$$\mathcal{K}^{-1}(x,y) \;=\; \Delta_{\mathrm{F}}(x-y). \tag{3.3.23}$$

This is known as the *Feynman propagator* and we shall now show that it really does only depend on the difference $(x-y)$. From the definition of $\mathcal{K}$, we can write

$$\mathcal{K}(x,y) \;=\; \int \text{đ}^4 p\; \mathrm{e}^{\mathrm{i}p\cdot(x-y)} \left[p^2 - m^2 + \mathrm{i}\varepsilon \right], \tag{3.3.24}$$

from which it immediately follows that

$$\mathcal{K}^{-1}(x,y) \;=\; \int \text{đ}^4 p\; \mathrm{e}^{\mathrm{i}p\cdot(x-y)} \left[p^2 - m^2 + \mathrm{i}\varepsilon \right]^{-1}. \tag{3.3.25}$$

This is indeed the Feynman propagator (or Green function), as may be seen by considering that from the above we have

$$\left[\Box_x + m^2 \right] \Delta_{\mathrm{F}}(x-y) \;=\; -\int \text{đ}^4 p\; \mathrm{e}^{\mathrm{i}p\cdot(x-y)} \;=\; -\delta^4(x-y). \tag{3.3.26}$$

And thus Δ_{F} is associated with the propagation of solutions to the Klein–Gordon

equation

$$\left[\Box + m^2\right] \varphi(x) \;=\; 0. \tag{3.3.27}$$

One should note that the "$\mathrm{i}\varepsilon$" prescription, as applied here (via the additional $\mathrm{i}\varepsilon\varphi^2$ term in the Lagrangian), cannot be readily generalised to the fermion or gauge-field cases (owing to the presence of negative-energy solutions in the former case and the preclusion of a mass term in the latter). We shall thus, in general, simply insert an $\mathrm{i}\varepsilon$ into the propagator by hand. The final form of the free-field generating functional is then

$$Z_0[J] \;=\; \exp\left\{-\mathrm{i}\int \mathrm{d}^4x\,\mathrm{d}^4y\,\tfrac{1}{2}J(x)\,\Delta_{\mathrm{F}}(x-y)\,J(y)\right\}, \tag{3.3.28}$$

with Δ_{F} defined as above.

3.3.3 Free-field Green functions

As already seen, functional derivatives of $Z[J]$ with respect to J generate the n-point Green functions of the theory. The formula just derived then provides an explicit method to calculate them. First of all, note that an odd number of derivatives always leaves one factor of J, which causes the entire term to vanish on setting $J=0$; we need thus only consider n even. We shall present $n=2$ and 4, the continuation to $n\geq 6$ being straightforward and $\mathcal{G}_0^{(0)}$ is, of course, unity.

$$\mathcal{G}_0^{(2)}(x_1,x_2) \;=\; \mathrm{i}\,\Delta_{\mathrm{F}}(x_1-x_2), \tag{3.3.29a}$$

$$\begin{aligned}\mathcal{G}_0^{(4)}(x_1,x_2,x_3,x_4) \;=\; \mathrm{i}\big[&\Delta_{\mathrm{F}}(x_1-x_2)\,\Delta_{\mathrm{F}}(x_3-x_4)\\ &+\Delta_{\mathrm{F}}(x_1-x_3)\,\Delta_{\mathrm{F}}(x_2-x_4)\\ &+\Delta_{\mathrm{F}}(x_1-x_4)\,\Delta_{\mathrm{F}}(x_2-x_3)\big].\end{aligned} \tag{3.3.29b}$$

Note that the full interacting theory will also generate terms for n odd (even $n=1$ may occur). This simple mathematical structure leads to a natural diagrammatical representation, in which, first of all, the free-field propagator or two-point (Green) function is indicated symbolically by a line:

$$\mathcal{G}_0^{(2)}(x_1,x_2) \;=\; x_1 \text{ ---- } x_2 \tag{3.3.30a}$$

and the free-field four-point function is then

$$\mathcal{G}_0^{(4)}(x_1,x_2,x_3,x_4) \;=\; \begin{array}{l} x_1 \quad\text{- - - -}\quad x_3 \\ \\ x_2 \quad\text{- - - -}\quad x_4 \end{array} \;+ (x_2 \leftrightarrow x_3) + (x_3 \leftrightarrow x_4). \qquad (3.3.30b)$$

Given that $\mathcal{G}^{(2)}(x_1,x_2)$ is associated with the propagation of one field between the space–time points x_1 and x_2, then it is natural to assume that $\mathcal{G}^{(4)}(x_1,x_2,x_3,x_4)$ will have to do with the propagation of two fields between the space–time points x_1, x_2, x_3 and x_4. We shall later show how this comes about.

It turns out that calculations are more easily performed in momentum space; it is therefore useful to apply a Fourier transform:

$$\mathcal{G}^{(n)}(p_1,\ldots,p_n)\,\tilde{\delta}^4(p_1+\cdots+p_n) := \int \mathrm{d}^4x_1\,\ldots\mathrm{d}^4x_n\;\mathrm{e}^{\mathrm{i}(p_1\cdot x_1+\cdots+p_n\cdot x_n)}\;\mathcal{G}^{(n)}(x_1,\ldots,x_n). \quad (3.3.31)$$

The energy–momentum conserving δ-function has been written explicitly since translation invariance implies that Green functions only depend on space–time differences. Moreover, we do not distinguish symbolically between $\mathcal{G}^{(n)}$ for ordinary space–time and momentum space since, given the arguments, there is evidently no ambiguity. For the free-field two-point function, we thus write

$$\mathcal{G}_0^{(2)}(p,-p) \;=\; \mathrm{i}\Delta_{\mathrm{F}}(p) \;=\; \frac{\mathrm{i}}{p^2-m^2+\mathrm{i}\varepsilon}. \qquad (3.3.32)$$

The corresponding diagrammatic representation is just

$$\mathcal{G}_0^{(2)}(p,-p) \;=\; \mathrm{i}\Delta_{\mathrm{F}}(p) \;=\; \underset{p}{\text{- - - - -}}\ . \qquad (3.3.33)$$

Note that the energy–momentum here has a specific diagrammatic direction of flow and so sometimes an explicit arrow is attached to the propagator line (although this is only strictly necessary when there is transport of conserved quantum numbers).

3.3.4 Connected Green functions

As shown in Eq. (3.3.29b), the Green functions generated by differentiating $Z[J]$ are *disconnected*. However, there can evidently be no interesting dynamics contained in such processes; *e.g.* $\mathcal{G}_0^{(4)}$ merely describes the independent propagation of two non-interacting fields. We shall now show how to exclude such contributions

and leave only *connected* processes.

To this end, we define a new generating functional $Z_c[J]$ by

$$\exp\left\{\mathrm{i}\,Z_c[J]\right\} \;:=\; Z[J], \tag{3.3.34a}$$

or equivalently

$$\mathrm{i}\,Z_c[J] \;:=\; \ln Z[J]. \tag{3.3.34b}$$

We then see that

$$\begin{aligned}\frac{\mathrm{i}\delta^n Z_c[J]}{\delta J_1\delta J_2\ldots\delta J_n} \;=\; & \frac{1}{Z[J]}\,\frac{\delta^n Z[J]}{\delta J_1\delta J_2\ldots\delta J_n} \\ & -\left\{\frac{1}{Z[J]}\,\frac{\delta Z[J]}{\delta J_1}\,\frac{1}{Z[J]}\,\frac{\delta^{n-1}Z[J]}{\delta J_2\ldots\delta J_n}+\text{perms}\right\} \\ & -\left\{\frac{1}{Z[J]}\,\frac{\delta^2 Z[J]}{\delta J_1\delta J_2}\,\frac{1}{Z[J]}\,\frac{\delta^{n-2}Z[J]}{\delta J_3\ldots\delta J_n}+\text{perms}\right\}-\ldots \\ & +\left\{\frac{1}{Z[J]}\,\frac{\delta Z[J]}{\delta J_1}\,\frac{1}{Z[J]}\,\frac{\delta Z[J]}{\delta J_2}\,\frac{1}{Z[J]}\,\frac{\delta^{n-2}Z[J]}{\delta J_3\ldots\delta J_n}+\text{perms}\right\}+\ldots,\end{aligned} \tag{3.3.35}$$

where we have adopted the shorthand $J_n := J(x_n)$. It is almost immediate that in a free-field theory the only connected Green function is the two-point function. To see this, let us express $Z_c[J]$ as a power series, of which the coefficients should now be connected Green functions:

$$\mathrm{i}\,Z_c[J] \;=\; \sum_{n=0}^{\infty}\frac{\mathrm{i}^n}{n!}\int \mathrm{d}^4x_1\,\mathrm{d}^4x_2\,\ldots\mathrm{d}^4x_n\,\mathcal{G}_c^{(n)}(x_1,x_2,\ldots,x_n)\,J(x_1)J(x_2)\ldots J(x_n). \tag{3.3.36}$$

Note that, since we require $\mathrm{e}^{\mathrm{i}\,Z_c[J]}|_{J=0}=1$ for a free-field theory, the $n=0$ term (vacuum diagrams) vanishes, as too the $n=1$ term (so-called tadpole diagrams). They have been retained, however, for complete generality.

Comparison of Eq. (3.3.34a) and Eq. (3.3.28) immediately leads to

$$Z_{c0}[J] \;=\; -\tfrac{1}{2}\int \mathrm{d}^4x\,\mathrm{d}^4y\,J(x)\,\Delta_{\mathrm{F}}(x-y)\,J(y) \tag{3.3.37}$$

and thus

$$\mathcal{G}_{0\,c}^{(2)}(x_1-x_2) \;=\; \mathcal{G}_0^{(2)}(x_1-x_2) \;=\; \mathrm{i}\Delta_{\mathrm{F}}(x_1-x_2). \tag{3.3.38}$$

Once we move over to the full interacting theory, there will also be connected diagrams with $n\neq 2$ external fields. Moreover, even the simple two-point function will have higher-order contributions (quantum corrections) and removal of the disconnected terms will then become important. We shall, however, postpone

further examination of this question until later.

What may not be so obvious is that the terms following the first on the right-hand side of Eq. (3.3.35) indeed have the function of removing the disconnected contributions. We shall now demonstrate this. First let us comment on the $Z[J]$ factors in the denominators. Although the free-field version of $Z[J]$ for $J=0$ is unity, the interacting theory generates higher-order (in perturbation theory) contributions. These are purely vacuum diagrams (bubbles), which are effectively an overall renormalisation with no physical importance. Precisely the same sum of bubble graphs appears multiplying all higher-order Green functions. The denominator $Z[J]$ associated with each set of derivatives of $Z[J]$ in Eq. (3.3.35) then simply cancels this irrelevant overall renormalisation factor in all cases.

The simplest proof of the connectedness of $\mathcal{G}_c^{(n)}$ proceeds via induction. We shall therefore first examine the lowest-order derivatives:

$$\frac{\mathrm{i}\delta Z_c[J]}{\delta J_1} = \frac{1}{Z[J]}\frac{\delta Z[J]}{\delta J_1}, \tag{3.3.39a}$$

$$\begin{aligned}\frac{\mathrm{i}\delta^2 Z_c[J]}{\delta J_1\delta J_2} &= \frac{1}{Z[J]}\frac{\delta^2 Z[J]}{\delta J_1\delta J_2} - \frac{1}{Z[J]}\frac{\delta Z[J]}{\delta J_1}\frac{1}{Z[J]}\frac{\delta Z[J]}{\delta J_2} \\ &= \frac{1}{Z[J]}\frac{\delta^2 Z[J]}{\delta J_1\delta J_2} - \frac{\mathrm{i}\delta Z_c[J]}{\delta J_1}\frac{\mathrm{i}\delta Z_c[J]}{\delta J_2}.\end{aligned} \tag{3.3.39b}$$

The first of these shows that, apart from the renormalisation discussed above, the connected and disconnected one-point functions are (rather obviously) identical. The second explicitly displays cancellation of the possible disconnected two-point contribution formed by the product of two (connected) one-point diagrams.

Exercise 3.3.1. *Evaluate the third-order derivative, show explicitly by regrouping of terms and substitution that the connected three-point function is the disconnected three-point function with the three possible products of connected one- and two-point functions together with the product of three one-point functions subtracted.*

Consider now the n-th derivative of Eq. (3.3.34a): dividing out a factor of $Z[J]$, we obtain

$$\begin{aligned}\frac{1}{Z[J]}\frac{\delta^n Z[J]}{\delta J_1\delta J_2\ldots\delta J_n} &= \frac{\mathrm{i}\delta^n Z_c[J]}{\delta J_1\delta J_2\ldots\delta J_n} + \left\{\frac{\mathrm{i}\delta Z_c[J]}{\delta J_1}\frac{\mathrm{i}\delta^{n-1} Z_c[J]}{\delta J_2\ldots\delta J_n} + \text{perms}\right\} \\ &\quad + \left\{\frac{\mathrm{i}\delta^2 Z_c[J]}{\delta J_1\delta J_2}\frac{\mathrm{i}\delta^{n-2} Z_c[J]}{\delta J_3\ldots\delta J_n} + \text{perms}\right\} + \ldots \\ &\quad + \left\{\frac{\mathrm{i}\delta Z_c[J]}{\delta J_1}\frac{\mathrm{i}\delta Z_c[J]}{\delta J_2}\frac{\mathrm{i}\delta^{n-2} Z_c[J]}{\delta J_3\ldots\delta J_n} + \text{perms}\right\} + \ldots\end{aligned} \tag{3.3.40}$$

By hypothesis, all of the first $n-1$ orders of derivatives appearing on the right-hand side generate connected functions. The terms after the first on the right-hand side thus generate all possible disconnected contributions to the n-point function. It therefore follows that the first term generates the connected n-point contribution. Since $\mathcal{G}_c^{(1)}$ and $\mathcal{G}_c^{(2)}$ have been shown explicitly to be connected, the inductive proof is complete.

3.3.5 The effective action

We have

$$Z[J] \;=\; \mathrm{e}^{\,\mathrm{i}\,Z_c[J]} \;=\; \int \mathcal{D}\varphi\, \exp\!\left\{ \mathrm{i} \int \mathrm{d}^4 y\, \Big[\mathcal{L}(\varphi, \partial_\mu \varphi) + J(y)\, \varphi(y) \Big] \right\}. \tag{3.3.41}$$

Now, taking the first functional derivative of $Z_c[J]$, we obtain

$$\frac{\delta Z_c[J]}{\delta J(x)} \;=\; \frac{1}{Z[J]} \int \mathcal{D}\varphi\, \varphi(x)\, \exp\!\left\{ \mathrm{i} \int \mathrm{d}^4 y\, \Big[\mathcal{L}(\varphi, \partial_\mu \varphi) + J(y)\, \varphi(y) \Big] \right\}. \tag{3.3.42}$$

This object is just the ground-state expectation value of $\varphi(x)$ in the presence of the source $J(x)$. In other words, it is just the *classical field* (but more on this shortly). The classical field so-defined has some importance, as we shall see later, when discussing *spontaneous symmetry breaking*. Note though that it is a *functional* of $J(x)$. We therefore define

$$\varphi_c(x) \;:=\; \frac{\delta Z_c[J]}{\delta J(x)}, \tag{3.3.43}$$

where the index c stands for *classical* on the left-hand side but *connected* on the right-hand side. And writing $Z_c[J] = \langle 0|0\rangle_J$, we thus have

$$\varphi_c(x) \;=\; \frac{\langle 0|\varphi(x)|0\rangle_J}{\langle 0|0\rangle_J}, \tag{3.3.44}$$

where $|0\rangle$ is used to indicate the vacuum or ground state.

Let us now perform a Legendre transform on $Z_c[J]$:

$$\Gamma[\varphi_c] \;:=\; Z_c[J] - \int \mathrm{d}^4 x\, J(x)\, \varphi_c(x). \tag{3.3.45}$$

Such a transform is well-known in classical mechanics and statistical thermodynamics; it is analogous here, *e.g.*, to the relation $F = E - TS$ relating the Helmholtz free energy F to the entropy S.

Exercise 3.3.2. *Take the functional derivative of* $\Gamma[\varphi_c]$ *with respect to* $J(x)$ *and*

thus show that $\Gamma[\varphi_c]$ *indeed only depends on* $\varphi_c(x)$ *and not on* $J(x)$.

The symmetry of the equations is evident and $J(x)$ can now be obtained by taking a functional derivative of $\Gamma[\varphi_c]$ with respect to $\varphi_c(x)$:

$$J(x) \;=\; -\frac{\delta\Gamma[\varphi_c]}{\delta\varphi_c(x)}. \tag{3.3.46}$$

The significance of the dual pair of equations (3.3.43 & 46) is that in the absence of J (*i.e.* $J=0$) Eq. (3.3.43) gives precisely the true classical field, which in turn is seen to extremise (usually minimise) the functional $\Gamma[\varphi_c]$. We shall thus be led to interpret $\Gamma[\varphi_c]$ as the *effective action*.

From the definition of φ_c, in the free-field case we have

$$\varphi_c(x) \;=\; -\int \mathrm{d}^4y\, \Delta_{\mathrm{F}}(x-y)\, J(y). \tag{3.3.47}$$

Now, since $\Delta_{\mathrm{F}}(x-y)$ is precisely the Green function for the Klein–Gordon equation, $\varphi_c(x)$ then satisfies

$$\left[\Box + m^2\right] \varphi_c(x) \;=\; J(x), \tag{3.3.48}$$

which, indeed, is none other than the *classical* field equation in the presence of an external source $J(x)$.

With these definitions in hand, we may now proceed with the calculation of the effective action $\Gamma_0[\varphi_c]$ for the free-field case. Let us first add one more definition or shorthand for the Klein–Gordon operator (*i.e.* the same symbol as used earlier but now having only *one* argument; so there should be no confusion):

$$\mathcal{K}(x) \;:=\; \Box_x + m^2. \tag{3.3.49}$$

We may start from Eq. (3.3.37) and use (3.3.48) to eliminate $J(x)$:

$$\begin{aligned} Z_c[J] \;&=\; -\tfrac{1}{2}\int \mathrm{d}^4x\, \mathrm{d}^4y\, J(x)\, \Delta_{\mathrm{F}}(x-y)\, J(y) \\ &=\; -\tfrac{1}{2}\int \mathrm{d}^4x\, \mathrm{d}^4y\, \left[\mathcal{K}(x)\, \varphi_c(x)\right] \Delta_{\mathrm{F}}(x-y)\, \left[\mathcal{K}(y)\, \varphi_c(y)\right]. \end{aligned} \tag{3.3.50}$$

where the square brackets are used to delimit the action of the operator. Thus,

$$\begin{aligned} \Gamma_0[\varphi_c] \;=\; -\tfrac{1}{2}\int \mathrm{d}^4x\, \mathrm{d}^4y\, \left[\mathcal{K}(x)\, \varphi_c(x)\right] \Delta_{\mathrm{F}}(x-y)\, \left[\mathcal{K}(y)\, \varphi_c(y)\right] \\ -\int \mathrm{d}^4x\, \left[\mathcal{K}(x)\varphi_c(x)\right] \varphi_c(x) \end{aligned}$$

$$
\begin{aligned}
&= +\tfrac{1}{2}\int \mathrm{d}^4x\,\varphi_c(x)\,[\mathcal{K}(x)\,\varphi_c(x)] - \int \mathrm{d}^4x\,\,[\mathcal{K}(x)\varphi_c(x)]\,\varphi_c(x) \\
&= -\int \mathrm{d}^4x\,\tfrac{1}{2}\left[\partial^\mu\partial_\mu\varphi_c(x) + m^2\varphi_c(x)\right]\varphi_c(x) \\
&= \int \mathrm{d}^4x\,\tfrac{1}{2}\left[\partial^\mu\varphi_c(x)\,\partial_\mu\varphi_c(x) - m^2\varphi_c^2(x)\right] \\
&= \int \mathrm{d}^4x\,\,\mathcal{L}_0(\varphi_c,\partial_\mu\varphi_c),
\end{aligned}
\tag{3.3.51}
$$

where in the various steps we have integrated by parts and exploited the relation $\mathcal{K}^{-1}(x) = \Delta_{\mathrm{F}}(x)$. The final expression is just the classical action corresponding to the classical free-field; hence the previous choice of the term *effective action*.

Such an interpretation takes on a particularly useful meaning in the full interacting theory. While interactions generally render the theory insoluble in closed form (owing to the induced quantum corrections), we may still perform a formal functional expansion in powers of the field $\varphi_c(x)$:

$$
\Gamma[\varphi_c] \;=\; \sum_{n=1}^{\infty}\frac{\mathrm{i}^n}{n!}\int \mathrm{d}^4x_1\,\ldots\mathrm{d}^4x_n\,\Gamma^{(n)}(x_1,\ldots,x_n)\,\varphi_c(x_1)\ldots\varphi_c(x_n). \tag{3.3.52}
$$

It will turn out that the coefficients $\Gamma^{(n)}$ correspond not only to n-point functions that are connected, but that are also what are termed *one-particle irreducible* (1PI); that is, they cannot be rendered disconnected by cutting just a single *internal* line. Some examples of 1PI and non-1PI functions are shown in Fig. 3.1.

Figure 3.1: Examples of (a–c) 1PI and (d) non-1PI or one-particle reducible diagrams.

In the free-field case, as already shown, there is only one such non-vanishing n-point 1PI function: that for $n=2$,

$$
\Gamma_0^{(2)}(x,y) \;=\; \mathcal{K}(x)\,\delta^4(x-y). \tag{3.3.53}
$$

Just as for $\mathcal{G}^{(n)}$, we may now perform a Fourier transform and define the momentum-

space versions:

$$\Gamma^{(n)}(p_1, p_2, \ldots, p_n)\ \tilde{\delta}^4(p_1 + p_2 + \cdots + p_n)$$
$$:= \int \mathrm{d}^4x_1\ \ldots \mathrm{d}^4x_n\ \mathrm{e}^{\mathrm{i}(p_1\cdot x_1 + \cdots + p_n\cdot x_n)}\ \Gamma^{(n)}(x_1, x_2, \ldots, x_n). \quad (3.3.54)$$

The free-field two-point momentum-space 1PI function is then

$$\Gamma_0^{(2)}(p, -p)\ =\ -(p^2 - m^2). \quad (3.3.55)$$

One can also make a low-energy or Taylor expansion, which is equivalent to an expansion in powers of momenta:

$$\Gamma[\varphi_c]\ =\ \int \mathrm{d}^4x\ \left[-\mathcal{V}(\varphi_c) + \tfrac{1}{2}\mathcal{A}(\varphi_c)\,\partial^\mu\varphi_c\,\partial_\mu\varphi_c + \ldots\right], \quad (3.3.56)$$

where now the coefficients $\mathcal{V}(\varphi_c)$, $\mathcal{A}(\varphi_c)$, ... are simple functions of φ_c and hence of x. The first term $\mathcal{V}(\varphi_c)$ is then called the *effective potential*. Note that in the case where the classical field is space–time independent (so that only this term survives), from Eq. (3.3.46) we have

$$\frac{\partial\mathcal{V}(\varphi_c)}{\partial\varphi_c(x)}\ =\ J(x). \quad (3.3.57)$$

Setting $J=0$, we then see that the classical field configuration φ_c is that which extremises* the effective potential. In other words, once we know $\mathcal{V}(\varphi_c)$, then we can immediately obtain the vacuum expectation value of the field φ_c. Of course, if φ_c is not constant, then we must minimise the full expression for $\Gamma[\varphi_c]$.

We should perhaps remark here that one can show that higher derivatives of $Z_c[J]$ for $J=0$ are actually the n-point connected Green functions for the true minimum; *i.e.* for the *shifted* field $\tilde{\varphi} := \varphi - \nu$, where

$$\nu\ :=\ \left.\frac{\delta Z_c[J]}{\delta J(x)}\right|_{J=0}. \quad (3.3.58)$$

That is, for $\tilde{\varphi}$ with vanishing vacuum expectation value. Note that, since the right-hand side is translationally invariant, ν is necessarily space–time independent.

Exercise 3.3.3. *By differentiating* (3.3.42) *with respect to* J *(remembering to*

* Since one can show that $\mathcal{V}(\varphi_c)$ corresponds to an energy density, then the ground state (lowest-energy state) must, in fact, evidently be a minimum.

differentiate the denominator too) and then setting $J=0$, show, for example, that

$$\left.\frac{\delta^2 Z_c[J]}{\delta J(x_1)\,\delta J(x_2)}\right|_{J=0} = \mathrm{i}\frac{\int \mathcal{D}\varphi\,\tilde{\varphi}(x_1)\tilde{\varphi}(x_2)\;\mathrm{e}^{\mathrm{i}\int \mathrm{d}^4x\;\mathcal{L}(x)}}{\int \mathcal{D}\varphi\;\mathrm{e}^{\mathrm{i}\int \mathrm{d}^4x\;\mathcal{L}(x)}}.$$

By induction, this can be extended to the general n-point function.

In general, as would appear natural from the foregoing observation, we shall use the shifted fields to construct the quantum field theory. That is, we shall always consider quantum fluctuations (or perturbations) around the *ground state* of the theory. This becomes especially important in the presence of spontaneously broken symmetries, whence fields may acquire non-vanishing vacuum expectation values owing to the particular form of the potential occurring in the Lagrangian. It would then be incorrect to perturb around the vacuum (zero-field) configuration as this is unstable by construction. In such cases the shift must, however, be performed by hand before the quantisation procedure.

3.4 Bibliography

Abers, E.S. and Lee, B.W. (1973), *Phys. Rep.* **9**, 1.

Dirac, P.A.M. (1933), *Phys. Z. Sowjetunion* **3**, 64; *reprinted in*, in *Feynman's Thesis: A New Approach to Quantum Theory*, Brown, L.M., ed. (World Sci., 2005), p. 111.

Feynman, R.P. (1948), *Rev. Mod. Phys.* **20**, 367.

Feynman, R.P. (1950), *Phys. Rev.* **80**, 440.

Feynman, R.P. and Hibbs, A.R. (1965), *Quantum Mechanics and Path Integrals* (McGraw–Hill).

Gordon, W. (1926), *Z. Phys.* **40**, 117.

Klein, O. (1927), *Z. Phys.* **41**, 407.

Lee, T.-D. and Yang, C.-N. (1962), *Phys. Rev.* **128**, 885.

Schwinger, J.S. (1951), *Proc. Nat. Acad. Sci.* **37**, 452.

Chapter 4

Scattering Amplitudes and Feynman Rules

In this chapter we shall build on what we have constructed so far and learn how to calculate scattering amplitudes in a quantum field theory. Up to now what we have directly shown how to calculate are vacuum transition amplitudes and vacuum expectation values of products of fields. A scattering amplitude is, however, a little more complex. The typical process we shall attempt to describe is that of scattering between two particles (or fields) that at some time in the distant past are sufficiently separated as to be considered effectively non-interacting (so-called *asymptotic* states). The two particles will then approach one another and will interact either directly via some local interaction simultaneously involving the two or indirectly via the exchange of some "force" field (such as the photon). The final-stare products (after any further decays or interactions) will then move apart and at some point in the distant future will be sufficiently separated as to be again considered effectively non-interacting. Naturally, we should also consider the other physical possibility: namely, the decay of an unstable particle into two or more final, lighter, states; there is, however, no substantial difference in the calculational procedure.

4.1 Asymptotic states

As mentioned earlier, all the interactions we shall consider are *local*: that is, they involve fields evaluated at precisely the *same* space–time point or, in other words, they are truly *point-like*. The historic notion of action-at-a-distance is then superceded by point-like interactions giving rise to the exchange of fields, which may transport energy, momentum and quantum numbers between different space–time points.

After the interaction phase we may find that the two original particles are no longer present and have been replaced by any number of new fields—this is an essential difference with respect to quantum mechanics, where particle number and type are always conserved. The final state must, however, be allowed to attain the non-interacting regime; that is, the particles must all again become sufficiently well separated as to be considered independent. We shall, in fact, see that in the case of long range interactions such as QED this ideal situation is not strictly attainable; it is, however, possible to make corrections for all residual effects.

An important, indeed central, aspect of such so-called *asymptotic* states is that they obey the usual *mass-shell* condition:

$$p^2 = m^2, \tag{4.1.1}$$

where m, the physical mass, need not necessarily correspond, for example, to the parameter appearing in the $m^2\varphi^2$ term of the Lagrangian. Obviously, any residual interaction would invalidate such a condition owing to the contribution of the interaction energy. In the canonical approach we should now derive the so-called *reduction formalism* of Lehmann, Symanzik and Zimmermann (1955) to relate the desired physical scattering amplitudes to the time-ordered field correlation functions we have been studying; the path-integral approach adopted here allows a somewhat more straight-forward derivation.

4.2 Scattering amplitudes

4.2.1 Scattering amplitudes in quantum mechanics

We shall first return to the simpler case of non-relativistic quantum mechanics. So far, the configurations contributing to the path integral have been determined by requiring that they satisfy some simple boundary conditions, such as $\varphi(t,\boldsymbol{x}) \to \varphi^{\pm}(\boldsymbol{x})$ as $t \to \pm\infty$, where $\varphi^{\pm}$ may or may not be zero. Now, even if we should decide to use fields that vanish in the distant past and future, there is no guarantee that the above physical mass-shell condition be satisfied. Thus, our prescription for computing scattering amplitudes must also encode this requirement independently.

Let us consider the simple case of non-relativistic quantum mechanics in one space dimension. The generalised transition amplitude to pass from some initial state $|i\rangle$ to some final state $|f\rangle$ is

$$\mathcal{M}_{fi}(t_f,t_i) := \langle f|i\rangle = \int \mathrm{d}q_i\,\mathrm{d}q_f\,\langle f|q_f,t_f\rangle\langle q_f,t_f|q_i,t_i\rangle\langle q_i,t_i|i\rangle. \tag{4.2.1}$$

Note that the initial and final states are not necessarily eigenstates of generalised

position while the intermediate states used here are defined to be such. The first and last bra–ket pairs represent the initial- and final-state wave-functions:

$$\langle q_i,t_i|i\rangle \;=\; \phi_i(q_i,t_i) \quad \text{and} \quad \langle q_f,t_f|f\rangle \;=\; \phi_f(q_f,t_f). \tag{4.2.2}$$

These may be expanded in terms of momentum eigenstates $|p\rangle$, in the Schrödinger picture:

$$\begin{aligned}\phi_a(q,t) \;&=\; \int \mathrm{d}p\, \langle q,t|p\rangle\langle p|a\rangle \\ &=\; \int \mathrm{d}p\, \langle q|\mathrm{e}^{-\mathrm{i}Ht}|p\rangle\, c_a(p),\end{aligned} \tag{4.2.3}$$

where the Fourier coefficient is $c_a(p)\equiv\langle p|a\rangle$ for $a=i,f$.

The asymptotic condition may be most easily attained by requiring the scattering potential to be short range; *i.e.*, $V(q)\sim 0$ for $|q|\gtrsim R_0$. The initial and final wave-functions will thus be taken as only being non-zero for $|q_{i,f}|\gg R_0$. We shall also take the infinite limits of the initial and final times. Thus, the matrix elements are

$$\mathcal{M}_{fi} \;=\; \lim_{\substack{t_i\to-\infty\\ t_f\to+\infty}} \int \mathrm{d}q_i\,\mathrm{d}q_f\, \phi_f^*(q_f,t_f)\langle q_f,t_f|q_i,t_i\rangle\phi_i(q_i,t_i). \tag{4.2.4}$$

The wave-functions only appear for times such that they are localised in positions effectively isolated from the scattering potential V and therefore only depend on the free Hamiltonian. We thus have

$$\begin{aligned}\langle q|\mathrm{e}^{-\mathrm{i}Ht}|p\rangle \;&\simeq\; \langle q|\mathrm{e}^{-\mathrm{i}P^2t/2m}|p\rangle \\ &=\; \mathrm{e}^{-\mathrm{i}p^2t/2m}\,\langle q|p\rangle \\ &=\; \tfrac{1}{\sqrt{2\pi}}\,\mathrm{e}^{-\mathrm{i}p^2t/2m}\,\mathrm{e}^{\mathrm{i}pq}\,.\end{aligned} \tag{4.2.5}$$

Substituting this back into the wave-function, we may then write ($a=i,f$)

$$\phi_a(q_a,t_a) \;=\; \tfrac{1}{\sqrt{2\pi}}\int \mathrm{d}p\, c_a(p)\;\mathrm{e}^{-\mathrm{i}p^2t_a/2m}\;\mathrm{e}^{\mathrm{i}pq_a}\,. \tag{4.2.6}$$

We may now make an asymptotic expansion of the wave function (*i.e.* for $|t|$ and therefore $|q|$ large). Since the interaction is inoperative in this region, we may assume that the momentum or velocity of the particle tends to some constant and that the ratio q/t will thus tend asymptotically to some finite value. Now, the

exponent above may be rewritten:

$$\frac{p^2 t}{2m} - pq \;=\; \frac{t}{2m}\left(p - \frac{mq}{t}\right)^2 - \frac{mq^2}{2t}. \tag{4.2.7}$$

Defining then a rescaled momentum

$$k \;:=\; \left(\frac{|t|}{2m}\right)^{1/2}\left(p - \frac{mq}{t}\right), \tag{4.2.8}$$

in the limit $|t| \to \infty$ we have

$$\phi_a(q,t) \;=\; \tfrac{1}{\sqrt{2\pi}}\left(\frac{|t|}{2m}\right)^{-1/2} \mathrm{e}^{\mathrm{i}mq^2/2t}\int \mathrm{d}k\, c_a(p)\; \mathrm{e}^{\mathrm{i}\epsilon(t)k^2}, \tag{4.2.9}$$

where, inverting (4.2.8),

$$p \;=\; \frac{mq}{t} + \left(\frac{2m}{|t|}\right)^{1/2} k \tag{4.2.10}$$

and

$$\epsilon(t) \;=\; \begin{cases} -1 & \text{for } t < 0, \\ +1 & \text{for } t > 0. \end{cases} \tag{4.2.11}$$

Performing a Taylor expansion of c_a in the small quantity

$$\left(\frac{2m}{|t|}\right)^{1/2} k \;\ll\; \frac{mq}{t}, \tag{4.2.12}$$

for $|t| \to \infty$, we have

$$c_a(p) \;\sim\; c_a(mq/t)\left[1 + O(|t|^{-1/2})\right]. \tag{4.2.13}$$

The integral in k may now be performed to give

$$\phi_a(q,t) \;=\; \left(\frac{|t|}{2m}\right)^{-1/2} \mathrm{e}^{\mathrm{i}mq^2/2t}\; c_a(mq/t)\left[1 + O(|t|^{-1/2})\right] \mathrm{e}^{-\mathrm{i}\epsilon(t)\pi/4} \quad (\text{for } |t| \;\to\; \infty). \tag{4.2.14}$$

Keeping just the non-vanishing terms in the limit and substituting back into

the expression for the transition matrix element, we have

$$\mathcal{M}_{fi} \;=\; \lim_{\substack{t_i\to-\infty\\ t_f\to+\infty}} \frac{\mathrm{i}m}{\sqrt{|t_i t_f|}} \int \mathrm{d}q_i\,\mathrm{d}q_f \times c_f^*\Big(\frac{mq_f}{t_f}\Big)\; \mathrm{e}^{-\mathrm{i}mq_f^2/2t_f}\; \langle q_f,t_f|q_i,t_i\rangle\; c_i\Big(\frac{mq_i}{t_i}\Big)\; \mathrm{e}^{+\mathrm{i}mq_i^2/2t_i}\,. \quad (4.2.15)$$

The arguments of $c_{i,f}$ are to be substituted using

$$p_a \;:=\; \frac{mq_a}{t_a} \quad (\text{for } a \;=\; i,f), \quad (4.2.16)$$

leading to

$$\mathcal{M}_{fi} \;=\; \lim_{\substack{t_i\to-\infty\\ t_f\to+\infty}} \frac{\mathrm{i}\sqrt{|t_i t_f|}}{m} \int \mathrm{d}p_i\,\mathrm{d}p_f \times c_f^*(p_f)\; \mathrm{e}^{-\mathrm{i}p_f^2 t_f/2m}\; \langle \tfrac{p_f t_f}{m},t_f|\tfrac{p_i t_i}{m},t_i\rangle\; c_i(p_i)\; \mathrm{e}^{+\mathrm{i}p_i^2 t_i/2m}\,. \quad (4.2.17)$$

Thus, the asymptotic boundary condition on the path integral is now

$$q(t) \;\sim\; \begin{cases} \dfrac{p_i t}{m} & \text{for } t \;\to\; -\infty \\[2ex] \dfrac{p_f t}{m} & \text{for } t \;\to\; +\infty, \end{cases} \quad (4.2.18)$$

which evidently corresponds to the equations of motion for a free particle of mass m and therefore velocity $v = p/m$. This is in stark contrast with the earlier requirement of constant boundary conditions on q and we shall now need to define a new generating functional with just this asymptotic behaviour.

4.2.2 Scattering amplitudes in quantum field theory

We now generalise the foregoing argument to the case of a quantum field theory. We have already know how to calculate the transition amplitude $\langle\varphi_f(\boldsymbol{x}),t_f|\varphi_i(\boldsymbol{x}),t_i\rangle^J$ in quantum field theory, when the boundary conditions are some generic initial and final field configurations. Now, however, it is necessary to impose conditions such that the asymptotic states have free-field behaviour, that is, they must satisfy the relevant free-field equation of motion (in our case just the Klein–Gordon equation).

Thus, for

$$\varphi(t,\boldsymbol{x}) \sim \begin{cases} \varphi_{\text{in}}(t,\boldsymbol{x}) & \text{for } t \to -\infty \\ \varphi_{\text{out}}(t,\boldsymbol{x}) & \text{for } t \to +\infty, \end{cases} \tag{4.2.19}$$

where $\varphi_{\text{in,out}}(t,\boldsymbol{x})$ satisfy the free-field equation derived from the free Lagrangian:

$$\left[\Box + m^2\right] \varphi_{\text{in,out}}(t,\boldsymbol{x}) = 0. \tag{4.2.20}$$

Now, as discussed in the section on classical fields, the free field may be expressed as a Fourier expansion:

$$\varphi_0(x) = \int \frac{\text{đ}^3\boldsymbol{k}}{2k_0} \left[b(k)\, \mathrm{e}^{-\mathrm{i}k\cdot x} + b^*(k)\, \mathrm{e}^{+\mathrm{i}k\cdot x} \right], \tag{4.2.21}$$

subject to the mass-shell condition $k^2 = m^2$. The asymptotic fields $\varphi_{\text{in,out}}(t,\boldsymbol{x})$ are then just special cases. We have seen in the foregoing that the initial state is associated with the momentum wave-function $c_i(p)$ and thus we require

$$\varphi_{\text{in}}(t,\boldsymbol{x}) = \int \frac{\text{đ}^3\boldsymbol{k}}{2k_0}\, b(k)\, \mathrm{e}^{-\mathrm{i}k\cdot x} \tag{4.2.22a}$$

and

$$\varphi_{\text{out}}(t,\boldsymbol{x}) = \int \frac{\text{đ}^3\boldsymbol{k}}{2k_0}\, b^*(k)\, \mathrm{e}^{+\mathrm{i}k\cdot x}. \tag{4.2.22b}$$

That is, $\varphi_{\text{in,out}}(t,\boldsymbol{x})$ are to be associated with the "positive" and "negative" energy pieces respectively. In other words, for the initial state we need to "create" particles from the vacuum (hence the positive energy) while for the final state we must "destroy" all the states to return to the vacuum (hence the negative energy).

Such behaviour can be arranged via the already familiar "$\mathrm{i}\varepsilon$" prescription: we add a small imaginary part to the energy. This has the result of giving the states effectively finite lifetimes since now

$$\mathrm{e}^{\pm\mathrm{i}k_0 t} \to \mathrm{e}^{\pm\mathrm{i}k_0 t}\, \mathrm{e}^{\pm\varepsilon t} \to 0 \quad \text{for} \quad t \to \mp\infty. \tag{4.2.23}$$

We thus obtain the desired behaviour:

$$\varphi_0(x) \sim \begin{cases} \varphi_{\text{in}}(x) & \text{for} \quad t \to -\infty, \\ \varphi_{\text{out}}(x) & \text{for} \quad t \to +\infty. \end{cases} \tag{4.2.24}$$

In this case the asymptotic conditions may be simply rewritten as

$$\varphi(x) \sim \varphi_0(x) \quad \text{for} \quad |t| \to \infty. \tag{4.2.25}$$

The next step is to discover the form of the generating functional for the relevant *scattering-matrix* (or S-matrix) elements. Consider the path integral

$$S[J,\varphi_0] \;:=\; \int \mathcal{D}\varphi \;\exp\!\left\{ \mathrm{i}\int \mathrm{d}^4x \;\left(\mathcal{L}+J\varphi\right)\right\}\Bigg|_{\varphi(|t|\to\infty)\sim\varphi_0}, \tag{4.2.26}$$

where the argument φ_0 indicates precisely the boundary conditions just defined. We now wish to consider the full interacting theory and shall therefore separate the Lagrangian as follows:

$$\mathcal{L} \;=\; \mathcal{L}_0+\mathcal{L}_{\mathrm{I}}\,. \tag{4.2.27}$$

The first piece is just the kinetic contribution discussed in the free-field case,

$$\mathcal{L}_0 \;=\; \tfrac{1}{2}\big[(\partial^\mu\varphi)(\partial_\mu\varphi)-m^2\varphi^2\big]\,, \tag{4.2.28}$$

while $\mathcal{L}_{\mathrm{I}}$ is intended to contain all remaining interaction contributions. Two common example choices are

$$\mathcal{L}_{\mathrm{I}} \;=\; -\tfrac{1}{3!}\lambda\varphi^3 \quad \text{or} \quad -\tfrac{1}{4!}\lambda\varphi^4. \tag{4.2.29}$$

The first has a certain similarity to the interaction of quantum electrodynamics but unfortunately pays the price of representing an inherently unstable potential (*i.e.* not bounded from below) while the second is stable but somewhat peculiar in only describing four-point interactions (similar to Fermi's theory of the weak interaction). For the following derivation, however, it will fortunately not be necessary to make any specific choice.

