235

Cours/Lecture Series

1990-1991 ACADEMIC TRAINING PROGRAMME

LECTURE SERIES FOR POSTGRADUATE STUDENTS

SPEAKER	: B. DE WIT / Institute for Theoretical Physics, Utrecht
TITLE	: Quantum field theory
DATES	: 15, 17, 19, 23 & 25 October
TIME	: 10.30 to 12.00 hrs – Auditorium
PLACE	: Auditorium

ABSTRACT

After a brief and practical introduction to field theory and the use of Feynman diagrams, we discuss the main concepts in gauge theories and their application in elementary particle physics. We present all the ingredients necessary for the construction of the standard model.

SIS/H/346
Div. DG/PU
Distr. int. & ext.

Secretariat : tel. 2844 – 3674

GAUGE THEORIES AND APPLICATIONS

B. de Wit
Institute for Theoretical Physics
P.O.B. 80.006, 3508 TA Utrecht
The Netherlands

Contents

The purpose of these lectures is to give a practical introduction to field theory, in particular gauge theory, and the use of Feynman diagrams. Of course, it is really not possible to give you a self-contained and thorough treatment of these topics in just six lectures, but the hope is that, also with the help of the discussion sessions, you will get acquainted with the basic principles and ideas of gauge fields and acquire some experience in Feynman diagram calculations. The latter will be put to a test in Hollik's lectures on precision tests of the electroweak theory later on in this school.

As you can see from the above table of contents, we will start with generic field theories and discuss the rules needed for the calculation of Feynman diagrams. After some applications we proceed to introduce simple abelian gauge theories and explain their properties. Then we consider the nonabelian version of these theories and discuss the ingredients that are necessary for the standard model.

1. The action

Field theories are usually defined in terms of a Lagrangian, or an action. The action, which has the dimension of Planck's constant $\hbar$, and the Lagrangian are well-known concepts in classical mechanics. For instance, for a point-particle subject to a conservative force

$$\mathbf{F} = -\frac{\partial V(\mathbf{r})}{\partial \mathbf{r}}, \tag{1.1}$$

the Lagrangian is defined as the difference of the kinetic and the potential energy,

$$L(\mathbf{r}, \dot{\mathbf{r}}) = \tfrac{1}{2} m\, \dot{\mathbf{r}}^2 - V(\mathbf{r}). \tag{1.2}$$

Consider now some particle trajectory $\mathbf{r}(t)$, which does not necessarily satisfy the equation of motion, with fixed endpoints given by

$$\mathbf{r}_1 = \mathbf{r}(t_1), \qquad \mathbf{r}_2 = \mathbf{r}(t_2). \tag{1.3}$$

The action corresponding to this trajectory is then defined as

$$S[\mathbf{r}(t)] = \int_{t_1}^{t_2} \mathrm{d}t\, L(\mathbf{r}(t), \dot{\mathbf{r}}(t)). \tag{1.4}$$

For each trajectory satisfying the boundary conditions (1.3) the action defines a number. According to Hamilton's principle, the extremum of (1.4) (usually a minimum) is acquired for those trajectories that satisfy the equation of motion,

$$m\, \ddot{\mathbf{r}} = -\frac{\partial V(\mathbf{r})}{\partial \mathbf{r}}. \tag{1.5}$$

This is precisely Newton's law.

Let us now generalize to a field theory. The system is now described in terms of fields, say, $\phi(x)$, which are functions of the four-vector of space-time,

$$x^\mu = (x_0, \mathbf{x}), \tag{1.6}$$

where $x_0 \equiv ct$ and c is the velocity of light (henceforth, we will use units such that $c = 1$). Fields such as ϕ may be used to describe the degrees of freedom of certain physical systems. For instance, they could describe the local displacement in a continuous medium like a violin string or the surface of a drum, or some force field such as an electric or a magnetic field.