Recalling then the particular property of the path integral that it is expressed in terms of classical (and therefore commuting) fields and not operators, we may write it as

$$S[J,\varphi_0] \;:=\; \int \mathcal{D}\varphi \;\exp\!\left\{ \mathrm{i}\int \mathrm{d}^4x \;\left[\mathcal{L}_0+J\varphi\right]\right\}\exp\!\left\{ \mathrm{i}\int \mathrm{d}^4x \;\mathcal{L}_{\mathrm{I}}\right\}\Bigg|_{\varphi(|t|\to\infty)\sim\varphi_0}. \tag{4.2.30}$$

In the presence of a non-trivial interaction term, the functional integral cannot generally be performed in closed form and one must either find some numerical (non-perturbative) method or make a perturbative (power-series) expansion of the interaction piece. We have no space here for a discussion of numerical methods, which at the time of writing have yet to produce solid and reliable results in *e.g.*, quantum chromodynamics (QCD). However, as and when such might become available they would evidently represent a more complete (though hermetic) solution. In any case, we shall now turn to the perturbative approach. Before doing so let us stress that such an approximation immediately implies an incomplete

solution to the full quantum theory and the risk of missing important behaviour and properties. It is, though, the only theoretically consolidated approach available a present and certainly the only approach allowing clear and unambiguous comparison between theory and experiment, where applicable.

The idea is simply to expand the exponential of the interaction term as a power series in the exponent:

$$\exp\left\{ \mathrm{i} \int \mathrm{d}^4x \; \mathcal{L}_\mathrm{I}(x) \right\} \; = \; \sum_{n=0}^{\infty} \frac{1}{n!} \left(\mathrm{i} \int \mathrm{d}^4x \; \mathcal{L}_\mathrm{I}(x) \right)^n . \tag{4.2.31}$$

This is the heart of the perturbative expansion: it is evidently an expansion in powers of the interaction potential, considered a small perturbation applied to the free system. We have already learnt that the vacuum expectation value of powers of fields may be generated from the *free-field* generating functional in the presence of an external source J by taking derivatives with respect to J. Thus, inasmuch as the perturbing term may be assumed expressible as simple products of fields, we may write

$$\mathcal{L}_\mathrm{I}\big(\varphi(x)\big) \exp\left\{ \mathrm{i} \int \mathrm{d}^4y \, J(y) \, \varphi(y) \right\} \; = \; \mathcal{L}_\mathrm{I}\left(-\mathrm{i} \frac{\delta}{\delta J(x)} \right) \exp\left\{ \mathrm{i} \int \mathrm{d}^4y \, J(y) \, \varphi(y) \right\} . \tag{4.2.32}$$

Since the term in $\mathcal{L}_\mathrm{I}$ does not now depend explicitly on $\varphi(x)$, we may extract it from the functional integral in Eq. (4.2.30) and thus write

$$S[J, \varphi_0] \; = \; \exp\left\{ \mathrm{i} \int \mathrm{d}^4x \; \mathcal{L}_\mathrm{I}\left(-\mathrm{i} \frac{\delta}{\delta J(x)} \right) \right\} S_0[J, \varphi_0], \tag{4.2.33}$$

where we have defined the free-field generating functional in the presence of an external source:

$$S_0[J, \varphi_0] \; := \; \int \mathcal{D}\varphi \, \exp\left\{ \mathrm{i} \int \mathrm{d}^4x \; \Big[\mathcal{L}_0 + J(x) \, \varphi(x) \Big] \right\} . \tag{4.2.34}$$

We now make a shift in the field $\varphi(x)$ and define

$$\tilde{\varphi}(x) \; := \; \varphi(x) - \varphi_0(x). \tag{4.2.35}$$

That is, $\tilde{\varphi}(x) \to 0$ asymptotically and thus represents the fluctuations around the asymptotic solution. Substituting this into the expression for the free Lagrangian plus source term, we have

$$\mathcal{L}_0(\varphi) + J\varphi \; = \; \mathcal{L}_0(\tilde{\varphi}) + J\tilde{\varphi} + \mathcal{L}_0(\varphi_0) + J\varphi_0 + \partial_\mu \tilde{\varphi} \, \partial^\mu \varphi_0 - m^2 \tilde{\varphi} \varphi_0 . \tag{4.2.36}$$

Integrating by parts and assuming surface terms to vanish, we obtain

$$\mathcal{L}_0(\varphi) + J\varphi \;=\; \mathcal{L}_0(\tilde{\varphi}) + J\tilde{\varphi} + J\varphi_0 - (\varphi_0 + \tilde{\varphi})E(\varphi_0), \tag{4.2.37}$$

where

$$E\big(\varphi_0(x)\big) \;:=\; \left[\Box_x + m^2\right]\varphi_0(x) \;=\; 0, \tag{4.2.38}$$

is a free-field equation-of-motion term. We can thus write

$$S_0[J,\varphi_0] \;=\; \int \mathcal{D}\tilde{\varphi}\, \exp\left\{ \mathrm{i}\int \mathrm{d}^4x\, \left[\mathcal{L}_0(\tilde{\varphi}) + J(x)\tilde{\varphi}(x)\right]\right\} \exp\left\{ \mathrm{i}\int \mathrm{d}^4x\, J(x)\varphi_0(x)\right\}. \tag{4.2.39}$$

Here the functional integral, now over $\tilde{\varphi}$, is subject to the asymptotic boundary condition $\tilde{\varphi}(x)\to 0$, which is just a special case of the type considered earlier in Chap. 3. The functional integral in the above is therefore none other than the free-field disconnected generating functional derived in Eq. (3.3.28) and we may thus rewrite $S_0[J,\varphi_0]$ as

$$S_0[J,\varphi_0] \;=\; Z_0[J]\exp\left\{ \mathrm{i}\int \mathrm{d}^4x\, J(x)\varphi_0(x)\right\}. \tag{4.2.40}$$

Exercise 4.2.1. *Perform explicitly the integration by parts and necessary algebraic manipulations described to arrive at Eq. (4.2.39).*

Differentiating $Z_0[J]$ functionally with respect to $J(x)$ once leads to

$$\frac{\delta Z_0[J]}{\delta J(x)} \;=\; -Z_0[J]\int \mathrm{d}^4y\, \mathrm{i}\Delta_{\mathrm{F}}(x-y)\, J(y). \tag{4.2.41}$$

The Feynman propagator, or two-point Green function, of the Klein–Gordon theory may be written as an inverse Fourier transform:

$$\Delta_{\mathrm{F}}(x-y) \;=\; \int \text{đ}^4p\; \mathrm{e}^{\mathrm{i}p\cdot(x-y)}\,[p^2 - m^2 + \mathrm{i}\varepsilon]^{-1} \tag{4.2.42}$$

and is just the inverse of Klein–Gordon operator, which we have called $\mathcal{K}(x,y)$. Therefore, recalling Eq. (3.3.26), we have

$$\left[\Box_x + m^2\right]\frac{\delta}{\delta J(x)}\, Z_0[J] \;=\; \mathrm{i}\,J(x)\, Z_0[J], \tag{4.2.43}$$

which may be applied repeatedly to provide a formula for any power of J. And thus

$$S_0[J,\varphi_0] \;=\; \exp\left\{\int \mathrm{d}^4x\, \varphi_0(x)\left[\Box_x + m^2\right]\frac{\delta}{\delta J(x)}\right\} Z_0[J]. \tag{4.2.44}$$

This then allows us to re-write the full interacting expression as

$$S[J,\varphi_0] \;=\; \exp\!\left\{\int \mathrm{d}^4x\,\varphi_0(x)\,\left[\Box_x+m^2\right]\frac{\delta}{\delta J(x)}\right\} \times \exp\!\left\{\mathrm{i}\int \mathrm{d}^4x\;\mathcal{L}_{\mathrm{I}}\!\left(-\mathrm{i}\,\frac{\delta}{\delta J(x)}\right)\right\} Z_0[J]. \quad (4.2.45)$$

Exercise 4.2.2. *One can now retrace the steps taken for $S[J,\varphi_0]$ but apply them to $Z[J]$—the only difference lies in the boundary conditions, which are irrelevant for this purpose—and thus show that*

$$Z[J] \;=\; \exp\!\left\{\mathrm{i}\int \mathrm{d}^4x\;\mathcal{L}_I\!\left(-\mathrm{i}\,\frac{\delta}{\delta J(x)}\right)\right\} Z_0[J]. \quad (4.2.46)$$

Finally then, we have

$$S[J,\varphi_0] \;=\; \exp\!\left\{\int \mathrm{d}^4x\,\varphi_0(x)\,\left[\Box_x+m^2\right]\frac{\delta}{\delta J(x)}\right\} Z[J] \quad (4.2.47)$$

which we may rewrite as

$$=\; \exp\!\left\{\int \mathrm{d}^4x\,\varphi_0(x)\;\mathcal{K}(x)\,\frac{\delta}{\delta J(x)}\right\} Z[J]. \quad (4.2.48)$$

We have thus derived the *ground-state-to-ground-state transition amplitude* in the presence of an external source. Of course, the physical amplitudes will be true ground-state amplitudes and we shall therefore finally set $J(x)=0$ to give

$$S[\varphi_0] \;=\; S[J,\varphi_0]\Big|_{J=0}. \quad (4.2.49)$$

The importance of expression (4.2.48) lies in the presence of the Klein–Gordon operator $\mathcal{K}$ as a pre-factor. Once again, integrating by parts, we may allow it to act on the field $\varphi_0(x)$, which leads to a factor (k^2-m^2) for each Fourier component. Since, by choice, $\varphi_0(x)$ satisfies the Klein–Gordon equation and thus $k^2-m^2=0$, this kills all contributions from $Z[J]$ except those containing a *pole* in (k^2-m^2). In other words, the effect is to eliminate all off-shell contributions, leaving only those satisfying the mass-shell condition. This is just the LSZ result, obtained in the canonical approach (Lehmann, Symanzik and Zimmermann, 1955), with Eq. (4.2.48) taking the place of the so-called LSZ reduction formula. In practice, when evaluating the functional derivatives acting on $Z[J]$, we shall see that all external propagators (*i.e.* those representing initial- or final-state particles) remain and these then are *truncated* (or *amputated*) by the Klein–Gordon pre-factor.

4.3 Feynman rules for $\lambda\varphi^4$ theory

At this point we have all that is necessary to derive the Feynman rules for a simple $\lambda\varphi^4$ theory. From Eq. (4.2.48), we see that the principal object to be evaluated is just $Z[J]$ (or, as one might suspect, the closely related $Z_c[J]$). As already remarked, this cannot be achieved in a general closed form, except for the trivial free-field case, and so we must resort to perturbative techniques.

4.3.1 Space–time formulation

For definiteness then, we now consider a scalar field theory with the following interaction term in the Lagrangian:

$$\mathcal{L}_{\mathrm{I}} \;=\; -\frac{\lambda}{4!}\varphi^4. \tag{4.3.1}$$

The power-series expansion of Eq. (4.2.31) then becomes

$$\exp\!\left\{\mathrm{i}\int \mathrm{d}^4x\;\mathcal{L}_{\mathrm{I}}\right\} \;=\; 1-\frac{\mathrm{i}\,\lambda}{4!}\int \mathrm{d}^4x\,\frac{\delta^4}{\delta J(x)^4}+O(\lambda^2), \tag{4.3.2}$$

keeping just terms to first order in λ, which leads to

$$Z[J] \;=\; Z_0[J]-\frac{\mathrm{i}\,\lambda}{4!}\int \mathrm{d}^4x\,\frac{\delta^4 Z_0[J]}{\delta J(x)^4}+O(\lambda^2). \tag{4.3.3}$$

Recall that from Eq. (3.3.28) we have

$$\frac{\delta Z_0[J]}{\delta J(x)} \;=\; -\mathrm{i}\int \mathrm{d}^4y\,\Delta_{\mathrm{F}}(x-y)\,J(y)\,Z_0[J]. \tag{4.3.4}$$

It is then just a matter of applying this formula the four times necessary at this order in the approximation:

$$\begin{aligned} Z[J] \;=\; \Bigg\{1-\frac{\mathrm{i}\,\lambda}{4!}\int \mathrm{d}^4x\,\Bigg[&3\left[\mathrm{i}\,\Delta_{\mathrm{F}}(x-x)\right]^2 \\ &-3!\,\mathrm{i}\,\Delta_{\mathrm{F}}(x-x)\int \mathrm{d}^4x_1\,\mathrm{d}^4x_2\,\mathrm{i}\,\Delta_{\mathrm{F}}(x-x_1)\,\mathrm{i}\,\Delta_{\mathrm{F}}(x-x_2)\,J(x_1)\,J(x_2) \\ &+\int \mathrm{d}^4x_1\,\mathrm{d}^4x_2\,\mathrm{d}^4x_3\,\mathrm{d}^4x_4\,\mathrm{i}\,\Delta_{\mathrm{F}}(x-x_1)\,\mathrm{i}\,\Delta_{\mathrm{F}}(x-x_2)\,\mathrm{i}\,\Delta_{\mathrm{F}}(x-x_3)\,\mathrm{i}\,\Delta_{\mathrm{F}}(x-x_4) \\ &\times J(x_1)\,J(x_2)\,J(x_3)\,J(x_4)\Bigg]+O(\lambda^2)\Bigg\}\,Z_0[J]. \end{aligned} \tag{4.3.5}$$

Recall Eq. (3.3.10), and the standard normalisation:

$$Z_0[0] := 1. \tag{4.3.6}$$

An immediate consequence is that the Green functions will now include new contributions arising from the interaction. We thus see that already even the purely vacuum term is non-trivial:

$$\mathcal{G}^{(0)} = Z[0] = 1 - \tfrac{1}{8}\mathrm{i}\lambda \int \mathrm{d}^4x \left[\mathrm{i}\Delta_{\mathrm{F}}(x-x)\right]^2 + O(\lambda^2). \tag{4.3.7}$$

Moving on to the first interesting Green function, we have

$$\mathcal{G}^{(2)}(x,y) = \mathcal{G}_0^{(2)}(x,y) - \tfrac{1}{2}\mathrm{i}\lambda \int \mathrm{d}^4z\, \mathrm{i}\Delta_{\mathrm{F}}(x-z)\, \mathrm{i}\Delta_{\mathrm{F}}(z-z)\, \mathrm{i}\Delta_{\mathrm{F}}(z-y) + O(\lambda^2), \tag{4.3.8}$$

where

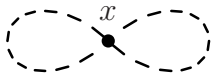
$$\mathcal{G}_0^{(2)}(x,y) = \mathrm{i}\Delta_{\mathrm{F}}(x-y) \tag{4.3.9}$$

is just the free-field two-point Green function already derived (*i.e.* it is the Feynman propagator).

Before examining the full interaction term or four-point function itself, let us try to give a diagrammatic interpretation of the expansion thus far. We have already suggested assigning a line to the free-field propagator or two-point function; we now add to this a vertex, corresponding to the factor $-\mathrm{i}\lambda$:

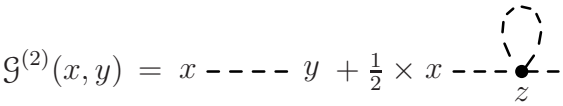
$$= -\mathrm{i}\lambda. \tag{4.3.10}$$

The first-order vacuum diagram for $\mathcal{G}^{(0)}$ is then

$\frac{1}{8} \times$ 

, (4.3.11)

where x is just the integration variable above (we shall comment on the numerical coefficient shortly). Up to first order, the two-point function is

$\mathcal{G}^{(2)}(x,y) = x$ 

$+ O(\lambda^2).$ (4.3.12)

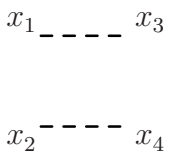
The four-point function is

$$\mathcal{G}^{(4)}(x_1,x_2,x_3,x_4) = \mathcal{G}_0^{(4)}(x_1,x_2,x_3,x_4)$$

$$
\begin{aligned}
&- \tfrac{1}{2}\mathrm{i}\lambda\Big\{ \mathrm{i}\Delta_{\mathrm{F}}(x_1-x_2)\int \mathrm{d}^4z\, \mathrm{i}\Delta_{\mathrm{F}}(x_3-z)\,\mathrm{i}\Delta_{\mathrm{F}}(z-z)\,\mathrm{i}\Delta_{\mathrm{F}}(z-x_4) \\
&\qquad\qquad + \text{perms}(x_1,x_2,x_3,x_4)\Big\} \\
&- \mathrm{i}\lambda\int \mathrm{d}^4z\, \mathrm{i}\Delta_{\mathrm{F}}(x_1-z)\,\mathrm{i}\Delta_{\mathrm{F}}(x_2-z)\,\mathrm{i}\Delta_{\mathrm{F}}(x_3-z)\,\mathrm{i}\Delta_{\mathrm{F}}(x_4-z) + O(\lambda^2). \qquad (4.3.13)
\end{aligned}
$$

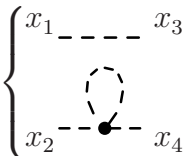
The first term $\mathcal{G}_0^{(4)}$ is just the free-field four-point function already encountered (recall that it contains three disconnected diagrams). The terms in braces corresponds to the free-field function with a loop correction attached to each of the possible free propagators in turn (giving $3\times 2=6$ diagrams). Finally, we have the only connected diagram at this order, with four propagators attached to a single four-point vertex. Diagrammatically this becomes

$\mathcal{G}^{(4)}(x_1,x_2,x_3,x_4) =$



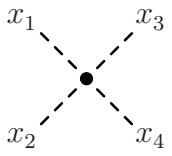
$+ \text{perms}(x_2,x_3,x_4)$

$+ \tfrac{1}{2}\Big\{$



$+ \text{perms}(x_1,x_2,x_3,x_4)\Big\}$

$+$



$+ O(\lambda^2).$ (4.3.14)

Exercise 4.3.1. *Derive these diagrammatic expressions and, in particular, examine the precise origin of the numerical factors in front of the diagrams containing loops. In this way attempt to derive a general rule.*

This is now a good point to discover the utility of the *connected* Green functions derived from $Z_c[J]$:

$$\mathrm{i}Z_c[J] \;=\; \ln Z[J]. \qquad (4.3.15)$$

Using the approximation $\ln(1+\varepsilon)=\varepsilon+O(\varepsilon^2)$ and the explicit expression for $Z_0[J]$ given in Eq. (3.3.28), we obtain

$$\mathrm{i}Z_c[J] \;=\; \text{constant} - \tfrac{1}{2}\int \mathrm{d}^4x_1\,\mathrm{d}^4x_2\, J(x_1)\,\mathrm{i}\Delta_{\mathrm{F}}(x_1-x_2)\,J(x_2)$$

$$
\begin{aligned}
&- \tfrac{1}{8}\,\mathrm{i}\lambda \int \mathrm{d}^4x \; \big[\mathrm{i}\Delta_{\mathrm{F}}(x-x)\big]^2 \\
&+ \tfrac{1}{4}\,\mathrm{i}\lambda \int \mathrm{d}^4x\, \mathrm{d}^4x_1\, \mathrm{d}^4x_2\; \mathrm{i}\Delta_{\mathrm{F}}(x_1-x)\; \mathrm{i}\Delta_{\mathrm{F}}(x-x)\; \mathrm{i}\Delta_{\mathrm{F}}(x-x_2)\; J(x_1)\, J(x_2) \\
&- \tfrac{1}{4!}\,\mathrm{i}\lambda \int \mathrm{d}^4x\, \mathrm{d}^4x_1\, \mathrm{d}^4x_2\, \mathrm{d}^4x_3\, \mathrm{d}^4x_4\; \mathrm{i}\Delta_{\mathrm{F}}(x-x_1)\; \mathrm{i}\Delta_{\mathrm{F}}(x-x_2) \\
&\qquad\qquad \times\, \mathrm{i}\Delta_{\mathrm{F}}(x-x_3)\; \mathrm{i}\Delta_{\mathrm{F}}(x-x_4)\; J(x_1)\, J(x_2)\, J(x_3)\, J(x_4) \\
&+ O(\lambda^2). \qquad (4.3.16)
\end{aligned}
$$

The additive constant term, reflecting the overall normalisation, will vanish on taking derivatives. The definition of the connected Green functions (3.3.36), gives

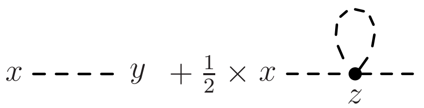
$$
\mathrm{i}^n\, \mathcal{G}_c^{(n)}(x_1, x_2, \ldots, x_n) \;=\; \mathrm{i}\, \frac{\delta^n Z_c[J]}{\delta J(x_1)\, \delta J(x_2) \ldots \delta J(x_n)}\bigg|_{J=0}. \qquad (4.3.17)
$$

We see that, except for the zero-point function, vacuum diagrams are eliminated:

$$
\mathcal{G}_c^{(0)} \;=\; Z_c[0] \;=\; 1 - \tfrac{1}{8}\,\mathrm{i}\lambda \int \mathrm{d}^4x\; \big[\mathrm{i}\Delta_{\mathrm{F}}(x-x)\big]^2. \qquad (4.3.18)
$$

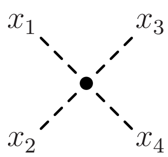
The first non-trivial connected function is the same as its disconnected version:

$$
\mathcal{G}_c^{(2)}(x,y) \;=\; \mathcal{G}_0^{(2)}(x,y) - \tfrac{1}{2}\,\mathrm{i}\lambda \int \mathrm{d}^4z\; \mathrm{i}\Delta_{\mathrm{F}}(x-z)\; \mathrm{i}\Delta_{\mathrm{F}}(z-z)\; \mathrm{i}\Delta_{\mathrm{F}}(z-y) + O(\lambda^2)
$$

$=$ 

$+\, O(\lambda^2)$. (4.3.19)

To this order the connected four-point function is now non-vanishing although only the vertex term itself survives:

$\mathcal{G}_c^{(4)}(x_1, x_2, x_3, x_4) \;=$ 

$+\, O(\lambda^2)$. (4.3.20)

Exercise 4.3.2. *Show that the interaction generates a non-trivial connected six-point function and derive its graphical representation.*

4.3.2 Momentum-space formulation

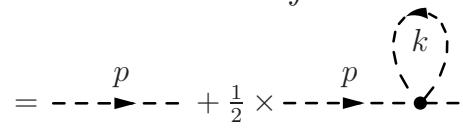
We have seen that, in general, the expressions for the n-point functions involve *convolutions* of simpler pieces (*e.g.* propagators and lowest-order vertices). Thus, if we make a Fourier transform to momentum space, we may expect these to become simple *products*. This, coupled to the fact that the propagator has its simplest form there, leads us to prefer the momentum-space formulation. The connected Green functions are then

$$\begin{aligned} \mathcal{G}_c^{(n)}(p_1, p_2, \ldots, p_n)\ \delta^4(p_1 + p_2 + \cdots + p_n) \\ = \int \mathrm{d}^4x_1\ \ldots \mathrm{d}^4x_n\ \mathrm{e}^{\mathrm{i}(p_1\cdot x_1 + \cdots + p_n\cdot x_n)}\ \mathcal{G}_c^{(n)}(x_1, x_2, \ldots, x_n), \end{aligned} \quad (4.3.21)$$

where the energy–momentum conserving δ-function has again been explicitly factorised.

Taking the Fourier transform of (4.3.19), we obtain

$$\begin{aligned} \mathcal{G}_c^{(2)}(p,-p) &= \mathcal{G}_0^{(2)}(p,-p) - \tfrac{1}{2}\mathrm{i}\lambda \int \mathrm{d}\!\!^{-4}k\ \mathrm{i}\Delta_{\mathrm{F}}(p)\ \mathrm{i}\Delta_{\mathrm{F}}(k)\ \mathrm{i}\Delta_{\mathrm{F}}(p) + O(\lambda^2) \\ &= \quad + \tfrac{1}{2} \times \quad + O(\lambda^2), \end{aligned} \quad (4.3.22)$$



where the momentum-space free-field two-point function is

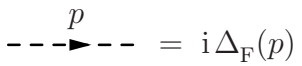
$$\mathcal{G}_0^{(2)}(p,-p) = \mathrm{i}\Delta_{\mathrm{F}}(p) = \frac{\mathrm{i}}{p^2 - m^2 + \mathrm{i}\varepsilon}. \quad (4.3.23)$$

We can now write down the momentum-space Feynman rules. The basic form should now be obvious. A little examination reveals that the space–time integration at each vertex turns into a four-momentum conservation δ-function and so a number of four-momentum integrations can then be performed immediately. At the end, we shall find just a single, overall, four-momentum conservation δ-function (already factored out in the definition of the Green functions) together with a number of remaining unconstrained four-momentum integrals, one for each closed loop. Note too that the above-defined Green functions automatically have propagators associated with each external leg. These will then be *truncated* by the corresponding Klein–Gordon operator associated with external particles in the reduction formula (4.2.48) for the vacuum-to-vacuum transition amplitudes.

In short then, the Feynman rules are:

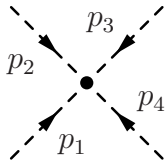
1. Internal lines, corresponding to propagators of momentum p, are represented

as



$$= \mathrm{i}\,\Delta_{\mathrm{F}}(p) = \frac{\mathrm{i}}{p^2 - m^2 + \mathrm{i}\varepsilon}. \tag{4.3.24}$$

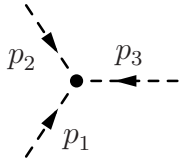
2. In φ^4 theory vertices are represented as



$$= -\mathrm{i}\,\lambda, \tag{4.3.25}$$

with the constraint that $\sum_{i=1}^{4} p_i = 0$.

2′. In φ^3 theory vertices are represented as



$$= -\mathrm{i}\,\lambda, \tag{4.3.26}$$

with a similar total-momentum constraint.

3. Each topologically distinct closed loop is associated with an unconstrained four-momentum integral over $\mathrm{d}\!\!^{-4}k$.

4. Symmetry factors are associated with equivalent internal propagators: n lines that leave the diagram unaltered under interchange carry a factor $1/n!$ and any propagator with both extremities connected to the same vertex carries an additional factor $1/2$.

We should note that these Feynman rules are specific to scalar theory and will change quite substantially for theories describing fermions and/or spin-one bosons. Evidently too, the loop-momenta integrals may and indeed often will lead to infinities and thus require further study.

4.3.3 The loop expansion

To better understand the nature of the perturbative expansion, it will now be useful to restore the factor $\hbar$ in the various expressions leading to the Feynman rules. That is, we should start from an expression such as

$$Z[J] = \exp\left\{\mathrm{i}\,Z_c[J]\right\} = \int \mathcal{D}\varphi\, \exp\left\{\frac{\mathrm{i}}{\hbar}\int \mathrm{d}^4x\, \Big[\mathcal{L}(\varphi, \partial_\mu\varphi) + J(x)\,\varphi(x)\Big]\right\}. \tag{4.3.27}$$

Let us first rescale both $\varphi(x)$ and $J(x)$ in the following manner:

$$\tilde{\varphi}(x) \;:=\; \tfrac{1}{\sqrt{\hbar}}\,\varphi(x) \tag{4.3.28a}$$

and

$$\tilde{J}(x) \;:=\; \tfrac{1}{\sqrt{\hbar}}\,J(x). \tag{4.3.28b}$$

This has the effect of removing all trace of $\hbar$ from the free theory (*i.e.* from both the free Lagrangian *and* the source term) while multiplying the interaction term in, *e.g.*, a $\lambda\varphi^n$ theory by $\hbar^{\frac{n}{2}-1}$. Since $n>2$, this is necessarily a non-vanishing, positive power.

Consider then our model $\lambda\varphi^4$ theory and take a generic diagram with L loops, E external fields, I internal propagators (*i.e. excluding* those that will be amputated) and V vertices. Since each propagator has two ends (either internal and attached to vertices or external and unattached) and each vertex connects to four such ends, we have

$$E + 2I \;=\; 4V. \tag{4.3.29}$$

For the number of *free* loop-momenta integrations, we also find that

$$L \;=\; I - (V-1) \;=\; 1 + V - E/2. \tag{4.3.30}$$

Thus, for fixed initial and final states (*i.e.* fixed E), the number of free loop integrations is directly determined by the number of vertices. Since each vertex carries a factor $\hbar$, we may, in fact, view the loop or perturbative expansion (in λ) as equivalently a quantum expansion. That is to say, while the leading-order diagram (usually a tree diagram corresponding to the Born approximation) for a given process may or may not (though, indeed, it most likely does) lead to a positive power of $\hbar$ and thus may or may not already be considered a quantum-mechanical effect, all higher-order contributions most certainly contain *extra* powers of $\hbar$ and therefore do correspond to quantum effects or corrections.

4.3.4 1PI n-point functions

In Sec. 3.3.5 we introduced the effective action $\Gamma[\varphi_c]$ as the generator of 1PI n-point functions via their definition (3.3.52). Thus, in the presence of interactions, Eq. (3.3.47) becomes

$$\begin{aligned}\varphi_c(x) \;&:=\; \frac{\delta Z_c[J]}{\delta J(x)} \\ &=\; -\int \mathrm{d}^4y\,\Delta_{\mathrm{F}}(x-y)\,J(y)\end{aligned}$$

$$
\begin{aligned}
&+ \tfrac{1}{2}\lambda \int \mathrm{d}^4y\, \mathrm{d}^4z\, \mathrm{i}\Delta_{\mathrm{F}}(x-y)\, \mathrm{i}\Delta_{\mathrm{F}}(y-y)\, \mathrm{i}\Delta_{\mathrm{F}}(y-z)\, J(z) \\
&- \tfrac{1}{3!}\lambda \int \mathrm{d}^4y\, \mathrm{d}^4z_1\, \mathrm{d}^4z_2\, \mathrm{d}^4z_3\, \mathrm{i}\Delta_{\mathrm{F}}(x-y) \\
&\qquad\qquad \times \mathrm{i}\Delta_{\mathrm{F}}(y-z_1)\, \mathrm{i}\Delta_{\mathrm{F}}(y-z_2)\, \mathrm{i}\Delta_{\mathrm{F}}(y-z_3)\, J(z_1)\, J(z_2)\, J(z_3) \\
&+ O(\lambda^2). \qquad (4.3.31)
\end{aligned}
$$

For the effective action, or the generator of 1PI n-point functions, $\Gamma[\varphi_c]$, using (4.3.16), we have then

$$
\begin{aligned}
\mathrm{i}\Gamma[\varphi_c] \;&:=\; \mathrm{i}Z_c[J] - \mathrm{i}\int \mathrm{d}^4x\, J(x)\, \varphi_c(x) \\
&=\; \text{constant} + \tfrac{1}{2}\int \mathrm{d}^4x\, \mathrm{d}^4y\, J(x)\, \mathrm{i}\Delta_{\mathrm{F}}(x-y)\, J(y) \\
&\quad - \tfrac{1}{8}\lambda \int \mathrm{d}^4x\, \left[\mathrm{i}\Delta_{\mathrm{F}}(x-x)\right]^2 \\
&\quad - \tfrac{1}{4}\lambda \int \mathrm{d}^4x\, \mathrm{d}^4y\, \mathrm{d}^4z\, \mathrm{i}\Delta_{\mathrm{F}}(x-y)\, \mathrm{i}\Delta_{\mathrm{F}}(y-y)\, \mathrm{i}\Delta_{\mathrm{F}}(y-z)\, J(x)\, J(z) \\
&\quad + \tfrac{1}{8}\lambda \int \mathrm{d}^4y\, \mathrm{d}^4z_1\, \mathrm{d}^4z_2\, \mathrm{d}^4z_3\, \mathrm{d}^4z_4\, \mathrm{i}\Delta_{\mathrm{F}}(y-z_1)\, \mathrm{i}\Delta_{\mathrm{F}}(y-z_2) \\
&\qquad\qquad \times \mathrm{i}\Delta_{\mathrm{F}}(y-z_3)\, \mathrm{i}\Delta_{\mathrm{F}}(y-z_4)\, J(z_1)\, J(z_2)\, J(z_3)\, J(z_4) \\
&\quad + O(\lambda^2). \qquad (4.3.32)
\end{aligned}
$$

The expression for $\varphi_c(x)$ as a functional of $J(x)$ may be inverted order-by-order in perturbation theory. We first apply the Klein–Gordon operator to (4.3.31):

$$
\begin{aligned}
\left[\Box + m^2\right]\varphi_c(x) \;&=\; J(x) - \tfrac{1}{2}\lambda \int \mathrm{d}^4z\, \mathrm{i}\Delta_{\mathrm{F}}(x-x)\, \mathrm{i}\Delta_{\mathrm{F}}(x-z)\, J(z) \\
&+ \tfrac{1}{3!}\lambda \int \mathrm{d}^4z_1\, \mathrm{d}^4z_2\, \mathrm{d}^4z_3\, \mathrm{i}\Delta_{\mathrm{F}}(x-z_1)\, \mathrm{i}\Delta_{\mathrm{F}}(x-z_2)\, \mathrm{i}\Delta_{\mathrm{F}}(x-z_3)\, J(z_1)\, J(z_2)\, J(z_3) \\
&\qquad\qquad + O(\lambda^2). \qquad (4.3.33)
\end{aligned}
$$

To this order in λ, we may evidently substitute $J = \left[\Box + m^2\right]\varphi_c(x) + O(\lambda)$ in the terms already multiplied by λ on the right-hand side (*cf.* the standard Dyson expansion technique) and thus obtain

$$
J(x) \;=\; \left[\Box + m^2\right]\varphi_c(x) + \tfrac{1}{2}\lambda \mathrm{i}\Delta_{\mathrm{F}}(0)\, \varphi_c(x) + \tfrac{1}{3!}\lambda\varphi_c^3(x) + O(\lambda^2), \qquad (4.3.34)
$$

where an integration by parts has been performed, to turn the Klein–Gordon operators around to act on the Feynman propagators and thus also exploit Eq. (3.3.26).

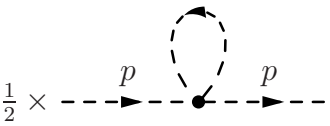
The equation derived differs from the classical equation of motion (2.4.4) on two accounts. Firstly, having introduced a source term into the Lagrangian used to construct the generating functional, we quite naturally find it here. Secondly, there is a quantum correction, in the term containing $\Delta_{\rm F}(0)$. Indeed, had we retained the $\hbar$ factors, as discussed earlier, we would have found such a term to be proportional to a positive power of $\hbar$.

The expansion for $J(x)$ may in turn be used to obtain a perturbative expansion for the effective action as an explicit functional of $\varphi_c(x)$:

$$\begin{aligned}\mathrm{i}\Gamma[\varphi_c] \;=\; \text{constant} - \tfrac{1}{8}\int \mathrm{d}^4x\, \left[\mathrm{i}\Delta_{\rm F}(0)\right]^2 - \tfrac{1}{2}\int \mathrm{d}^4x\, \varphi_c(x)\left[\Box + m^2\right]\varphi_c(x) \\ - \tfrac{1}{4}\lambda\,\mathrm{i}\Delta_{\rm F}(0)\int \mathrm{d}^4x\,\varphi_c^2(x) - \tfrac{1}{4!}\lambda\int \mathrm{d}^4x\,\varphi_c^4(x) + O(\lambda^2). \quad (4.3.35)\end{aligned}$$

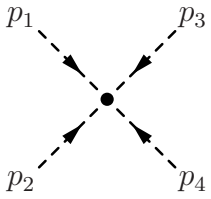
And from this we may derive the one-particle irreducible Green functions, which in momentum space are (the first term is the *inverse* propagator)

$$\begin{aligned}\mathrm{i}\Gamma^{(2)}(p,-p) \;&=\; \mathrm{i}\left[p^2 - m^2\right] - \tfrac{1}{2}\mathrm{i}\lambda\,\mathrm{i}\Delta_{\rm F}(0) + O(\lambda^2) \\ &=\; -\Big[\;\overset{p}{\longrightarrow}\;\Big]^{-1} \;+\; \tfrac{1}{2}\times \quad (4.3.36a)\end{aligned}$$



and

$$\mathrm{i}\Gamma^{(4)}(p_1,p_2,p_3,p_4) \;=\; \quad , \quad (4.3.36b)$$



where now the external lines do *not* have an associated propagator—the Green function has been effectively *amputated* by the Klein–Gordon operator associated with each external particle in (4.2.48)

4.3.5 *S*-matrix elements

We have already shown that the functional $S[\varphi_0]$ in (4.2.48) generates the scattering amplitudes:

$$S[\varphi_0] \;=\; \exp\left\{\int \mathrm{d}^4x\,\varphi_0(x)\;\mathcal{K}(x)\,\frac{\delta}{\delta J(x)}\right\} Z[J]\bigg|_{J=0}. \quad (4.3.37)$$

Recall that $Z[J]$ is the generating functional for disconnected Green functions:

$$\begin{aligned} \mathcal{G}^{(n)}(x_1,\ldots,x_n) &:= (-\mathrm{i})^n \frac{\delta^n Z[J]}{\delta J(x_1)\ldots\delta J(x_n)}\bigg|_{J=0} \\ &= \langle 0|\mathrm{T}\big[\varphi(x_1)...\varphi(x_n)\big]|0\rangle, \end{aligned} \tag{3.3.10}$$

Performing now an operator expansion:

$$\begin{aligned} \exp\biggl\{\int \mathrm{d}^4x\,\varphi_0(x)\;\mathcal{K}(x)\,\frac{\delta}{\delta J(x)}\biggr\} \;=\; \sum_{n=0}^{\infty}\frac{1}{n!}\int \mathrm{d}^4x_1\;\ldots\mathrm{d}^4x_n\,\varphi_0(x_1)\ldots\varphi_0(x_n) \\ \times\,\mathcal{K}(x_1)\ldots\mathcal{K}(x_n)\,\frac{\delta^n}{\delta J(x_1)\ldots\delta J(x_n)}, \end{aligned} \tag{4.3.38}$$

we obtain the following expansion for $S[\varphi_0]$ in terms of n-point Green functions:

$$\begin{aligned} S[\varphi_0] \;=\; \sum_{n=0}^{\infty}\frac{\mathrm{i}^n}{n!}\int \mathrm{d}^4x_1\;\ldots\mathrm{d}^4x_n\,\varphi_0(x_1)\ldots\varphi_0(x_n) \\ \times\,\mathcal{K}(x_1)\ldots\mathcal{K}(x_n)\,\mathcal{G}^{(n)}(x_1,\ldots,x_n). \end{aligned} \tag{4.3.39}$$

We thus see explicitly, as mentioned earlier, that for each of the n external particles associated with $\mathcal{G}^{(n)}(x_1,\ldots,x_n)$ there is a corresponding Klein–Gordon operator. The effect is to amputate each of the external legs and replace them with a free field φ_0, which will then have the role of creating (destroying) an incoming (outgoing) state in the distant past (future).

Physical processes studied in the laboratory usually involve plane-wave states (incoming beams and outgoing, detected, particles); moreover, as already noted, it is easier computationally to work in momentum space. We shall therefore now move over to the momentum-space representation. Inverting the Fourier transform (3.3.31), we have

$$\begin{aligned} \mathcal{G}^{(n)}(x_1,\ldots,x_n) \;=\; \int \mathrm{đ}^4p_1\;\ldots\mathrm{đ}^4p_n\,\delta^4(p_1+\cdots+p_n) \\ \times\,\mathrm{e}^{\,\mathrm{i}\,(p_1\cdot x_1+\cdots+p_n\cdot x_n)}\;\mathcal{G}^{(n)}(p_1,\ldots,p_n). \end{aligned} \tag{4.3.40}$$

Inserting this into the previous expression for $S[\varphi_0]$ leads to

$$\begin{aligned} S[\varphi_0] \;=\; & \sum_{n=0}^{\infty} \frac{\mathrm{i}^n}{n!} \int \mathrm{d}^4x_1 \,\ldots \mathrm{d}^4x_n \, \varphi_0(x_1) \ldots \varphi_0(x_n) \; \mathcal{K}(x_1) \ldots \mathcal{K}(x_n) \\ & \times \int \text{đ}^4p_1 \,\ldots \text{đ}^4p_n \, \text{đ}^4(p_1 + \cdots + p_n) \; \mathrm{e}^{\mathrm{i}(p_1\cdot x_1 + \cdots + p_n \cdot x_n)} \; \mathcal{G}^{(n)}(p_1, \ldots, p_n) \\ = & \sum_{n=0}^{\infty} \frac{\mathrm{i}^n}{n!} \int \text{đ}^4p_1 \,\ldots \text{đ}^4p_n \, \text{đ}^4(p_1 + \cdots + p_n) \\ & \times (m^2 - p_1^2) \cdots (m^2 - p_n^2) \, \mathcal{G}^{(n)}(p_1, \ldots, p_n) \\ & \times \int \mathrm{d}^4x_1 \,\ldots \mathrm{d}^4x_n \, \varphi_0(x_1) \ldots \varphi_0(x_n) \; \mathrm{e}^{\mathrm{i}(p_1\cdot x_1 + \cdots + p_n \cdot x_n)} \,. \end{aligned} \tag{4.3.41}$$

The last set of integrals above are nothing other than products of Fourier transforms of the free field φ_0:

$$\begin{aligned} \int \mathrm{d}^4x \; \mathrm{e}^{\mathrm{i}p\cdot x} \; \varphi_0(x) \;=\; & \int \mathrm{d}^4x \; \mathrm{e}^{\mathrm{i}p\cdot x} \int \frac{\text{đ}^3\boldsymbol{k}}{2k_0} \Big[b(k) \; \mathrm{e}^{-\mathrm{i}k\cdot x} + b^*(k) \; \mathrm{e}^{+\mathrm{i}k\cdot x} \Big] \Big|_{k^2=m^2} \\ = & \int \frac{\text{đ}^3\boldsymbol{k}}{2k_0} \Big[b(k) \, \text{đ}^4(p-k) + b^*(k) \, \text{đ}^4(p+k) \Big] \Big|_{k^2=m^2} \,. \end{aligned} \tag{4.3.42}$$

Applying the identity

$$\frac{\text{đ}^3\boldsymbol{k}}{2k_0}\bigg|_{k^2=m^2} \;\equiv\; \text{đ}^4k \; \text{đ}^{(+)}(k^2 - m^2), \tag{4.3.43}$$

where $\text{đ}^{(+)}(k^2 - m^2)$ is a short-hand for $\theta(k_0)\text{đ}(k^2 - m^2)$, this becomes

$$\int \mathrm{d}^4x \; \mathrm{e}^{\mathrm{i}p\cdot x} \; \varphi_0(x) \;=\; \Big[b(p) + b^*(-p) \Big] \, \text{đ}^{(+)}(p^2 - m^2). \tag{4.3.44}$$

Here we see the origin of the Feynman interpretation: creation of a particle with four-momentum p^μ is equivalent to the destruction of a particle* with four-

* Note that we do not yet have antiparticles; simple scalars are their own antiparticles.

momentum $-p^\mu$ and *vice versa*. Finally then, we obtain

$$S[\varphi_0] \;=\; \sum_{n=0}^{\infty} \frac{1}{n!} \int \text{đ}^4 p_1 \,\ldots \text{đ}^4 p_n \,\delta\!\!\!^{-4}(p_1 + \cdots + p_n)\, \widehat{\mathcal{G}}^{(n)}(p_1, \ldots, p_n) \\ \times \prod_{i=1}^{n} \Big[b(p_i) + b^*(-p_i) \Big]\, \delta\!\!\!^{-(+)}(p_i^2 - m^2), \quad (4.3.45)$$

where the so-called *amputated* n-point Green function,

$$\widehat{\mathcal{G}}^{(n)}(p_1, \ldots, p_n) \;:=\; \prod_{i=1}^{n} \left[-\mathrm{i}\,(p_i^2 - m^2) \right] \times \mathcal{G}^{(n)}(p_1, \ldots, p_n), \quad (4.3.46)$$

has been introduced to explicitly perform the cancellations between the pre-factors $-\mathrm{i}\,(p_i^2 - m^2)$ and the propagators for the external legs in $\mathcal{G}^{(n)}$. We thus see that either a particle with energy–momentum $+p^\mu$ is created by $b(p)$ or with $-p^\mu$ is destroyed by $b^*(p)$, but in any case the mass-shell condition $p^2 = m^2$ is guaranteed by the presence of the $\delta\!\!\!^{-(+)}(p_i^2 - m^2)$ factors. Note also that overall energy–momentum conservation is guaranteed by the factor $\delta\!\!\!^{-4}(p_1 + \cdots + p_n)$, which also implies that at least one of the p_i^μ's has negative energy, *i.e.* is outgoing, *and*, likewise, that at least one of the E_i's is positive, *i.e.* incoming.

Let us now construct the S-matrix element or scattering amplitude S_{fi} for a process involving n external fields, of which m are initial (or incoming) and $n-m$ final (or outgoing). All momenta must satisfy mass-shell conditions $p_i^2 = m^2$ and we must also have

$$p_1 + \cdots + p_m \;=\; p_{m+1} + \cdots + p_n. \quad (4.3.47)$$

Note that, in writing this constraint, we have explicitly flipped the sign in front of the outgoing momenta and thus too in all the following expressions. The particular matrix element we wish to single out from the generating functional $S[\varphi_0]$ then involves m factors of $b(p_i)$ and $n-m$ factors of $b^*(p_i)$. We therefore take the following set of functional derivatives:*

$$S_{fi} \;=\; \left. \frac{\delta^n S[\varphi_0]}{\delta b(p_1) \ldots \delta b(p_m)\, \delta b^*(p_{m+1}) \ldots \delta b^*(p_n)} \right|_{b=0=b^*}, \quad (4.3.48)$$

Note that b and b^* must be set to zero after the derivatives in order that only the required n-point contribution should survive. Using (4.3.45) and performing the

* The Lorentz-covariance pre-factor appearing in, *e.g.*, Bailin and Love (1993) is absent here as we shall use a more explicitly covariant definition of the matrix element.

functional derivatives we obtain

$$S_{fi} \;=\; \widehat{\mathcal{G}}^{(n)}(p_1,\ldots,p_m,-p_{m+1},\ldots,-p_n)\,\tilde{\delta}^4(p_1+\cdots+p_m-p_{m+1}-\cdots-p_n). \quad (4.3.49)$$

We thus discover that the Feynman rules for calculating scattering amplitudes are just those already derived by considering the connected n-point Green functions, with the exception that external lines here have *no* associated propagator and are also constrained to lie on their mass shell. Moreover, here we see that *all* diagrams, both connected (one-particle *reducible* and not) *and* disconnected must be included. This is because real physical processes do not distinguish the paths by which the final states are reached. We shall usually be interested in processes with, however, just one or two particles in the initial state. And for those in which there are two initial particles, we shall usually require that either the energies and momenta in the final state be different or that the particles themselves should have undergone some change (*e.g.* $e^+e^-\to\mu^+\mu^-$). The case of decay amplitudes trivially satisfies such requirements. Therefore, the disconnected diagrams will, in general, be excluded on physical grounds. The question of one-particle irreducibility will become important in dealing systematically with quantum corrections and the ensuing problem of *renormalisation*.

4.3.6 Cross-sections

As the last task in this chapter, we shall now turn the Feynman rules just described into a complete calculation of a physical scattering cross-section. Our somewhat elementary $\lambda\varphi^4$ toy model does not allow for any particularly exotic processes (although with sufficient energy, more particles could be produced, *e.g.* in $2\to n$ processes with $n>2$). We shall thus consider the simplest process: namely, $2\to 2$. That is, two initial particles having momenta $p^\mu_{1,2}$ scattering into a final state with momenta $p^\mu_{3,4}$, satisfying

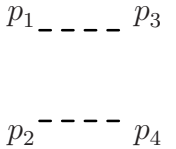
$$p_1^\mu + p_2^\mu \;=\; p_3^\mu + p_4^\mu. \quad (4.3.50)$$

The relevant S-matrix element is then

$$S_{fi} \;=\; \widehat{\mathcal{G}}^{(4)}(p_1,p_2,-p_3,-p_4)\,\tilde{\delta}^4(p_1+p_2-p_3-p_4). \quad (4.3.51)$$

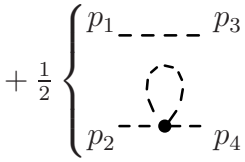
The (amputated) diagrams contributing to S_{fi} are

$$\widehat{\mathcal{G}}^{(4)}(p_1,p_2,-p_3,-p_4) \;=$$



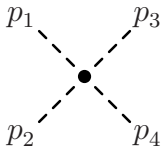
$$+\,\text{perms}(p_2,p_3,p_4)$$

$$+\tfrac{1}{2}\left\{\begin{array}{l}\\ \\ \\ \end{array}\right.$$



$$+\text{perms}(p_1,p_2,p_3,p_4)\left.\begin{array}{l}\\ \\ \\ \end{array}\right\}$$

$$+$$



$$+\,O(\lambda^2). \tag{4.3.52}$$

There are thus ten diagrams in total. However, the first nine are excluded once we require non-trivial scattering, *i.e.* that all individual initial and final momenta be different.

We thus have the rather simple result that the transition amplitude is

$$S_{fi} \;=\; -\mathrm{i}\,\lambda\,\bar{\delta}^4(p_1+p_2-p_3-p_4). \tag{4.3.53}$$

As usual, the transition probability is given by the squared modulus of the amplitude:

$$P_{fi} \;=\; |S_{fi}|^2 \;=\; \lambda^2\,\bar{\delta}^4(p_1+p_2-p_3-p_4)\,\bar{\delta}^4(0). \tag{4.3.54}$$

The apparently infinite factor $\bar{\delta}^4(0)$ arises from an integral of the following type:

$$\int \mathrm{d}^4x\;\mathrm{e}^{\mathrm{i}P\cdot x}\,\Big|_{P=0}. \tag{4.3.55}$$

This may be regulated, as in the derivation of Fermi's golden rule, by limiting the integration four-volume to a finite box of, say, time interval T and three-volume V. We then have

$$\bar{\delta}^4(0) \;\to\; VT. \tag{4.3.56}$$

Thus, since we really require the transition probability per unit volume per unit time, we divide P_{fi} by VT to obtain the transition rate

$$R_{fi} \;=\; P_{fi}/VT \;=\; \lambda^2\,\bar{\delta}^4(p_1+p_2-p_3-p_4). \tag{4.3.57}$$

Now, the flux of particles scattered into a given element of solid angle $\mathrm{d}\Omega$ by a single target particle is given by

$$\mathrm{d}N_{fi} \;=:\; \Phi\,\mathrm{d}\sigma_{fi}\,, \tag{4.3.58}$$

where Φ is the beam flux and the differential cross-section $\mathrm{d}\sigma_{fi}$ is defined precisely by this equation. Let us first calculate the incident flux. It is easiest to consider

a fixed target (particle 1, say), and to work in its rest frame (often called the laboratory frame); we may boost to, say, the centre–of–mass frame later.* In the target rest frame the incident flux is simply given by the product of the beam density and velocity.

We make the natural choice of the covariant three-momentum integration measure already given in (4.3.43). And so the states must be normalised as

$$\langle p|p'\rangle \;=\; 2p_0\,\delta^3(\boldsymbol{p}-\boldsymbol{p}'), \tag{4.3.59}$$

which is equivalent to $2p_0$ particles per unit volume. The incident flux is thus

$$\Phi \;=\; 2E_2\,v_2. \tag{4.3.60}$$

We must also divide $\mathrm{d}N_{fi}$ by $2E_1$ to obtain the transition probability per target particle, giving

$$\mathrm{d}N_{fi} \;=\; \frac{1}{2E_1}\,\lambda^2\,\delta^4(p_1+p_2-p_3-p_4)\,\mathrm{đ}^4p_3\,\delta^{(+)}(p_3^2-m^2)\,\mathrm{đ}^4p_4\,\delta^{(+)}(p_4^2-m^2), \tag{4.3.61}$$

Rewriting this all to give the differential cross-section, we have

$$\mathrm{d}\sigma_{fi} \;=\; \frac{1}{2E_1\,2E_2\,v_2}\,\lambda^2\,\delta^4(p_1+p_2-p_3-p_4)\,\mathrm{đ}^4p_3\,\delta^{(+)}(p_3^2-m^2)\,\mathrm{đ}^4p_4\,\delta^{(+)}(p_4^2-m^2), \tag{4.3.62}$$

We can now combine the flux factor with the target-density factor, calculated in the target rest frame, and rewrite them in a manifestly covariant form exploiting the relation $p_1{\cdot}p_2 = mE_2$, which leads to

$$2E_1\,2E_2\,v_2 \;=\; 4m|\boldsymbol{p}_2| \;=\; 4\sqrt{m^2(E_2^2-m^2)} \;=\; 4\sqrt{(p_1{\cdot}p_2)^2-m^4}. \tag{4.3.63}$$

One of the four-momentum integrals may be performed using the δ-function and the other may be directly reduced to a three-momentum integral using (4.3.43):

$$\mathrm{d}\sigma_{fi} \;=\; \frac{\lambda^2}{4\sqrt{(p_1{\cdot}p_2)^2-m^4}}\,\delta^{(+)}\left((p_1+p_2-p_3)^2-m^2\right)\frac{\mathrm{đ}^3\boldsymbol{p_3}}{2E_3}, \tag{4.3.64}$$

Transforming now to the centre–of–mass frame for simplicity, we may write

$$p_1 \;=\; (E,+\boldsymbol{p}) \quad\text{and}\quad p_2 \;=\; (E,-\boldsymbol{p}). \tag{4.3.65}$$

* In fact, our explicitly Lorentz-covariant formulation make such considerations superfluous.

The three-momentum integral over $\boldsymbol{p}_3$ may be rewritten as

$$\frac{\text{đ}^3\boldsymbol{p}_3}{2E_3} \;=\; \frac{|\boldsymbol{p}_3|^2}{2E_3}\,\text{đ}|\boldsymbol{p}_3|\,\text{đ}^2\Omega \;=\; \frac{|\boldsymbol{p}_3|}{16\pi^3}\,\mathrm{d}E_3\,\mathrm{d}^2\Omega\,, \tag{4.3.66}$$

where the scattering angles subsumed in $\text{đ}^2\Omega$ are measured with respect to to the direction of particle 1. In the centre–of–mass frame

$$\begin{aligned}(p_1+p_2-p_3)^2-m^2 \;&=\; (p_1+p_2)^2-2(p_1+p_2)\cdot p_3+p_3^2-m^2\\ &=\; s-4EE_3,\end{aligned} \tag{4.3.67}$$

where $s:=(p_1+p_2)^2=4E^2$. The remaining δ-function may thus be used to eliminate the integration in E_3 to finally obtain

$$\frac{\mathrm{d}\sigma_{fi}}{\mathrm{d}\Omega} \;=\; \frac{\lambda^2}{64\pi^2 s}. \tag{4.3.68}$$

The full set of (invariant) so-called Mandelstam variables are defined as follows:

$$s \;:=\; (p_1+p_2)^2 \;\equiv\; (p_3+p_4)^2, \tag{4.3.69a}$$

$$t \;:=\; (p_1-p_3)^2 \;\equiv\; (p_2-p_4)^2, \tag{4.3.69b}$$

$$u \;:=\; (p_1-p_4)^2 \;\equiv\; (p_2-p_3)^2, \tag{4.3.69c}$$

and satisfy

$$s+t+u \;=\; 4m^2. \tag{4.3.70}$$

Exercise 4.3.3. *Derive Eq. (4.3.70), find t in terms of s and the centre–of–mass angle θ_{CM}, and determine the kinematical limits on t.*

Exercise 4.3.4. *Integrate out the azimuthal angle ϕ and substitute t as the second independent variable in place of $\cos\theta$ (where θ is the usual scattering angle), to show that we may recast the two-body differential cross-section in the following Lorentz-invariant form:*

$$\frac{\mathrm{d}\sigma_{fi}}{\mathrm{d}t} \;=\; \frac{\lambda^2}{16\pi s^2}. \tag{4.3.71}$$

Exercise 4.3.5. *Explicitly evaluate the second-order,* i.e. *$O(\lambda^2)$, perturbative contributions to $Z[J]$ and $Z_c[J]$. Using the results, verify that the Feynman rules already outlined are indeed correct to this order.*

Exercise 4.3.6. *Explicitly evaluate the $O(\lambda^2)$ contributions to the two- and four-point Green functions $\mathcal{G}^{(2)}(p,-p)$ and $\mathcal{G}^{(4)}(p_1,p_2,p_3,p_4)$. In particular, obtain the numerical multiplicative factors and verify their designation as symmetry factors according to the Feynman rules given earlier.*

4.4 Bibliography

Bailin, D. and Love, A. (1993), *Introduction to Gauge Field Theory* (IOP Pub.), revised edition.

Gordon, W. (1926), *Z. Phys.* **40**, 117.

Klein, O. (1927), *Z. Phys.* **41**, 407.

Lehmann, H., Symanzik, K. and Zimmermann, W. (1955), *Nuovo Cim.* **1**, 205.

Chapter 5

Renormalisation

5.1 The need for renormalisation

As we have already begun to see, an effect of quantisation is to "correct" the values of classical functions or variables and to shift or *renormalise* them with respect to to their *naïve* values. The first examples encountered were the bubble or loop diagrams for the vacuum and two-point functions shown in (4.3.11 & 12). In fact, analogous effects have already been encountered in perturbation theory applied to non-relativistic quantum mechanics.* In particular, we know that the position of the propagator pole is shifted in the complex plane. That is, the physical mass is shifted by a real amount and may also acquire an imaginary contribution, which corresponds to a finite width or decay rate. There should be no mystery in such effects: having introduced an interaction, the free-field mass, for example, gains a contribution due to the (self-) interaction energy.

More specifically, if we examine the perturbative expansion for the effective action in Eq. (4.3.35), the new term quadratic in φ_c,

$$-\tfrac{1}{4}\lambda\,\mathrm{i}\Delta_{\mathrm{F}}(0)\int \mathrm{d}^4x\,\varphi_c^2(x), \tag{5.1.1}$$

should evidently be interpreted as precisely a shift in the mass of the particle. The same may be deduced from the quantum-corrected equation of motion (4.3.34) with the external source term set to zero:

$$0 \;=\; \left[\Box + m^2\right]\varphi_c(x) + \tfrac{1}{2}\lambda\mathrm{i}\Delta_{\mathrm{F}}(0)\,\varphi_c(x) + \tfrac{1}{3!}\lambda\varphi_c^3(x) + O(\lambda^2). \tag{5.1.2}$$

* See, for example, the previous course on advanced quantum mechanics by the same author.

We are thus led to the following identification:

$$m_R^2 \;=\; m^2 + \tfrac{1}{2}\lambda\,\mathrm{i}\,\Delta_{\mathrm{F}}(0) + O(\lambda^2), \tag{5.1.3}$$

where the subscript "R" stands for *renormalised*. In fact, m_R^2 is what we shall call the *physical* mass, since it corresponds to the mass that would be measured experimentally (*e.g.* as the position of a Breit–Wigner resonance pole) while m^2 is therefore simply a dummy parameter having no real *physical* significance in the full interacting theory. Henceforth, we shall also indicate such so-called "*bare*" quantities with the subscript "B".*

Indeed, the entire Lagrangian (including its parameters and variables), from which we have so far derived the Feynman rules, should be denoted as *bare*:

$$\mathcal{L}_B \;=\; \tfrac{1}{2}\big[(\partial^\mu\varphi_B)(\partial_\mu\varphi_B) - m_B^2\varphi_B^2\big] - \tfrac{1}{4!}\lambda_B\varphi_B^4. \tag{5.1.4}$$

Moreover, all quantities so far constructed, such as the Green functions of the theory, should be termed *bare* since they are calculated in terms of the bare parameters. Proceeding further, we shall find a similar phenomenon of renormalisation of the coupling constant λ. And, since the factor $\frac{1}{2}$ multiplying the derivative and mass terms is, in fact, arbitrary, we should also expect to find that the field φ itself is renormalised by quantum effects.† Finally, the self-interactions of our scalar theory may even lead to the generation of *effective* couplings of higher-order in the number of external fields.

We are therefore faced with the problem that the (bare) quantities appearing in the Lagrangian are not the actual parameters that can be measured and therefore do not have any intrinsic meaning while, on the other hand, the quantum theory generates the physical (renormalised) quantities via what are known as *quantum corrections*, which we shall soon discover are often functions of kinematical variables, such as the centre–of–mass energy or the external field momenta. We shall thus need to *define*, say, the physical coupling constant as, *e.g.*, the value of the *renormalised* four-point function $\mathcal{G}_B^{(4)}(p_1,p_2,p_3,p_4)$ for some specific values of the momenta $p_1, \ldots, p_4$. This quantum-induced energy or *scale* dependence of the coupling constant (and all other physical parameters) is perhaps the single most important aspect of a quantum field theory. Indeed, the very expression coupling *constant* will immediately reveal itself to be a misnomer.

This apparently simple fact has repercussions of profound importance. Consider for a moment a massless theory, *i.e.* one that contains no reference to any mass or energy scale. The *classical* theory then possesses a symmetry known as *con-*

* When no ambiguity should ensue, we shall omit either or both of the new subscripts.

† Recall indeed that in non-relativistic perturbation theory the normalisation of the states has to be recalculated.

formal invariance; that is, apart from the natural (dimensional) scale dependence arising from the dimensionality of, say, cross-sections or other measurable quantities, there should be no variation if we rescale all the momenta in the process by some common factor. However, quantisation of the theory *induces* such a scale dependence through the renormalisation just discussed. In massless, two-body scattering, the only truly independent dimensional parameter is the Mandelstam variable s but the coupling λ is *dimensionless*. It can therefore only be a function of other dimensionless quantities. The renormalisation process will be found then to introduce a new scale: the so-called *renormalisation scale*. This phenomenon, *i.e.* the mechanism by which a dimensionless parameter is transformed into a dimensional (physical) parameter, is known as *dimensional transmutation*. This is perhaps the first example of what is usually known as an *anomaly*, that is the breaking of a classical symmetry by quantum effects.

The immediate consequence is that while we write the Lagrangian in terms of bare or unrenormalised quantities (including the fields themselves), all physical quantities, such as cross-sections or decay rates must eventually be expressed purely in terms of physical or renormalised quantities. The systematic theoretical procedure constructed to achieve this reparametrisation is called *renormalisation*, the subject of the present chapter. At this point one might justifiably question the necessity of a systematic procedure. After all, the previous equations provide the connection between the renormalised and bare parameters of the theory. However, there are two important observations:

1. The new terms, *e.g.* as in Eq. (5.1.2), generally turn out to be *infinite* and thus so too must the bare quantities be if the physical parameters are to remain finite.
2. The quantum corrections may effectively *generate* new, again apparently infinite, effective couplings and thus new parameters.

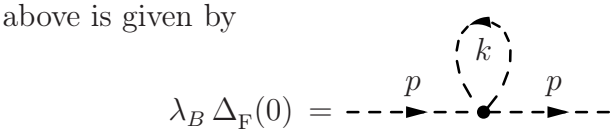
The first difficulty absolutely demands attention and we shall see in the course of the next few lessons how to deal with it. The second turns out to provide a discriminating criterion in the choice of theories that may be seriously consider as *fundamental*.

In our toy φ^4 theory it turns out that the number of new couplings increases with increasing perturbation order and thus one needs ever more measured processes to determine them—this severely limits the predictive power of some theories to the point of rendering them essentially useless. As we shall see, QED and QCD, for example, belong a class of theories that do *not* suffer this limitation: the (gauge) symmetries they possess protect against such proliferation of couplings and renders them *renormalisable*. This term is reserved for precisely those theories that do not generate new (infinite) parameters as we move to higher perturbative orders and which can therefore be entirely defined in term of a finite number of physical

parameters. Electrodynamics is the paradigm: the theory describing the electron and its electromagnetic interactions requires just two physical renormalised parameters, corresponding to the electron mass m_e and the fine-structure constant α. In addition, the electron wave-function is also renormalised, but the new renormalisation constants Z_e is entirely determined by the other two via the requirements of gauge invariance. Even this is not the complete problem: gravity may be described by a zero-mass field, the graviton, but owing to its essential non-linearity, the quantum theory of gravitation is non-renormalisable. Indeed, to date no fully consistent quantum theory of gravitation has been constructed.

5.2 Divergences in loop diagrams

As a first step, let us examine the nature of the infinities already encountered. In the momentum-space representation we see that the infinities are simply the result of an unlimited integration over a loop four-momentum. For example, the mass correction above is given by

$\lambda_B \, \Delta_{\rm F}(0) \;=$ 

$$= \lambda_B \int \frac{\mathrm{đ}^4 k}{\left(k^2 - m_B^2 + \mathrm{i}\varepsilon\right)}. \tag{5.2.1}$$

The presence of the "$\mathrm{i}\varepsilon$" term shifts the pole in the propagator away from the real axis and thus we immediately see that the pole singularity is *not* the cause of the infinity (which would be what is called an infrared divergence). The problem lies in the ultraviolet behaviour.

To calculate the integral, we first separate the energy and three-momentum components. Consider the energy integral in the complex k_0 plane. For $\varepsilon = 0^+$ the poles in k_0 appear below (above) the real axis to the right (left) of the imaginary axis, see Fig. 5.1. We may therefore rotate the k_0 contour from the real to the imaginary axis (the arcs contribute zero at infinite radius). This is equivalent to the substitution $k_0 \to \mathrm{i}\, k_4$ with the space part unchanged and the variable becomes Euclidean $k \to k_E$:

$$\tfrac{1}{2}\mathrm{i}\,\lambda_B \, \Delta_{\rm F}(0) \;=\; \tfrac{1}{2}\mathrm{i}\,\lambda_B \int \frac{\mathrm{i}\mathrm{đ}^4 k_E}{\left(-k_E^2 - m_B^2\right)}, \tag{5.2.2}$$

where now $k_E^2 = \boldsymbol{k}^2 + k_4^2$. It is evident that, while there is actually then no singu-

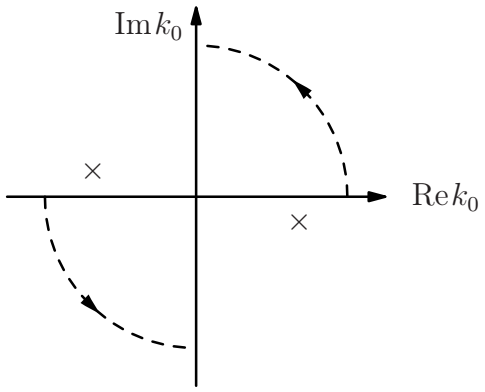


Figure 5.1: The Wick rotation to the complex k_0 plane; the poles lie at the two points $k_0 = \pm[\sqrt{(\boldsymbol{k}^2 + m_B^2)} - \mathrm{i}\varepsilon]$, indicated × in the figure.

larity for small values of $|k_E|$,* the integral diverges as $|k_E| \to \infty$. In this particular case, writing $\mathrm{đ}^4 k_E = k_E^3 \mathrm{đ}k_E \, \mathrm{d}\Omega_4$ (where $\mathrm{d}\Omega_4$ is the Euclidean four-space solid-angle element), we see that the integral diverges quadratically. That is, if we introduce a cut-off $|k_E| < \Lambda$ for some very large scale Λ, then the integral is proportional to Λ^2.

Using the same approach, let us examine one more case before attempting to generalise. Consider the one-loop, order λ_B^2, correction to the four-point Green function, $\mathcal{G}_B^{(4)}$ shown in Fig. 5.2.† The Feynman rules lead to the following integral

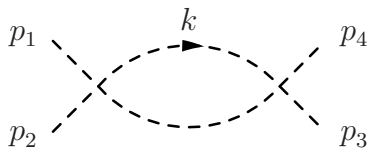


Figure 5.2: A one-loop, order λ_B^2, correction to the four-point Green function, $\mathcal{G}_B^{(4)}$. To this must be added the same diagram with the other two possible permutations of the connections between the momenta $p_{2,3,4}$.

$$-(\mathrm{i}\lambda_B)^2 \int \frac{\mathrm{i}\,\mathrm{đ}^4 k_E}{\left[-k_E^2 - m_B^2\right]\left[-(p_E - k_E)^2 - m_B^2\right]}, \tag{5.2.3}$$

* We have though a hint of the infrared (IR) problems to be expected in the case of massless field theories.

† Note that while the subscript "B" attached to parameters indicates those of the original Lagrangian, in the case of derived quantities it means those expressed in terms of the bare parameters and which therefore do not necessarily correspond immediately to physical quantities.

where p_E is the Euclidean version of $p := p_1 + p_2$. In this case the ultraviolet behaviour of the integrand is such that using a cut-off as before leads to a logarithmic divergence.

Higher-order Feynman diagrams generally lead to integrals over products of propagators. The degree of divergence of a single integral is evidently $4-2I$ where I is the number of internal lines contributing to the integral. The case $I=1$ leads to a quadratic divergence, $I=2$ a logarithmic divergence and $I>2$ will be convergent. We can thus define the "*naïve*" or superficial degree of divergence of an entire, multiloop diagram (with L loops) as

$$D \;:=\; 4L - 2I. \tag{5.2.4}$$

We say *naïve* or superficial since it is only a sort of average divergence and the various sub-integrals may be relatively more or less divergent.

As already noted, there exist relations between L, V (the number of vertices) *etc.*, such as $L = 1 - V + I$ and $E + 2I = 4V$, thus (for a $\lambda\varphi^4$)

$$D \;=\; 2I - 4(V-1) \;=\; 4 - E. \tag{5.2.5}$$

Remarkably then, we find that the superficial degree of divergence only depends on the number of external legs and not on the number of loops or vertices. The actual value of D depends on the particular form of the theory and would be different for, say, $\lambda\varphi^3$ or a theory describing fermions *etc.* However, the important consequence is that only a finite and very limited number of Green functions are superficially divergent (for *any* given theory). In our case we see that only $\mathcal{G}^{(0)}$, $\mathcal{G}^{(2)}$ and $\mathcal{G}^{(4)}$ are superficially divergent (there are no odd n-point functions in a φ^4 theory). However, we have termed this "superficial" degree of divergence for a good reason: it is easy to imagine diagrams that are superficially convergent but have at least one divergent sub-diagram. On the other hand, if $D<0$ for the entire diagram *and* for each and every sub-diagram, then it is convergent (Weinberg, 1960)

Exercise 5.2.1. *By considering theories with a single fundamental N-point vertex and thus generalising the previous discussion, show that the requirements of renormalisability determine the restriction $N \leq 4$.*

The simplest example, in our theory, of a superficially convergent diagram that nevertheless diverges is given in Fig. 5.3. The right-hand loop is convergent while that on the left diverges. However, it is also evident that the divergence of the left-hand loop is intimately related to the divergence already encountered in the four-point function at a lower order in perturbation theory. This leads almost directly to the idea that lies at the heart of the concept of renormalisability: if we can cure the divergence in, *e.g.* the one-loop four-point, then we might hope

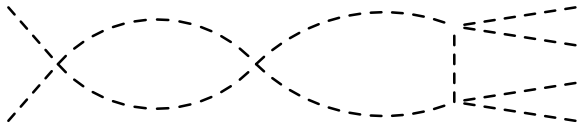


Figure 5.3: The simplest, two-loop, order λ_B^4, divergent correction to the six-point scalar-field Green function $\mathcal{G}_B^{(6)}$.

that the related divergence will be automatically removed from all higher-order diagrams in which it appears as a sub-diagram. More precisely, a theory will be renormalisable if (and only if) *all* of the divergences, to *all orders* in perturbation theory, may be removed by simple redefinitions of the parameters present in the bare theory. That this occurs is by no means guaranteed *a priori* for any given classical theory and must be verified explicitly for each new construction.

One further observation is that any given diagram (and therefore the corresponding Green function) may have a true degree of divergence *inferior* to the value given in Eq. (5.2.5). This can be for a variety of reasons; the common denominator though is almost always a symmetry. The effect may be to entirely exclude certain diagrams or to ensure cancellations between different diagrams. The classic and most important example is the gauge symmetry possessed by QED and QCD. Indeed, we shall find that the gauge principle is central to the construction of renormalisable theories. It must be stressed that renormalisability is not merely a fanciful mathematical requirement. A theory that is *not* renormalisable will usually generate new *effective* interactions at higher orders. That is, new coupling constants that were not present at the classical level will need to be determined. In general, with increasing perturbative order, the number of new parameters grows rapidly and *without bound*. In this case, to achieve higher accuracy, ever more *distinct* experimentally measurable quantities are needed simply to define the theory, which therefore inevitably loses true predictive power. An example of such a case is the effective theory describing the nucleon–nucleon interaction via exchange of a pseudoscalar field representing the pion. If low precision is sufficient, then one can obtain real predictions, but as soon as higher precision and therefore higher orders are required, the number of independent data available are insufficient to determine all the new parameters introduced.

5.3 Regularisation prescriptions

In order to proceed, we now need a practical *regularisation* prescription for performing the ill-defined or divergent integrals. The idea is that by the time we

arrive at the expression for a *physical* quantity (such as a cross-section) all of the divergences will have been eliminated in favour of the finite physical parameters. Whatever method has been adopted, the regularisation may then be removed to provide the final unambiguous result.

5.3.1 Cut-off regularisation

The conceptually simplest method of renormalisation is to place an upper bound or *cut-off* on the (Euclidean) momentum-space integrals, thus: $k_E^2 < \Lambda^2$, where Λ is some very large mass or energy scale. This is what has already been suggested in previous sections and it allows us to define the concept of degree of divergence. For the case of a φ^4 theory, this is quite sufficient and once we have prescribed a subtraction procedure for the resulting Λ^2 and $\ln\Lambda$ terms, then the theory may be rendered finite and well-behaved. Obviously, the final step here is to let $\Lambda \to \infty$. Such a procedure has, however, serious drawbacks: it manifestly lacks Lorentz invariance and also turns out to violate any gauge symmetry present. Thus, for example, in those situations where relativistic invariance or, more typically, gauge symmetry is vital, we can expect problems.

5.3.2 Pauli–Villars regularisation

As already noted, the superficial divergence of a diagram or set of diagrams nay be cured by the presence of a symmetry that leads to cancellations among such diagrams. This leads to the idea of another simple method of regularisation, due Pauli and Villars (1949). We introduce a new field with very similar properties the original, but such that the propagator takes on the following form (note the minus sign):

$$\mathrm{i}\Delta_{\mathrm{F}}(p) \;=\; -\frac{\mathrm{i}}{p^2 - \Lambda^2 + \mathrm{i}\varepsilon}. \tag{5.3.1}$$

For each internal loop then there will be two contributions, leading to terms of the following form

$$\frac{\mathrm{i}}{p^2 - m^2 + \mathrm{i}\varepsilon} - \frac{\mathrm{i}}{p^2 - \Lambda^2 + \mathrm{i}\varepsilon} \;=\; \frac{\mathrm{i}(m^2 - \Lambda^2)}{(p^2 - m^2 + \mathrm{i}\varepsilon)(p^2 - \Lambda^2 + \mathrm{i}\varepsilon)}. \tag{5.3.2}$$

The effect is to introduce an extra factor of p^2 into the denominator in the high-energy limit. An equivalent procedure would be to simply multiply all divergent diagrams by a suitable power of $-\Lambda^2/(p^2 - \Lambda^2 + \mathrm{i}\varepsilon)$.

The field introduced is, of course, very non-physical. The propagator and therefore the kinetic and mass terms have the "wrong" sign. However, at the end of the calculations, when again we let $\Lambda \to \infty$, the extra field is essentially excluded

or decoupled from the theory as it becomes infinitely massive. While this approach fully respects Lorentz invariance, it still manifestly violates any gauge symmetry present—a fake *massive* photon must be introduced to regularise divergent photon loops.

5.3.3 Dimensional regularisation

The method most widely adopted today is based on varying the number of space–time dimensions in the loop integrals. Consider the logarithmically divergent integral d^4k_E/k_E^4 in the ultraviolet (UV) limit, but now replace the number of space–time dimensions with $D=4-2\varepsilon<4$, where ε is a positive, infinitesimal quantity (the choice of the coefficient 2 is merely for later convenience). The behaviour is evidently rendered less divergent and, in fact, the integral will now converge. Such a method turns out to be not only to respect gauge symmetries, when present in the classical theory, but also relatively simple to implement. While our $\lambda\varphi^4$ theory does not possess gauge symmetry and indeed any method would work equally well, we shall, for didactic purposes, apply the dimensional method immediately and for all calculations in this volume. We shall also see that this method automatically regularises even quadratically divergent integrals.

Before examining any specific integrals, there is one further highly non-trivial consideration to be made. The dimensional method consists, in principle, of varying the number of space–time dimensions in *all* aspects of the theory, at least until the renormalisation procedure has been completed. The path-integral formulation is based on the principle of stationary action and thus on expressions of the form $\exp\{\mathrm{i}\int \mathrm{d}^4x\,\mathcal{L}(x)\}$. Now, if we simply change the dimensionality of the integral, then the action will no longer be dimensionless. Therefore, to maintain the action dimensionless, we must also modify the continuation to a D-dimensional space–time measure:*

$$\mathrm{d}^4x \;\to\; \mu^{D-4}\,\mathrm{d}^Dx\,. \tag{5.3.3}$$

Correspondingly, we must define the continuation to a D-dimensional energy–momentum measure as

$$đ^4k \;\to\; \mu^{4-D}\,đ^Dk\,, \tag{5.3.4}$$

where μ is an arbitrary mass parameter.

We thus see that a new scale (μ in dimensional regularisation or Λ in the previous two methods) is *always* introduced even in theories that possess *no* natural scale at the classical level, such as our massless $\lambda\varphi^4$ theory. We shall see that this occurrence has deep and far-reaching repercussions for all quantised field theories.

* Note that this modification is often presented in the form of a modification of the coupling constant alone: $\lambda^2\to\mu^{4-D}\lambda^2$. While it is indeed always present in the divergent integrals with the same power, the definition presented here appears more generally self-consistent.

5.3.4 A first integral in dimensional regularisation

Let us now examine the simplest of divergent integrals: the mass correction derived earlier,

$$\int \frac{đ^4k}{\left[k^2 - m_B^2 + \mathrm{i}\varepsilon\right]} \;\to\; I(D;1) \;:=\; \int \frac{\mu^{4-D} đ^D k}{\left[k^2 - m_B^2 + \mathrm{i}\varepsilon\right]}$$
$$\to\; \int \frac{\mu^{4-D}\,\mathrm{i}\,đ^D k_E}{\left[-k_E^2 - m_B^2\right]}, \tag{5.3.5}$$

where in the second line the standard Wick rotation has been applied to continue over to Euclidean space. The integrand has no angular dependence and therefore the integral may be evaluated in spherical polar coordinates using

$$đ^D k_E \;=\; \frac{|k_E|^{D-1}\mathrm{d}|k_E|\,\mathrm{d}\Omega_D}{(2\pi)^D}, \tag{5.3.6}$$

where $\mathrm{d}\Omega_D$ is a $(D-1)$-dimensional integral over the D-dimensional Euclidean-space solid angle. Now,

$$\int \mathrm{d}\Omega_D \;=\; \frac{\pi^{D/2}}{\Gamma(1+D/2)}, \tag{5.3.7}$$

where the Euler Gamma function, $\Gamma(m+1)=m!$, has singularities for m a negative integer. Next, substitute $|k_E|^2 = m_B^2 x$; then, provided $D<4$, the integral is well-defined and gives just a so-called beta function, or product of Gamma functions, and we obtain

$$I(D;1) \;=\; -\frac{\mathrm{i}}{(4\pi)^{D/2}}\,\mu^{4-D}\,m_B^{D-2}\,\Gamma(1-D/2). \tag{5.3.8}$$

In fact, we see that, provided $D\neq 4$ (or any larger integer), the integral is well-defined and finite. The dimensional regularisation procedure thus regulates *all* possible UV-divergent integrals.

We shall, of course, need to separate out the singularity; we therefore expand the Gamma function around $D=4$ (or equivalently $\varepsilon=0$):

$$\Gamma(1-D/2) \;=\; \frac{\Gamma(3-D/2)}{(1-D/2)(2-D/2)}$$
$$=\; \frac{\Gamma(1+\varepsilon)}{(\varepsilon-1)\varepsilon} \;=\; -\left(\frac{1-\gamma_{\mathrm{E}}\varepsilon}{1-\varepsilon}\right)\frac{1}{\varepsilon}, \tag{5.3.9}$$

where the Euler–Mascheroni constant is $\gamma_{\rm E} \simeq 0.5772157$. Since there is a pole at $\varepsilon = 0$, we must take care to expand *all* terms involving D. Putting this all back into the one-loop two-point correction and then expanding in ε, we finally have

$$\begin{aligned}\Delta_{\rm F}(0) \;=\; I(D;1) \;&=\; {\rm i}m_B^2\,\frac{1}{(4\pi)^2}\left[\frac{4\pi\mu^2}{m_B^2}\right]^{\varepsilon}\left(\frac{1-\gamma_{\rm E}\varepsilon}{1-\varepsilon}\right)\frac{1}{\varepsilon} \\ &=\; {\rm i}m_B^2\,\frac{1}{(4\pi)^2}\left[\frac{1}{\varepsilon}+1+\ln 4\pi-\gamma_{\rm E}+\ln\frac{\mu^2}{m_B^2}+O(\varepsilon)\right]. \quad (5.3.10)\end{aligned}$$

Comments are now in order on some of the terms in the previous expression:

The $1/\varepsilon$ pole

The singularity of the integral has been simply and conveniently transformed into a pole in ε, which we note is constant, *i.e.* independent of the specific details of the process considered. We shall later show that this allows us to remove it systematically by absorption into a redefinition of the (unphysical) bare mass. Note too that, at one loop, all orders of divergence (*e.g.* quadratic and logarithmic alike) manifest themselves as simple poles in ε.