Again one can define the action, which can now be written as the integral over space-time of the Lagrangian *density* (which is also commonly referred to as the Lagrangian),

$$S[\phi(x)] = \int \mathrm{d}^4x\, \mathcal{L}(\phi(x), \partial_\mu \phi(x)), \tag{1.7}$$

where the values of the fields at the boundary of the integration domain are fixed just as we did for the particle trajectory in (1.3). As indicated in (1.7), the Lagrangian is usually a function of the fields and their first-order derivatives. The action thus assigns a number to every field configuration, and it is again possible to invoke Hamilton's principle and verify that the action has an extremum for fields that satisfy the classical equations of motion.

Let us now discuss a few examples of field theories that one encounters in particle physics. The simplest theory is that of a single scalar field $\phi(x)$. This field is called a *scalar* field because it transforms trivially under Lorentz transformations: $\phi'(x') = \phi(x)$, where x' and x are related by a Lorentz transformation. It can be used to describe spinless particles. A standard Lagrangian is

$$\mathcal{L} = -\tfrac{1}{2}(\partial_\mu\phi)^2 - \tfrac{1}{2}m^2\phi^2 - \lambda\,\phi^3 - g\,\phi^4, \tag{1.8}$$

where

$$(\partial_\mu\phi)^2 = \left(\frac{\partial\phi}{\partial\mathbf{x}}\right)^2 - \left(\frac{\partial\phi}{\partial x_0}\right)^2. \tag{1.9}$$

The quadratic part of (1.8) is called the Klein-Gordon Lagrangian, and the corresponding field equation the Klein-Gordon equation. Equation (1.9) shows that the time-derivatives appear with positive sign in the Lagrangian, just as in (1.2). Observe that we have introduced the Lorentz-invariant inner product of two four-vectors, defined by

$$x\cdot y = x_\mu\, y^\mu = x^\mu\, y_\mu = \mathbf{x}\cdot\mathbf{y} - x_0\, y_0. \tag{1.10}$$

where four-vector indices are lowered (raised) with a metric $\eta_{\mu\nu}$ ($\eta^{\mu\nu}$), with $\mu,\nu = 0,1,2,3$, which is a diagonal matrix with eigenvalues $(-,+,+,+)$. In the literature also a metric with opposite sign is used. Alternatively, we may use indices $\mu,\nu = 1,2,3,4$ and define $x_4 \equiv i\,x_0$. In that case there is no difference between upper and lower indices and we do not need a metric in order to contract four-vectors. This convention is convenient when dealing with gamma matrices, to be introduced shortly.

Later we will consider the fields in the momentum representation, defined by the Fourier transform

$$\phi(k) = \int d^4x\, e^{ik\cdot x}\phi(x). \tag{1.11}$$

The inverse of this relation is

$$\phi(x) = (2\pi)^{-4}\int d^4k\, e^{-ik\cdot x}\phi(k). \tag{1.12}$$

For real fields, as in (1.8), the fields in the momentum representation satisfy the condition

$$\phi^*(k) = \phi(-k). \tag{1.13}$$

For complex fields there is no such condition. Complex fields are convenient if the theory is invariant under phase transformations. For instance, the Lagrangian

$$\mathcal{L} = -|\partial_\mu\phi|^2 - m^2\,|\phi|^2 - g\,|\phi|^4 \tag{1.14}$$

is invariant under

$$\phi \to \phi' = e^{i\xi}\,\phi, \tag{1.15}$$

with ξ an arbitrary parameter. Of course, it is only a matter of convenience to use complex fields. One always has the option to decompose a complex field into its real and its imaginary part, and to write the Lagrangian (1.14) in terms of two real fields. The invariance (1.15) then takes the form of a rotation between these two real fields. Observe the normalization factors in (1.14) which differ from those in (1.8).

In principle, it is only a small step to consider Lagrangians for fields that transform nontrivially under the Lorentz group. For instance, one has *spinor* fields (which transform as spinors under the Lorentz transformation), which describe the fermions. In spite of the fact that they are extremely important, we will largely ignore the spinor fields due to lack of time. However, in the tutorial sessions we will discuss several applications with fermions. Here we give a typical Lagrangian for spin-$\frac{1}{2}$ fermions interacting with a scalar and a pseudoscalar field, ϕ_s and ϕ_p, respectively,

$$\mathcal{L} = -\overline{\psi}\partial\!\!\!/\psi - m\,\overline{\psi}\psi + G_s\,\phi_s\overline{\psi}\psi + iG_p\,\phi_p\overline{\psi}\gamma_5\psi, \tag{1.16}$$