The finite constant terms

There are also a few finite terms that are constant with respect to ε, of which some are again independent of the details of the process. These too may be reabsorbed although, since the combination $(\ln 4\pi - \gamma_{\rm E})$ is recurrent and is associated with the D-dimensional extension of the measure, it is often more convenient to redefine the measure so as to avoid it. Moreover, the term $\ln\mu^2/m_B^2$ may be eliminated by the choice $\mu^2 = m_B^2$, for the arbitrary parameter introduced.

The $O(\varepsilon)$ terms

Finally, there are terms of higher order in ε. However, provided there are no further poles in the diagram, these may safely be neglected as they will vanish in the limit $\varepsilon \to 0$, to be taken at the end of the calculation.

We should also comment that all simple integrals, such as that evaluated above, are well-known and tabulated in many textbooks dealing with the calculation of Feynman diagrams (see, for example, Bailin and Love, 1993, p. 81). Let us just provide one warning: the change in the number of space–time dimensions has numerous repercussions and great care must be taken when first using this technique. For example, the metric tensor now has trace $g^{\mu}{}_{\mu} = D$. There will also

be differences in the Dirac algebra relevant for spinor fields and the matrix γ_5 presents particular problems.

5.4 Evaluation of loop diagrams

The loop diagram just evaluated represents the simplest possible case. We shall now examine a more general type of diagram, in order to pave the way to calculation of the renormalisation constants applicable to a quantised field theory. The particular simplicity of the previous calculation lay in the presence of just a single Feynman propagator. In general, loop integrals involve more than one propagator. However, except in rather extreme theories, divergent diagrams do not involve more than three propagators at the one-loop level. The three-propagator case is rather particular and occurs in the so-called triangle diagram. Since it has special importance in gauge theories, we shall examine it closely later in the volume.

So, for the time-being the most we shall have to tackle is a two-propagator diagram such as that occurring in the one-loop correction to the four-point function, see Fig. 5.2. The integral to evaluate was given in Euclidean space in Eq. (5.2.3):

$$\begin{aligned} I(D;1,1) &:= \int \frac{\mu^{4-D}\text{đ}^D k}{\left[k^2 - m_B^2 + \mathrm{i}\varepsilon\right]\left[(p-k)^2 - m_B^2 + \mathrm{i}\varepsilon\right]} \\ &= \int \frac{\mathrm{i}\mu^{4-D}\text{đ}^D k_E}{\left[-k_E^2 - m_B^2\right]\left[-(p_E - k_E)^2 - m_B^2\right]}. \end{aligned} \tag{5.4.1}$$

A trick, due to Feynman, allows the two denominators to be combined into one. Consider the following identity:

$$\frac{1}{a\,b} \equiv \int_0^1 \mathrm{d}x\, \frac{1}{\left[ax + b(1-x)\right]^2}. \tag{5.4.2}$$

Exercise 5.4.1. *Show that this identity can be generalised to an arbitrary product of arbitrary powers of denominators:*

$$\prod_{i=1}^m a_i^{-n_i} = \frac{\Gamma\left(\sum_{i=1}^m n_i\right)}{\prod_{i=1}^m \Gamma(n_i)} \int_0^1 \prod_{i=1}^m \left[\mathrm{d}x_i\, x_i^{n_1 - 1}\right] \delta\left(1 - \sum_{i=1}^m x_i\right) \left[\sum_{i=1}^m a_i x_i\right]^{-\sum_{i=1}^m n_i}. \tag{5.4.3}$$

We then have (rearranging the denominator slightly)

$$I(D;1,1) = \int \mathrm{i}\mu^{4-D}\text{đ}^D k_E \int_0^1 \mathrm{d}x\, \left[(k_E - x p_E)^2 + x(1-x)p_E^2 + m_B^2\right]^{-2}. \tag{5.4.4}$$

Since the integrals converge, we may interchange the order and then also make a shift in the momentum variable to $k'_E = k_E - xp_E$ to obtain

$$I(D;1,1) \;=\; \int_0^1 \mathrm{d}x \int \mathrm{i}\mu^{4-D}\mathrm{đ}^D k'_E \,\left[k'^2_E + x(1-x)s_E + m_B^2\right]^{-2}, \qquad (5.4.5)$$

where we have also introduced $s_E = p_E^2$, the Euclidean version of the standard Mandelstam variable $s = p^2$. It is easy to see that the momentum integral here may be obtained by differentiating the previous single-propagator integral with respect to m_B^2, thus:

$$\begin{aligned}\int \mathrm{i}\mu^{4-D}\mathrm{đ}^D k'_E \,\left[k'^2_E + M^2\right]^{-2} \;&=\; -\frac{\mathrm{d}}{\mathrm{d}M^2}\int \mathrm{i}\mu^{4-D}\mathrm{đ}^D k'_E \,\left[k'^2_E + M^2\right]^{-1} \\ &=\; \frac{\mathrm{d}}{\mathrm{d}M^2}\left\{\frac{\mathrm{i}M^2}{(4\pi)^2}\left[\frac{1}{\varepsilon} + 1 - \gamma_\mathrm{E} + \ln\frac{4\pi\mu^2}{M^2} + O(\varepsilon)\right]\right\} \\ &=\; \frac{\mathrm{i}}{(4\pi)^2}\left[\frac{1}{\varepsilon} - \gamma_\mathrm{E} + \ln\frac{4\pi\mu^2}{M^2} + O(\varepsilon)\right]. \qquad (5.4.6)\end{aligned}$$

We then set $M^2 = x(1-x)s_E + m_B^2$ and insert this result into the expression for $I(D;1,1)$ to obtain

$$\begin{aligned}I(D;1,1) \;&=\; \frac{\mathrm{i}}{(4\pi)^2}\int_0^1 \mathrm{d}x \,\left[\frac{1}{\varepsilon} - \gamma_\mathrm{E} + \ln\frac{4\pi\mu^2}{x(1-x)s_E + m_B^2} + O(\varepsilon)\right] \\ &=\; \frac{\mathrm{i}}{(4\pi)^2}\int_0^1 \mathrm{d}x \,\left[\frac{1}{\varepsilon} + \ln 4\pi - \gamma_\mathrm{E} + \ln\frac{\mu^2}{m_B^2}\right. \\ &\qquad\qquad \left. - \ln\left(1 - x(1-x)\frac{s}{m_B^2}\right) + O(\varepsilon)\right], \qquad (5.4.7)\end{aligned}$$

where in the last line we have returned to Minkowski space for the external momentum p and variable s.

To perform the final integral in x, we must now specify whether s is greater or less than $4m_B^2$. In the latter case, since $4x(1-x) \leq 1$, the integral may be performed as-is to give

$$\begin{aligned}I(D;1,1) \;=\; \frac{\mathrm{i}}{(4\pi)^2}\Bigg[\, &\frac{1}{\varepsilon} + \ln 4\pi - \gamma_\mathrm{E} + \ln\frac{\mu^2}{m_B^2} + 2 \\ &- 2\left(\frac{4m_B^2}{s} - 1\right)^{1/2} \operatorname{arccot}\left(\frac{4m_B^2}{s} - 1\right)^{1/2} + O(\varepsilon)\,\Bigg]. \qquad (5.4.8)\end{aligned}$$

Exercise 5.4.2. *Calculate the final integral for* $I(D;1,1)$ *in the case* $s>4m_B^2$. Hint: *split the* x *integration into two regions given by* $x(1-x)s\lessgtr m_B^2$ *and use the fact that* $\ln(-1)=\mathrm{i}\pi$.

The physical origin of the difference between the two cases $s\lessgtr 4m_B^2$ lies in the fact that for $s<4m_B^2$ the intermediate two-particle state is constrained to be off-shell and is therefore always virtual while for $s>4m_B^2$ the intermediate pair may go on-shell and thus propagate indefinitely. The latter case will provide an imaginary or so-called *absorptive* contribution to the amplitude and indeed corresponds to an imaginary shift in the physical mass of an unstable particle. We shall discuss this more in depth later.

Before concluding this section let us make one more important observation. The expressions derived also remain well-defined and finite in the infrared region. This would not have been the case for a massless theory, as should be obvious from the expressions obtained. The reason can be traced to the fact that the probability, or amplitude, for emission of a real particle is inversely proportional to the its energy. Evidently though, a real, massive particle cannot be emitted with zero energy. In many cases, it is sufficient to include a (fictitious) non-zero mass for the purposes of performing calculations, with the certainty that once the *complete* expression for a physical quantity is obtained the IR behaviour will have been tamed. Unfortunately, in general, the procedure is not so simple: even a fictitious mass violates gauge symmetry and therefore such a trick cannot be easily applied to any of the standard theories. This is another problem we shall examine in more detail later. We might just mention though that dimensional regularisation can deal with the IR problem too: if we consider $D>4$ in the Feynman parameter integral, *i.e.* after the UV-divergent integrations, then again the divergences are transformed into (extra) simple ε poles.

5.5 One-loop renormalisation

Having performed the singular but regularised integrals and having thus identified the divergences, we must now set about defining the procedure by which they are absorbed into redefinitions of the bare, unphysical, parameters of the theory. As already suggested, this may or may not involve the introduction of new bare parameters to cancel divergent higher-order effective couplings generated by the quantum corrections.

5.5.1 Definition of the renormalisation constants

The idea then is to define a set of physical *renormalised* parameters, in terms of which all physical quantities (such as cross-sections, decay rates *etc.*) may be

expressed with no singular ε poles appearing. Experience shows that all such renormalisation effects are essentially multiplicative; we shall therefore begin by defining the so-called *renormalisation constants*:

$$\sqrt{Z_\varphi}\ \varphi_R \ = \ \varphi_B, \tag{5.5.1a}$$
$$Z_\varphi^{-1} Z_m\ m_R^2 \ = \ m_B^2, \tag{5.5.1b}$$
$$Z_\varphi^{-2} Z_\lambda\ \lambda_R \ = \ \lambda_B. \tag{5.5.1c}$$

The reason Z_φ appears on the left-hand side of the second a third lines is to explicitly take into account the renormalisation of the wave-function associated with the mass and coupling terms appearing in the Lagrangian. These may be rewritten as

$$Z_\varphi \ = \ 1 + \delta Z_\varphi, \tag{5.5.2a}$$
$$Z_m\, m_R^2 \ = \ m_R^2 + \delta m_R^2, \tag{5.5.2b}$$
$$Z_\lambda\, \lambda_R \ = \ \lambda_R + \delta\lambda_R. \tag{5.5.2c}$$

The next step is to define a renormalised Lagrangian via

$$\mathcal{L}_B \ = \ \mathcal{L} + \delta\,\mathcal{L}, \tag{5.5.3}$$

where $\mathcal{L}$ has the form of the original, classical, Lagrangian expressed though in terms of *renormalised* quantities (here and henceforth, when unnecessary, we omit the subscript "R" as superfluous):

$$\mathcal{L} \ := \ \tfrac{1}{2}\big[(\partial^\mu\varphi)(\partial_\mu\varphi) - m^2\varphi^2\big] - \tfrac{1}{4!}\,\lambda\varphi^4. \tag{5.5.4}$$

Naturally, $\mathcal{L}_B$ is the same, but expressed in terms of the bare quantities. Thus, $\delta\mathcal{L}$ is just the difference and contains what are commonly known as *counter terms*:

$$\delta\,\mathcal{L} \ = \ \tfrac{1}{2}\big[\delta Z_\varphi(\partial^\mu\varphi)(\partial_\mu\varphi) - \delta m^2\varphi^2\big] - \tfrac{1}{4!}\,\delta\lambda\varphi^4. \tag{5.5.5}$$

The trick will now be to treat *all* of $\delta\mathcal{L}$ as an interaction contribution (*i.e.*, quadratic terms included).

Exercise 5.5.1. *Verify that the definitions given above for the renormalised quantities do indeed lead to the form of the counter-term Lagrangian shown here.*

5.5.2 Counter-term Feynman rules

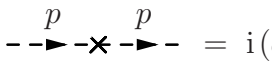
The Feynman rules to be derived will now arise from two different sources: V_R and $\delta\mathcal{L}$. They will have precisely the same form as before but in the former

case they contain only renormalised quantities and in the latter represent counter terms, which will be chosen so as to cancel the divergences. However, as already mentioned, the counter-term piece will be interpreted as contributing only to the interaction potential. It will thus be treated perturbatively, giving rise to a set of new vertices. In other words, we shall write

$$\mathcal{L}_I \;=\; -\tfrac{1}{4!}\,\lambda\varphi^4 + \tfrac{1}{2}\big[\delta Z_\varphi(\partial^\mu\varphi)(\partial_\mu\varphi) - \delta m^2\varphi^2\big] - \tfrac{1}{4!}\,\delta\lambda\varphi^4. \tag{5.5.6}$$

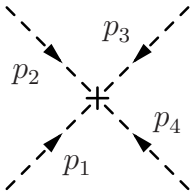
From the procedure already followed to derive the Feynman rules, it is easy to see that the extra set will be just:

5. A two-point vertex, represented as



$$= \mathrm{i}(\delta Z_\varphi\, p^2 - \delta m^2). \tag{5.5.7}$$

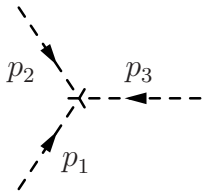
6. In φ^4 theory, a four-point vertex, represented as



$$= -\mathrm{i}\,\delta\lambda, \tag{5.5.8}$$

with the usual constraint that $\sum_{i=1}^4 p_i = 0$.

6′. In φ^3 theory, a three-point vertex, represented as



$$= -\mathrm{i}\,\delta\lambda, \tag{5.5.9}$$

with the same overall momentum constraint.

Exercise 5.5.2. *Following the standard procedure, derive the Feynman rules given above arising for the new counter-terms.*

5.5.3 Calculation of the renormalisation constants

If we were to switch off the interaction (*i.e.* by setting $\lambda = 0$), then evidently all the counter terms should vanish. It is therefore reasonable to write perturbative

expansions for all three of the renormalisation constants:

$$\delta Z_\varphi = \sum_{n=1}^{\infty} c_\varphi^{(n)} \lambda^n, \tag{5.5.10a}$$

$$\frac{\delta m^2}{m^2} = \sum_{n=1}^{\infty} c_m^{(n)} \lambda^n, \tag{5.5.10b}$$

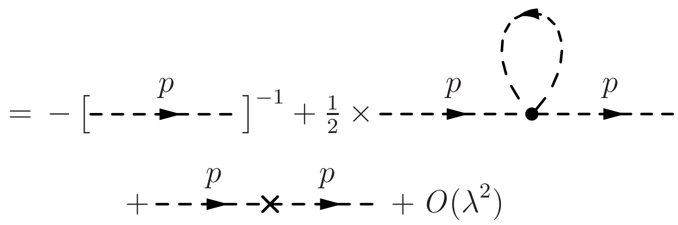
$$\frac{\delta\lambda}{\lambda} = \sum_{n=1}^{\infty} c_\lambda^{(n)} \lambda^n. \tag{5.5.10c}$$

Note that the series for $\delta\lambda$ must evidently begin at $O(\lambda^2)$ and this has been taken into account by the explicit factor λ^{-1} on the left-hand side. Likewise, we shall find, as already mentioned, that renormalisation is multiplicative and thus $\delta m^2 \propto m^2$; again, this has been rendered explicitly. Our task now is to calculate the perturbative coefficients $c_i^{(n)}$. In reality, we have already calculated the necessary Feynman diagrams; what has now to be found is a definition of reference physical quantities that will determine the so-called *subtraction procedure* and thus determine the parameters.

The 1PI two-point Green function

Consider first the 1PI two-point Green function:

$$\mathrm{i}\,\Gamma^{(2)}(p,-p) = \mathrm{i}\left[p^2 - m^2\right] - \tfrac{1}{2}\mathrm{i}\lambda\,\mathrm{i}\Delta_{\mathrm{F}}(0) + \mathrm{i}\left[\delta Z_\varphi\, p^2 - \delta m^2\right] + O(\lambda^2)$$



(5.5.11)

Regrouping the various contributions according to type (mass, momentum *etc.*) and using the expansion in perturbative coefficients, we have

$$\begin{aligned}\Gamma^{(2)}(p,-p) &= p^2\left[1 + c_\varphi^{(1)}\lambda\right] - \left[m^2 + \delta m^2 + \tfrac{1}{2}\lambda\,\mathrm{i}\Delta_{\mathrm{F}}(0)\right] + O(\lambda^2) \\ &= p^2\left[1 + c_\varphi^{(1)}\lambda\right] - m^2\left[1 + \left(c_m^{(1)} + \tfrac{1}{2}c_\Delta\right)\lambda\right] + O(\lambda^2),\end{aligned} \tag{5.5.12}$$

where we have introduced c_Δ defined by

$$\mathrm{i}\lambda\Delta_{\mathrm{F}}(0) = \lambda c_\Delta. \tag{5.5.13}$$

Using the result of the one-loop correction to the propagator already obtained in (5.3.10), we have

$$c_\Delta \;=\; -\frac{1}{(4\pi)^2}\left[\frac{1}{\varepsilon}+1+\ln 4\pi-\gamma_{\rm E}-\ln\frac{m^2}{\mu^2}+O(\varepsilon)\right]. \tag{5.5.14}$$

The precise values of the two constants $c_\varphi^{(1)}$ and $c_m^{(1)}$ will be fixed by the chosen boundary conditions on the two-point function itself, which is effectively a physical observable. This choice is what defines the so-called *renormalisation scheme*. The main point though is that the pole pieces must evidently cancel between the two terms $c_m^{(1)}$ and $\frac{1}{2}c_\Delta$ while $c_\varphi^{(1)}$ must immediately be finite. In other words, we require that

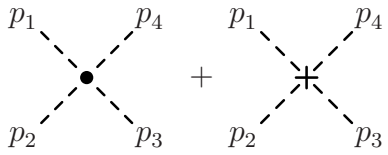
$$c_m^{(1)} \;=\; \frac{1}{2}\,\frac{1}{(4\pi)^2}\,\frac{1}{\varepsilon}+\text{finite constant}. \tag{5.5.15}$$

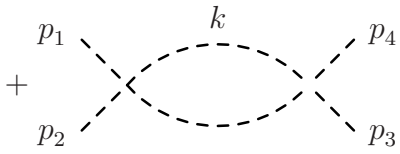
The finite constant here is entirely arbitrary and, together with the other remaining constant terms, will be fixed precisely by the chosen renormalisation or so-called *subtraction* scheme. One possibility is merely to cancel the poles and leave all constant pieces as calculated, a second possibility is to choose specific (finite) values for the relevant subset of n-point Green functions for some specific set of values of the external momenta. Different choices will lead to different values for the constant terms. However, we shall find that the arbitrariness disappears when calculating any physical quantity. The various possibilities will be examined in some detail shortly.

The 1PI four-point Green function

Let us now examine the case of the 1PI four-point Green function (all momenta are taken as ingoing):

$\mathrm{i}\Gamma^{(4)}(p_1,p_2,p_3,p_4) \;=$





$$+\,\text{perms}(p2,p3,p4)+O(\lambda^3). \tag{5.5.16}$$

The diagrams in the first row are $O(\lambda)$ and are the leading-order (or zeroth-order*) contributions. The remaining terms are order λ^2 or higher. Thus, the second row (containing three diagrams) constitutes the first-order correction. Note that since the diagrams are amputated, we do not have to consider bubbles attached to the external legs and the loop diagram shown (together with its permutations) is the only first-order contribution.

We have already, in effect, performed the integrals arising in these loop diagrams. Of course, only one need actually be calculated, the others being obtained by suitable exchange of the external momenta, or more easily, the Mandelstam variables. We thus obtain

$$\Gamma^{(4)}(p_1,p_2,p_3,p_4) \;=\; -\lambda - c_\lambda^{(1)}\lambda^2 + \frac{\lambda^2}{32\pi^2}\left[\frac{3}{\varepsilon} + 3(\ln 4\pi - \gamma_{\mathrm{E}}) - 3\ln\frac{m^2}{\mu^2} + A(s,\mu^2) + A(t,\mu^2) + A(u,\mu^2)\right] + O(\lambda^3), \quad (5.5.17)$$

where the amplitudes $A(P^2,\mu^2)$ for $P^2 = s$, u and t represent the three different loop diagrams with "incoming" momenta giving s, u and t respectively.† They are calculated to be

$$A(P^2,\mu^2) \;=\; 2(1-\beta\,\mathrm{arccot}\,\beta), \quad \text{with} \quad \beta \;=\; \left(\frac{4m^2}{P^2}-1\right)^{1/2}, \qquad (5.5.18)$$

where, as usual,

$$s \;:=\; (p_1+p_2)^2 \;\equiv\; (p_3+p_4)^2, \qquad (5.5.19a)$$

$$t \;:=\; (p_1+p_3)^2 \;\equiv\; (p_2+p_4)^2, \qquad (5.5.19b)$$

$$u \;:=\; (p_1+p_4)^2 \;\equiv\; (p_2+p_3)^2 \qquad (5.5.19c)$$

Note the signs: the momenta are all considered as *incoming*. Let us remark that once again we see the ubiquitous term $\ln 4\pi - \gamma_{\mathrm{E}}$. Note also the $\ln\mu^2$ accompanying the ε pole—they are indeed strictly related. Finally, note the problem already encountered of needing to decide whether $4m^2 \lessgtr P^2$.

As before, we shall determine the counter terms in order to cancel (at least) the ε poles and thus render $\Gamma^{(4)}$ finite. In other words, we shall require

$$c_\lambda^{(1)} \;=\; \frac{1}{32\pi^2}\,\frac{3}{\varepsilon} + \text{finite constant} \qquad (5.5.20)$$

* We shall always use the expression *zeroth*-order to indicate the *leading* effect and thus a *first*-order effect will be the lowest-order *correction*

† The symmetry under interchange of external legs is generally known as *crossing* symmetry.

And, as before, the finite constant, being arbitrary, will be fixed by the boundary conditions chosen for the four-point function.

We have already seen that $\Gamma_B^{(2)}$ and $\Gamma_B^{(4)}$ are the only two divergent bare Green functions at one-loop order in a φ^4 theory. On the other hand, we have also seen how the counter terms, introduced via redefinition of the Lagrangian, render finite the corresponding renormalised or physical functions. It is thus interesting to relate them.

First of all, note that if we wish to add source terms to both sides of Eq. (5.5.3), then we are also forced to renormalise the source $J_B(x)$: since we require

$$J_B(x)\,\varphi_B(x) \;=\; J(x)\,\varphi(x), \tag{5.5.21}$$

we therefore take

$$J_B(x) \;=\; Z_\varphi^{-1/2}\,J(x). \tag{5.5.22}$$

The generating functionals $Z_B[J_B]$ and $Z[J]$ of the bare and renormalised theories respectively can thus be related. Naturally, the bare functional is that already defined, written now in terms of $\mathcal{L}_B$, φ_B and J_B; the renormalised functional is that written using $\mathcal{L}+\delta\mathcal{L}$ as the Lagrangian together with φ and J. Then, since we have

$$\mathcal{L}_B + J_B(x)\,\varphi_B(x) \;=\; \mathcal{L} + \delta\mathcal{L} + J(x)\,\varphi(x), \tag{5.5.23}$$

we see that

$$Z_B[J_B] \;=\; Z[J]. \tag{5.5.24}$$

From this it immediately follows that

$$\mathcal{G}^{(n)}(p_1,\ldots,p_n) \;=\; \left(Z_\varphi^{-1/2}\right)^n\,\mathcal{G}_B^{(n)}(p_1,\ldots,p_n). \tag{5.5.25}$$

Exercise 5.5.3. *Verify the previous relation between bare and renormalised Green functions.*

Following the same procedure, we may relate the bare and renormalised one-particle irreducible generating functionals. Given Eq. (5.5.24), one evidently has

$$Z_c[J] \;=\; Z_{cB}[J_B], \tag{5.5.26}$$

which leads to

$$\varphi_c(x) \;:=\; \frac{\delta Z_c[J]}{\delta J(x)} \;=\; Z_\varphi^{-1/2}\,\frac{\delta Z_{cB}[J_B]}{\delta J_B(x)} \;=\; Z_\varphi^{-1/2}\,\varphi_{c\,B}(x) \tag{5.5.27}$$

and therefore

$$\Gamma[\varphi_c] \;=\; \Gamma_B[\varphi_{c\,B}]. \tag{5.5.28}$$

Finally, from the expansion of $\Gamma[\varphi_c]$ in terms of the one-particle irreducible Green functions, we have

$$\Gamma^{(n)}(p_1,\ldots,p_n) \;=\; \left(Z_\varphi^{1/2}\right)^n \,\Gamma_B^{(n)}(p_1,\ldots,p_n). \tag{5.5.29}$$

For the two-point function, setting $n=2$, the right-hand side is

$$\begin{aligned} &Z_\varphi\,\Gamma_B^{(2)}(p,-p) \\ &\quad=\; Z_\varphi\left\{p^2-m_B^2+\frac{\lambda_B m_B^2}{32\pi^2}\left[\frac{1}{\varepsilon}+1+\ln 4\pi-\gamma_{\rm E}+\ln\frac{\mu^2}{m_B^2}+O(\varepsilon)\right]+O(\lambda_B^2)\right\}, \end{aligned} \tag{5.5.30}$$

which is the $\Gamma^{(2)}$ calculated earlier. In other words, as promised, the divergences have all been absorbed into redefinitions or renormalisations of the bare parameters of the theory. This constitutes a verification, to one-loop order, that the φ^4 theory is renormalisable. It is indeed possible to prove renormalisability to all orders (see, for example, Itzykson and Zuber, 1980; Abers and Lee, 1973). However, the proof is not simple and we shall not go further into this aspect.

Exercise 5.5.4. *Compare Eq. (5.5.30) with the previous calculation and verify that the divergence has indeed been cancelled.*

5.5.4 Renormalisation schemes

As already mentioned, we need to specify the precise manner in which the renormalisation constants are to be determined order-by-order in perturbation theory. The particular procedure is known as the *renormalisation scheme*. A particularly obvious scheme consists of simply removing or subtracting the ε pole and nothing else. This is what is commonly called the minimal subtraction (MS) scheme. We have already seen, however, that there are also two omnipresent constant terms. These can be removed order-by-order by suitably redefining the continuation of the measure to D dimensions. For example:

$$đ^4k \;\to\; \frac{\mu^{4-D}}{\Gamma(1+\varepsilon)\,(4\pi)^\varepsilon}\,đ^Dk\,. \tag{5.5.31}$$

Often one quite simply defines the subtraction procedure to include the common constant pieces. That is, we subtract a term proportional to $\left(\frac{1}{\varepsilon}+\ln 4\pi-\gamma_{\rm E}\right)$. Evidently, in any case, this has to be refined order-by-order. Such a scheme is usually denoted modified minimal subtraction ($\overline{\rm MS}$).

To be more precise, we have seen that the application of dimensional regularisation replaces the divergences arising in the loop integrals by inverse powers of ε. Whatever scheme we choose, these must be removed by subtraction of the counter terms, which may also be expanded in inverse powers of ε:

$$\delta Z_\varphi \;=\; a_\varphi^{(0)}(\lambda, \mu/m, \varepsilon) + \sum_{n=1}^{\infty} a_\varphi^{(n)}(\lambda, \mu/m)\, \varepsilon^{-n}, \tag{5.5.32a}$$

$$\frac{\delta m^2}{m^2} \;=\; a_m^{(0)}(\lambda, \mu/m, \varepsilon) + \sum_{n=1}^{\infty} a_m^{(n)}(\lambda, \mu/m)\, \varepsilon^{-n}, \tag{5.5.32b}$$

$$\frac{\delta \lambda}{\lambda} \;=\; a_\lambda^{(0)}(\lambda, \mu/m, \varepsilon) + \sum_{n=1}^{\infty} a_\lambda^{(n)}(\lambda, \mu/m)\, \varepsilon^{-n}. \tag{5.5.32c}$$

The zeroth-order term in each series contains all the non-singular dependence on ε while the higher-order terms render explicit possible multiple poles. Note, of course, that each of the coefficients $a_i^{(n)}$ also possesses an expansion in powers of λ, as defined earlier. It should also be evident that order-by-order in perturbation theory the inverse powers of ε are limited: at one-loop order a massive φ^4 theory only has simple poles, at two-loop order double poles appear and so on.

A further observation on the coefficients is that they have been deliberately defined so as to be dimensionless. This means that they can in turn only be functions of dimensionless quantities. Thus, in particular, for $n \geq 1$ they can only depend on the coupling constant λ and the dimensionless ratio μ/m, as explicitly indicated.

If we limit ourselves to the MS scheme then, by definition, we have

$$a_\varphi^{(0)} \;=\; a_m^{(0)} \;=\; a_\lambda^{(0)} \;=\; 0. \tag{5.5.33}$$

On the other hand, in the $\overline{\text{MS}}$ scheme and according to the method of implementation they may be non-zero, but simply proportional to the combination $\ln 4\pi - \gamma_\text{E}$. Moreover, the coefficients of the *singular* terms turn out to be independent of the dimensionless ratio μ/m and are therefore mass independent:

$$a_\varphi^{(n)}(\lambda, \mu/m) \;=\; a_\varphi^{(n)}(\lambda) \quad \textit{etc.} \quad (\text{for } n \;\geq\; 1). \tag{5.5.34}$$

We have already seen this at first perturbative order, see Eq. (5.5.17):

$$\left.\frac{\delta \lambda}{\lambda}\right|_{\overline{\text{MS}}} \;=\; \frac{3\lambda}{32\pi^2}\left[\frac{1}{\varepsilon} + \ln 4\pi - \gamma_\text{E}\right] + O(\lambda^2), \tag{5.5.35a}$$

implying

$$a_\lambda^{(1)}\Big|_{\overline{\rm MS}} = \frac{3\lambda}{32\pi^2} + O(\lambda^2). \tag{5.5.35b}$$

Similarly, for the other renormalisation constants, from Eq. (5.5.14), we have

$$\frac{\delta\varphi}{\varphi}\Big|_{\overline{\rm MS}} = O(\lambda^2) \quad \text{and} \quad \frac{\delta m^2}{m^2}\Big|_{\overline{\rm MS}} = \frac{\lambda}{32\pi^2}\left[\frac{1}{\varepsilon} + \ln 4\pi - \gamma_{\rm E}\right] + O(\lambda^2), \tag{5.5.36a}$$

leading to

$$a_\varphi^{(1)}\Big|_{\overline{\rm MS}} = O(\lambda^2) \quad \text{and} \quad a_m^{(1)}\Big|_{\overline{\rm MS}} = \frac{\lambda}{32\pi^2} + O(\lambda^2). \tag{5.5.36b}$$

The evident absence of mass-parameter dependence here simplifies the dimensional or so-called *renormalisation-group* analysis of the scale breaking introduced by the renormalisation procedure and embodied in the appearance of the new scale μ. We shall discuss this important concept in more detail later.

Once the counter terms have been determined and the ε poles have been subtracted, since the remaining expressions for the Green functions will then be finite, we may take the limit $\varepsilon \to 0$. Putting everything together, in the $\overline{\rm MS}$ scheme we obtain

$$\Gamma^{(2)}_{\overline{\rm MS}}(p,-p) = p^2\left\{1 + O(\lambda^2)\right\} - m^2\left\{1 - \frac{\lambda}{32\pi^2}\left[1 - \ln\frac{m^2}{\mu^2}\right] + O(\lambda^2)\right\}, \tag{5.5.37a}$$

$$\Gamma^{(4)}_{\overline{\rm MS}}(p_1,p_2,p_3,p_4) = -\lambda\left\{1 - \frac{\lambda}{32\pi^2}\left[-3\ln\frac{m^2}{\mu^2} + A(s,m^2) + A(t,m^2) + A(u,m^2)\right] + O(\lambda^2)\right\}, \tag{5.5.37b}$$

where A is defined as in Eq. (5.5.18).

Among other things, it should be noticed that to first order in λ (in this particular theory) there is no renormalisation of the wave-function. Furthermore, and more importantly, we see how the energy or scale dependence has crept in: the effective renormalised coupling constant will *no longer be constant* owing to the momentum dependence acquired by the 1PI four-point function, with which it will be defined operationally.

In the MS and $\overline{\rm MS}$ schemes the subtraction terms are defined independently of the particular n-point Green function considered. However, alternative renormalisation schemes exist (and are often used) in which the counter terms are determined by imposing boundary conditions on a sufficient number of Green functions, naturally taken to be $\Gamma^{(2)}$ and $\Gamma^{(4)}$. Since there are three renormalisation con-

stants, we evidently need three conditions. Given that the two-point function actually contains two functionally independent pieces (p^2 and m^2), we may define two conditions on $\Gamma^{(2)}$ (*e.g.* one on $\Gamma^{(2)}$ and one on its derivative with respect to p^2) and a third on $\Gamma^{(4)}$.

One such scheme, known as the momentum subtraction (MOM) scheme, applies the conditions for chosen values of the external momenta; *i.e.* by setting the values of the Green functions at the (non-singular) point $p^2 = -\mu_R^2$, where μ_R^2 is as arbitrary external momentum scale. The fact that this is an unphysical point is irrelevant for the present purpose. It thus takes advantage of the momentum scale already introduced, but sets the conditions at a Euclidean point, *i.e.* well away from the natural propagator pole. The conditions on the two-point function are then

$$\Gamma^{(2)}(p,-p)\Big|_{p^2=-\mu_R^2} \;=\; -m^2 - \mu_R^2 \tag{5.5.38a}$$

and

$$\frac{\partial}{\partial p^2}\,\Gamma^{(2)}(p,-p)\Big|_{p^2=-\mu_R^2} \;=\; 1. \tag{5.5.38b}$$

This implies

$$\Gamma^{(2)}(p,-p) \;=\; p^2 - m^2 + O\big((p^2+\mu_R^2)^2\big)\,. \tag{5.5.39}$$

In terms of the renormalisation constant δZ_φ and δZ_m, this becomes

$$\delta Z_\varphi \;=\; O(\lambda^2) \tag{5.5.40a}$$

and

$$\delta Z_m \;=\; \frac{\lambda}{32\pi^2}\left[\frac{1}{\varepsilon} + (\ln 4\pi - \gamma_{\mathrm{E}}) - \ln\frac{m^2}{\mu^2}\right] O(\lambda^2). \tag{5.5.40b}$$

Similarly, $\delta\lambda$ is fixed here by placing a boundary condition on $\Gamma^{(4)}$. If we require a symmetric condition, *i.e.* that does not privilege any of the external legs, then we should first set

$$p_i^2 \;=\; -\mu_R^2 \quad (i,j \;=\; 1,\ldots,4), \tag{5.5.41}$$

just as for the external legs in $\Gamma^{(2)}$. Moreover, since this implies $s+t+u=-4\mu_R^2$, a symmetric condition is

$$s \;=\; t \;=\; u \;=\; -\tfrac{4}{3}\,\mu_R^2. \tag{5.5.42}$$

Again, this is an unphysical point, but this is not important here. The condition on $\Gamma^{(4)}$ can then be written as

$$\Gamma^{(4)}(p_1,p_2,p_3,p_4) \;=\; -\lambda \quad \text{for} \quad p_i{\cdot}p_j \;=\; \mu_R^2\frac{(1-4\delta_{ij})}{3} \quad (i,j \;=\; 1,\ldots,4). \tag{5.5.43}$$

We therefore have

$$\delta Z_\lambda \;=\; \frac{\lambda}{32\pi^2}\left[\frac{3}{\varepsilon}+3(\ln 4\pi-\gamma_{\rm E})-3\ln\frac{m^2}{\mu^2}+3A(-\tfrac{4}{3}\mu_R^2,m^2)\right]+O(\lambda^2). \quad (5.5.44)$$

Finally then, in the MOM scheme, sending $\varepsilon\to 0$, for generic values of the external momenta, leads to

$$\Gamma^{(2)}(p,-p) \;=\; p^2-m^2+O(\lambda^2) \quad (5.5.45a)$$

and

$$\Gamma^{(4)}(p_1,p_2,p_3,p_4) \;=\; -\lambda\left\{1-\frac{\lambda}{32\pi^2}\Big[A(s,m^2)+A(t,m^2)+A(u,m^2) - 3A(-\tfrac{4}{3}\mu_R^2,m^2)\Big]+O(\lambda^2)\right\}. \quad (5.5.45b)$$

We immediately notice that in this scheme the dependence on μ has disappeared from the n-point Green functions. However, upon examining Eqs. (5.5.40b) and (5.5.44), we see that the counter terms are now manifestly mass and therefore μ dependent.

The last scheme we shall examine is the so-called "on-shell" subtraction scheme. It is in this scheme that the parameters m and λ actually assume their canonical physical significance of particle mass and coupling constant. It is important to realise that this was not true for the arbitrary definitions of the subtractions in the previous schemes. Here instead they will acquire the values that might be directly measured experimentally.

For the mass, we simply take the position of pole in the two-point Green function $\mathcal{G}^{(2)}$, or equivalently the zero in $\Gamma^{(2)}$. The conditions to be applied are then

$$\Gamma^{(2)}(p,-p)\Big|_{p^2=m^2} \;=\; 0 \quad (5.5.46a)$$

and

$$\frac{\partial}{\partial p^2}\,\Gamma^{(2)}(p,-p)\Big|_{p^2=m^2} \;=\; 1. \quad (5.5.46b)$$

So far this is just equivalent to MOM with subtraction point $\mu_R^2=-m^2$. In other words, we may calculate the Green functions by simply taking those defined in MOM and making the previous substitution.

Again, examining Eqs. (5.5.40b) and (5.5.44), bearing in mind that $\mu_R^2=-m^2$, we see that the counter terms are mass dependent in this scheme too. However, all mass dependence is encoded in m, the now physical mass. Indeed, since, by definition, all renormalisation schemes must render the Green function finite, the effective difference between any two schemes must be finite. Moreover, one can

derive a table of correspondence to permit the translation from any given scheme to another—the coefficients being always finite.

Finally, note that there is no scheme in which the momentum dependence in $\mathcal{G}^{(4)}$ may be removed and thus too λ_R will *always* have momentum dependence, which must then be characterised by some extra scale, μ_R say, as λ_R itself is adimensional. This observation leads to an important basis for analysis of the general scaling behaviour, known as the *renormalisation group*.

5.6 Bibliography

Abers, E.S. and Lee, B.W. (1973), *Phys. Rep.* **9**, 1.

Bailin, D. and Love, A. (1993), *Introduction to Gauge Field Theory* (IOP Pub.), revised edition.

Itzykson, C. and Zuber, J.-B. (1980), *Quantum Field Theory* (McGraw–Hill).

Pauli, W. and Villars, F. (1949), *Rev. Mod. Phys.* **21**, 434.

Weinberg, S. (1960), *Phys. Rev.* **118**, 838.

Chapter 6

Spinor and Gauge Fields

If we wish to describe the interactions of the matter that makes up the visible universe, then we are forced to include spin-half fields such as electrons, muons and quarks *etc*. Moreover, their interactions will be described by exchange gauge fields such as the photon. Let us then open this chapter by highlighting the fundamental differences with respect to the simple scalar field theory studied in the previous chapter.

The obvious differences regard the spin and the mass of the fields involved. However, as we shall see, the nature of the charge of the gauge fields, Abelian or non-Abelian, is also of great relevance. In any case, the introduction of a spin-half field requires a construction capable of incorporating the Fermi statistics it must obey. As to the inclusion of gauge fields, we shall find that the gauge symmetry, which has the effect of reducing the number of physical degrees of freedom from the three normally associated with a massive vector field down to two, requires careful treatment and some additional techniques.

6.1 Spinor-field quantisation

6.1.1 Spinor path integrals

It was already shown in Sec. 2.4.2 that the correct formulation of a spinor field theory requires the introduction of anticommuting or Grassmann variables. This implies that we now need to deal with Feynman path integrals over such objects. Using the results given in the appendix for a discrete set of Grassmann variables, we can immediately write down the corresponding result for a continuum of com-

ponents, in other words, for a Grassmann field $\psi(x)$ and A anti-Hermitian:

$$\int \mathcal{D}\psi \, \exp\Big\{ -\tfrac{1}{2} \int \mathrm{d}^4x \, \mathrm{d}^4y \, \psi(x) \, A(x,y) \, \psi(y) \Big\} \;=\; \Big[\det A \Big]^{1/2} \;=\; \exp\Big\{ \tfrac{1}{2} \operatorname{Tr} \ln A \Big\} , \tag{6.1.1}$$

where for the moment we consider a field $\psi(x)$ over a simple one-dimensional space. Note that it is not yet a spinor. Once again, since we shall also need to include source terms, the extension to include a linear term in the exponential is useful:

$$\begin{aligned} \int \mathcal{D}\psi \, \exp\Big\{ -\tfrac{1}{2} \int \mathrm{d}^4x \, \mathrm{d}^4y \, \psi(x) \, A(x,y) \, \psi(y) + \int \mathrm{d}^4x \, \sigma(x) \, \psi(x) \Big\} \\ = \; \exp\Big\{ \tfrac{1}{2} \operatorname{Tr} \ln A \Big\} \exp\Big\{ -\tfrac{1}{2} \int \mathrm{d}^4x \, \mathrm{d}^4y \, \sigma(x) \, A^{-1}(x,y) \, \sigma(y) \Big\} . \end{aligned} \tag{6.1.2}$$

Extending this to a complex field, we have

$$\int \mathcal{D}\psi \, \mathcal{D}\psi^* \, \exp\Big\{ - \int \mathrm{d}^4x \, \mathrm{d}^4y \, \psi^*(x) \, A(x,y) \, \psi(y) \Big\} \;=\; \det A \;=\; \exp\Big\{ \operatorname{Tr} \ln A \Big\} . \tag{6.1.3}$$

And the inclusion of linear source terms gives

$$\begin{aligned} \int \mathcal{D}\psi \, \mathcal{D}\psi^* \, \exp\Big\{ - \int \mathrm{d}^4x \, \mathrm{d}^4y \, \psi^*(x) \, A(x,y) \, \psi(y) \\ + \int \mathrm{d}^4x \, \Big[\sigma^*(x) \, \psi(x) - \psi^*(x) \, \sigma(x) \Big] \Big\} \\ = \; \exp\Big\{ \operatorname{Tr} \ln A \Big\} \exp\Big\{ - \int \mathrm{d}^4x \, \mathrm{d}^4y \, \sigma^*(x) \, A^{-1}(x,y) \, \sigma(y) \Big\} . \end{aligned} \tag{6.1.4}$$

Note the necessary sign difference between the two linear terms.

However, our general formulation has always led to path integrals in which the exponent is $\mathrm{i} \int \mathrm{d}^4x \, \mathcal{L}$ in Minkowski space. Therefore, we shall be interested in integrals of the form

$$\begin{aligned} \int \mathcal{D}\psi \, \mathcal{D}\psi^* \, \exp\Big\{ \mathrm{i} \int \mathrm{d}^4x \, \mathrm{d}^4y \, \psi^*(x) \, B(x,y) \, \psi(y) \\ + \mathrm{i} \int \mathrm{d}^4x \, \Big[\sigma^*(x) \, \psi(x) + \psi^*(x) \, \sigma(x) \Big] \Big\} \\ = \; \exp\Big\{ \operatorname{Tr} \ln(\mathrm{i} B) \Big\} \exp\Big\{ - \mathrm{i} \int \mathrm{d}^4x \, \mathrm{d}^4y \, \sigma^*(x) \, B^{-1}(x,y) \, \sigma(y) \Big\} , \end{aligned} \tag{6.1.5}$$

where now the operator B is *Hermitian* to compensate the explicit factor "i" in front of the integrals and there is no longer a sign difference between the two source terms.

6.1.2 The spinor generating functional

In analogy with the scalar case, the natural choice for the generating functional to quantise a free spinor field is

$$Z_0[\sigma,\overline{\sigma}] \;\propto\; \int \mathcal{D}\psi\,\mathcal{D}\bar{\psi}\,\exp\!\left\{\mathrm{i}\int \mathrm{d}^4x\,\left[\mathcal{L}+\overline{\sigma}\psi+\bar{\psi}\sigma\right]\right\}, \tag{6.1.6}$$

where now ψ and σ are intended to be a spinor field and source respectively; as usual, $\bar{\psi}:=\psi^\dagger\gamma^0$ and $\overline{\sigma}:=\sigma^\dagger\gamma^0$. The normalisation of $Z_0[\sigma,\overline{\sigma}]$ is chosen, as always, to be such that it is unity for $\sigma=0=\overline{\sigma}$. We cannot use the $\mathrm{i}\varepsilon$ trick here owing to the anticommuting nature of the objects involved and we shall simply resort to inserting it by hand into the final propagators.

The spinor Green functions are defined by

$$\mathcal{G}^{(2n)}(x_1,\ldots,x_n;y_1,\ldots,y_n) \;=\; \langle 0|\mathrm{T}\left[\psi(x_1)\ldots\psi(x_n)\bar{\psi}(y_1)\ldots\bar{\psi}(y_n)\right]|0\rangle, \tag{6.1.7}$$

where the time ordering must now also introduce a minus sign for every pair-wise commutation of fermion fields. Conservation of fermion number is made explicit by inserting the same number of fields $\psi(x_i)$ and $\bar{\psi}(y_i)$. These are indeed the only non-vanishing products. As usual, the Green functions are generated by functionally differentiating $Z_0[\sigma,\overline{\sigma}]$ a suitable number of times:

$$\mathrm{i}^{\,2n}\,\mathcal{G}^{(2n)}(x_1,\ldots,x_n;y_1,\ldots,y_n) \;=\; \frac{\delta^{2n}Z_0[\sigma,\overline{\sigma}]}{\delta\overline{\sigma}(x_1)\ldots\delta\overline{\sigma}(x_n)\,\delta\sigma(y_1)\ldots\delta\sigma(y_n)}. \tag{6.1.8}$$

First of all, note that it is no longer necessary to set $\sigma=0=\overline{\sigma}$ explicitly since at most only linear terms are present in $Z_0[\sigma,\overline{\sigma}]$. Moreover, owing to the anticommutation of any individual pair of derivatives, $\mathcal{G}^{(2n)}(x_1,\ldots,x_n;y_1,\ldots,y_n)$ is antisymmetric under pair-wise interchange of the coordinates $\{x_i,y_j\}$, as required by Fermi–Dirac statistics.

6.1.3 The free-field Dirac propagator

Since the free-field generating functional is quadratic, we can, as usual, evaluate it exactly. The relevant Lagrangian has already been encountered, see Eq. (2.4.31). Henceforth, we shall suppress the explicit unit Dirac matrix $\mathbb{1}$. We thus substitute

the following operator into (6.1.5)

$$\mathrm{B}(x,y) \;=\; \delta^4(x-y)\left[\mathrm{i}\not{\partial}_x - m\right]. \tag{6.1.9}$$

Using this and our new results on Gaussian functional integrals, we obtain

$$Z_0[\sigma,\overline{\sigma}] \;=\; \exp\!\left\{-\mathrm{i}\int \mathrm{d}^4x\,\mathrm{d}^4y\,\overline{\sigma}(y)\;\mathrm{B}^{-1}(x,y)\,\sigma(x)\right\}. \tag{6.1.10}$$

We now rewrite the inverse of $\mathrm{B}(x,y)$ in the following suggestive form:

$$\mathrm{B}^{-1}(x,y) \;=\; S_{\mathrm{F}}(x-y), \tag{6.1.11}$$

the Feynman propagator for a fermion field. From the definition of B, we have

$$\mathrm{B}(x,y) \;=\; \int \text{đ}^4p\;\mathrm{e}^{-\mathrm{i}p\cdot(x-y)}\left[\not{p} - m + \mathrm{i}\varepsilon\right], \tag{6.1.12}$$

from which it immediately follows that

$$\mathrm{B}^{-1}(x,y) \;=\; S_{\mathrm{F}}(x-y) \;=\; \int \text{đ}^4p\;\mathrm{e}^{-\mathrm{i}p\cdot(x-y)}\left[\not{p} - m + \mathrm{i}\varepsilon\right]^{-1}. \tag{6.1.13}$$

We can now easily move over to the momentum-space representation and write

$$S_{\mathrm{F}}(p) \;=\; \frac{1}{\not{p} - m + \mathrm{i}\varepsilon} \;=\; \frac{\not{p} + m}{p^2 - m^2 + \mathrm{i}\varepsilon'}, \tag{6.1.14}$$

where, since $\varepsilon' = 2m\varepsilon$ and m is constant, the prime is irrelevant.

6.1.4 Renormalisability

We can now address the question of renormalisability. Consider a theory involving both boson and fermion fields with vertices attached to N_B boson and N_F fermion lines. The form of the fermion propagator just derived, taken together with that for bosons derived earlier leads to the following superficial degree of divergence for an L-loop diagram with I_B (I_F) internal boson (fermion) propagators:

$$D \;=\; 4L - 2I_B - I_F \tag{6.1.15}$$

If a diagram contains $V(N_B,N_F)$ vertices with N_B (N_F) boson (fermion) legs and E_B (E_F) external boson (fermion) legs, then

$$E_B + 2I_B = \sum_{N_B,\,N_F} N_B\, V(N_B,N_F) \tag{6.1.16a}$$

and

$$E_F + 2I_F = \sum_{N_B,\,N_F} N_F\, V(N_B,N_F). \tag{6.1.16b}$$

The total numbers of internal and external lines and vertices are

$$I = I_B + I_F, \tag{6.1.17a}$$

$$E = E_B + E_F, \tag{6.1.17b}$$

$$V = \sum_{N_B,\,N_F} V(N_B,N_F) \tag{6.1.17c}$$

and it is still true that

$$L = 1 - V + I. \tag{6.1.18}$$

Putting all this together we have

$$D = 4 - E_B - \tfrac{3}{2}E_F + \sum_{N_B,\,N_F} (N_B + \tfrac{3}{2}N_F - 4)\, V(N_B,N_F). \tag{6.1.19}$$

If the coefficient multiplying $V(N_B,N_F)$ in the sum above is positive, then the degree of divergence grows with growing perturbation order and the theory will not be renormalisable. We therefore require

$$N_B + \tfrac{3}{2}N_F \leq 4. \tag{6.1.20}$$

For $N_F=0$, this is just the condition derived in the previous chapter. We can exclude $N_F=1$ on the grounds of angular-momentum conservation and thus we need only consider $N_F=2$ with $N_B=0$ or 1. The case $N_B=0$ is uninteresting since it corresponds to a non-interacting purely fermionic theory. Note that in this way we have excluded the famous Fermi theory of nuclear β-decay. Indeed, its non-renormalisability is symptomatic of incompleteness; once the weak gauge bosons are included in the correct manner, there are no fermionic four-point functions and the theory is renormalisable. Note too that we have not considered vertices with derivative terms since these would have an even poorer degree of convergence and thus no chance of being renormalisable.

6.1.5 Yukawa theory

The only tenable theory is then that containing fermion–boson (three-point) vertices. This corresponds (for a spin-zero boson) to the theory developed by Yukawa to explain the strong nuclear interaction. Such a theory was intended to describe the interactions between nucleons as mediated by pions. There are just two possible couplings or interaction terms:

$$\mathcal{L}_{\mathrm{I}} \;=\; -g\,\varphi\,\bar{\psi}\,\Gamma\,\psi \quad \text{with} \quad \Gamma \;=\; \mathbb{1} \text{ or } \gamma_5. \tag{6.1.21}$$

In fact, the pion turns out to be a pseudoscalar particle and therefore the second is the correct form in this case. We shall write the full Lagrangian keeping the interaction term separate from the free part:

$$\mathcal{L} \;=\; \mathcal{L}_0 + \mathcal{L}_{\mathrm{I}}(\varphi,\psi,\bar{\psi}). \tag{6.1.22}$$

The generating functional will then naturally take the following form

$$Z[J,\sigma,\overline{\sigma}] \;\propto\; \int \mathcal{D}\varphi\,\mathcal{D}\psi\,\mathcal{D}\bar{\psi}\,\exp\!\left\{\mathrm{i}\int \mathrm{d}^4x\,\left[\mathcal{L} + J\varphi + \overline{\sigma}\psi + \bar{\psi}\sigma\right]\right\}, \tag{6.1.23}$$

with $\mathcal{L}$ depending on both φ and ψ. For the free-field case Z_0 factorises into separate scalar and fermion pieces and has therefore already been covered in the previous discussions; we have

$$\begin{aligned} Z_0[J,\sigma,\overline{\sigma}] \;=\; &\exp\!\left\{-\tfrac{1}{2}\mathrm{i}\int \mathrm{d}^4x\,\mathrm{d}^4y\,J(x)\,\Delta_{\mathrm{F}}(x-y)\,J(y)\right\} \\ &\times \exp\!\left\{-\mathrm{i}\int \mathrm{d}^4x\,\mathrm{d}^4y\,\overline{\sigma}(x)\,S_{\mathrm{F}}(x-y)\,\sigma(y)\right\}. \end{aligned} \tag{6.1.24}$$

In the full interacting theory the path integral may be factorised as follows:

$$\begin{aligned} &\exp\!\left\{\mathrm{i}\int \mathrm{d}^4x\,\left[\mathcal{L} + J\varphi + \overline{\sigma}\psi + \bar{\psi}\sigma\right]\right\} \\ &\qquad = \exp\!\left\{\mathrm{i}\int \mathrm{d}^4x\;\mathcal{L}_{\mathrm{I}}\right\}\exp\!\left\{\mathrm{i}\int \mathrm{d}^4x\,\left[\mathcal{L}_0 + J\varphi + \overline{\sigma}\psi + \bar{\psi}\sigma\right]\right\}. \end{aligned} \tag{6.1.25}$$

Note that the presence of Grassmann variables in pairs is crucial for this factorisation. To expand the resulting interaction exponential as a power series in φ, ψ and $\bar{\psi}$, we shall need the following functional derivatives with respect to the external

spinor sources:

$$\frac{\delta}{\delta\sigma(x)}\,\exp\!\Big\{\mathrm{i}\int \mathrm{d}^4y\,\Big[\mathcal{L}_0+J\varphi+\overline{\sigma}\psi+\bar{\psi}\sigma\Big]\Big\} = -\mathrm{i}\,\bar{\psi}(x)\,\exp\!\Big\{\mathrm{i}\int \mathrm{d}^4y\,\Big[\mathcal{L}_0+J\varphi+\overline{\sigma}\psi+\bar{\psi}\sigma\Big]\Big\}, \quad (6.1.26a)$$

$$\frac{\delta}{\delta\overline{\sigma}(x)}\,\exp\!\Big\{\mathrm{i}\int \mathrm{d}^4y\,\Big[\mathcal{L}_0+J\varphi+\overline{\sigma}\psi+\bar{\psi}\sigma\Big]\Big\} = \mathrm{i}\,\psi(x)\,\exp\!\Big\{\mathrm{i}\int \mathrm{d}^4y\,\Big[\mathcal{L}_0+J\varphi+\overline{\sigma}\psi+\bar{\psi}\sigma\Big]\Big\}. \quad (6.1.26b)$$

Note the difference in sign between the two cases.

Putting this all together, we have

$$\mathcal{L}_{\mathrm{I}}\big(\varphi,\psi,\bar{\psi}\big)\exp\!\Big\{\mathrm{i}\int \mathrm{d}^4y\,\Big[\mathcal{L}_0+J\varphi+\overline{\sigma}\psi+\bar{\psi}\sigma\Big]\Big\} = \mathcal{L}_{\mathrm{I}}\Big(-\mathrm{i}\frac{\delta}{\delta J},-\mathrm{i}\frac{\delta}{\delta\overline{\sigma}},\mathrm{i}\frac{\delta}{\delta\sigma}\Big)\exp\!\Big\{\mathrm{i}\int \mathrm{d}^4y\,\Big[\mathcal{L}_0+J\varphi+\overline{\sigma}\psi+\bar{\psi}\sigma\Big]\Big\}, \quad (6.1.27)$$

from which we obtain

$$Z[J,\sigma,\overline{\sigma}] = \exp\!\Big\{\mathrm{i}\int \mathrm{d}^4x\;\mathcal{L}_{\mathrm{I}}\Big(-\mathrm{i}\frac{\delta}{\delta J},-\mathrm{i}\frac{\delta}{\delta\overline{\sigma}},\mathrm{i}\frac{\delta}{\delta\sigma}\Big)\Big\}\,Z_0[J,\sigma,\overline{\sigma}]. \quad (6.1.28)$$

The functional derivative of $Z_0[J,\sigma,\overline{\sigma}]$ with respect to to J has already been derived in (4.2.41); the two new cases are:

$$\frac{\delta Z[J,\sigma,\overline{\sigma}]}{\delta\sigma(x)} = Z[J,\sigma,\overline{\sigma}]\int \mathrm{d}^4y\,\overline{\sigma}(y)\,\mathrm{i}S_{\mathrm{F}}(x-y) \quad (6.1.29a)$$

and

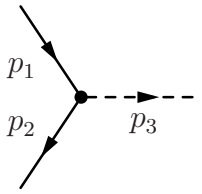
$$\frac{\delta Z[J,\sigma,\overline{\sigma}]}{\delta\overline{\sigma}(x)} = -Z[J,\sigma,\overline{\sigma}]\int \mathrm{d}^4y\,\mathrm{i}S_{\mathrm{F}}(x-y)\,\sigma(y). \quad (6.1.29b)$$

We now have all the necessary definitions to proceed with the derivation of the Feynman rules. Steps analogous to those taken in the scalar case lead to a new propagator, a new vertex and a new *anti*symmetry factor:

1. Internal lines, corresponding to fermion propagators of momentum p, are represented as

$$\xrightarrow{\;p\;} = \mathrm{i}S_{\mathrm{F}}(p) = \frac{\mathrm{i}(\not{p}+m)}{p^2-m^2+\mathrm{i}\varepsilon}. \quad (6.1.30)$$

2. In a theory with scalar fields there are Yukawa vertices, represented as



$$= -\mathrm{i}g\Gamma \quad (\Gamma = \mathbb{1} \text{ or } \gamma_5), \tag{6.1.31}$$

   with the constraint $p_1 - p_2 = p_3$.

3. Unconstrained loop four-momenta k are integrated over đ^4k.

4. A fermion loop leads to a trace of the Dirac matrices involved.

5. Any closed fermion loop carries an additional factor -1.

None of the symmetry factors associated with scalar fields apply to fermions in this theory. We shall explain the -1 for closed fermion loops shortly. Note that the matrix multiplications associated with the fermion propagators and vertices are to be taken in the *reverse* direction with respect to the propagator arrows.

The purely scalar vertices are as already derived in the previous chapter. Note, however, that even if we start with a scalar field having only a Yukawa interaction with the fermions in the theory and no explicit cubic or quartic term ion the Lagrangian, loop diagrams such as that in Fig. 6.1 will generate higher-order *effective* self-couplings.

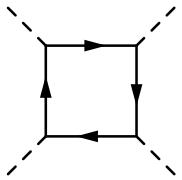


Figure 6.1: An effective quartic scalar interaction term generated at one-loop order via Yukawa interaction with a fermion field.

Let us now examine the case of closed fermion loops. Note first that, since fermion fields are neither created nor destroyed individually, but only in pairs, then any diagram will contain a number of continuous fermion lines. There are just two possibilities for each line: it is either open, *i.e.* it represents a particle (or antiparticle) entering the diagram and in some other point leaving; or it is a closed loop (representing a virtual state). The simplest case of the latter is the one-loop correction to the scalar two-point function. The relevant diagram is shown in Fig. 6.2. Identifying the corresponding term in the expansion of $Z[J,\sigma,\overline{\sigma}]$, we have

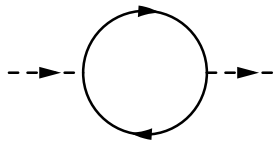


Figure 6.2: A fermion-loop diagram correction to the scalar two-point function.

(as already stated, $\Gamma = \mathbb{1}$ or γ_5):

$$\begin{aligned} Z[J,\sigma,\overline{\sigma}] \;=\; & \ldots \\ & + \frac{\mathrm{i}^2}{2!}\int \mathrm{d}^4x\,\mathrm{d}^4y\,\left[\left(-\mathrm{i}\frac{\delta}{\delta J(x)}\right)\left(\mathrm{i}\frac{\delta}{\delta\sigma(x)}\right)\left(-\mathrm{i}\,g\Gamma\right)\left(-\mathrm{i}\frac{\delta}{\delta\overline{\sigma}(x)}\right)\right] \\ & \quad\times\left[\left(-\mathrm{i}\frac{\delta}{\delta J(y)}\right)\left(\mathrm{i}\frac{\delta}{\delta\sigma(y)}\right)\left(-\mathrm{i}\,g\Gamma\right)\left(-\mathrm{i}\frac{\delta}{\delta\overline{\sigma}(y)}\right)\right] Z_0[J,\sigma,\overline{\sigma}] \\ & \qquad\qquad + \ldots \end{aligned} \tag{6.1.32}$$

And, suppressing the scalar part (since it causes no problems with respect to anti-commutation), the relevant term in the expansion of $Z_0[J,\sigma,\overline{\sigma}]$ to be differentiated is

$$\begin{aligned} & \frac{1}{2!}\left\{-\mathrm{i}\int \mathrm{d}^4z_1\,\mathrm{d}^4z_2\,\overline{\sigma}(z_1)\,S_{\mathrm{F}}(z_1-z_2)\,\sigma(z_2)\right\} \\ & \qquad\qquad\times\left\{-\mathrm{i}\int \mathrm{d}^4z_3\,\mathrm{d}^4z_4\,\overline{\sigma}(z_3)\,S_{\mathrm{F}}(z_3-z_4)\,\sigma(z_4)\right\}. \end{aligned} \tag{6.1.33}$$

Notice first of all that the expressions in braces in both Eqs. (6.1.32 & 32) may be reordered without incurring sign changes as they contain even numbers of Grassmann variables. Consider differentiating the sources at z_1 and z_4 with the derivatives at x; the ordering is already correct and no minus sign is introduced and thus the sign difference between Eqs. (6.1.29a & b) leaves one overall minus sign. This leaves the sources at z_2 and z_3 to be differentiated with the derivatives at y; the ordering is wrong here and a minus sign is introduced, which is then cancelled by the sign difference between Eqs. (6.1.29a & b). It should not be difficult to convince oneself that for any number of similarly connected propagators, none of the source derivatives but those acting on the outermost pair (*e.g.* z_1 and z_4 above) introduce minus signs, leaving just one overall change of sign due to this last pair if derivatives. Moreover, in the case of open fermion lines, which, as remarked above, is the only other case, this never happens since precisely the outermost pair of derivatives is then absent.

To complete the rules for calculation of full scattering amplitudes, we should

construct the analogue of $S[J,\varphi_0]$, also including the fermion fields and their sources. Following the procedure already established it is not difficult to see that the following new rules apply to external fermionic legs:

6. With each incoming or outgoing fermion line with momentum p and spin s we associate a factor $u(p,s)$ or $\bar{u}(p,s)$ respectively; for antifermions we use $v(p,s)$ or $\bar{v}(p,s)$.

7. There is an additional factor -1 between diagrams that differ by exchange of external identical fermion (or antifermion) legs.

A few cautionary words are in order. Some care must be exercised with massless fermions, such as the standard model neutrino. Although in general they can be taken to have the same propagator as derived here with the mass simply set to zero, in some cases this may lead to the appearance of "wrong" helicity states. Such problems can be avoided by use of Weyl spinors.

We shall not pursue Yukawa's theory further here. We now know that neither the nucleons nor the pions may be treated as elementary states (except in some low-energy approximation) and such an approach is of little practical use nowadays. For more discussions on the details of Yukawa's theory and examples of calculations performed, see Bjorken and Drell (1965). We now turn to what is considered the paradigm of elementary particle interactions, QED.

6.2 Gauge-field quantisation

All modern-day fundamental theories of particle interactions are based on the notion of *gauge symmetry*, which rests on the masslessness of the spin-one field it describes. It turns out that such a symmetry guarantees, above all, renormalisability. The prime example of a gauge theory is QED, which describes the electromagnetic interactions of electrons (spin-half, fermions) via the exchange of (spin-one, massless) photons. In its most elementary form then it contains just two fields and a single charge. With the advent of the unified theory of the electroweak interaction and also the fundamental theory of the strong interaction QCD, this basic notion has been extended to the case of a multiple charges with more complex gauge symmetries based on non-Abelian groups. We shall start with the easier Abelian case of QED.

6.2.1 Abelian gauge fields

To construct the fermion Lagrangian in the presence of a gauge field we apply the well-known principle of minimal coupling to the basic Lagrangian describing

a single fermion, which requires the following change

$$\partial^\mu \;\to\; D^\mu \;:=\; \partial^\mu + \mathrm{i}gA^\mu(x), \tag{6.2.1}$$

where $A^\mu(x)$ is the gauge field and g the charge of the fermion. The pure gauge part of the Lagrangian can be constructed from

$$F^{\mu\nu} \;:=\; \partial^\mu A^\nu - \partial^\nu A^\mu \;=\; (\mathrm{i}g)^{-1}\left[D^\mu, D^\nu\right], \tag{6.2.2}$$

which is manifestly covariant. The full Lagrangian is then

$$\mathcal{L} \;=\; \bar{\psi}\left(\mathrm{i}\not{D} - m\right)\psi - \tfrac{1}{4}F^{\mu\nu}F_{\mu\nu}. \tag{6.2.3}$$

From this it is easy to see that a mass term, which can only be $-\frac{1}{2}m^2A^\mu A_\mu$, would violate gauge invariance, *i.e.* invariance under the following transformation:

$$\psi(x) \;\to\; \mathrm{e}^{-\mathrm{i}g\Lambda(x)}\,\psi(x) \tag{6.2.4a}$$

and

$$A^\mu \;\to\; A^\mu + \partial^\mu\Lambda(x), \tag{6.2.4b}$$

for any scalar function $\Lambda(x)$.

6.2.2 Non-Abelian gauge fields

Let us now suppose that we wish to introduce a new conserved *multi-valued* quantum number (*e.g.* colour) and that we wish to *gauge* it, *i.e.* render local the global symmetry implied by conservation.* In this case we shall have, say, p fermion fields, where p is the number of different charges. Note that we do not merely wish to have p copies of the Abelian theory previously described—this would hardly be interesting. The point here is that there should be complete symmetry between the different charges and that there should only be a single charge *strength*, say, g. In other words, we require symmetry under the following more general infinitesimal transformation:

$$\delta\psi_i(x) \;=\; -\mathrm{i}\,g\,\delta\Lambda^a(x)\,\mathsf{T}^a_{ij}\,\psi_j(x) \qquad (i,j \;=\; 1,\ldots,p), \tag{6.2.5}$$

where the usual summation convention over repeated indices is understood, the infinitesimal functions $\delta\Lambda^a(x)$ are arbitrary and the form of the matrices T^a (for, say, $a=1,\ldots,r$) is to be determined. If the matrices T^a were merely multiples of the identity matrix then all we would have would be the case just mentioned: p

* Such a case was first consider by Yang and Mills (1954) and such theories are therefore often denoted with their names.

Abelian theories each with a possibly different charge strength.

The situation will be more interesting if the matrices are non-trivial. Evidently though, there are some restrictions they must satisfy. Let us introduce the natural extension of the covariant derivative to a $p\times p$ matrix (which may then act on the p-valued field ψ_i):

$$D^\mu \;:=\; \partial^\mu + \mathrm{i}g\mathsf{T}^a A^{a\mu}(x), \tag{6.2.6}$$

where the term ∂^μ is understood to be multiplied by the unit $p\times p$ matrix. The infinitesimal transformation of $\partial^\mu\psi_i(x)$ may be written as

$$\delta(\partial^\mu\psi_i(x)) \;=\; -\mathrm{i}g\,\mathsf{T}^a_{ij}\,\delta\Lambda^a(x)\,\partial^\mu\psi_j(x) - \mathrm{i}g\,\mathsf{T}^a_{ij}\,\psi_j(x)\,\partial^\mu\delta\Lambda^a(x). \tag{6.2.7}$$

As usual, it is the appearance of the second term on the right-hand side that requires the introduction of a field $A^{a\mu}(x)$ with specific transformation properties. If we write the more general

$$\delta A^{a\mu}(x) \;=\; \partial^\mu\,\delta\Lambda^a(x) + gf^{abc}\delta\Lambda^b(x)A^{c\mu}(x), \tag{6.2.8}$$

where the constants f^{abc} are to be determined, then the necessary cancellations will take place if and only if

$$\Big[\mathsf{T}^a\delta\Lambda^a(x), \mathsf{T}^b A^{b\mu}(x)\Big] \;=\; \mathrm{i}\,f^{abc}\,\mathsf{T}^a\delta\Lambda^b(x)A^{c\mu}(x). \tag{6.2.9}$$

Since both $\Lambda^a(x)$ and $A^{b\mu}(x)$ are arbitrary functions, this implies

$$\Big[\mathsf{T}^a, \mathsf{T}^b\Big] \;=\; \mathrm{i}\,f^{abc}\,\mathsf{T}^c. \tag{6.2.10}$$

We thus see that the matrices T^a form a representation of the (non-Abelian) Lie algebra defined by the structure constants f^{abc}. Hence, we refer to this as a non-Abelian gauge-field theory or Yang–Mills (YM) theory (Yang and Mills, 1954).

Exercise 6.2.1. *Derive the commutation relation in Eq. (6.2.9), as suggested above, by explicitly considering the effect of a gauge transformation on $D^\mu\psi$.*

To construct a suitable kinetic term for the gauge-field Lagrangian, it is easiest to follow the same approach as used in (6.2.2) above. Define first

$$A^\mu(x) \;:=\; \mathsf{T}^a A^{a\mu}(x), \tag{6.2.11}$$

so that $A^\mu(x)$ is now implicitly a matrix in the new space. The non-Abelian field

strength may then be defined naturally as

$$F^{\mu\nu} := (\mathrm{i}g)^{-1}\left[D^\mu, D^\nu\right] = \partial^\mu A^\nu - \partial^\nu A^\mu + \mathrm{i}g\left[A^\mu, A^\nu\right] \tag{6.2.12}$$

or, more explicitly,

$$F^{a\mu\nu} = \partial^\mu A^{a\nu} - \partial^\nu A^{a\mu} - g f^{abc} A^{b\mu} A^{c\nu}, \tag{6.2.13}$$

where the last term is a non-linear term, which will lead to the self-interactions characteristic of non-Abelian theories. We can thus finally construct the full gauged Lagrangian:

$$\mathcal{L} = \bar{\psi}\left(\mathrm{i}\not{D} - m\right)\psi - \tfrac{1}{4}F^{a\mu\nu}F^a_{\mu\nu}. \tag{6.2.14}$$

Again, it is manifestly impossible to maintain gauge invariance and at the same time insert a mass term, $-\frac{1}{2}m^2 A^{a\mu}A^a_\mu$, for the field $A^{a\mu}(x)$. Later on we shall discuss a mechanism to bestow the gauge field a non-vanishing effective mass via interaction with another field. Note that if the gauge group is a *simple* (*i.e.* irreducible) Lie algebra, then there is only one coupling constant; however, if it is semi-simple (*i.e.* a product of subgroups), then there may be as many different couplings as there are distinct subgroups. In the case of the electroweak theory, for example, the group is SU(2)×U(1) and there are two independent couplings.

We shall now examine the problem of quantising general gauge theories, both Abelian and non-Abelian. While the problem of quantising Abelian theories is somewhat simpler, the clearest and most intuitive approach is that which is generally applied to non-Abelian theories. We shall therefore jump in at the deep end and consider non-Abelian quantum field theories from the start.

6.2.3 Difficulties in gauge-field quantisation

Let us first consider the case of a pure gauge theory; *i.e.* not coupled to a fermion field as in full QED or QCD. If we *naïvely* follow the practice so far developed and consolidated, then we immediately arrive at the following form for the generating functional:

$$Z_0[J^{a\mu}] \propto \int \mathcal{D}A \exp\left\{\mathrm{i}\int \mathrm{d}^4x \left[\mathcal{L}_0(A^{a\mu}) + J^a_\mu A^{a\mu}\right]\right\}, \tag{6.2.15}$$

where $\mathcal{D}A$ implies a path integral running over space–time coordinates *and* charges (*i.e.* over *all* the indices of $A^{a\mu}$) and $\mathcal{L}_0$ is just the gauge-field kinetic term* in the Lagrangian (6.2.14), or

$$\mathcal{L}_0 = -\tfrac{1}{4}F^{a\mu\nu}F^a_{\mu\nu}. \tag{6.2.16}$$

* As we have seen, in the case of Yang–Mills (YM) (*i.e.* non-Abelian) theories, it is not, however, purely quadratic and contains higher-order interaction terms.

We shall now see that this is insufficient, owing principally to the gauge symmetry of the field and the fact that this property induces a reduction in the three natural degrees of freedom of a massive vector field down to the two transverse states of, say, a photon.

To first glimpse the nature of the problem, let us imagine performing the integral of the part quadratic in the fields $A^{a\mu}$ (*i.e.* just the QED-like piece) in the standard manner to arrive at an exact form for the generating functional:

$$Z_0[J] \;=\; \exp\!\left\{-\tfrac{1}{2}\,\mathrm{i}\int \mathrm{d}^4x\,\mathrm{d}^4y\,J^{a\mu}(x)\,D^{ab}_{\mathrm{F}\mu\nu}(x,y)\,J^{b\nu}(y)\right\}, \tag{6.2.17}$$

where $D^{ab}_{\mu\nu}(x,y)$ should be the Feynman propagator or inverse of what we might call the Maxwell operator (the equivalent of the Klein–Gordon or Dirac operators in the scalar or spinor cases already discussed). First we make explicit the conservation of the charge:

$$D^{ab}_{\mathrm{F}\mu\nu}(x,y) \;=\; \delta^{ab}\,D_{\mathrm{F}\mu\nu}(x,y). \tag{6.2.18}$$

Next we define the Fourier transform:

$$D_{\mathrm{F}\mu\nu}(x,y) \;=\; \int \text{đ}^4q\;\mathrm{e}^{-\mathrm{i}\,q\cdot(x-y)}\,D_{\mathrm{F}\mu\nu}(q). \tag{6.2.19}$$

From the Maxwell equations (2.4.45a), we see that in momentum space the inverse propagator (or Maxwell operator) is

$$D^{-1}_{\mathrm{F}\mu\nu}(q) \;=\; -q^2\,g_{\mu\nu} + q_\mu q_\nu. \tag{6.2.20}$$

We now require the gauge-field propagator or inverse of the previous expression. The most general expression possible is

$$D^{\mu\nu}_{\mathrm{F}}(q) \;=\; q^2\,g^{\mu\nu}A(q^2) + q^\mu q^\nu B(q^2). \tag{6.2.21}$$

The only other tensor available $\varepsilon^{\mu\nu\rho\sigma}$ is antisymmetric and is therefore excluded in this case. We thus simply require

$$D^{\mu\nu}_{\mathrm{F}}(q)\,D^{-1}_{\mathrm{F}\nu\rho}(q) \;=\; \delta^\mu_\rho. \tag{6.2.22}$$

Now, it is not difficult to see that there is, in fact, *no* solution to this equation.

Exercise 6.2.2. *Define the transverse and longitudinal projection operators*

$$P^{\mu\nu}_T(q) \;:=\; g^{\mu\nu} - q^\mu q^\nu/q^2, \tag{6.2.23a}$$

$$P^{\mu\nu}_L(q) \;:=\; q^\mu q^\nu/q^2 \tag{6.2.23b}$$

and use these to show that since $D^{-1}_{F\mu\nu}(q)$ has zero longitudinal component, its inverse $D_F^{\mu\nu}(q)$ would have an infinite longitudinal component.

Note that the projection operators defined in Ex. 6.2.2 have the following properties:

$$P_X^{\mu\rho}(q)\, P_{X\rho}{}^{\nu}(q) \;\equiv\; P_X^{\mu\nu}(q) \qquad \text{for } X \;=\; \mathrm{T}, \mathrm{L} \tag{6.2.24a}$$

and

$$P_{\mathrm{T}}^{\mu\nu}(q) + P_{\mathrm{L}}^{\mu\nu}(q) \;\equiv\; g^{\mu\nu} \qquad \text{(the identity operator)}. \tag{6.2.24b}$$

In other words, the operator $D_{\mathrm{F}}^{\mu\nu}(q)$ is itself a projector onto the transverse components of A^μ. This implies, in particular, that in the Euclidean formulation the Gaussian integrals along the longitudinal directions are not damped and therefore do not converge.

We see then that the problem lies with the longitudinal components. Indeed, we might have expected such a problem since the functional measure, $\mathcal{D}A$, used implies a product of path integrals over each of the four space–time components of the gauge field $A^{a\mu}$. In other words, the full integral runs over subspaces of field configurations that are equivalent under gauge transformation. Each gauge trajectory in this equivalence space has infinite length and the Lagrangian is invariant, *i.e.* constant, therefore the contribution to the path integral is infinite.

The situation might be likened to attempting minimisation with respect to redundant or dummy variables. Intuitively, one might resolve such a problem by inserting a δ-function to select one particular (arbitrary) value of each redundant variable. In the path-integral this would mean restricting the region to a hypersurface of (arbitrary) constant gauge. This, in essence, is the technique developed by Fadde'ev and Popov (1967). Note that one could, in principle, fix the gauge from the start and avoid such a problems (for example, see Abers and Lee, 1973, p. 76).* However, gauge covariance would then be evidently lost; indeed, for many choices of the gauge-fixing condition, even Lorentz covariance would no longer be manifest. The obvious difficulty is the definition of the required gauge *δ-functional*.

6.2.4 Fadde'ev–Popov ghosts

Consider the full purely Yang–Mills action

$$S[A] \;=\; \int \mathrm{d}^4x\; \mathcal{L}_0(A). \tag{6.2.25}$$

By construction, this is gauge invariant; that is, it does not vary for fields lying along the trajectory $\mathsf{U}(\lambda)A^a_\mu$ where A^a_μ is some arbitrary fixed reference point, λ an

* This would also be closer to the canonical approach to gauge-field quantisation.

element of the gauge group $\mathcal{G}$ and $\mathsf{U}(\lambda)$ the corresponding gauge-transformation operator. That is,

$$S[A^\lambda] \;=\; S[A], \tag{6.2.26}$$

where

$$A^\lambda_\mu \;\equiv\; \mathsf{T}^a A^{a\lambda}_\mu \;=\; \mathsf{U}(\lambda)\left[\mathsf{T}^a A^a_\mu + (\mathrm{i}g)^{-1}\,\mathsf{U}^{-1}(\lambda)\partial_\mu\,\mathsf{U}(\lambda)\right]\mathsf{U}^{-1}(\lambda). \tag{6.2.27}$$

Such behaviour evidently makes nonsense of the stationary-phase condition, as implemented so far.

The idea is now to fix a gauge via a δ-functional and thus limit the path integral to a hypersurface in the manifold of field configurations. That is, we shall apply a condition of the type

$$\mathcal{F}^a(A_\mu) \;=\; 0, \qquad (a \;=\; 1,\ldots,N), \tag{6.2.28}$$

where N is the dimension of the gauge group. For this to single out just one point on each gauge trajectory, the equation

$$\mathcal{F}^a(A^\lambda_\mu) \;=\; 0, \tag{6.2.29}$$

must have a unique solution in λ for any given A^a_μ. The condition corresponding, *e.g.*, to the Coulomb gauge would be

$$\mathcal{F}^a(A_\mu) \;=\; \boldsymbol{\nabla}\cdot\boldsymbol{A}^a. \tag{6.2.30}$$

In the neighbourhood of the identity a general gauge transformation may be parametrised with

$$\mathsf{U}(\lambda) \;\simeq\; 1 - \mathrm{i}g\mathsf{T}^a\Lambda^a(x). \tag{6.2.31}$$

And an invariant measure $\mathrm{d}\lambda$ may be defined over the gauge group such that

$$\mathrm{d}\lambda' \;=\; \mathrm{d}(\lambda\lambda') \tag{6.2.32}$$

and it may always be written locally as

$$\mathrm{d}\lambda \;=\; \mathrm{d}\Lambda\,, \qquad (\lambda \;\approx\; 1). \tag{6.2.33}$$

Consider now the path integral

$$1 \;=\; \Delta_{\mathcal{F}}[A]\int \mathcal{D}\lambda\,\delta[\mathcal{F}(A^\lambda)], \tag{6.2.34}$$

where $\delta[\mathcal{F}(A^\lambda)]$ is the δ-functional alluded to previously. This equation simply defines the normalisation factor $\Delta_{\mathcal{F}}[A]$. Let us insert this into the standard path

integral for the vacuum-to-vacuum amplitude:

$$Z[J]|_{J=0} \propto \int \mathcal{D}A\, \Delta_{\mathcal{F}}[A] \int \mathcal{D}\lambda\, \delta[\mathcal{F}(A^{\lambda})] \exp\left\{ \mathrm{i} \int \mathrm{d}^4x\ \mathcal{L}(A) \right\}. \tag{6.2.35}$$

We now can perform a gauge transformation on the integrand:

$$A^a_\mu \ \rightarrow\ A^{a(\lambda^{-1})}_\mu. \tag{6.2.36}$$

Observe though that not only is $\mathcal{L}(A)$ invariant, but so too is $\Delta_{\mathcal{F}}[A]$. To see the latter invariance, write

$$\begin{aligned} \Delta^{-1}_{\mathcal{F}}[A] \ :=\ & \int \mathcal{D}\lambda''\, \delta[\mathcal{F}(A^{\lambda''})] \ =\ \int \mathcal{D}(\lambda\lambda')\, \delta[\mathcal{F}(A^{\lambda\lambda'})] \\ =\ & \int \mathcal{D}(\lambda')\, \delta[\mathcal{F}(A^{\lambda\lambda'})] \ =\ \Delta^{-1}_{\mathcal{F}}[A^{\lambda}]. \end{aligned} \tag{6.2.37}$$

We thus obtain

$$Z[J]|_{J=0} \propto \int \mathcal{D}A\, \Delta_{\mathcal{F}}[A] \int \mathcal{D}\lambda\, \delta[\mathcal{F}(A)] \exp\left\{ \mathrm{i} \int \mathrm{d}^4x\ \mathcal{L}(A) \right\}, \tag{6.2.38}$$

in which the integrand is now manifestly gauge invariant. This is precisely the observation of Fadde'ev and Popov: namely, that $\int \mathcal{D}\lambda$ is just a constant factor (albeit infinite) that is now independent of the fields and that may therefore be separated out and absorbed into the normalisation. We may thus write

$$Z[J]|_{J=0} \propto \int \mathcal{D}A\, \Delta_{\mathcal{F}}[A]\, \delta[\mathcal{F}(A)] \exp\left\{ \mathrm{i} \int \mathrm{d}^4x\ \mathcal{L}(A) \right\}. \tag{6.2.39}$$

One can show that the inclusion of source terms involving *conserved* currents does not change this result.