Such couplings of the (pseudo)scalar fields to fermions are called Yukawa couplings. The quadratic terms in (1.16) constitute the Dirac Lagrangian. The corresponding field equation is the Dirac equation. The spinor fields have 4 independent components.* An important ingredient are the Dirac gamma matrices γ^μ, which are 4×4 matrices satisfying the anticommutation relations

$$\gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu = 2\eta^{\mu\nu}\,\mathbf{1}. \qquad (\mu,\nu = 0,1,2,3) \tag{1.17}$$

We should caution you that that exist many different conventions for gamma matrices and spinors in the literature. Adopting (1.17) as a starting point, it is convenient to define $\gamma^4 \equiv i\gamma^0$, so that all gamma matrices can be chosen hermitean,

$$(\gamma^\mu)^\dagger = \gamma^\mu. \qquad (\mu = 1,2,3,4) \tag{1.18}$$

In this notation there is no difference between upper and lower four-vector indices. The conjugate spinor field $\overline{\psi}$ is then defined by

$$\overline{\psi} = \psi^\dagger\gamma_4, \quad \text{or, in components,} \quad \overline{\psi}_\alpha = \psi^\dagger_\beta(\gamma_4)_{\beta\alpha}. \tag{1.19}$$

Another Lagrangian which is relevant is based on *vector* fields (i.e., fields that transform as vectors under Lorentz transformations). For the description of massive spin-1 particles one uses the Proca Lagrangian,

$$\mathcal{L} = -\tfrac{1}{4}(\partial_\mu V_\nu - \partial_\nu V_\mu)^2 - \tfrac{1}{2}M^2 V_\mu^2. \tag{1.20}$$

Modulo a total divergence, which we can drop because it contributes only a boundary term to the action (this follows from applying Gauss' law), this Lagrangian can be written as

$$\mathcal{L} = -\tfrac{1}{2}(\partial_\mu V_\nu)^2 + \tfrac{1}{2}(\partial_\mu V^\mu)^2 - \tfrac{1}{2}M^2 V_\mu^2. \tag{1.21}$$

The first and the last term in (1.21) are obvious generalization of the first two terms of (1.8). The second term is required, with precisely the coefficient $\frac{1}{2}$, in order that the Lagrangian describes

* The fact that in a four-dimensional space-time, spinors have also four components, should be regarded as a coincidence. In a d-dimensional space-time spinors have in general $2^{[d/2]}$ components.

pure spin-1 particles, and no additional spinless particles. From a simple counting argument one can already see that some care is required here. Massive particles with spin s have in general $2s+1$ independent polarizations. So a Lagrangian for spin-1 particles should give rise to 3 independent polarization states, whereas the field V_μ on which the Proca Lagrangian is based has 4 independent components. It is this discrepancy which forces us to include the second term in (1.21). We have implicitly assumed that V_μ is a real field, i.e. $V_\mu^* = V_\mu$, but it is perfectly possible to extend (1.20-21) to complex fields, analogous to what we did previously for scalar fields.

The $M \to 0$ limit of (1.20) describes massless spin-1 particles such as photons, and will play an important role in these lectures. The Lagrangian is called the Maxwell Lagrangian, and reads

$$\begin{aligned}\mathcal{L} &= -\tfrac{1}{4}(\partial_\mu A_\nu - \partial_\nu A_\mu)^2 \\ &= -\tfrac{1}{2}(\partial_\mu A_\nu)^2 + \tfrac{1}{2}(\partial_\mu A^\mu)^2,\end{aligned} \tag{1.22}$$

where we have again suppressed a total divergence in the second line. Massless partices with spin have precisely 2 independent polarizations, irrespective of the value of the spin. We will discuss this in section 3. An important ingredient in the proof of this is the invariance of (1.22) under so-called *gauge transformations*,

$$A_\mu(x) \to A'_\mu(x) = A_\mu(x) + \partial_\mu \xi(x), \tag{1.23}$$

where $\xi(x)$ is an arbitrary function of x.