We now have the task of evaluating $\Delta_{\mathcal{F}}[A]$. The integral already contains the δ-functional and so we may evaluate $\Delta_{\mathcal{F}}$ for fields satisfying the gauge condition (6.2.28). We may therefore restrict ourselves to gauge transformations in the neighbourhood of the identity and write

$$\mathcal{D}\lambda \ =\ \mathcal{D}\Lambda \ =\ \mathcal{D}\mathcal{F}\, \det\left(\frac{\delta \Lambda^a}{\delta \mathcal{F}^b} \right) \Bigg|_{\mathcal{F}^b=0}. \tag{6.2.40}$$

We thus finally have

$$\Delta^{-1}_{\mathcal{F}}[A] \ =\ \int \mathcal{D}\lambda\, \delta[\mathcal{F}(A^{\lambda})] \ =\ \det\left(\frac{\delta \Lambda^a}{\delta \mathcal{F}^b} \right) \Bigg|_{\mathcal{F}^b=0}, \tag{6.2.41}$$

the inverse of which is

$$\Delta_{\mathcal{F}}[A] \;=\; \det\left(\frac{\delta\mathcal{F}^b}{\delta\Lambda^a}\right)\Bigg|_{\mathcal{F}^b=0}. \tag{6.2.42}$$

All that remains is to find useful representations of both the determinant and δ-functional. The solution is to convert them to exponential (Gaussian) form.

The δ-functional $\delta[\mathcal{F}(A)]$

A typical choice for the gauge-fixing condition might be

$$\mathcal{F}^a(A) \;=\; \partial^\mu A^a_\mu(x) - f^a(x) \;=\; 0 \tag{6.2.43}$$

for some set of arbitrary functions $f^a(x)$. (We shall examine a further possibility shortly.) Now, the following path integral is evidently just a constant:

$$\text{constant} \;=\; \int \mathcal{D}f\, \exp\left\{-\frac{\mathrm{i}}{2\xi}\int \mathrm{d}^4x\,[f^a(x)]^2\right\}, \tag{6.2.44}$$

where summation over the index a is understood in $[f^a(x)]^2$. We can therefore insert it into the previous expression for the vacuum-to-vacuum amplitude:

$$\begin{aligned} Z[J]|_{J=0} \;\propto& \int \mathcal{D}f\, \exp\left\{-\frac{\mathrm{i}}{2\xi}\int \mathrm{d}^4x\,[f^a(x)]^2\right\} \\ &\qquad\times \int \mathcal{D}A\,\Delta_{\mathcal{F}}[A]\,\delta[\mathcal{F}(A)]\exp\left\{\mathrm{i}\int \mathrm{d}^4x\;\mathcal{L}(A)\right\} \\ =& \int \mathcal{D}A\, \exp\left\{-\frac{\mathrm{i}}{2\xi}\int \mathrm{d}^4x\,[\partial^\mu A^a_\mu(x)]^2\right\} \\ &\qquad\times \Delta_{\mathcal{F}}[A]\exp\left\{\mathrm{i}\int \mathrm{d}^4x\;\mathcal{L}(A)\right\}, \end{aligned} \tag{6.2.45}$$

where the last equality is obtained by interchanging the order of integration and using the δ-functional to perform the path integral in $\mathcal{D}f$. The result is none other than the technique of Lagrange multipliers applied to path integrals.

The functional determinant $\det(\delta\Lambda^a/\delta\mathcal{F}^b)$

Noting that we have a determinant raised to a positive power, we recall that just such an object was previously obtained as the result of path integration over

complex Grassmann variables, see Eq. (6.1.3). Thus, we may write

$$\int \mathcal{D}\theta\,\mathcal{D}\theta^*\,\exp\left\{\mathrm{i}\int \mathrm{d}^4x\,\mathrm{d}^4y\,\theta^{a*}(x)\,M^{ab}(x,y)\,\theta^b(y)\right\}\,\propto\,\det(-M), \tag{6.2.46}$$

where gauge invariance requires that the complex scalar Grassmann fields belong to the adjoint representation, in addition a factor "i" has been introduced into the right-hand side for correct damping in the integral representation to be introduced next. If we now set

$$M^{ab}(x,y)\;=\;-\frac{\delta\mathcal{F}^a(x)}{\delta\Lambda^b(y)}, \tag{6.2.47}$$

then we may write

$$\det\left(\frac{\delta\mathcal{F}^a}{\delta\Lambda^b}\right)\;\propto\;\int \mathcal{D}\theta\,\mathcal{D}\theta^*\,\exp\left\{\mathrm{i}\int \mathrm{d}^4x\,\mathrm{d}^4y\,\theta^{a*}(x)\,\frac{\delta\mathcal{F}^a(x)}{\delta\Lambda^b(y)}\,\theta^b(y)\right\}. \tag{6.2.48}$$

From the infinitesimal form of the gauge transformation given in Eq. (6.2.8) we immediately have

$$\frac{\delta A^{a\mu}(x)}{\delta\Lambda^b(y)}\;=\;\left[\delta^{ab}\,\partial_x^\mu+gf^{abc}A^{c\mu}(x)\right]\delta^4(x-y) \tag{6.2.49}$$

and

$$\frac{\delta\mathcal{F}^a(x)}{\delta\Lambda^b(y)}\;=\;\partial_{y\mu}\left\{\left[\delta^{ab}\,\partial_x^\mu+gf^{abc}A^{c\mu}(x)\right]\delta^4(x-y)\right\}. \tag{6.2.50}$$

Inserting this into the exponential in the Grassmann path integral for the determinant gives

$$\begin{aligned}\int \mathrm{d}^4x\,\mathrm{d}^4y\;\theta^{a*}(x)\,&\frac{\delta\mathcal{F}^a(x)}{\delta\Lambda^b(y)}\,\theta^b(y)\\ &=\;\int \mathrm{d}^4x\,\partial^\mu\theta^{a*}(x)\left[\partial_\mu\theta^a(x)+gf^{abc}\theta^b(x)\,A^{c\mu}(x)\right].\end{aligned} \tag{6.2.51}$$

Finally, putting all this together in the full generating functional, we have

$$\begin{aligned}Z_0[J_\mu^a]\;\propto\;&\int \mathcal{D}A\,\mathcal{D}\theta\,\mathcal{D}\theta^*\\ &\times\exp\left\{\mathrm{i}\int \mathrm{d}^4x\left[\mathcal{L}_0(A)+\mathcal{L}_{\mathrm{GF}}(A)+\mathcal{L}_{\mathrm{FP}}(A,\theta,\theta^*)+J_\mu^a\,A^{a\mu}\right]\right\},\end{aligned} \tag{6.2.52}$$

where

$$\mathcal{L}_0(A)\;=\;-\tfrac{1}{4}\,F^{\mu\nu}\,F_{\mu\nu}, \tag{6.2.53a}$$

$$\mathcal{L}_{\rm GF}(A) \;=\; -\tfrac{1}{2}\xi^{-1}\left(\partial^\mu A^a_\mu\right)^2, \tag{6.2.53b}$$

$$\mathcal{L}_{\rm FP}(A,\theta,\theta^*) \;=\; \partial^\mu\theta^{a*}(x)\left[\partial_\mu\theta^a(x) + gf^{abc}\theta^b(x)\,A^{c\mu}(x)\right]. \tag{6.2.53c}$$

The first term is just the original Yang–Mills Lagrangian, the second is a gauge-fixing term while the third contains the so-called Fadde'ev–Popov ghosts. The purpose of these last is quite simply to cancel the unwanted extra degree of freedom of the gauge field (hence their Grassmannian nature). Note that the second and third term go hand-in-hand and cannot be considered as independent.*

Indeed, inspecting the ghost-field term, we see that there is a quadratic contribution, which will lead to Feynman rules corresponding to an *anticommuting* scalar with therefore a minus sign for each ghost loop. There is also an interaction term involving the gauge field itself. However, we immediately see that for an Abelian gauge theory (for which effectively $f^{abc}=0$) the ghosts actually decouple. That is, they do not contribute to any physical process and may therefore be ignored. In a non-Abelian theory, with this choice of gauge, they are necessary. Note, of course, that they only contribute as internal propagators and never appear as physical states.

One might wonder if there is a gauge choice that simply avoids the spurious degrees of freedom and thus too such unphysical fields. There do indeed exists such physical (or axial) gauges, but there is a price to be paid. The required gauge condition is

$$\mathcal{F}^a(A) \;=\; n^\mu A^a_\mu(x) - f^a(x) \;=\; 0, \tag{6.2.54}$$

where n^μ is some arbitrary external four-vector. An immediately obvious drawback is the loss of manifest Lorentz invariance (due to the introduction of an external vector). Of course, Lorentz invariance is only lost in individual diagrams; in any final physical quantity all reference to the vector n^μ must (and does) disappear. However, a further problem is that while the ghosts then decouple even in a non-Abelian theory, the gauge-field propagator takes on a rather complicated form. It may be simplified to some extent by a suitable choice of gauge-fixing parameter ξ and of a light-like vector for n^μ. Such a physical gauge is vital in certain applications; for example, those in which a diagram-by-diagram probabilistic interpretation is desired.

Exercise 6.2.3. *Derive the form of the generating functional corresponding to the gauge choice* (6.2.54) *and show that, even for a non-Abelian gauge theory, the ghosts do indeed decouple.*

* In principle, in order to derive the Feynman rules, a ghost source term should also be added.

6.3 Feynman rules for QED and YM theories

To the Lagrangian (6.2.53c) given towards the end of the last section we may also add a minimally coupled fermion contribution:

$$\mathcal{L}_{\mathrm{F}} \;=\; \bar{\psi}\left(\mathrm{i}\not{D} - m\right)\psi, \tag{6.3.1}$$

with the covariant derivative defined by

$$D^{\mu} \;:=\; \partial^{\mu} + \mathrm{i}g\mathsf{T}^{a}A^{a\mu}(x). \tag{6.3.2}$$

The Abelian case is obtained by replacing the fundamental representation matrices T^{a} with a single unit matrix—while still allowing for possible copies of the fermion field.* We can now move directly to the derivation of the Feynman rules corresponding to such a Lagrangian.

6.3.1 Gauge-field propagator

Carrying out the path integral on the entire quadratic part in $A^{a\mu}$ leads to the usual exact form for the free-field generating functional (6.2.17) with now

$$D^{ab}_{\mathrm{F}\mu\nu}(q) \;=\; \delta^{ab}\,D_{\mathrm{F}\mu\nu}(q) \tag{6.3.3a}$$

and

$$D^{-1}_{\mathrm{F}\mu\nu}(q) \;=\; -q^{2}\,g_{\mu\nu} + (1-\xi^{-1})\,q_{\mu}q_{\nu}, \tag{6.3.3b}$$

which, rewritten in terms of the transverse and longitudinal projectors defined in Eqs. (6.2.23), is

$$= \; -q^{2}\left[P_{\mathrm{T}}(q) + \xi^{-1}P_{\mathrm{L}}(q)\right]. \tag{6.3.3c}$$

This can now be readily inverted to give

$$\begin{aligned} D^{\mu\nu}_{\mathrm{F}}(q) \;&=\; -(q^{2}+\mathrm{i}\varepsilon)^{-1}\left[P^{\mu\nu}_{\mathrm{T}}(q) + \xi\,P^{\mu\nu}_{\mathrm{L}}(q)\right] \\ &=\; (q^{2}+\mathrm{i}\varepsilon)^{-1}\left[-g^{\mu\nu} + (1-\xi)\,q^{\mu}q^{\nu}/q^{2}\right]. \end{aligned} \tag{6.3.4}$$

The gauge-fixing parameter ξ is arbitrary and may be suitably chosen so as to simplify calculations. Two common choices are due to Feynman $\xi = 1$ and to Landau $\xi = 0$. Note that while the propagator is at its absolute simplest for the Feynman choice, in the Landau gauge the propagator is purely transverse.

* That is, for example, to admit a number of different lepton and/or quark flavours—distinguished via a *global* quantum number.

Exercise 6.3.1. *Find the explicit Lorentz structure of $D_F^{\mu\nu}(q)$ for the Landau gauge $\xi=0$ in the reference frame for which the gauge-field momentum is $q^\mu \propto (1,0,0,1)$.*

Exercise 6.3.2. *Derive the form of the propagator in the case of an axial gauge and examine the possible simplifications due to specific choices of external vector and gauge-fixing parameter.*

The Feynman rule for the propagator for an Abelian field (*e.g.* the photon) is traditionally represented with a wavy line, thus:

$$\mu \overset{q}{\sim\!\sim\!\sim} \nu \quad = \mathrm{i}\, D_{\mathrm{F}}^{\mu\nu}(q). \tag{6.3.5}$$

In contrast, a gluon (or YM) propagator is usually represented with a curly line:

$$a,\mu \overset{q}{\sim\!\sim\!\sim} b,\nu = \mathrm{i}\,\delta^{ab}\, D_{\mathrm{F}}^{\mu\nu}(q). \tag{6.3.6}$$

In both cases an arrow is only really necessary if the momentum flow in the diagram is ambiguous.

6.3.2 The gauge-field vertices

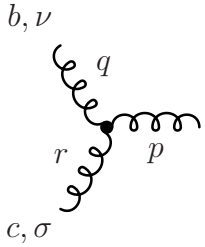
There are now a number of vertices. First of all, there is a cubic gauge-field term in the Lagrangian:

$$\int \mathrm{d}^4x\; \mathcal{L}(A^{\mathrm{cubic}}) \;=\; -\mathrm{i}\, g f^{abc} \int \text{đ}^4p\, \text{đ}^4q\, \text{đ}^4r\, \delta^4(p{+}q{+}r)\, p^\nu\, g^{\mu\sigma}\, A_\mu^a(p)\, A_\nu^b(q)\, A_\sigma^c(r). \tag{6.3.7}$$

Remembering that the fields A_μ^a *etc.* will be replaced by functional derivatives of the corresponding source or current and that these will act in all possible permutations, the previous expression may rendered more manifestly symmetric with respect to the incoming legs by suitably cycling the colour and space–time indices (exploiting both the *symmetry* of $g^{\mu\sigma}$ and the *antisymmetry* of f^{abc}):

$$-\,\mathrm{i}\, g f^{abc}\, p^\mu\, g^{\nu\sigma} \;\mathrel{\widehat{=}}\; -\tfrac{1}{3!}\,\mathrm{i}\, g f^{abc}\big[(p-r)^\nu\, g^{\mu\sigma} + (q-p)^\sigma\, g^{\nu\mu} + (r-q)^\mu\, g^{\sigma\nu}\big]. \tag{6.3.8}$$

Note that all three momenta are understood to be incoming. This then leads to the following Feynman rule:



$$a,\mu \;=\; -\mathrm{i}\, g\, f^{abc}\big[(p{-}r)^\nu\, g^{\mu\sigma} + (q{-}p)^\sigma\, g^{\nu\mu} + (r{-}q)^\mu\, g^{\sigma\nu}\big], \tag{6.3.9}$$

with the constraint $p+q+r=0$.

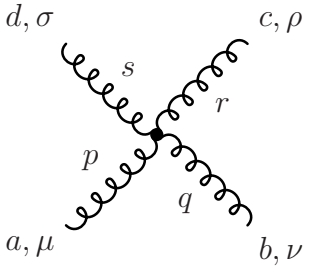
Next, there is a quartic gauge-field term in the Lagrangian:

$$\int \mathrm{d}^4x\ \mathcal{L}(A^{\text{quartic}}) \;=\; -\tfrac{1}{4}g^2 f^{abe} f^{cde} \int \text{đ}^4p\,\text{đ}^4q\,\text{đ}^4r\,\text{đ}^4s\,\delta^4(p+q+r+s)\; g^{\mu\rho}\,g^{\nu\sigma}\,A^a_\mu(p)\,A^b_\nu(q)\,A^c_\rho(r)\,A^d_\sigma(s). \quad (6.3.10)$$

Again, the functional derivatives of the corresponding source (or current) will act in all possible permutations and thus the previous expression may be rendered more manifestly symmetric:

$$-\tfrac{1}{4}g^2 f^{abe} f^{cde} g^{\mu\rho}\,g^{\nu\sigma} \;\widehat{=}\; -\tfrac{1}{4!}g^2\big[f^{abe} f^{cde}(g^{\mu\rho}\,g^{\nu\sigma} - g^{\nu\rho}\,g^{\mu\sigma}) + \text{cyclic perms.}\big]. \quad (6.3.11)$$

As before, all four momenta are understood to be incoming. This then leads to the following Feynman rule:



$$= -g^2\big[f^{abe} f^{cde}(g^{\mu\rho}\,g^{\nu\sigma} - g^{\nu\rho}\,g^{\mu\sigma}) + \text{cyclic perms.}\big], \quad (6.3.12)$$

with the constraint $p+q+r+s=0$.

Finally, we note that the related Abelian theory is recovered by simply setting the structure constants f^{abe} to zero. In which case we see that both the cubic and quartic self-interaction terms disappear. Turning this the other way around, we also see that, in contrast to all the theories studied so far, a free-field version of a non-Abelian theory *does not exist*. In other words, a non-Abelian or Yang–Mills theory is inevitably non-trivial: one might imagine bound states constructed purely from the gauge field itself. In QCD these are known as glueballs and although there are no known states that can be identified as such with 100% confidence, there exist several states that are probably at least superpositions of a glueball and some $q\bar{q}$-like object.

6.3.3 The Fadde'ev–Popov ghost rules

We now turn to the remaining terms in the interaction Lagrangian. The quadratic term in the Lagrangian for the Fadde'ev–Popov ghosts is the same as that already derived for a scalar (with *no* mass term) and the corresponding propagator is

usually indicated with a dotted line:

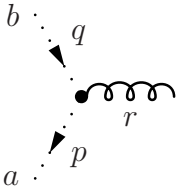
$$a \cdots\!\!\blacktriangleright^{q}\!\!\cdots b \quad = \mathrm{i}\Delta_{\mathrm{F}}(q) \,=\, \delta^{ab}\,\frac{\mathrm{i}}{q^2+\mathrm{i}\varepsilon}. \tag{6.3.13}$$

Here the arrow is always necessary as the vertex has explicit momentum dependence.

The ghost fields have an interaction term connecting them to the gauge field:

$$\int \mathrm{d}^4x\; \mathcal{L}(A^{\mathrm{ghost}}) \,=\, \mathrm{i}gf^{abc}\int \text{đ}^4p\,\text{đ}^4q\,\text{đ}^4r\,p^\mu\,\theta^{a*}(p)\,\theta^b(q)\,A^{c\mu}(r)\,\text{đ}^4(q+r-p), \tag{6.3.14}$$

where we have chosen to explicitly express one of the ghost lines as outgoing. The corresponding Feynman rule is as follows:



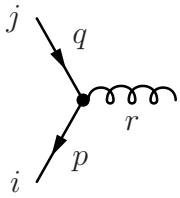
$$= -g\,f^{abc}\,p^\mu, \tag{6.3.15}$$

with the constraint that now $q+r=p$.

The fermion propagator is almost the same as before, with just the minor addition of colour conservation:

$$i \xrightarrow{\;p\;} j \;=\; \mathrm{i}\delta_{ji}\,S_{\mathrm{F}}(p). \tag{6.3.16}$$

The fermion–gauge-field interaction term leads to the following Feynman rule:



$$= -\mathrm{i}\,g\,(\mathsf{T}^a)_{ij}\,\gamma^\mu, \tag{6.3.17}$$

with the constraint $p+r=q$. For an Abelian theory one merely replaces the T^a with a unit matrix. Note that the matrix structure (in both Dirac and colour space) implies multiplication from left to right starting from an outgoing (incoming) fermion (antifermion), or simply running *backwards* in a fermion loop, where again a trace is then taken.

In order to calculate scattering amplitudes, we also need to represent possible external gauge fields (recall that ghosts obviously never appear as external physical states). From Eqs. (2.4.55) and (2.4.63), considering the functional derivatives with

respect to $a(k,\lambda)$ and $a^*(k,\lambda)$, we immediately derive the following factors $\varepsilon_\mu(k,\lambda)$ and $\varepsilon^*_\mu(k,\lambda)$ for incoming and outgoing states respectively, where λ indicates either a specific helicity or linear polarisation state.

Exercise 6.3.3. *From the form of the generating functional corresponding to the gauge choice* (6.2.54)*, derived in a previous exercise, extract the related Feynman rules. Consider the possible choices of gauge-fixing parameter and vector that simplify the expression.*

Exercise 6.3.4. *Consider the choice* $\xi = 0$ *and* n^μ *light-like, use the reference frame in which the gauge-field momentum is* $q^\mu \propto (1,0,0,1)$ *and also* $n^\mu \propto (1,0,0,-1)$*. Find the explicit Lorentz structure of* $D_F^{\mu\nu}(q)$ *and compare it to the case of the Landau gauge.*

6.4 Renormalisation in QED and YM theories

In this section we turn our attention to the question of renormalisation in gauge theories, such as QED and Yang–Mills. Since the technicalities are very much the same, we shall continue to concentrate on the discussion on the more general case of Yang–Mills. It will, however, be interesting to compare the specific results in both Abelian and non-Abelian theories.

6.4.1 Renormalisability

Following the previous arguments on superficial degrees of divergence and noting that the high-momentum behaviour of the gauge propagator is the same as for a scalar field, we can derive a renormalisability condition for vertices in the gauge theories discussed in the previous section:

$$N_B + \tfrac{3}{2}N_F + N_D \;\leq\; 4, \tag{6.4.1}$$

where the additional term N_D reflects the number of derivatives associated with the interaction (or vertex). Considering the vertices just derived, we see that a gauge theory is, in general, just renormalisable (that is, the previous inequality is precisely saturated for all vertices).

Note, however, that even in the pure gauge theory (*i.e.* with no fermions) there are two vertices involving the same coupling constant g and the ghost fields (for certain gauge choices) imply a third. Moreover, the addition of fermions leads to a fourth. At first sight it, would thus seem that there is a risk of different renormalisation constants for the single coupling, arising from consideration of the different vertices. Fortunately, it can be shown, to *all orders* in perturbation

theory that this does *not* happen and that the gauge symmetry somehow protects the coupling by guaranteeing that, whichever vertex we choose to consider, the renormalisation is the same (this is partially discussed in Sec. 7.3). That is, only a single subtraction constant is necessary. One can, of course, also verify explicitly that this is indeed true order-by-order in perturbation theory.

6.4.2 Counter terms

The full Lagrangian for the general case of a non-Abelian gauge theory based on a simple Lie group and coupled to N_f fermion fields* is

$$\mathcal{L} \;=\; \bar{\psi}\left(\mathrm{i}\not{D} - m\right)\psi - \tfrac{1}{4}F^{a\mu\nu}F^{a}_{\mu\nu} - \tfrac{1}{2}\xi^{-1}\left(\partial^\mu A^a_\mu\right)^2 + \partial^\mu\theta^{a*}\left(\partial_\mu\theta^a + g f^{abc}\theta^b A^{c\mu}\right). \tag{6.4.2}$$

As before, we now wish to rewrite this as

$$\mathcal{L}_B \;=\; \mathcal{L} + \delta\mathcal{L}, \tag{6.4.3}$$

where the bare and renormalised lagrangians ($\mathcal{L}_B$ and $\mathcal{L}$) should have exactly the same form, with the first being simply written in terms of bare quantities; $\delta\mathcal{L}$ contains only counter terms, which *must* remove all infinities and *may* also subtract (irrelevant) constant terms. The objects to be renormalised are then the three fields ψ, $A^{a\mu}$ and θ^a (the ghosts); the coupling constant g; the fermion mass m and finally too the gauge-fixing parameter ξ.

The counter-term Lagrangian will then take the following form:

$$\begin{aligned}\delta\mathcal{L} \;=\;& \delta Z_\psi\,\bar{\psi}\mathrm{i}\not{\partial}\psi - g\,\delta Z_2\,\bar{\psi}\,\mathsf{T}^a\not{A}^a\psi - m\,\delta Z_m\,\bar{\psi}\psi \\ & - \tfrac{1}{4}\delta Z_A\left(\partial^\mu A^\nu - \partial^\nu A^\mu\right)\left(\partial_\mu A_\nu - \partial_\nu A_\mu\right) - \tfrac{1}{2}\delta Z_\xi\,\xi^{-1}\left(\partial^\mu A^a_\mu\right)^2 \\ & + g\,\delta Z_3\,f^{abc}\left(\partial_\mu A^a_\nu\right)A^{b\mu}A^{c\nu} - \tfrac{1}{4}g^2\,\delta Z_4\,f^{abc}f^{ade}\,A^{b\mu}A^{c\nu}A^d_\mu A^e_\nu \\ & + \delta Z_\theta\,\partial^\mu\theta^{a*}\,\partial_\mu\theta^a + g\,\delta Z_1 f^{abc}\,\partial^\mu\theta^{a*}\theta^b A^{c\mu}.\end{aligned} \tag{6.4.4}$$

These counter terms imply the following relations between the bare and renormalised quantities:

$$A^{a\mu}_B \;=\; Z_A^{1/2}A^{a\mu}, \tag{6.4.5a}$$

$$\theta^a_B \;=\; Z_\theta^{1/2}\theta^a, \tag{6.4.5b}$$

* In the case of more than one fermion, the possibility arises for a mixing matrix between fermion states since the eigenstates of mass need not necessarily correspond to those of the interaction.

$$\psi_B = Z_\psi^{1/2}\psi, \tag{6.4.5c}$$

$$\xi_B^{-1} = Z_\xi\, Z_A^{-1}\xi^{-1}, \tag{6.4.5d}$$

$$m_B = Z_m\, Z_\psi^{-1} m, \tag{6.4.5e}$$

$$g_B = Z_1\, Z_\theta^{-1} Z_A^{-1/2} g, \tag{6.4.5f}$$

$$= Z_2\, Z_\psi^{-1} Z_A^{-1/2} g, \tag{6.4.5g}$$

$$= Z_3\, Z_A^{-3/2} g, \tag{6.4.5h}$$

$$g_B^2 = Z_4\, Z_A^{-2} g^2, \tag{6.4.5i}$$

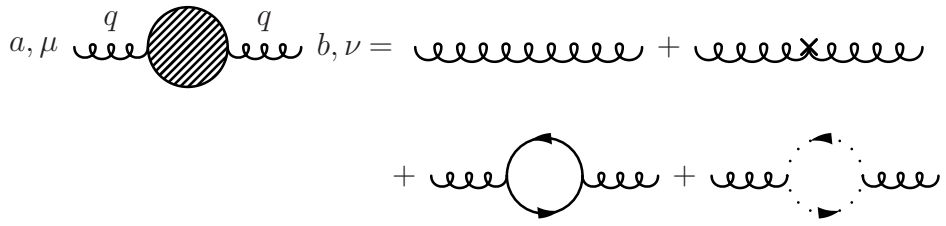
where, as usual, we have defined $Z_i := 1+\delta Z_i$ for $i = A,\theta,\psi,\xi,m,1,\ldots,4$. In the case of the last four of the previous relations gauge symmetry then implies the following relations between the renormalisation constants:

$$\frac{Z_1}{Z_\theta} = \frac{Z_2}{Z_\psi} = \frac{Z_3}{Z_A} = \frac{\sqrt{Z_4}}{\sqrt{Z_A}}. \tag{6.4.6}$$

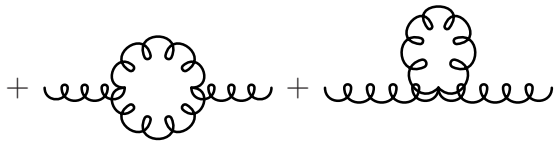
6.4.3 Calculating the renormalisation constants

6.4.4 Gauge-field propagator

The first renormalisation constant we shall examine is that of the gauge-field propagator Z_A—also known as the vacuum polarisation.* This is calculated from the one-loop correction to the two-point Green function for the gauge field:



* The expression vacuum polarisation, introduced in the early 1930's by Dirac, is used in analogy with the corresponding effect in a dielectric medium, where the presence of charges alters the electromagnetic field; here the charges are e^+e^- and YM gauge-boson pairs produced spontaneously in the vacuum.



(6.4.7)

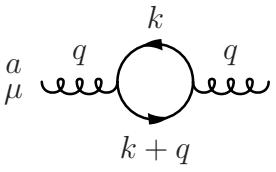
The first term is just the bare propagator, the second with the cross is the counter term, the third term contains a fermion loop. The first three are common to both Abelian and non-Abelian theories. The fourth and fifth terms are ghost and gauge-field loops while the sixth is a so-called *tadpole* diagram or gauge-field loop on a single vertex. These last three only appear in non-Abelian theories.

The sixth diagram is similar to one already calculated in a scalar theory. The result, given in Eq. (5.3.10), shows that it is proportional to the square of the boson mass (the extra gauge structure does not alter this) and since the gauge-field mass vanishes, so too does this diagram. Note that this is a case where the Pauli–Villars technique, if applied *naïvely*, would fail, since it introduces a fictitious mass, and extra *ad hoc* counter terms are required.

We shall now calculate just the ε-pole pieces for each of the other diagrams. These are precisely the pieces to be subtracted in the case of the MS scheme and (at one loop) are sufficient to determine the $\overline{\text{MS}}$ subtractions.

Fermion loop

We shall start with the fermion loop. We are only interested in the UV ε-pole terms and so may immediately set the fermion mass to zero. We have thus to calculate the following diagram:



$$= -g^2 \left(\mu^2\right)^{\varepsilon} N_{\text{f}}\, \text{Tr}[\mathsf{T}^a \mathsf{T}^b] \int \text{đ}^D k\, \text{Tr}[\gamma^\mu S_{\text{F}}(k) \gamma^\nu S_{\text{F}}(k+q)], \tag{6.4.8}$$

where the momentum flow follows the fermion arrow and the factor N_{f} is the number of fermion species; it arises because we must sum over all possible fermions running around the loop. The colour-matrix trace is given in Eq. (A.21). Since we are not interested in non-pole terms, we may now set $D=4$ (or $\varepsilon=0$) everywhere except in the integration measure $\text{đ}^D k$. The Dirac-matrix trace is then

$$\text{Tr}[\gamma^\mu \not{k} \gamma^\nu (\not{k}+\not{q})] = 4\big[k^\mu (k+q)^\nu + k^\nu (k+q)^\mu - k\cdot(k+q) g^{\mu\nu}\big]. \tag{6.4.9}$$

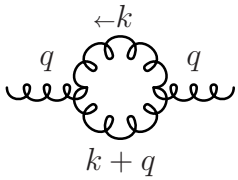
The integral is readily obtained from those already performed and, retaining just the pole term, we find

$$\text{fermion loop} = \frac{\text{i}\, g^2 N_{\text{f}} C_1\, \delta^{ab}}{32\pi^2} \frac{1}{\varepsilon} \frac{8}{3} \left(-q^2 g^{\mu\nu} + q^\mu q^\nu\right). \tag{6.4.10}$$

Note that, had we kept the fermion mass different from zero, we would have found this contribution to be suppressed for masses larger than the natural energy scale in the Green function considered. Thus, the sum over fermion species is understood to include only those energetically accessible.* Note too that it is transverse—had it not been so, then the quantisation procedure would have broken the gauge symmetry. We shall find the same behaviour for the other contributions, when combined. Indeed, one can show that this must be the case, provided the regularisation procedure respects gauge invariance (*i.e.* does not introduce spurious mass terms).

Gluon loop

The gluon contribution to the self-energy is



$$= \tfrac{1}{2} g^2 \left(\mu^2\right)^{\varepsilon} f^{acd} f^{dcb} \int \mathrm{d}\!\!^{-D} k \, I^{\mu\nu}(k,q), \tag{6.4.11}$$

where

$$I^{\mu\nu}(k,q) = \Big[-(q+2k)_\mu g_{\rho\sigma} + (k-q)_\sigma g_{\rho\mu} + (2q+k)_\rho g_{\mu\sigma}\Big] D_{\mathrm{F}}^{\sigma\tau}(q+k)\, D_{\mathrm{F}}^{\lambda\rho}(k)$$
$$\times \Big[-(q+2k)_\nu g_{\lambda\tau} + (k-q)_\tau g_{\lambda\nu} + (2q+k)_\lambda g_{\nu\tau}\Big] \tag{6.4.12}$$

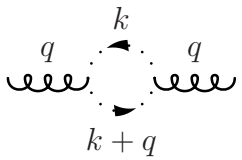
and $f^{acd} f^{dcb} = C_{\mathrm{A}} \delta^{ab}$ (see App. A.2); $C_{\mathrm{A}} = N$ for SU(N). The integral then leads to the following pole term:

$$\text{gluon loop} = \frac{\mathrm{i} g^2 C_{\mathrm{A}} \delta^{ab}}{32\pi^2} \frac{1}{\varepsilon} \left\{ -\left(\tfrac{25}{6} - \xi\right)\left(-q^2 g^{\mu\nu} + q^\mu q^\nu\right) - \left(\tfrac{3}{2} - \xi\right) q^\mu q^\nu \right\}. \tag{6.4.13}$$

* A theorem, due to Appelquist and Carazzone (1975), states that all fields heavier than the relevant energy scale effectively decouple from the theory. This may be understood in terms of time scale: to be important, intermediate virtual states should have lifetimes comparable to (*i.e.* not much smaller than) the time scale of the process under consideration and therefore their masses should be comparable to (*i.e.* not much larger than) the process energy scale

Ghost loop

The ghost contribution to the self-energy is



$$= \tfrac{1}{2} g^2 \left(\mu^2\right)^{\varepsilon} f^{acd} f^{dcb} \int \mathrm{đ}^D k \, \frac{(k+q)^\mu k^\nu}{\left[(k+q)^2 + \mathrm{i}\varepsilon\right](k^2 + \mathrm{i}\varepsilon)}. \quad (6.4.14)$$

The pole term in the integral is then

$$\text{ghost loop} \;=\; \frac{\mathrm{i} g^2 C_{\mathrm{A}} \delta^{ab}}{32\pi^2} \frac{1}{\varepsilon} \left\{ -\tfrac{1}{6} \left(-q^2 g^{\mu\nu} + q^\mu q^\nu \right) + \tfrac{1}{2} q^\mu q^\nu \right\}. \quad (6.4.15)$$

Gathering the contributions (for just the pole pieces), we can equate these to the gauge-field and gauge-parameter counter terms:

$$\begin{aligned} &\mathrm{i}\delta Z_A \left(-q^2 g^{\mu\nu} + q^\mu q^\nu \right) - \mathrm{i}\xi^{-1} \delta Z_\xi \, q^\mu q^\nu \\ &\quad = \frac{\mathrm{i} g^2}{32\pi^2} \frac{1}{\varepsilon} \left\{ \left[\left(-\tfrac{13}{3} + \xi \right) C_{\mathrm{A}} + \tfrac{8}{3} N_{\mathrm{f}} C_1 \right] \left(-q^2 g^{\mu\nu} + q^\mu q^\nu \right) - (1-\xi) C_{\mathrm{A}} \, q^\mu q^\nu \right\}, \end{aligned} \quad (6.4.16)$$

from which we obtain

$$\delta Z_A \;=\; -\frac{g^2}{32\pi^2} \frac{1}{\varepsilon} \left[\left(-\tfrac{13}{3} + \xi \right) C_{\mathrm{A}} + \tfrac{8}{3} N_{\mathrm{f}} C_1 \right], \quad (6.4.17a)$$

and

$$\delta Z_\xi \;=\; -\frac{g^2}{32\pi^2} \frac{1}{\varepsilon} \xi (1-\xi) C_{\mathrm{A}}. \quad (6.4.17b)$$

First of all, notice that choice of the Feynman gauge, or $\xi = 1$, eliminates the renormalisation of the gauge parameter itself at this order, as would too the choice $\xi = 0$. Secondly, note that the contribution of the gauge-field and related ghost loops give a combined contribution of *opposite* sign with respect to the fermion loop. This is a first signal of the antiscreening that is characteristic of non-Abelian gauge theories. That is, while the vacuum polarisation in QED tends to cancel any applied field, in pure Yang–Mills theories precisely the opposite happens. This has far-reaching and profound implications for the applicability of perturbation theory in strong interactions. It is also particularly important in relation to to grand unification, where the different *running* of the coupling constants allows them to be the same at some (very high) energy scale. We shall discuss further the effects of the vacuum polarisation shortly.

6.4.5 Ghost and fermion propagators

The same calculations can be performed for the ghost and fermion propagators. We simply list the results:

$$\delta Z_\theta = -\frac{g^2}{32\pi^2}\,\frac{1}{\varepsilon}\left(\tfrac{3}{2}-\tfrac{5}{2}\,\xi\right) N_{\rm f} C_1, \tag{6.4.17c}$$

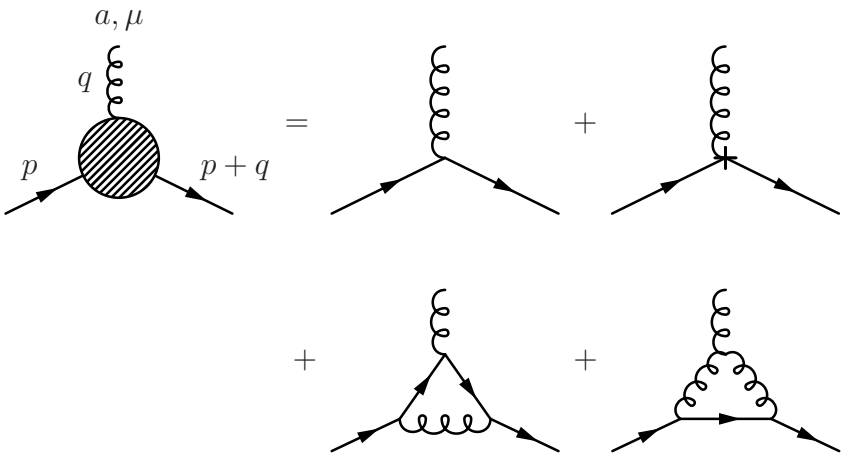
$$\delta Z_\psi = -\frac{g^2}{32\pi^2}\,\frac{1}{\varepsilon}\,2\,\xi\, C_{\rm F}, \tag{6.4.17d}$$

$$\delta Z_m = -\frac{g^2}{32\pi^2}\,\frac{1}{\varepsilon}\,2\,(3+\xi)\, C_{\rm F}. \tag{6.4.17e}$$

Notice how in the Landau gauge ($\xi = 0$) the fermion wave function is not renormalised. In fact, an important point to note is that the individual renormalisations are *not* gauge independent, but then neither are they physical quantities. It will only be when combined in physical quantities that the gauge dependence will disappear.

6.4.6 Coupling constant

There now only remains the coupling constant to renormalise. As already remarked, gauge invariance implies that it is sufficient to calculate the loop contributions to any single one of the three three-point Green functions or to the four-point gluon function. The simplest to calculate is that labelled Z_2, *i.e.* the fermion–gauge vertex. The diagrams to consider are as follows:



(6.4.18)

The first term is just the bare vertex, the second (with the cross) is the counter term and the third represents a gauge-field correction. The first three are common to both Abelian and non-Abelian theories. The fourth term involves a three-point

gauge-field vertex and therefore only appears in non-Abelian theories. The result of the calculation is

$$\delta Z_2 \;=\; -\frac{g^2}{32\pi^2}\,\frac{1}{\varepsilon}\left\{2\,C_{\rm F}\xi + C_{\rm A}\left(\tfrac{3}{2}+\tfrac{1}{2}\,\xi\right)\right\}. \tag{6.4.19}$$

Once again, we see that the answer is gauge dependent. Note too the presence of both Casimir constants, $C_{\rm F}$ and $C_{\rm A}$. That is, there are both fermion- and gauge-like contributions here (it is easy to see that all vertices have both fermion- and gauge-loop corrections).

6.5 Bibliography

Abers, E.S. and Lee, B.W. (1973), *Phys. Rep.* **9**, 1.

Appelquist, T. and Carazzone, J. (1975), *Phys. Rev.* **D11**, 2856.

Bjorken, J.D. and Drell, S.D. (1965), *Relativistic Quantum Fields* (McGraw–Hill).

Fadde'ev, L.D. and Popov, V.N. (1967), *Phys. Lett.* **B25**, 29.

Yang, C.-N. and Mills, R.L. (1954), *Phys. Rev.* **96**, 191.

Chapter 7

Advanced Topics

7.1 The renormalisation group

7.1.1 The renormalisation group equation

A fundamental aspect of renormalisation is the inevitable introduction of a new mass scale. In the previous chapter we saw how in dimensional regularisation (for any subtraction scheme) we are forced to introduce a parameter μ as a dimensional object to maintain the correct dimensionality of the action. After the renormalisation procedure is performed, dependence on this parameter *remains* in the Green functions of the theory. Evidently though, as it is not a physical parameter and is indeed *arbitrary*, no *physical* quantity may depend on μ.

It is however interesting to examine the dependence of the renormalised Green functions on μ. Since the bare Green functions have no knowledge of μ, the dependence in the renormalised functions is encoded in the renormalisation constants. We shall denote a generic n-point renormalised Green function, having n_A external gauge legs and n_ψ fermion legs ($n = n_A + n_\psi$), by

$$\Gamma^{(n)}(\{p_i\}, g, \xi, m, \mu), \tag{7.1.1}$$

where $\{p_i\}$ $(i = 1, \ldots, n)$ are the external momenta, g and ξ remain dimensionless parameters and m, the renormalised fermion mass, is the only dimensional parameter in the bare theory. We have already shown, see Eq. (5.5.29), that the relation between renormalised and bare Green functions is

$$\Gamma^{(n)}(\{p_i\}, g, \xi, m, \mu) \;=\; Z_A^{\frac{1}{2}n_A}\, Z_\psi^{\frac{1}{2}n_\psi}\, \Gamma_B^{(n)}(\{p_i\}, g_B, \xi_B, m_B), \tag{7.1.2}$$

where Z_A and Z_ψ are the gauge- and fermion-field renormalisation constants, defined in Eqs. (6.4.5a & c).