In relativistic quantum field theory it is convenient to use units such that the velocity of light in vacuum and Planck's constant are dimensionless and equal to unity: $c = \hbar = 1$. With this convention there is only one dimensional unit; for example, one may choose length, in which case mass parameters have dimension $[length]^{-1}$, or mass can be adopted as the basic unit so that length and time have dimension $[mass]^{-1}$. The action is then dimensionless, so that Lagrangians have dimension $[mass]^4$. It is then easy to see that scalar and vector fields have dimension $[mass]$, whereas spinor fields have dimension $[mass]^{3/2}$. Observe that all parameters that we have introduced in the above Lagrangians have *positive* mass dimension. This fact is important for the quantum mechanical properties of these theories. If quantum field theories have parameters with *negative* mass dimension, then the theory is not *renormalizable*. Usually this implies that the theory does not lead to sensible predictions. An example of such a theory is Einstein's theory of gravitation, general relativity, which is very succesful as a classical field theory, but cannot be quantized consistently.

Let us end this section by defining the field equations corresponding to a given Lagrangian. As we have already mentioned above, Hamilton's principle implies that the field configurations for which the action has an extremum, must satisfy the equations of motion. These field equations are the so-called Euler-Lagrange equations. For a general Lagrangian $\mathcal{L}$, defined in terms of fields ϕ and first-order derivatives of fields $\partial_\mu\phi$ only, these equations read

$$\partial_\mu\left(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}\right) - \frac{\partial\mathcal{L}}{\partial\phi} = 0. \tag{1.24}$$

Problem 1:
Verify that the field equations corresponding to the Lagrangians (1.8), (1.14), (1.16) and (1.20-21) are given by

$$(\partial^2 - m^2)\phi = 3\lambda\,\phi^2 + 4g\,\phi^3, \tag{1.25}$$

$$(\partial^2 - m^2)\phi = 2g\,|\phi|^2\phi, \tag{1.26a}$$

$$(\partial^2 - m^2)\phi^* = 2g\,|\phi|^2\phi^*, \tag{1.26b}$$

$$(\not\partial + m)\psi = G_s\,\phi_s\,\psi + iG_p\,\varphi_p\,\gamma_5\psi, \tag{1.27a}$$

$$\overline{\psi}(\,\overleftarrow{\not\partial} + m) = G_s\,\phi_s\,\overline{\psi} + iG_p\,\phi_s\,\overline{\psi}\gamma_5, \tag{1.27b}$$

$$\partial^\nu(\partial_\mu V_\nu - \partial_\nu V_\mu) + M^2\,V_\mu = 0, \tag{1.28}$$

where $\partial^2 \equiv \partial^\mu\partial_\mu$.

Problem 2:
Add an extra source term $J^\mu A_\mu$ to the Lagrangian (1.22) and show that the corresponding field equations coincide with the inhomogeneous Maxwell equations (in the relativistic formulation)

$$\partial_\nu F^{\mu\nu} = J^\mu, \tag{1.29}$$

where the electromagnetic fields $F_{\mu\nu}$ are defined by $F_{\mu\nu} \equiv \partial_\mu A_\nu - \partial_\nu A_\mu$.

Problem 3:
Consider plane wave solutions, i.e., solutions proportional to $\exp(ik\cdot x)$, for the free field equations found in problem 1. Find the conditions for the momentum k_μ. Verify that the Proca Lagrangian gives rise to only *three* independent polarizations.

2. Feynman rules

In the previous section we have presented field theories in terms of an action or a Lagrangian. Such theories can be studied as classical field theories, and this is often done in perturbation theory. Ultimately we are interested in the quantum mechanical scattering amplitudes for elementary particles. Those amplitudes can also be evaluated in perturbation theory, for which there exists a convenient graphical representation in terms of so-called Feynman diagrams. Some of these Feynman diagrams will correspond to the same contributions that one would find for a classical field theory. Such diagrams have the structure of tree diagrams. This in contradistinction with diagrams that contain closed loops. Their contributions do not follow from classical field theory, but can only be understood within the context of *quantum* field theory.