Let us now take the logarithmic derivative of (7.1.2) with respect to μ. First though, we move all the factors involving Z_A and Z_ψ to the left-hand side, thus leaving a right-hand side that is entirely independent of μ. The important point to realise is that $\Gamma^{(n)}$ contains μ dependence not only explicitly, but also implicitly through the renormalised parameters g, ξ and m. We may thus write

$$\left[\mu\frac{\partial}{\partial\mu}+\beta\frac{\partial}{\partial g}+\gamma_\xi\frac{\partial}{\partial\xi}+\gamma_m\, m\frac{\partial}{\partial m}-n_A\gamma_A-n_\psi\gamma_\psi\right]\Gamma^{(n)}(\{p_i\},g,\xi,m,\mu)\;=\;0, \tag{7.1.3}$$

where the dimensionless coefficients introduced are: the so-called *β-function*,

$$\beta\;:=\;\mu\frac{\partial g(\mu)}{\partial\mu}, \tag{7.1.4a}$$

and the various so-called *anomalous dimensions* of the theory,

$$\gamma_\xi\;:=\;\mu\frac{\partial\xi(\mu)}{\partial\mu}, \tag{7.1.4b}$$

$$\gamma_m\;:=\;\frac{\mu}{m(\mu)}\frac{\partial m(\mu)}{\partial\mu}, \tag{7.1.4c}$$

$$\gamma_A\;:=\;\frac{\mu}{Z_A^{1/2}(\mu)}\frac{\partial Z_A^{1/2}(\mu)}{\partial\mu}, \tag{7.1.4d}$$

$$\gamma_\psi\;:=\;\frac{\mu}{Z_\psi^{1/2}(\mu)}\frac{\partial Z_\psi^{1/2}(\mu)}{\partial\mu}, \tag{7.1.4e}$$

Naturally, for each parameter differentiated in the above, it is understood that the others are held constant. Equation (7.1.3) is known as the renormalisation group equation (RGE); it is based on the almost trivial observation that the *bare* Green functions cannot depend on μ, but sill embodies the scale dependence of the Green functions and all parameters of the theory.

The coefficients (the β-function and anomalous dimensions γ_i) evidently have perturbative expansions and therefore naturally depend on g. However, in a subtraction scheme with no external scale dependence, such as MS or $\overline{\text{MS}}$, since they therefore do not depend on μ (by definition, they only contain ε poles and at most finite constants), they cannot depend on m either. In general schemes, this is not the case (*cf.* the MOM scheme), but it is an important simplification to be exploited.

In any case, we have already performed the necessary calculations to extract the one-loop coefficients. Indeed, recalling that the ε pole is always accompanied

by a factor $\left(\mu^2\right)^{\varepsilon}$, we see that the β-function and anomalous dimensions γ_i may be read off directly from the renormalisation coefficients just calculated. We thus find:

$$\beta \;=\; -\frac{g^3}{16\pi^2}\left[\tfrac{11}{3}\,C_{\rm A} - \tfrac{4}{3}\,N_{\rm f}\,C_1\right], \tag{7.1.5a}$$

$$\gamma_\xi \;=\; \frac{g^2}{16\pi^2}\left[\left(\tfrac{13}{3} - \xi^2\right)C_{\rm A} - \tfrac{8}{3}N_{\rm f}\,C_1\right]\xi, \tag{7.1.5b}$$

$$\gamma_m \;=\; -\frac{g^2}{16\pi^2}\;6\,C_{\rm F}, \tag{7.1.5c}$$

$$\gamma_A \;=\; -\frac{g^2}{16\pi^2}\left[\left(\tfrac{13}{3} - \tfrac{1}{2}\xi\right)C_{\rm A} - \tfrac{4}{3}N_{\rm f}\,C_1\right], \tag{7.1.5d}$$

$$\gamma_\psi \;=\; \frac{g^2}{16\pi^2}\;\xi\,C_{\rm F}, \tag{7.1.5e}$$

where we have considered the simple case of $N_{\rm f}$ replicas of a fermion field all belonging to the same representation; should there be fermions in different representations, then the term $N_{\rm f}C_1$ would be substituted by a weighted sum over the relevant Casimir constants. The first important point to note is the sign of the β-function. The contributions are *negative* for non-Abelian gauge fields (an Abelian gauge field makes no contribution) while they are *positive* for fermions. Note too that while the various anomalous dimensions start at $O(g^2)$ the first non-zero term in the β-function is $O(g^3)$.

Exercise 7.1.1. *Using the results obtained earlier for the renormalisation constants Z_i ($i = \psi$, $A^{a\mu}$, g), calculate the β-function given above.*

7.1.2 Asymptotic freedom

The β-function describes the scale variation of the coupling constant. Let us write to one-loop order

$$\beta \;=\; \mu\,\frac{\partial g(\mu)}{\partial \mu} \;\simeq\; -\beta^{(0)}\,g^3(\mu). \tag{7.1.6}$$

The coefficient $\beta^{(0)}$ is positive (negative) for non-Abelian (Abelian) theories. This equation can be solved rather easily by changing variable to $t := \ln(\mu/\mu_0)$, whence

$$\frac{\partial g(t)}{\partial t} \;\simeq\; -\beta^{(0)}\,g^3(t) \tag{7.1.7}$$

and therefore

$$g^2(t) \;\simeq\; \frac{1}{\beta^{(0)}\,t} \;=\; \frac{1}{\beta^{(0)}\,\ln(\mu/\mu_0)}, \tag{7.1.8}$$

where the constant of integration has been absorbed into the definition of μ_0. Note that the sign of $\beta^{(0)}$ determines the region of validity of the above solution: for $\beta^{(0)} > 0$ ($\beta^{(0)} < 0$) we must take $\mu > \mu_0$ ($\mu < \mu_0$).

Indeed, one thus sees that for $\beta^{(0)} > 0$, $g(t)$ *decreases* as t (or μ) grows. This gives rise to the notion of *asymptotic freedom*. That is, the theory tends to a free-field theory at high energy. Conversely, as $t \to 0$ or $\mu \to \mu_0$ the coupling explodes and at some point perturbation theory will therefore no longer be valid. This point is known as the Landau pole (Landau and Pomeranchuk, 1955).* The regions of validity for the two types of theory are then low-energy (below the Landau pole) for QED – and any other Abelian theory – but high energy (sufficiently larger than μ_0 in the case of YM or non-Abelian theories.

Now, in QED the Landau pole occurs at very high energy indeed, well beyond the Planck scale, and is therefore only of academic interest since quantum gravity would presumably make itself felt first. However, the coupling remains perfectly finite (and small) as we go to lower and lower scales. For a Yang–Mills theory the coupling tends to zero for growing energy and thus perturbation theory is applicable in the high-energy regime. However, at low energies the perturbative coupling grows without bound thus invalidating the perturbative approach at some point. Note, moreover, that for a Yang–Mills theory the one-loop β-function changes sign (*i.e.* the theory becomes QED-like) for $\frac{11}{3} C_{\mathrm{A}} = \frac{4}{3} N_{\mathrm{f}} C_1$, which occurs in SU(3) for a number of fermion species $N_{\mathrm{f}} = 17$. Presumably then, one can safely say that QCD is asymptotically free—at least in the energy range so far probed.† Indeed, the β-function in QCD is now known to four-loop order (van Ritbergen *et al.*, 1997; Czakon, 2005) and all coefficients are positive.

Exercise 7.1.2. *Solve the β-function equation for the QED case and show, indeed, that the singular point for the QED coupling constant lies at energies well above the Planck scale.*

7.1.3 Scaling behaviour

To see how the RGE (7.1.3) describes scale dependence, consider the behaviour of a generic n-point Green function when the external momenta $p_1, \ldots, p_n$ are rescaled by some common dimensionless factor λ. In D-dimensional space–time the Green function has energy dimensions

$$d_\Gamma \;=\; D - \tfrac{1}{2}(D-2)n_A - \tfrac{1}{2}(D-1)n_\psi. \tag{7.1.9}$$

* This asymptotic behaviour generates a second pole in the propagator, located at very large p^2, which implies the existence of another state in the spectrum, known as the Landau ghost.

† There is reasonable experimental indications that the number of fermion generations is limited to three and that there are therefore only six quark flavours.

Now, since $\Gamma^{(n)}$ is homogeneous and of degree d_Γ in the parameters $p_1, \ldots, p_n$, m and μ, we have

$$\left[\lambda\frac{\partial}{\partial\lambda} + m\,\frac{\partial}{\partial m} + \mu\,\frac{\partial}{\partial\mu}\right]\Gamma^{(n)}(\{\lambda p_i\}, g, \xi, m, \mu) \;=\; d_\Gamma\Gamma^{(n)}\big(\{\lambda p_i\}, g, \xi, m, \mu\big). \tag{7.1.10}$$

We can now use the RGE (7.1.3) to eliminate the term in μ:

$$\left[-\lambda\frac{\partial}{\partial\lambda} + \beta\,\frac{\partial}{\partial g} + \gamma_\xi\,\frac{\partial}{\partial\xi} - (1-\gamma_m)\,m\,\frac{\partial}{\partial m} - n_A\gamma_A - n_\psi\gamma_\psi + d_\Gamma\right] \times\,\Gamma^{(n)}\big(\{\lambda p_i\}, g, \xi, m, \mu\big) \;=\; 0. \tag{7.1.11}$$

This equation may be solved with aid of a so-called *running* mass $\widetilde{m}(\mu)$ and gauge parameter $\xi(\mu)$, defined in a similar manner to the coupling $g(\mu)$:

$$\frac{\partial\xi(t)}{\partial t} \;=\; \gamma_\xi\big(g(t),\xi(t)\big)\,, \tag{7.1.12a}$$

$$\frac{\partial\widetilde{m}(t)}{\partial t} \;=\; -\Big[1-\gamma_m\big(g(t),\xi(t)\big)\Big]\,\widetilde{m}(t). \tag{7.1.12b}$$

Note that the running mass $\widetilde{m}(t)$ so defined is *not* now the physical mass. The solution may then be expressed as

$$\begin{aligned}\Gamma^{(n)}&\big(\{\lambda p_i\}, g(\mu), \xi(\mu), m(\mu), \mu\big)\\ &=\;\lambda^{d_\Gamma}\,\exp\left\{-\int_1^\lambda\frac{\mathrm{d}\lambda'}{\lambda'}\Big[n_A\,\gamma_A\big(g(\lambda'\mu),\xi(\lambda'\mu)\big) + n_\psi\,\gamma_\psi\big(g(\lambda'\mu),\xi(\lambda'\mu)\big)\Big]\right\}\\ &\qquad\times\,\Gamma^{(n)}\big(\{p_i\}, g(\lambda\mu), \xi(\lambda\mu), \widetilde{m}(\lambda\mu), \mu\big).\end{aligned} \tag{7.1.13}$$

Exercise 7.1.3. *Verify by explicit substitution that this is indeed a solution to the previous equation.*

Now, we see that, with all the parameters running, the expressions are likely to become rather complicated. A useful simplification is obtained by setting $\xi(\mu)=0$ (the Landau gauge); from the equation governing its scale dependence, we see that if ξ vanish for some scale then it vanishes at *all* scales, *i.e.* $\gamma_\xi=0$. Moreover, such a choice also sets $\gamma_\psi=0$. Thus, we shall continue assuming this choice of gauge and omit the parameter from all expressions. Finally, we see from the differential equations and the anomalous dimensions that $\widetilde{m}(\lambda\mu)$, as defined here, decreases as a *power* of λ while g only falls *logarithmically*. To a reasonable approximation,

we may thus also set $\widetilde{m}=0$. The Green function is then approximately

$$\Gamma^{(n)}\big(\{\lambda p_i\}, g(\mu), \mu\big) \;\approx\; \lambda^{d_\Gamma}\, \exp\!\left\{-n_A \int_1^\lambda \frac{\mathrm{d}\lambda'}{\lambda'}\, \gamma_A\big(g(\lambda'\mu)\big)\right\} \Gamma^{(n)}\big(\{p_i\}, g(\lambda\mu), \mu\big). \tag{7.1.14}$$

The integral on the right-hand side may be re-expressed as follows:

$$\begin{aligned}\int_1^\lambda \frac{\mathrm{d}\lambda'}{\lambda'}\, \gamma_A\big(g(\lambda'\mu)\big) \;&=\; \int_{g(\mu)}^{g(\lambda\mu)} \mathrm{d}g(\lambda'\mu)\, \frac{1}{\lambda'} \frac{\partial\lambda'}{\partial g(\lambda'\mu)}\, \gamma_A\big(g(\lambda'\mu)\big) \\ &=\; \int_{g(\mu)}^{g(\lambda\mu)} \mathrm{d}g\, \frac{\gamma_A(g)}{\beta(g)}. \end{aligned} \tag{7.1.15}$$

For an explicit solution, we may insert the one-loop expression for γ_A into the last integral:

$$\gamma_A \;=\; -\gamma_A^{(0)}\, g^2(\lambda\mu), \tag{7.1.16}$$

which, together with the one-loop β-function from Eq. (7.1.6), gives

$$\begin{aligned}\int_{g(\mu)}^{g(\lambda\mu)} \mathrm{d}g\, \frac{\gamma_A(g)}{\beta(g)} \;&=\; \frac{\gamma_A^{(0)}}{\beta^{(0)}} \int_{g(\mu)}^{g(\lambda\mu)} \frac{\mathrm{d}g}{g} \\ &=\; \frac{\gamma_A^{(0)}}{\beta^{(0)}}\, \ln\!\left[\frac{g(\lambda\mu)}{g(\mu)}\right]. \end{aligned} \tag{7.1.17}$$

The Green function we are considering is then

$$\Gamma^{(n)}\big(\{\lambda p_i\}, g(\mu), \mu\big) \;\approx\; \lambda^{d_\Gamma} \left[\frac{g(\lambda\mu)}{g(\mu)}\right]^{-n_A \frac{\gamma_A^{(0)}}{\beta^{(0)}}} \Gamma^{(n)}\big(\{p_i\}, g(\lambda\mu), \mu\big), \tag{7.1.18}$$

where now all quantities should be considered in the one-loop approximation. Here we see the full power of the RGE and the significance of asymptotic freedom. The right-hand side has factored the principal scale dependence out of the Green function. In an asymptotically free theory at high energies (large λ) this last may be expanded around the point $g(\lambda\mu)=0$ and is therefore described in the free-field approximation.

The significance of this is that a physical quantity, a cross-section say, may be calculated or even *measured experimentally* at some relatively low energy scale μ and then *predicted* at a higher energy scale $\lambda\mu$ simply through multiplication by

something similar to the above pre-factor

$$\lambda^{d_\Gamma} \left[\frac{g(\lambda\mu)}{g(\mu)}\right]^{-\frac{\gamma^{(0)}}{\beta^{(0)}}}, \tag{7.1.19}$$

where the first factor is given by the *naïve* dimensionality of the quantity and the remaining *correction* is determined by the relevant anomalous dimension γ. If necessary, with a little complication, one can systematically reinstate the mass terms and consider other gauges. Note moreover that, provided the coupling constant is small, the approach is also useful for non-asymptotically free theories. The point is that the anomalous-dimension factor gathers all the dominant corrections, *i.e.* those that contain possibly large logarithms, greatly simplifying the expressions for various cross-sections as functions of energy scale. Indeed, nowadays such techniques are often even used in the treatment of radiative corrections to QED in, for example, high-energy e^+e^- and ep colliders.

7.2 spontaneous symmetry breaking

In this section we shall briefly describe the mechanism by which it is possible to retain a gauge symmetry (with all the consequent benefits) while allowing the physical gauge bosons to acquire significant masses. The concept of a broken symmetry, with the consequent generation of massless states was first discussed by Goldstone (1961).

Spontaneously broken symmetries (*i.e.* symmetries of the Lagrangian not respected by the vacuum or ground state) had already been discussed by Nambu (1960) in the context of superconductivity. That they could circumvent the Goldstone theorem and thus avoid the existence of massless states was proposed by Anderson (1963). Nambu and Jona-Lasinio (1961) then successively adapted the ideas to provide an effective theory of the nucleon, in which the pions emerge as so-called "would-be" Goldstone bosons. The spontaneous breaking of a gauge symmetry within the context of quantum field theory and resultant mass generation, known as the Higgs (or Higgs–Kibble) mechanism, was first discussed by Higgs (1964), Englert and Brout (1964) and Guralnik, Hagen and Kibble (1964). Though not covered in these notes, the the developments described here culminated in the unified theory of the electroweak interactions due to Glashow (1961), Salam *et al.* (1964) and Weinberg (1967).

7.2.1 A real scalar field

The simplest case is the scalar φ^4 theory with a mass term of the "wrong" sign:*

$$\mathcal{L} \;=\; \tfrac{1}{2}(\partial^\mu\varphi)(\partial_\mu\varphi) + \tfrac{1}{2}\,\mu^2\varphi^2 - \tfrac{1}{4!}\,\lambda\varphi^4. \tag{7.2.1}$$

The potential

$$V(\varphi) \;=\; -\tfrac{1}{2}\,\mu^2\varphi^2 + \tfrac{1}{4!}\,\lambda\varphi^4 \tag{7.2.2}$$

has the form shown in Fig. 7.1 and is evidently symmetric under the discrete transformation $\varphi \to -\varphi$. The point $\varphi = 0$ is no longer the minimum and there are

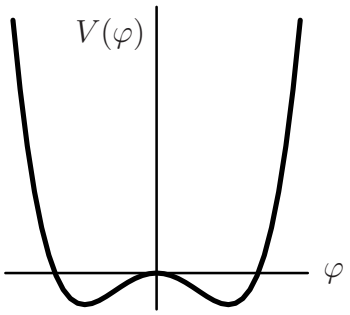


Figure 7.1: The form of the potential in the case of a negative mass term for a single real scalar field.

now two equivalent minima at $\varphi = \pm\sqrt{6\mu^2/\lambda}$. The vacuum or ground state of the theory may be one or the other, but not both, and will therefore break the original symmetry.

As found in Ex. 3.3.3, the Green functions are to be calculated for perturbations around the minimum. We must therefore shift to $\varphi' := \varphi - \upsilon$, where for the purposes of example we shall take $\upsilon = \sqrt{6\mu^2/\lambda}$. The Lagrangian then takes the form

$$\mathcal{L} \;=\; \tfrac{1}{2}(\partial^\mu\varphi')(\partial_\mu\varphi') - \mu^2\varphi'^2 - \tfrac{1}{3!}\,\lambda\varphi'^3 - \tfrac{1}{4!}\,\lambda\varphi'^4, \tag{7.2.3}$$

where an irrelevant constant term has been eliminated. Thus, the true physical state of the theory has (bare) mass $m = \sqrt{2}\mu$ and a cubic interaction has also appeared.

* Note that it is the term that has the "wrong" sign and not the mass parameter itself.

7.2.2 A complex scalar field

Now, the case of a continuous symmetry is rather more interesting. Consider a *complex* field φ with the following Lagrangian:

$$\mathcal{L} \;=\; (\partial^\mu \varphi^*)(\partial_\mu \varphi) + \mu^2 \varphi^* \varphi - \tfrac{1}{2}\lambda(\varphi^* \varphi)^2. \tag{7.2.4}$$

Note the choice of different numerical coefficients with respect to the real case. The potential

$$V(\varphi, \varphi^*) \;=\; -\mu^2 \varphi^* \varphi + \tfrac{1}{2}\lambda(\varphi^* \varphi)^2 \tag{7.2.5}$$

has the so-called Mexican-hat form shown in Fig. 7.2 and is evidently symmetric under the *global* U(1) gauge transformation $\varphi \to \mathrm{e}^{\mathrm{i}\phi}\varphi$. The possible ground or

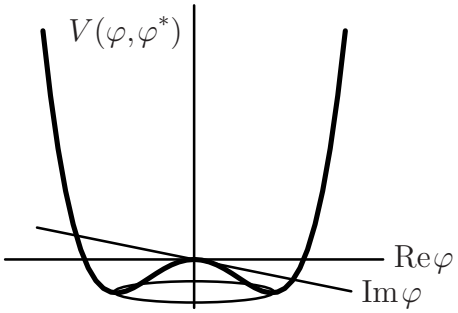


Figure 7.2: The so-called Mexican-hat form of the potential in the case of a complex scalar field with a negative mass-squared term.

vacuum states now belong to a continuum, corresponding to the variable ϕ in such a transformation. Without loss of generality, we may choose the ground state to be $\varphi = \sqrt{\mu^2/\lambda}$ and again make the shift to $\varphi' := \varphi - \upsilon$, with $\upsilon = \sqrt{\mu^2/\lambda}$ purely real. It will now be convenient to re-express the fields in terms of two real fields $\varphi_{1,2}$:

$$\varphi' \;=\; \tfrac{1}{\sqrt{2}}[\varphi_1 + \mathrm{i}\,\varphi_2]. \tag{7.2.6}$$

The Lagrangian then takes the form

$$\begin{aligned}\mathcal{L} \;=\; & \tfrac{1}{2}\left[(\partial^\mu \varphi_1)(\partial_\mu \varphi_1) - 2\mu^2 \varphi_1^2\right] + \tfrac{1}{2}(\partial^\mu \varphi_2)(\partial_\mu \varphi_2) \\ & - \lambda \upsilon \varphi_1(\varphi_1^2 + \varphi_2^2) - \tfrac{1}{4}\lambda(\varphi_1^2 + \varphi_2^2)^2.\end{aligned} \tag{7.2.7}$$

In this case we see that, together with interaction terms of various degrees, there is just a single non-zero mass term (for φ_1). That is, one state (φ_2) has remained massless. It is easy to see why: the original continuous symmetry implies that there is always a direction in which the derivative of the field is zero. This is the essence of the Goldstone theorem, which we shall now prove.

7.2.3 A non-Abelian global symmetry

Let us now generalise to a p-dimensional representation of some non-Abelian symmetry group G. Consider a real representation:

$$\boldsymbol{\varphi}(x) \;=\; \begin{pmatrix} \varphi_1(x) \\ \varphi_2(x) \\ \vdots \\ \varphi_p(x) \end{pmatrix} \tag{7.2.8}$$

The infinitesimal global transformation

$$\boldsymbol{\varphi}(x) \;\to\; \boldsymbol{\varphi}'(x) \;=\; \boldsymbol{\varphi}(x) + \delta\boldsymbol{\varphi}(x), \tag{7.2.9}$$

with

$$\delta\varphi_i(x) \;=\; -\,\mathrm{i}\,g\Lambda^a\,\mathsf{T}^a_{ij}\,\varphi_j(x), \tag{7.2.10}$$

then leaves the Lagrangian invariant. It follows from this invariance that there is a conserved Noether current

$$j^{a\mu}(x) \;=\; \mathrm{i}\,\pi^\mu_i(x)\,\mathsf{T}^a_{ij}\,\varphi_j(x), \tag{7.2.11}$$

where

$$\pi^\mu_i \;:=\; \frac{\delta\mathcal{L}}{\delta\partial_\mu\varphi_i}. \tag{7.2.12}$$

From the Euler–Lagrange equations, we have

$$\partial_\mu\pi^\mu_i \;=\; \frac{\partial\mathcal{L}}{\partial\varphi_i}. \tag{7.2.13}$$

Together with current conservation, this implies

$$\frac{\partial\mathcal{L}}{\partial\varphi_i}\,\mathsf{T}^a_{ij}\,\varphi_j + \pi^\mu_i\,\mathsf{T}^a_{ij}\,\partial^\mu\varphi_j \;=\; 0. \tag{7.2.14}$$

For a theory with Lagrangian of the form

$$\mathcal{L} \;=\; \tfrac{1}{2}(\partial^\mu\varphi_i)(\partial_\mu\varphi_i) - V(\varphi_i), \tag{7.2.15}$$

we have

$$\pi^\mu_i \;=\; \partial^\mu\varphi_i \tag{7.2.16}$$

and

$$\frac{\partial\mathcal{L}}{\partial\varphi_i} \;=\; -\frac{\partial V}{\partial\varphi_i}. \tag{7.2.17}$$

From (7.2.16) we see that the second term in Eq. (7.2.14) must vanish since the matrices T^a are antisymmetric. Using (7.2.17), we thus obtain

$$\frac{\partial V}{\partial \varphi_i}\,\mathsf{T}^a_{ij}\,\varphi_j \;=\; 0. \tag{7.2.18}$$

Now, as we have seen earlier, the field theory is to be considered as perturbed around the *minimum* of the potential and the particle masses are then determined by the second derivatives at the minimum. We shall thus write

$$\langle 0|\boldsymbol{\varphi}(x)|0\rangle \;=\; \boldsymbol{v}, \tag{7.2.19}$$

for which

$$\left.\frac{\partial V}{\partial \varphi_i}\right|_{\boldsymbol{\varphi}=\boldsymbol{v}} = 0. \tag{7.2.20}$$

Now the vacuum expectation value v breaks the symmetry and so

$$\left(1 - \mathrm{i}\,g\Lambda^a\,\mathsf{T}^a\right)\boldsymbol{v} \;\neq\; \boldsymbol{v}, \tag{7.2.21}$$

for some Λ^a and therefore for some a we have

$$\mathsf{T}^a\,\boldsymbol{v} \;\neq\; 0. \tag{7.2.22}$$

We can now make a Taylor expansion of $V(\varphi)$ in $\boldsymbol{\varphi}'$:

$$\boldsymbol{\varphi}' \;:=\; \boldsymbol{\varphi} - \boldsymbol{v}, \tag{7.2.23}$$

having, by definition, zero vacuum expectation value

$$\langle 0|\boldsymbol{\varphi}'(x)|0\rangle \;=\; 0. \tag{7.2.24}$$

Therefore, concentrating on the quadratic terms, which describe the masses, we find

$$V(\boldsymbol{\varphi}') \;=\; V(\boldsymbol{v}) + \varphi'_i\varphi'_j\,\left.\frac{\partial^2 V}{\partial\varphi'_i\,\partial\varphi'_j}\right|_{\varphi'_i=0} + O(\boldsymbol{\varphi}'^3). \tag{7.2.25}$$

The constant term $V(\boldsymbol{v})$ is irrelevant and may be omitted.

We can thus define a mass-squared matrix M^2, whose eigenvalues will be the (bare) masses of the physical states:

$$(\mathsf{M}^2)_{ij} \;:=\; \left.\frac{\partial^2 V}{\partial\varphi'_i\,\partial\varphi'_j}\right|_{\varphi'_i=0}. \tag{7.2.26}$$

Differentiating Eq. (7.2.18) with respect to φ'_i and setting $\boldsymbol{\varphi}' = 0$ leads to

$$\begin{aligned} 0 \;&=\; \frac{\partial}{\partial\varphi'_i}\left(\frac{\partial V}{\partial\varphi'_j}\,\mathsf{T}^a_{jk}\left[\varphi'_k + \upsilon_k\right]\right)\Bigg|_{\varphi'_j=0} \\ &=\; \frac{\partial^2 V}{\partial\varphi'_i\,\partial\varphi'_j}\Bigg|_{\varphi'_i=0} \mathsf{T}^a_{jk}\upsilon_k \;=\; (\mathsf{M}^2)_{ij}\,\mathsf{T}^a_{jk}\upsilon_k. \end{aligned} \tag{7.2.27}$$

Since M^2 may be diagonalised in this equation, there must be at least one zero eigenvalue. Indeed, if m of the n generators are left unbroken and the vectors $\mathsf{T}^a\boldsymbol{\upsilon}$ span an $(n-m)$-dimensional subspace, then corresponding to each of the $n-m$ broken generators (*i.e.* for which $\mathsf{T}^a\boldsymbol{\upsilon}\neq 0$) there exists a massless boson. This is the essence of the Goldstone theorem (Goldstone, 1961; Goldstone, Salam and Weinberg, 1962).

A formal proof may be given as follows. Construct the matrix

$$A^{ab} \;:=\; (\mathsf{T}^a\boldsymbol{\upsilon})^*\cdot(\mathsf{T}^b\boldsymbol{\upsilon}). \tag{7.2.28}$$

Then, since the T^a are Hermitian, we have

$$A^{ab} \;=\; (\boldsymbol{\upsilon})^*\cdot(\mathsf{T}^a\mathsf{T}^b\boldsymbol{\upsilon}) \;=\; (\boldsymbol{\upsilon})\cdot(\mathsf{T}^a\mathsf{T}^b\boldsymbol{\upsilon}) \tag{7.2.29}$$

and therefore

$$A^{ab} - A^{ba} \;=\; (\boldsymbol{\upsilon})\cdot([\mathsf{T}^a,\mathsf{T}^b]\boldsymbol{\upsilon}) \;=\; \mathrm{i}\,f^{abc}\,\upsilon_i\mathsf{T}^c_{ij}\upsilon_j \;=\; 0. \tag{7.2.30}$$

Where the last equality follows from the antisymmetry of the T. Therefore, A^{ab} is symmetric and can be diagonalised by some unitary matrix U. Now, define $\tilde{A}^{ab}$ as the $(n-m)\times(n-m)$ matrix restricted to those values of a and b for which $\mathsf{T}^a\boldsymbol{\upsilon}\neq 0$. We diagonalise $\tilde{A}^{ab}$ to

$$\tilde{A}'^{ab}_{\mathrm{diag}} \;=\; \left(U\tilde{A}U^{\mathsf{T}}\right)^{ab} \;=\; (U^{ac}\mathsf{T}^c\boldsymbol{\upsilon})^*\cdot(U^{bd}\mathsf{T}^d\boldsymbol{\upsilon}), \tag{7.2.31}$$

where, by construction, $U^{ac}\mathsf{T}^c$ belongs to the broken sector and therefore does not annihilate $\boldsymbol{\upsilon}$. All the diagonal elements of $\tilde{A}'$ are thus non-zero and, from the previous equation, they are positive. Finally, since they are independent, they indeed span the $(n-m)$-dimensional subspace not annihilated by $U^{ac}\mathsf{T}^c$ and so there must be precisely $(n-m)$ zero eigenvalues or massless bosons.

7.2.4 The Higgs–Kibble mechanism

Let us now examine the case of spontaneous symmetry breaking in the presence of a *local* gauge symmetry (Higgs, 1964; Englert and Brout, 1964; Guralnik, Hagen and Kibble,

1964). We shall thus find an exception to the Goldstone theorem, in that the massless bosons do not appear as true individual states of the theory, but are absorbed into the gauge boson fields to provide the third (longitudinal) components for the spin-one states, which thus become massive. The important point to realise though is that the underlying gauge symmetry survives and the theory is thus guaranteed to be renormalisable in the usual way.

Consider the simplest example of a single, charged, scalar field, minimally coupled to a U(1), Abelian, gauge field. The Lagrangian is taken to be

$$\mathcal{L}(A^\mu,\varphi) \;=\; (D^\mu\varphi^*)(D_\mu\varphi) + \mu^2\varphi^*\varphi - \lambda(\varphi^*\varphi)^2 - \tfrac{1}{4}F^{\mu\nu}F_{\mu\nu}, \tag{7.2.32}$$

where the covariant derivative D^μ is defined as

$$D^\mu\varphi \;:=\; \big(\partial^\mu + \mathrm{i}\,gA^\mu(x)\big)\,\varphi \tag{7.2.33a}$$

and

$$D^\mu\varphi^* \;:=\; \big(\partial^\mu - \mathrm{i}\,gA^\mu(x)\big)\,\varphi^*. \tag{7.2.33b}$$

The field-strength tensor $F^{\mu\nu}$ is as defined in Eq. (2.2.10). Recall that since we are considering an Abelian theory, although we must introduce a gauge-fixing term, there is no need for Fadde'ev–Popov ghost fields.

As before, we rewrite the field φ in terms of its real and imaginary parts $\varphi_{1,2}$, but (without loss of generality) immediately shift the real part φ_1:

$$\varphi \;=\; \tfrac{1}{\sqrt{2}}\,(\upsilon + \varphi_1 + \mathrm{i}\varphi_2), \tag{7.2.34}$$

where

$$\upsilon \;=\; \sqrt{\mu^2/\lambda}. \tag{7.2.35}$$

In the absence of the gauge field and related coupling, the field φ_2 would be the massless Goldstone boson associated with spontaneous breaking of a global U(1) symmetry. However, the locality of the gauge symmetry causes mixing between the φ_2 and A^μ fields. To see this, let us expand the Lagrangian in terms of $\varphi_{1,2}$:

$$\begin{aligned}\mathcal{L}(A^\mu,\varphi) \;=\; &-\tfrac{1}{4}F^{\mu\nu}F_{\mu\nu} + \tfrac{1}{2}g^2\upsilon^2A^\mu A_\mu + \tfrac{1}{2}\left[\partial^\mu\varphi_1\partial_\mu\varphi_1 - 2\mu^2\varphi_1^2\right] + \tfrac{1}{2}\partial^\mu\varphi_2\partial_\mu\varphi_2\\ &+ g\upsilon A^\mu\partial_\mu\varphi_2 + g\upsilon A^\mu(\varphi_1\overset{\leftrightarrow}{\partial}_\mu\varphi_2) + g^2\upsilon A^\mu A_\mu\varphi_1 + \tfrac{1}{2}g^2A^\mu A_\mu(\varphi_1^2+\varphi_2^2)\\ &- \lambda\upsilon\varphi_1(\varphi_1^2+\varphi_2^2) - \tfrac{1}{4}\lambda(\varphi_1^2+\varphi_2^2)^2,\end{aligned} \tag{7.2.36}$$

where, recall, $\overset{\leftrightarrow}{\partial} \equiv \overset{\rightarrow}{\partial} - \overset{\leftarrow}{\partial}$ and we have suppressed irrelevant constant terms. The higher-order (cubic and quartic) terms describe various interactions and are of interest for the full phenomenology; however, we are only interested here in the possible masses. At first sight, we seem to have the usual boson (φ_1) of mass $m=\sqrt{2}\mu$, a massless Goldstone boson (φ_2) and a massive gauge field (since a term

$A^\mu A_\mu$ has appeared). However, the presence of the term $gvA_\mu\partial^\mu\varphi_2$ directly mixes the φ_2 and A^μ fields and thus requires more care, as neither field can now represent a true asymptotic state of the theory.

First of all, notice that the three quadratic-order terms involving φ_2 and A^μ may be combined as follows

$$\tfrac{1}{2}g^2v^2A^\mu A_\mu+\tfrac{1}{2}\partial^\mu\varphi_2\partial_\mu\varphi_2+gvA^\mu\partial_\mu\varphi_2 \;=\; \tfrac{1}{2}(gvA^\mu+\partial^\mu\varphi_2)(gvA_\mu+\partial_\mu\varphi_2), \quad (7.2.37)$$

which is suggestive of the following gauge transformation:

$$A^\mu \;\to\; A'^\mu \;=\; A^\mu+\tfrac{1}{gv}\,\partial^\mu\varphi_2 \quad (7.2.38)$$

This would evidently remove the problematic mixing term *without*, however, eliminating the mass term for the field A'^μ. The new, massive, field would thus be seen to have effectively acquired a longitudinal component $\frac{1}{gv}\partial^\mu\varphi_2$.

Indeed, one might imagine that a judicious choice of gauge transformation could completely eliminate the field φ_2. To see this, it is more convenient to reparametrise the field φ exponentially in terms of two real fields H and ω:

$$\varphi(x) \;=\; \tfrac{1}{\sqrt{2}}\big[v+H(x)\big]\,\mathrm{e}^{\,\mathrm{i}\omega(x)}\,. \quad (7.2.39)$$

Evidently, the "radial" excitation H is somehow equivalent to φ_1 while the "angular" field ω replaces φ_2. We may now exploit the gauge invariance of the Lagrangian and apply the following gauge transformation:

$$\begin{aligned}\varphi(x) \;\to\; \varphi'(x) \;&=\; \mathrm{e}^{-\,\mathrm{i}\omega(x)}\,\varphi(x)\\ &=\; \tfrac{1}{\sqrt{2}}\,\big[v+H(x)\big] \qquad (7.2.40a)\end{aligned}$$

and

$$A_\mu(x) \;\to\; A'_\mu(x) \;=\; A_\mu(x)-\tfrac{1}{g}\,\partial_\mu\omega(x). \quad (7.2.40b)$$

The Lagrangian then becomes

$$\begin{aligned}\mathcal{L}(A'^\mu,H,\omega) \;=\; &-\tfrac{1}{4}F'^{\mu\nu}F'_{\mu\nu}+\tfrac{1}{2}g^2v^2A'^\mu A'_\mu+\tfrac{1}{2}\left(\partial^\mu H\,\partial_\mu H-2\mu^2H^2\right)\\ &+\tfrac{1}{2}g^2A'^\mu A'_\mu H(2v+H)-\lambda vH^3-\tfrac{1}{4}\lambda H^4. \quad (7.2.41)\end{aligned}$$

The gauge field A'_μ really has thus acquired a mass $m_A=gv$ and the scalar field H has mass $m_H=\sqrt{2}\mu$ while the field ω has simply disappeared.

The physical interpretation should now be evident: a *massive* vector field necessarily has three degrees of freedom whereas the original *massless* gauge field had only two; the third is provided by the *would-be* Goldstone boson ω, which, we

say figuratively, has thus been "eaten" by the gauge field.

7.2.5 Renormalisability

The problem now is to show that the theory is still renormalisable. As written, the fields used do not manifest the gauge symmetry possessed by the original construction and it is known that the propagator for a *massive* vector field has bad UV behaviour. In fact, the gauge chosen according to Eqs. (7.2.40) is known as the *unitary* gauge, since in this form the fields used correspond to the asymptotic states of the theory and the Green-function poles are located at the physical masses of real particles while all non-physical states have been eliminated.

For any other gauge choice, this clear physical picture is lost. In particular, 't Hooft (1971) defined the so-called R_ξ gauges, forgoing the physical interpretation in order to prove renormalisability. Indeed, spurious poles, both vector and scalar, will always be found but will, of course, disappear in the calculation of any physical quantity such as a cross-section. That this should be so is evident from the gauge symmetry of the original theory and the absence of such non-physical poles for particular gauge choices. Thus, while not manifestly unitary, the symmetry properties of the R_ξ gauge afford the usual protection against uncontrollable divergences in higher-order loop diagrams.

Recall that the previously chosen gauge-fixing term,

$$\mathcal{L}_{\rm GF}(A) \;=\; -\tfrac{1}{2}\xi^{-1}\left(\partial^\mu A_\mu\right)^2, \tag{7.2.42}$$

ensures the Lorentz condition

$$\partial^\mu A_\mu \;=\; 0. \tag{2.4.46}$$

In place of this, we now choose

$$\mathcal{L}_{\rm GF}(A) \;=\; -\tfrac{1}{2}\xi^{-1}\left(\partial^\mu A_\mu - \xi g\, \upsilon\varphi_2\right)^2, \tag{7.2.43}$$

which leads to the following condition:

$$\partial^\mu A_\mu \;=\; \xi g \upsilon \varphi_2. \tag{7.2.44}$$

We thus see that, for $\upsilon \neq 0$,

$$\varphi_2 \;=\; 0 \quad \text{for} \quad \xi \;\to\; \infty. \tag{7.2.45}$$

And so we may view the unitary gauge as a limiting case of the class of R_ξ gauges.

Now, the principal advantage of the 't Hooft gauge is that it eliminates the mixing term in A^μ and ω. Recall that this arises from the kinetic piece $(D^\mu\varphi^*)(D_\mu\varphi)$,

which contains a term $-g\upsilon A_\mu \partial^\mu \omega$. This may be rewritten as

$$-g\upsilon\, A_\mu\, \partial^\mu \omega \;=\; g\upsilon\, (\partial^\mu A_\mu)\, \omega - g\upsilon\, \partial^\mu (A_\mu\, \omega) \tag{7.2.46}$$

and then omit the total divergence since it does not contribute to the action. The remaining term on the right-hand side now exactly cancels the mixing term in Eq. (7.2.43). The Lagrangian is then

$$\begin{aligned}\mathcal{L}(A^\mu, \varphi) \;=\; & \tfrac{1}{2}\left[\partial^\mu \varphi_1 \partial_\mu \varphi_1 - 2\mu^2 \varphi_1^2\right] + \tfrac{1}{2}\left[\partial^\mu \varphi_2 \partial_\mu \varphi_2 - \xi m_A^2 \varphi_2^2\right] \\ & + \tfrac{1}{2}\left[(1-\xi^{-1})(\partial^\mu A_\mu)^2 - (\partial_\mu A_\nu)(\partial^\mu A^\nu) + m_A^2 A^\mu A_\mu\right] \\ & + g\upsilon A^\mu (\varphi_1 \overset{\leftrightarrow}{\partial}_\mu \varphi_2) + g^2 \upsilon A^\mu A_\mu \varphi_1 + \tfrac{1}{2} g^2 A^\mu A_\mu (\varphi_1^2 + \varphi_2^2) \\ & - \lambda \upsilon \varphi_1 (\varphi_1^2 + \varphi_2^2) - \tfrac{1}{4}\lambda (\varphi_1^2 + \varphi_2^2)^2, \end{aligned} \tag{7.2.47}$$

where the vector-field mass is $m_A = g\upsilon$. The field φ_1 appears with its usual mass while the would-be Goldstone boson φ_2 appears with a mass $\sqrt{\xi} m_A$.