In these lectures there will be no time to give a detailed derivation of the Feynman diagrams. We shall just define the Feynman rules, which tell you how to evaluate the complicated mathematical expressions corresponding to a Feynman diagram. We refer to the literature for explicit derivations. The rules are presented in a number of steps:

• *The theory*: Begin with a field theory defined in terms of an action, which is expressed as an integral over space-time of a Lagrangian.

Fig. 1. Propagator line

• *Propagators*: Calculate the propagators of the theory, which follow from the terms in the action that are quadratic in the fields. The quadratic terms define a matrix in momentum space which is diagonal in the momentum variables. Suppose we take the Lagrangian (1.8) as an example. The action is

$$\begin{aligned} S &= \int \mathrm{d}^4x\, \mathcal{L} \\ &= \int \mathrm{d}^4x\, \left[-\tfrac{1}{2}(\partial_\mu\phi)^2 - \tfrac{1}{2}m^2\phi^2 + \mathrm{O}(\phi^3)\right] \end{aligned} \tag{2.1}$$

Now express the action in terms of the Fourier transforms of the fields The terms quadratic in the fields are then equal to

$$S = -\tfrac{1}{2}(2\pi)^4 \int \mathrm{d}^4k\, \phi^*(k)\left[k^2 + m^2\right]\phi(k), \tag{2.2}$$

where we have made use of the fact that we are dealing with real fields (i.e. $\phi^*(k) = \phi(-k)$). Hence the elements of the diagonal matrix are just equal to $-\frac{1}{2}(2\pi)^4[k^2 + m^2]$. For real fields the propagator is defined as a factor $\frac{1}{2}i$ times the inverse of this matrix. For the case at hand we thus find

$$\Delta(k) = \frac{1}{i(2\pi)^4}\frac{1}{k^2 + m^2 - i\varepsilon}\,. \tag{2.3}$$

Its graphical representation is a line, with an arrow indicating the momentum flow, while the endpoints refer to two space-time points (see Fig. 1). The $i\varepsilon$-term defines how to deal with the pole at $k^2 = -m^2$; the limit $\varepsilon \downarrow 0$ should only be taken at the end of the calculations. This prescription for dealing with the propagator poles is crucial for the causality and the unitarity (i.e. probability conservation) of the resulting theory.

We have already pointed out that the normalization factors are different for complex fields. In that case the kinetic terms in the action are

$$\begin{aligned} S &= \int \mathrm{d}^4x\, \mathcal{L} \\ &= \int \mathrm{d}^4x\, \left[-\mid \partial_\mu\phi\mid^2 - m^2 \mid\phi\mid^2 + \mathrm{O}(\mid\phi\mid^4)\right], \end{aligned} \tag{2.4}$$

which, in terms of the Fourier transforms of the fields, leads to (we no longer have $\phi^*(-k) = \phi(-k)$)

$$S = -(2\pi)^4 \int \mathrm{d}^4k\, \phi^*(k)\left[k^2 + m^2\right]\phi(k). \tag{2.5}$$

The propagator is now defined by the inverse of $-(2\pi)^4[k^2 + m^2]$ multiplied by a factor i. This leads to the same diagram as for real fields, but now the arrow also indicates that the propagator

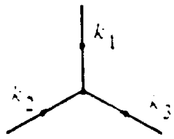
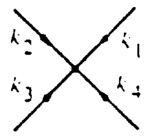
	$\frac{1}{i(2\pi)^4}\frac{1}{k^2+m^2-i\varepsilon}$
	$i(2\pi)^4\,\delta^4(k_1+k_2+k_3)(-\lambda)$
	$i(2\pi)^4\,\delta^4(k_1+k_2+k_3+k_4)(-g)$

Table 1. Feynman rules for the Lagrangian (1.8)

is oriented in the sense that the endpoints of the propagator lines refer to independent fields, namely φ and φ^*; the standard convention is that incoming arrows refer to φ, and outgoing ones to φ^*. Of course, complex fields can always be regarded as a linear combination of two real fields, and by decomposing $\varphi = \frac{1}{2}\sqrt{2}(\phi_1 + i\varphi_2)$ one makes contact with the description given for real fields.