Exercise 7.2.1. *Derive the Lagrangian shown above.*

The question of renormalisability is determined by the high-energy behaviour of the vector field A^μ. The standard procedure yields

$$D_{\mathrm{F}}^{\mu\nu}(q) \;=\; \frac{-g^{\mu\nu} + (1-\xi) q^\mu q^\nu / (q^2 - \xi m_A^2)}{q^2 - m_A^2 + \mathrm{i}\varepsilon}. \tag{7.2.48}$$

In the unitary limit $\xi \to \infty$ and in the ultraviolet region this behaves just as the propagator of a *massive* vector field. Moreover, the mass of the Goldstone boson φ_2 becomes infinite and thus it decouples from the theory. However, for any *finite* value of ξ the vector propagator has the desired behaviour of a *massless* gauge field. The complete proof, provided by 't Hooft (1971), also demonstrates that the non-physical poles at $q^2 = \xi m_A^2$, present in the propagators of both φ_2 and the massive vector field, do indeed cancel.

Exercise 7.2.2. *Derive the vector-boson propagator shown above.*

7.3 Ward identities

Gauge invariance leads to important relationships between the 1PI vertex function and fermion propagator. In QED these were first derived by Ward (1950) and later generalised by Takahashi (1957). The presence of ghosts in the Yang–Mills case complicates the analysis, but later Taylor (1971) and Slavnov (1972) showed that

the identities still hold. For simplicity, we shall only consider the Abelian case here.

The starting point is the observation that in QED the full generating functional $Z[J^\mu,\sigma,\overline{\sigma}]$ (*i.e.* including the gauge-fixing term) must itself still be gauge invariant although the exponential in the path integral is evidently not. The remaining symmetry leads to a differential equation for Z, which is finally expressed in terms of the 1PI vertex function and fermion propagator. Re-expressing this in turn in terms of the respective renormalisation constants Z_2 and Z_ψ leads to the all-orders relation $Z_2 = Z_\psi$. That is, the two constants are not independent quantities. This result is of key importance for the proof of renormalisability of the theory.

As stated, we start with the full generating functional for QED:

$$Z[J^\mu,\sigma,\overline{\sigma}] \;\propto\; \int \mathcal{D}A^\mu\,\mathcal{D}\psi\,\mathcal{D}\bar{\psi}\,\exp\!\left\{\mathrm{i}\int \mathrm{d}^4x\,\left[\mathcal{L}+J^\mu A_\mu+\overline{\sigma}\psi+\bar{\psi}\sigma\right]\right\}, \tag{7.3.1}$$

where the full, gauge-fixed, QED Lagrangian $\mathcal{L}$ is

$$\mathcal{L} \;=\; -\tfrac{1}{4}F^{\mu\nu}F_{\mu\nu}+\bar{\psi}\left(\mathrm{i}\not{D}-m\right)\psi-\tfrac{1}{2}\xi^{-1}\left(\partial^\mu A_\mu\right)^2. \tag{7.3.2}$$

As we have seen, with this choice of gauge fixing, the Fadde'ev–Popov ghost fields in QED decouple from all physical fields and their contribution to the path integral is just a multiplicative constant, which may be absorbed into the overall normalisation. We are thus justified in neglecting their presence.

Now, the above Lagrangian is manifestly not gauge invariant while all physical quantities (which may be expressed in terms of Z) must be gauge invariant and therefore so too Z itself. This non-trivial constraint may be turned into a differential equation for Z.

Consider the effect on Eq. (7.3.1) of an infinitesimal gauge transformation:

$$\delta\varphi(x) \;=\; -\mathrm{i}e\,\delta\omega(x)\,\varphi(x) \tag{7.3.3a}$$

and

$$\delta A^\mu(x) \;=\; \partial^\mu\delta\omega(x). \tag{7.3.3b}$$

The first two terms in (7.3.2) are invariant while the last is not and nor are the source terms in (7.3.1). This means that the transformation generates a contribution to the exponent in the integrand of Z:

$$\mathrm{i}\int \mathrm{d}^4x\,\left[-\xi^{-1}(\partial^\mu A_\mu)\,\Box\,\delta\omega+J^\mu\partial_\mu\delta\omega-\mathrm{i}e\delta\omega(\overline{\sigma}\psi-\bar{\psi}\sigma)\right] \;=$$
$$\mathrm{i}\int \mathrm{d}^4x\,\delta\omega\left[-\xi^{-1}\,\Box\,\partial^\mu A_\mu-\partial_\mu J^\mu-\mathrm{i}e(\overline{\sigma}\psi-\bar{\psi}\sigma)\right], \tag{7.3.4}$$

where, to obtain the right-hand side, we have integrated by parts.

Gauge invariance then implies that this operator acting on Z give zero. Since this must be true for any infinitesimal $\delta\omega(x)$, we obtain the following functional differential equation:

$$\left[\mathrm{i}\,\xi^{-1}\,\Box\,\partial^{\mu}\frac{\delta}{\delta J^{\mu}} - \partial_{\mu}J^{\mu} - e\left(\overline{\sigma}\frac{\delta}{\delta\overline{\sigma}} - \sigma\frac{\delta}{\delta\sigma}\right)\right] Z[\overline{\sigma},\sigma,J^{\mu}] \;=\; 0, \tag{7.3.5}$$

where we have applied the usual substitutions,

$$\psi \;\rightarrow\; -\mathrm{i}\,\frac{\delta}{\delta\overline{\sigma}}, \qquad \overline{\psi} \;\rightarrow\; \mathrm{i}\,\frac{\delta}{\delta\sigma}, \qquad A_{\mu} \;\rightarrow\; -\mathrm{i}\,\frac{\delta}{\delta J^{\mu}}. \tag{7.3.6}$$

We can transform the above into an equation for the connected functional Z_c, defined by $Z = \exp[\mathrm{i}\,Z_c]$:

$$-\,\xi^{-1}\,\Box\,\partial^{\mu}\frac{\delta Z_c}{\delta J^{\mu}} \;-\; \mathrm{i}\,\partial_{\mu}J^{\mu} \;-\; \mathrm{i}\,e\left(\overline{\sigma}\frac{\delta Z_c}{\delta\overline{\sigma}} - \sigma\frac{\delta Z_c}{\delta\sigma}\right) \;=\; 0, \tag{7.3.7}$$

We can further rewrite this in terms of the effective action Γ,

$$\Gamma[\psi,\overline{\psi},A^{\mu}] \;:=\; Z_c[\overline{\sigma},\sigma,J^{\mu}] - \int \mathrm{d}^4x\,[\overline{\sigma}\psi + \overline{\psi}\sigma + J^{\mu}A^{\mu}], \tag{7.3.8}$$

for which we have

$$\begin{aligned}
\frac{\delta\Gamma}{\delta\psi} \;&=\; -\overline{\sigma}, & \frac{\delta Z_c}{\delta\overline{\sigma}} \;&=\; \psi,\\
\frac{\delta\Gamma}{\delta\overline{\psi}} \;&=\; \sigma, & \frac{\delta Z_c}{\delta\sigma} \;&=\; -\overline{\psi},\\
\frac{\delta\Gamma}{\delta A_{\mu}} \;&=\; -J^{\mu}, & \frac{\delta Z_c}{\delta J_{\mu}} \;&=\; A^{\mu}.
\end{aligned} \tag{7.3.9}$$

We may thus write

$$-\,\xi^{-1}\,\Box\,\partial^{\mu}A_{\mu} + \mathrm{i}\,\partial_{\mu}\frac{\delta\Gamma}{\delta A_{\mu}} + \mathrm{i}\,e\left(\psi\frac{\delta\Gamma}{\delta\psi} - \overline{\psi}\frac{\delta\Gamma}{\delta\overline{\psi}}\right) \;=\; 0, \tag{7.3.10}$$

If we now functionally differentiate this equation with respect to ψ and $\overline{\psi}$, setting $\psi = \overline{\psi} = A^{\mu} = 0$ afterwards, we obtain ($\Gamma[0]$ indicates setting the arguments to zero

after differentiation)

$$\mathrm{i}\partial_z^\mu \frac{\delta^3\Gamma[0]}{\delta\bar\psi(x)\,\delta\psi(y)\,\delta A_\mu(z)} \;=\; \mathrm{i}e\,\delta(z-x)\,\frac{\delta^2\Gamma[0]}{\delta\bar\psi(x)\,\delta\psi(y)} - \mathrm{i}e\,\delta(z-y)\,\frac{\delta^2\Gamma[0]}{\delta\bar\psi(x)\,\delta\psi(y)}, \tag{7.3.11}$$

where, note, the term containing ξ has disappeared, as it should for a physically meaningful result. Now, the left-hand side is just the space–time derivative of the QED 1PI three-point vertex while each of the two terms on the right-hand side contain the full inverse electron propagator.

To see the content of this last equation more clearly, it is helpful, as usual, to transform to momentum space. For the Feynman propagator, we have

$$\int \mathrm{d}^4x\,\mathrm{d}^4y\;\mathrm{e}^{\mathrm{i}(p\cdot x - p'\cdot y)}\,\frac{\delta^2\Gamma[0]}{\delta\bar\psi(x)\,\delta\psi(y)} \;=\; \delta^4(p-p')\,\Pi(p), \tag{7.3.12}$$

where $\Pi(p)$ indicates the full inverse propagator, while we may define the proper vertex function by

$$\int \mathrm{d}^4x\,\mathrm{d}^4y\,\mathrm{d}^4z\;\mathrm{e}^{\mathrm{i}(p\cdot x - p'\cdot y + q\cdot z)}\,\frac{\delta^3\Gamma[0]}{\delta\bar\psi(x)\,\delta\psi(y)\,\delta A^\mu(z)} \;=\; \mathrm{i}e\,\bar\delta^4(p-p'+q)\,\Gamma_\mu(p,p',q), \tag{7.3.13}$$

where we have extracted the implicit "$\mathrm{i}e$" factor. Using these relations, we finally obtain

$$q^\mu\Gamma_\mu(p,p+q,q) \;=\; \Pi(p+q) - \Pi(p). \tag{7.3.14}$$

This result is precisely the Ward–Takahashi identity. In the limit $q^\mu \to 0$ it reduces to the simple Ward identity:

$$\Gamma_\mu(p,p,0) \;=\; \frac{\partial\Pi(p)}{\partial p^\mu}. \tag{7.3.15}$$

Note that since these expressions involve the *full* inverse propagator and vertex functions, they hold to *all* orders in perturbation theory.

Exercise 7.3.1. *Starting from Eq. (7.3.11) derive explicitly the final expression for the Ward–Takahashi identity* (7.3.14).

At leading order, we could have easily derived the Ward identity explicitly. Note first that at leading order the full inverse propagator and vertex functions simply correspond to their bare counterparts:

$$\Pi^{(0)}(p) \;=\; \not{p} - m \tag{7.3.16a}$$

and

$$\Gamma^{(0)}_{\mu}(p, p+q, q) \;=\; \gamma_{\mu}. \tag{7.3.16b}$$

Taking the derivative with respect to p^{μ} of the first, we immediately obtain the second.

Exercise 7.3.2. *Show that, to leading order, the full vertex function* Γ_{μ} *in QED is indeed just* γ_{μ}.

Such an explicit verification of the identity may be continued to the one-loop level without actually evaluating the loop integrals. Now, the question of higher orders brings to mind the problem of renormalisation. Indeed, the higher-order corrections to the inverse propagator are related to the wave-function renormalisation constant Z_{ψ} while those for the vertex function determine the coupling renormalisation constant Z_2. The Ward–Takahashi identity then evidently forges a connection between the two: since it is an all-orders expression, we find the simple but profound consequence that

$$Z_{\psi} \;=\; Z_2. \tag{7.3.17}$$

In other words, there is one less independent renormalisation constant than would have been expected *a priori*. This is of vital importance to the proof of renormalisability of the theory.

In the case of a YM theory the situation is rather more complex: the colour factors and the existence of other vertex functions render the entire procedure much more involved. However, as stated earlier, Taylor (1971) and Slavnov (1972) were able to show that such identities do indeed still hold. The interested reader is referred to any of the many texts dealing with this subject.

7.4 Bibliography

Anderson, P.W. (1963), *Phys. Rev.* **130**, 439.

Czakon, M. (2005), *Nucl. Phys.* **B710**, 485.

Englert, F. and Brout, R. (1964), *Phys. Rev. Lett.* **13**, 321.

Glashow, S.L. (1961), *Nucl. Phys.* **22**, 579; *see also*, *Rev. Mod. Phys.* **52** (1980) 539; *Science* **210** (1980) 1319.

Goldstone, J. (1961), *Nuovo Cim.* **19**, 154.

Goldstone, J., Salam, A. and Weinberg, S. (1962), *Phys. Rev.* **127**, 965.

Guralnik, G.S., Hagen, C.R. and Kibble, T.W.B. (1964), *Phys. Rev. Lett.* **13**, 585.

Higgs, P.W. (1964), *Phys. Lett.* **12**, 132; Higgs, P.W., *Phys. Rev. Lett.* **13**, 508.

Landau, L.D. and Pomeranchuk, I.Ya. (1955), *Dokl. Akad. Nauk. Ser. Fiz.* **102**, 489.

Nambu, Y. (1960), *Phys. Rev.* **117**, 648.

Nambu, Y. and Jona-Lasinio, G. (1961), *Phys. Rev.* **122**, 345; **124**, 246.

Salam, A. and Ward, J.C. (1964), *Phys. Lett.* **13**, 168; *see also*, Salam, A., *Rev. Mod. Phys.* (1980) 525; *Science* **210**, 723.

Slavnov, A.A. (1972), *Teor. Mat. Fiz.* **10**, 153; *transl.*, *Theor. Math. Phys.* **10**, 99; *reprinted*, in *Gauge Theory of Weak and Electromagnetic Interactions (selected papers)*, ed. C.H. Lai (World Sci., 1981), p. 264.

't Hooft, G. (1971), *Nucl. Phys.* **B35**, 167.

Takahashi, Y. (1957), *Nuovo Cim.* **6**, 371.

Taylor, J.C. (1971), *Nucl. Phys.* **B33**, 436.

van Ritbergen, T., Vermaseren, J.A.M. and Larin, S.A. (1997), *Phys. Lett.* **B400**, 379.

Ward, J.C. (1950), *Phys. Rev.* **78**, 182.

Weinberg, S. (1967), *Phys. Rev. Lett.* **19**, 1264; *see also*, *Rev. Mod. Phys.* **52** (1980) 515; *Science* **210** (1980) 1212.

Appendix A

Miscellaneous Topics

A.1 Grassmann variables

Here we shall introduce the notion of Grassmann or anticommuting variables and examine their properties under differentiation and integration. We thus start by defining two such quantities θ and θ' having the following anticommutation properties:

$$\{\theta, \theta\} \;=\; \{\theta, \theta'\} \;=\; \{\theta', \theta'\} \;=\; 0. \tag{A.1}$$

That is, $\theta^2 = 0 = \theta'^2$ and $\theta\theta' = -\theta'\theta$. Therefore, for any function $f(\theta, \theta')$ that admits a Taylor expansion, the series automatically truncates to

$$f(\theta, \theta') \;=\; f_0 + f_1\theta + f_1'\theta' + f_2\theta\theta', \tag{A.2}$$

where the f_i are just $\mathbb{C}$-number coefficients. Considering differentiation as an operation, we must have, *e.g.*,

$$\frac{\partial}{\partial\theta}\,\theta' \;=\; -\theta'\frac{\partial}{\partial\theta} \tag{A.3}$$

and therefore

$$\frac{\partial f}{\partial\theta} \;=\; f_1 + f_2\theta' \quad \text{and} \quad \frac{\partial f}{\partial\theta'} \;=\; f_1' - f_2\theta. \tag{A.4}$$

Moreover,

$$\frac{\partial^2 f}{\partial\theta\,\partial\theta'} \;=\; -\frac{\partial^2 f}{\partial\theta'\,\partial\theta} \;=\; -f_2. \tag{A.5}$$

The notion of an integral may also be defined: since it must be linear and the measures $\mathrm{d}\theta$, $\mathrm{d}\theta'$ must also be Grassmannian, we have

$$\{\mathrm{d}\theta, \theta\} \;=\; \{\mathrm{d}\theta, \theta'\} \;=\; \{\mathrm{d}\theta', \theta\} \;=\; \{\mathrm{d}\theta, \mathrm{d}\theta'\} \;=\; 0. \tag{A.6}$$

We evidently only need the two integrals $\int \mathrm{d}\theta$ and $\int \mathrm{d}\theta\theta$. Now,

$$\left[\int \mathrm{d}\theta\right]^2 = \left[\int \mathrm{d}\theta\right]\left[\int \mathrm{d}\theta'\right] = \int \mathrm{d}\theta\,\mathrm{d}\theta' = -\int \mathrm{d}\theta'\mathrm{d}\theta = -\left[\int \mathrm{d}\theta\right]^2. \quad \text{(A.7)}$$

and so

$$\int \mathrm{d}\theta = 0. \quad \text{(A.8)}$$

Finally, since θ^2 and all higher powers are zero there is no intrinsic scale to a Grassmann variable and so we may simply *define*

$$\int \mathrm{d}\theta\,\theta := 1. \quad \text{(A.9)}$$

By integrating the expansion of the function $f(\theta,\theta')$ above, we find that the operations of differentiation and integration are entirely equivalent when applied to Grassmann variables:

$$\int \mathrm{d}\theta\, f(\theta,\theta') = f_1 + f_2\theta' = \frac{\partial f(\theta,\theta')}{\partial\theta}. \quad \text{(A.10)}$$

These properties lead to various further useful identities and relations for functions of Grassmann variables. First of all it is easy to show that the δ-function defined over a space of Grassmann variables is just $\delta(\theta)=\theta$.

Exercise A.1.1. *Show that the above definition of the Grassmann δ-function is indeed correct; that is show that*

$$\int \mathrm{d}\theta\, f(\theta) = f(0).$$

Hint: *write $f(\theta)=f_0+f_1\theta$ and evaluate the integral.*

Exercise A.1.2. *Show further that if θ and η are both Grassmann variables, then an integral representation of the Grassmann δ-function is*

$$\int \mathrm{d}\eta\, \mathrm{e}^{\mathrm{i}\eta\theta} = \mathrm{i}\delta(\theta).$$

Exercise A.1.3. *As a final exercise, show that if we change integration variables thus $\theta'=\lambda\theta$ ($\lambda\neq 0$), then the Jacobian is actually the opposite of what we normally expect:*

$$\int \mathrm{d}\theta\, f(\theta) = \lambda \int \mathrm{d}\theta' f(\theta'/\lambda).$$

Show also that similarly

$$\delta\big(f(\theta)\big) \;=\; \frac{\partial f}{\partial\theta}\,\delta(\theta).$$

The extension to a finite set of Grassmann variables and vectors with Grassmann components is straightforward. The functional integral over a Grassmann field can then be defined in the usual manner, as the limit of an infinite number of discrete Grassmann variables. Moreover, the construction of a complex Grassmann variable is also straightforward. Note finally that, there being no intrinsic scale, there is no difference between definite and indefinite Grassmann integrals.

For the path-integral quantisation of a field, we need to perform Gaussian-type integrals. For just a single Grassmann variable, this is evidently a trivial operation since $\theta^2=0$. We shall thus naturally be forced to move to higher dimensional integrals. Consider therefore, a set of n Grassmann variables $\theta_1,\,\ldots,\,\theta_n$; that is, they all mutually anticommute. Again, a power-series expansion of any function will naturally terminate after the term containing the product of each variable $\theta_1\ldots\theta_n$. Consider next the matrix product

$$\boldsymbol{\Theta}^{\mathsf{T}}\mathsf{M}\,\boldsymbol{\Theta}, \tag{A.11}$$

where $\boldsymbol{\Theta}$ represents the real n-component vector $(\theta_1,\ldots,\theta_n)$ and M a real $n\times n$ matrix. We see that the symmetric piece trivially vanishes and so, since any matrix M may be decomposed into the sum of a symmetric and an antisymmetric matrix, only antisymmetric matrices need be considered.

The most general, real, Gaussian integral we shall meet then is

$$\int \mathrm{d}^n\boldsymbol{\Theta}\,\exp\Big\{-\tfrac{1}{2}\boldsymbol{\Theta}^{\mathsf{T}}\mathsf{A}\,\boldsymbol{\Theta}\Big\}, \tag{A.12}$$

where A is a real $n\times n$ antisymmetric matrix and the Grassmann measure is defined to be $\mathrm{d}^n\boldsymbol{\Theta}:=\mathrm{d}\theta_1\ldots\mathrm{d}\theta_n$. Each term in the power-series expansion of the exponential will only contain overall even powers of the θ_i and thus the integral will only be non-zero for n even. Moreover, since we need precisely one power of each variable θ_i (no more, no less) the only term leading to a non-vanishing integral is the $n/2$-th term. That is,

$$\int \mathrm{d}^n\boldsymbol{\Theta}\,\exp\Big\{-\tfrac{1}{2}\boldsymbol{\Theta}^{\mathsf{T}}\mathsf{A}\,\boldsymbol{\Theta}\Big\} \;=\; \int \mathrm{d}^n\boldsymbol{\Theta}\,\frac{1}{(n/2)!}\Big[-\tfrac{1}{2}\boldsymbol{\Theta}^{\mathsf{T}}\mathsf{A}\,\boldsymbol{\Theta}\Big]^{n/2}. \tag{A.13}$$

In the previous expression the only products of the $n/2$ terms that will survive are those containing each of the n variables just once and there will be just $(n/2)!$ such terms. If we now rearrange the θ_i into natural order, then the relative minus sign introduced is just that corresponding to the related determinant (considering

the product of elements of A involved). Finally, the factor $(-1)^{n/2}$ compensates a similar sign due to the rearrangement under integration and thus the final answer is just

$$\int \mathrm{d}^n\boldsymbol{\Theta}\, \exp\Big\{-\tfrac{1}{2}\boldsymbol{\Theta}^{\mathsf{T}}\mathsf{A}\,\boldsymbol{\Theta}\Big\} \;=\; (\det \mathsf{A})^{1/2} \;=\; \exp\Big\{\tfrac{1}{2}\,\mathrm{Tr}\ln \mathsf{A}\Big\}\,. \tag{A.14}$$

Note that, in contrast to the corresponding expression (3.1.2) for *commuting* fields, we find a *positive* power of the determinant. The difference in numerical coefficient is, of course, irrelevant for our purposes. Note also that, since the determinant of an $n\times n$ antisymmetric matrix is zero for n odd, we may extend the previous result to include both odd n and even.

One further useful result is obtained by including a linear term in $\boldsymbol{\Theta}$ in the exponent:

$$\int \mathrm{d}^n\boldsymbol{\Theta}\, \exp\Big\{-\tfrac{1}{2}\boldsymbol{\Theta}^{\mathsf{T}}\mathsf{A}\,\boldsymbol{\Theta} + \boldsymbol{\sigma}^{\mathsf{T}}\boldsymbol{\Theta}\Big\} \;=\; \exp\Big\{\tfrac{1}{2}\,\mathrm{Tr}\ln \mathsf{A}\Big\}\exp\Big\{-\tfrac{1}{2}\boldsymbol{\sigma}^{\mathsf{T}}\mathsf{A}^{-1}\boldsymbol{\sigma}\Big\}\,. \tag{A.15}$$

where $\boldsymbol{\sigma}$ is an external vector source of anticommuting Grassmann components $(\sigma_1,\ldots,\sigma_n)$ that also anticommute with the θ_i.

Exercise A.1.4. *Derive this last equation in the usual way,* i.e. *by completing the square in the exponent.*

Finally, we shall require the extension to *complex* Grassmann variables, which is easily obtained from the previous examples. The integration measure is naturally defined as

$$\mathrm{d}\theta\,\mathrm{d}\theta^{\,*} \;:=\; 2\,\mathrm{d}(\mathrm{Re}\,\theta)\,\mathrm{d}(\mathrm{Im}\,\theta)\,, \tag{A.16}$$

whence

$$\int \mathrm{d}^n\boldsymbol{\Theta}\,\mathrm{d}^n\boldsymbol{\Theta}^*\, \exp\Big\{-\boldsymbol{\Theta}^{\mathsf{T}}\mathsf{A}\,\boldsymbol{\Theta}\Big\} \;=\; (-1)^{n/2}(\det \mathsf{A})^{1/2}, \tag{A.17}$$

where now A is an $n\times n$ *anti*-Hermitian or *skew*-Hermitian matrix; that is,

$$\mathsf{A}^\dagger \;=\; -\mathsf{A}. \tag{A.18}$$

Again, the multiplicative factor, which here is $(-1)^{n/2}$, in front of the right-hand side above is irrelevant.

A.2 The SU(N) algebra

A Lie algebra is defined by a set of commutation relations:

$$\left[\mathsf{T}^a, \mathsf{T}^b\right] \;=\; \mathrm{i}\, f^{abc}\, \mathsf{T}^c, \tag{A.19}$$

where the f^{abc} are totally antisymmetric *structure constants*, characterising the particular group, and the matrices T^a belong to the *fundamental* representation. For SU(N) the fundamental representation has dimension N and the indices a,b,c run from 1 to N^2-1. In the case of SU(2) the structure constants are none other than the Levi-Civita alternating tensor ϵ_{ijk} and for SU(3) the non-zero elements are shown in Table A.1. For higher-rank groups the structure constants are more

Table A.1: The non-zero elements of the antisymmetric structure constants f^{abc} in the case of SU(3).

abc	f^{abc}
123	1
147	$\frac{1}{2}$
156	$-\frac{1}{2}$
246	$\frac{1}{2}$
257	$\frac{1}{2}$
345	$\frac{1}{2}$
367	$-\frac{1}{2}$
458	$\frac{\sqrt{3}}{2}$
678	$\frac{\sqrt{3}}{2}$

complicated and the reader should consult one of the many texts dealing with Lie algebras.

For SU(N) with $N \geq 2$, there is also a set of non-zero symmetric structure constants defined via

$$\left\{\mathsf{T}^a, \mathsf{T}^b\right\} \;=\; 2C_1\, \delta^{ab} + d^{abc}\, \mathsf{T}^c, \tag{A.20}$$

where $C_1 = \frac{1}{2}$ is a constant, determined by the normalisation of the matrices T^a and is therefore representation dependent. Taking the trace of this equation, we obtain

$$\mathrm{Tr}\left[\mathsf{T}^a\, \mathsf{T}^b\right] \;=\; C_1\, \delta^{ab}, \tag{A.21}$$

There exist various trace relations of which two important identities are (for clarity we write the sums over indices explicitly):

$$\sum_a \mathsf{T}^a \mathsf{T}^a \;=\; C_{\mathrm{F}}\, \mathbb{1} \tag{A.22a}$$

and

$$\sum_{c,d} f^{acd} f^{bcd} \;=\; C_{\mathrm{A}}\, \delta^{ab}, \tag{A.22b}$$

where C_{F} and C_{A} are constants, known as Casimir constants, which are characteristic of the particular group. In the case of SU(N) they are

$$C_{\mathrm{F}} \;=\; \frac{N^2-1}{2N} \tag{A.23a}$$

and

$$C_{\mathrm{A}} \;=\; N. \tag{A.23b}$$

A further useful trace-like relation is

$$\sum_a \mathsf{T}^a_{ij}\, \mathsf{T}^a_{kl} \;=\; \frac{1}{2}\left(\delta_{il}\delta_{jk} - \frac{1}{N}\,\delta_{ij}\delta_{kl}\right). \tag{A.24}$$

We conclude by pointing out that by making the identification

$$(\mathsf{F}^a)_{bc} \;:=\; -\mathrm{i}\, f^{abc}, \tag{A.25}$$

then, as a result of the Bianchi identity, we find that the matrices F^a satisfy the same algebra:

$$\left[\mathsf{F}^a, \mathsf{F}^b\right] \;=\; \mathrm{i}\, f^{abc}\, \mathsf{F}^c. \tag{A.26}$$

This is known as the *adjoint* or *regular* representation.

A.3 The Gamma function

The Euler Gamma function $\Gamma(z)$ is single valued and analytic over the entire complex z plane, except for the points $z=-n$ ($n=0,\ 1,\ 2,\ \ldots$), where it has simple poles with residues $(-1)^n/n!$.

There are numerous integral representations of the Gamma function; the best known is perhaps the Euler integral:

$$\Gamma(z) \;=\; \int_0^\infty \mathrm{d}t\, t^{z-1}\, \mathrm{e}^{-t} \qquad (\mathrm{Re}\, z \;>\; 0). \tag{A.27}$$

It is given simply by $\Gamma(n+1)=n!$ for $n\in\mathbb{Z}$. We thus have, for example, $\Gamma(1)=1$ while $\Gamma(\frac{1}{2})=\sqrt{\pi}$ and, as noted, $\Gamma(-n)=\infty$ for $n=0,\,1,\,2,\,\ldots$

Other useful formulæ involving the Gamma function are:

$$z\,\Gamma(z) \;=\; \Gamma(z+1), \tag{A.28a}$$

$$\Gamma(z)\,\Gamma(1-z) \;=\; \pi\,\mathrm{cosec}\,\pi z, \tag{A.28b}$$

$$\int_0^1 \mathrm{d}x\, x^{m-1}\,(1-x)^{n-1} \;=\; \frac{\Gamma(m)\Gamma(n)}{\Gamma(m+n)} \;=\; B(m,n), \tag{A.28c}$$

where the second is Euler's reflection formula and the last is also known in mathematics as the beta function or Euler integral of the first kind (such integrals commonly occur from the Feynman parametrisation of propagator products).

The derivative of the Gamma function is also useful and the Digamma function $\psi(z)$ is defined as follows:*

$$\psi(z) \;=\; \frac{\mathrm{d}}{\mathrm{d}z}\ln\Gamma(z) \;=\; \frac{\Gamma'(z)}{\Gamma(z)}. \tag{A.29}$$

For $n\in\mathbb{N}$, it is not difficult to show that

$$\psi(n) \;=\; S_n - \gamma_{\mathrm{E}}, \tag{A.30}$$

where $S_n := \sum_{m=1}^{n-1} 1/m$ (with $S_1=0$) and the Euler–Mascheroni constant γ_{E} is defined by

$$\begin{aligned}\gamma_{\mathrm{E}} \;&=\; \lim_{n\to\infty}\left[1+\tfrac{1}{2}+\tfrac{1}{3}+\tfrac{1}{4}+\cdots+\tfrac{1}{n}-\ln n\right]\\ &\simeq\; 0.57721\,56649\ldots\end{aligned} \tag{A.31}$$

We may thus perform an $O(\epsilon)$ Taylor expansion for the Gamma function near positive integer values of its argument:

$$\Gamma(n+\epsilon) \;=\; \Gamma(n)[1+(S_n-\gamma_{\mathrm{E}})\epsilon+O(\epsilon^2)], \tag{A.32}$$

from which we immediately have

$$\Gamma(\epsilon) \;=\; \frac{1}{\epsilon}\,\Gamma(1+\epsilon) \;=\; \frac{1}{\epsilon}-\gamma_{\mathrm{E}}+O(\epsilon). \tag{A.33}$$

* In some texts the definition is given as $\psi(z)=\mathrm{d}[\ln\Gamma(z+1)]/\mathrm{d}z$.

A.4 The Legendre transform

The Legendre transform appears in many diverse areas of physics: perhaps the best known is the the connection in classical mechanics between the Lagrangian $\mathcal{L}(q,\dot{q})$ and the Hamiltonian $\mathcal{H}(q,p)$, but it also relates the the Helmholtz free energy F to the entropy S in statistical mechanics; in the present notes it is seen to provide the relation between the the generating functional for connected Green functions $Z_c[J]$ and the effective action $\Gamma[\varphi]$. For a more complete and recent discussion, see Zia *et al.* (2008).

Just as the Fourier and Laplace transforms, the Legendre transform provides a method of re-encoding the dynamical information on a system. Consider a function $f(x)$ whose gradient is strictly monotonic (*i.e.*, it is either strictly convex or strictly concave). In such a case there exists a unique one-to-one relation between the variable x and the derivative $f'(x) = \mathrm{d}f(x)/\mathrm{d}x$. We may therefore change variables from x to $y(x) = f'(x)$ without loss of information. The usefulness of such a change lies in the possibility that y be more accessible than x. Note that the convexity condition guarantees that we may also invert the function $y(x)$ to obtain $x(y)$.

Of course, the Legendre transform is more than just a simple change of variables. The strict mathematical definition is as follows:

$$g(y) := \max_x \big[xy - f(x)\big] , \tag{A.34}$$

where the maximum is taken with respect to variations in x for y held fixed. A simple pictorial representation of this definition is shown in Fig. A.1. The

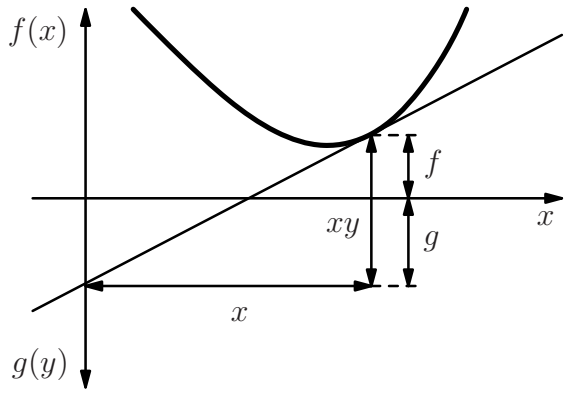


Figure A.1: A diagrammatic illustration of the Legendre transformation $g(y)$ of the function $f(x)$, as defined in the text.

function $-g = -[xy - f(x)]$ is given by the intercept with the vertical axis of the straight line having slope y and crossing the curve $f(x)$ at the point x. From the diagram, it is clear that, for a given slope y, the maximum of $[xy - f(x)]$ with

respect to variations in x is attained when the straight line is precisely tangent to $f(x)$ and thus at the point x for which $y = f'(x)$. The need for convexity is also obvious from the diagram. Moreover, the strict definition given above makes clear the independence of $g(y)$ from x. Note, of course, that there may also be other independent variables involved but these are mere spectators in the transform.

With the understanding that such a pictorial description provides, we can shift to a simpler definition, more commonly found in physics:

$$g(y) := xy - f(x) \quad \text{for} \quad y = \frac{\mathrm{d}f(x)}{\mathrm{d}x}, \tag{A.35}$$

where it is implicit that to evaluate $g(y)$ we must write x as a function of y by inverting the equation $y = f'(x)$. From this definition the symmetry of the Legendre transform is immediately manifest as we may thus write

$$xy = f(x) + g(y), \tag{A.36}$$

or the inverse transform

$$f(x) := xy - g(y) \quad \text{for} \quad x = \frac{\mathrm{d}g(y)}{\mathrm{d}y}. \tag{A.37}$$

Note that in all cases the variables x and y are not independent as each should be viewed as a function of the other.

Some simple properties of the transform are immediately obvious from Fig. A.1. Consider the minimum of the function $f(x)$ with respect to x: $f_{\min} := f(x_{\min})$. By definition, the derivative vanishes at the point $x_{\min}$; that is, $y = 0$. We therefore have

$$f_{\min} = -g(0) \quad \text{and} \quad g_{\min} = -f(0). \tag{A.38}$$

Now, by taking derivatives of the relations defining the variables x and y, we obtain

$$\frac{\mathrm{d}y}{\mathrm{d}x} = \frac{\mathrm{d}^2 f}{\mathrm{d}x^2} \quad \text{and} \quad \frac{\mathrm{d}x}{\mathrm{d}y} = \frac{\mathrm{d}^2 g}{\mathrm{d}y^2}, \tag{A.39}$$

from which we see that

$$\frac{\mathrm{d}^2 f}{\mathrm{d}x^2}\,\frac{\mathrm{d}^2 g}{\mathrm{d}y^2} = 1. \tag{A.40}$$

The above reciprocal relation between the curvatures highlights the convexity requirement: neither of the second derivatives may vanish. Moreover, it is somewhat reminiscent of the uncertainty relations $\Delta x \Delta p$, $\Delta\theta \Delta L \geq \hbar$ *etc.*

A.5 Bibliography

Zia, R.K.P., Redish, E.F. and McKay, S.R. (2008), arXiv:0806.1147.

Appendix B

Errata to Bailin and Love

In this appendix we provide a list of errata to the book:

Reading list

Bailin, D. and Love, A. (1993), *Introduction to Gauge Field Theory* (IOP Pub.), revised edition.

Page 22, *op. cit.*

The derivation of the formula for P_μ in Eq. (3.48) is erroneous in that the formula given in the last line of Eq. (3.46) is incorrect. The error lies in the addition of a so-called total divergence $\delta^0_\nu \frac{1}{2}\partial_\mu[\varphi\partial^\mu\varphi]$; while this would be applicable under a four-space integration, here there is only a three-space integration. Indeed, addition of the full four-divergence renders an incorrect expression for P_ν, which is not even time independent. The correct derivation is then as follows—beginning from the sentence following Eq. (3.45) the text should be replaced by something like:

In the free-field case $\lambda = 0$ and the "charge" is

$$\begin{aligned} P_\nu &\equiv \int \mathrm{d}^3\boldsymbol{x}\, T^0_\nu \\ &= \int \mathrm{d}^3\boldsymbol{x}\, [(\partial^0\varphi)(\partial_\nu\varphi) - \delta^0_\nu \tfrac{1}{2}(\partial_\lambda\varphi\,\partial^\lambda\varphi + \varphi\,\Box\,\varphi)] \\ &= \int \mathrm{d}^3\boldsymbol{x}\, [(\partial^0\varphi)(\partial_\nu\varphi) - \delta^0_\nu \tfrac{1}{2}(\partial_0\varphi\,\partial^0\varphi + \varphi\,\partial_0\,\partial^0\varphi)]. \qquad (3.46) \end{aligned}$$

In deriving this we have used the equations of motion in order to eliminate the term $-\delta^0_\nu \frac{1}{2} m^2 \varphi^2$ and subtracted the total spatial divergence $\delta^0_\nu \frac{1}{2} \nabla_r [\varphi \nabla^r \varphi]$ from the integrand in order to eliminate the pieces containing spatial derivatives in the second term.

Eq. (3.48) is then obtained by noting that, under symmetric integration in $\boldsymbol{k}$,

$$\int \mathrm{d}^3\boldsymbol{k}\, f(k_0)\, k_r\, a(k_0, \boldsymbol{k})\, a(k_0, -\boldsymbol{k}) \;=\; 0,$$

where $f(k_0)$ is an arbitrary function of k_0 only. This eliminates the time-dependent pieces in the spatial components of P_ν while those in the temporal component cancel between the terms of opposite sign in Eq. (3.46).

Page 83, *op. cit.*

In Eq. (7.31) the argument of the $\tan^{-1}$ should be raised to the power $\frac{1}{2}$.

Page 103, *op. cit.*

On the left-hand side of Eq. (8.52) the term $\sigma\varphi$ should be replaced with $J\varphi$.

Page 127, *op. cit.*

In Eq. (10.74) there is an extra term $(q-p)_\nu g_{\lambda\mu}$, which should be removed.

Page 133, *op. cit.*

In Eq. (11.22) the second term should be $Z_2 Z_\psi^{-1}$ and not $Z_1 Z_\theta^{-1}$.

Appendix C

Selected suitable examination topics

In this appendix we provide a short list of suitable topics for the course final oral examination. The student should prepare a short oral presentation in which a brief motivation is given followed by a suitably in-depth discussion of the salient points. Other topics may be suggested by the student, but should be communicated to the lecturer *prior* to the examination.

Note that while some have been covered quite thoroughly in the lectures others have not. In the latter case, prior consultation with the lecturer is advisable. Such topics would be of particular use to those interested in further study of the subject of this course.

Oral Examination Topics

- The calculation of the $O(\alpha)$ correction to the muon anomalous magnetic momentum (often known as the muon $g-2$).
- The calculation of the β-function in QCD using a three- or four-point function chosen freely by the student.
- The first-order perturbative corrections to a simple scattering process, such as $e^+e^- \to \mu^+\mu^-$.
- A complete discussion of the so-called *effective action* $\Gamma[\varphi]$, with particular reference to the two papers Coleman *et al.* (1973); Jackiw (1974).
- The derivation of the Feynman rules for the gauge propagator and the Fadde'ev–Popov ghost fields in the cases of both covariant and axial gauges.
- spontaneous symmetry breaking and the generation of the weak-interaction intermediate-boson masses in the standard model of elementary particle

physics. At least one concrete example of the consequences should be discussed.

- The proof of the renormalisability of QED.
- The Adler–Bell–Jackiw triangle anomaly and neutral-pion decay
- The Bloch–Nordsieck cancellation and the KLN theorem for IR divergences.

Bibliography

Coleman, S.R. and Weinberg, E. (1973), *Phys. Rev.* **D7**, 1888.

Jackiw, R. (1974), *Phys. Rev.* **D9**, 1686.

Appendix D

Glossary of Acronyms

1PI: one-particle irreducible

IR: infrared

MOM: momentum subtraction

MS: minimal subtraction

$\overline{\text{MS}}$: modified minimal subtraction

QCD: quantum chromodynamics

QED: quantum electrodynamics

RGE: renormalisation group equation

UV: ultraviolet

YM: Yang–Mills