• *Vertices*: The next step is to define the vertices of the graphs. We associate a vertex with n lines with every term in the Lagrangian that contains n field. A Lagrangian with a φ^3-term thus yields a vertex with three lines. Translational invariance ensures that the Fourier transform yields a delta-function in momentum space, thus guaranteeing energy-momentum conservation. Each vertex therefore has the structure

$$\text{vertex} = i(2\pi)^4\delta^4(\textstyle\sum_j k_j) \times (\text{coefficient of } \varphi^n \text{ in the Lagrangian}), \tag{2.6}$$

where our conventions are such that the k_j denote *incoming* momenta associated with each of the fields. For example, the Feynman rules for the theory described by the Lagrangian (1.8) are summarized in Table 1.

If the vertices in the Lagrangian contain derivatives then each differentiation of the fields contributes a factor ik_j to the vertex where k_j is the *incoming* momentum of the jth line. These momentum factors are part of the coefficient indicated in the generic definition (2.6). Thus the terms $g\varphi^3$ and $g\varphi(\partial_\mu\phi)^2$ both correspond to three-point vertices, but yield different factors: $i(2\pi)^4 g\,\delta^4(k_1+k_2+k_3)$ and $i(2\pi)^4 g\,(-k_2\cdot k_3)\,\delta^4(k_1+k_2+k_3)$, respectively. In the latter case, choosing the second and the third momentum as those corresponding to the differentiated fields, is arbitrary. A complete calculation must also include other possible line attachments, so that also factors $(-k_1\cdot k_3)$ and $(-k_1\cdot k_2)$ will contribute. The way in which these contributions must be summed will be discussed next, but it is rather obvious in this case that the total contribution of the second interaction becomes proportional to $(k_1\cdot k_2 + k_2\cdot k_3 + k_3\cdot k_1)$.

In the Feynman diagrams for complex fields the lines at the vertices carry an orientation: recall that fields correspond to lines with incoming arrows and their complex conjugates to lines with outgoing arrows. A formulation in terms of complex rather than real fields is useful if the theory is invariant under phase transformations, i.e.

$$\phi \to \phi' = e^{i\alpha}\phi.$$

In that case every interaction must contain an equal number of fields ϕ and their complex conjugates ϕ^*, so that each vertex has an equal number of incoming and outgoing lines. The lines coming from the vertices can now only be joined if their orientational arrows match (the orientation often corresponds to the flow of electric charge; obviously, charge will be conserved if the number of incoming and outgoing arrows is the same at each vertex).

• *Diagrams*: One now joins all the lines emanating from the vertices via propagators in order to form the various diagrams. The momentum flow through the various lines is determined by the momentum-conserving δ-functions at the vertices, and for real fields one may readjust the arrows in order to reflect this fact. If the arrow denotes more than just the momentum assignment of the line, but also the orientation (e.g. charge flow) then the lines cannot always be joined, and the number of possible diagrams will be reduced.

• *Summing and combinatorics*: Finally one sums over all possible diagrams with the same configuration of external lines. In order to do so one must determine the combinatorial weight factor associated with each of the diagrams. In principle this weight factor counts the number of ways in which a diagram can be formed by connecting vertices to propagators and external lines, but diagrams that only differ in the position of the vertices are counted as identical (because we ultimately integrate over all vertex positions in space-time).

There is only one exception to the above counting argument. If identical vertices occur in an indistinguishable way, i.e. not distinguished by their attachments to external lines, then one must avoid overcounting by dividing by $n!$, where n is the number of such indistinguishable vertices.

These rules also apply to diagrams with closed loops. However, in this case not all the momenta of the internal lines are fixed by momentum conservation, and one is left with one or more unrestricted momenta over which one should integrate. Likewise, we must also sum over all types of internal lines that are possible. Therefore, we have to sum over all components of vector and spinor fields that are possible.

From these rules it is in principle straightforward to write down Feynman diagrams and their corresponding mathematical expressions. To get acquainted with their use, I recommend that you start from a simple Lagrangian, such as (1.8), and calculate some diagrams. The following problems contain a few suggestions.

Problem 4:

Consider the tree diagrams with four external lines that follow from the Lagrangian (1.8), and calculate the corresponding expressions. There are four Feynman diagrams, each giving rise to a characteristic dependence on the momenta associated with the external lines.

Problem 5:

Determine the three possible one-loop diagrams following from the Lagrangian (1.8) with two external lines. Write down the corresponding expressions and verify whether the momentum integrals are well defined.

3. Photons

We would now like to derive the Feynman rules for theories that involve *masless* spin-1 fields. In this section we refer to the particles described by these fields as photons. As it turns out there are a number of technical complications for theories with photons. Let us start by recalling the Lagrangian for massless spin-1 fields,

$$\mathcal{L} = -\tfrac{1}{4}(\partial_\mu A_\nu - \partial_\nu A_\mu)^2, \tag{3.1}$$

which is invariant under *local* gauge transformations

$$A_\mu(x) \to A_\mu(x) + \partial_\mu \xi(x). \tag{3.2}$$

This transformation is familiar from Maxwell's theory of electromagnetism where the vector potential is subject to the same transformations. The electromagnetic field strength is then equal to

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \tag{3.3}$$

The main consequence of an invariance under *local* gauge transformations is that the theory depends on a smaller number of fields. Correspondingly the number of plane wave solutions is also reduced in comparison to the massive case. To see this explicitly consider the field equation following from (3.1),

$$\partial^\nu(\partial_\nu A_\mu - \partial_\mu A_\nu) = 0. \tag{3.4}$$

In order to examine plane wave solutions of this equation we take the Fourier transform of $A_\mu(x)$

$$A_\mu(k) = (2\pi)^{-4} \int d^4x\, A_\mu(x)\, e^{-ik\cdot x}. \tag{3.5}$$

Under gauge transformations $A_\mu(k)$ changes by a vector proportional to k_μ

$$A_\mu(k) \longrightarrow A'_\mu(k) = A_\mu(k) + \xi(k)\, k_\mu. \tag{3.6}$$

The field equation (3.4) now takes the form

$$k^2 A_\mu(k) - k_\mu k^\nu A_\nu(k) = 0. \tag{3.7}$$

which is manifestly invariant under the transformation (3.6). Decomposing $A_\mu(k)$ into four independent vectors, $\varepsilon_\mu(\mathbf{k},\lambda), k_\mu$ and $\tilde{k}_\mu$, defined by

$$\begin{aligned} k_\mu \varepsilon_\mu(\mathbf{k},\lambda) &= \varepsilon_0(\mathbf{k},\lambda) = 0, \qquad (\lambda = 1,2) \\ k_\mu &= (k_0, \mathbf{k}), \qquad \tilde{k}_\mu = (-k_0, \mathbf{k}). \end{aligned} \tag{3.8}$$

we may write

$$A_\mu(k) = a^\lambda(k)\,\varepsilon_\mu(\mathbf{k},\lambda) + b(k)\,\tilde{k}_\mu + c(k)\,k_\mu. \tag{3.9}$$

The field equation (3.7) then implies

$$k^2 a^\lambda(k)\,\varepsilon_\mu(\mathbf{k},\lambda) + b(k)\left[k^2\tilde{k}_\mu - (k\cdot\tilde{k})k_\mu\right] = 0, \tag{3.10}$$

from which we infer for the coefficient functions (note that $k\cdot\tilde{k}$ is positive)

$$k^2\,a^\lambda(k) = 0 \qquad \text{and} \qquad b(k) = 0. \tag{3.11}$$

The field equation does not lead to any restriction on $c(k)$. This should not come as a surprise because $c(k)$ can be changed arbitrarily by a gauge transformation, whereas the field equation is gauge invariant. Consequently the field equation cannot fix the value of $c(k)$. By means of a gauge transformation we may adjust $c(k)$ to zero, which shows that $c(k)$ has no physical meaning. We thus find that there are only *two* independent plane wave solutions characterized by lightlike momenta ($k^2 = 0$) and transverse polartizations.

The fact that massless particles have fewer polarization states than massive ones, can also be understood as follows. For a massive spin-s particle one can always choose to work in the rest frame, where the four momentum of the particle remains unchanged under ordinary spatial rotations, so that its spin degrees of freedom transform according to a $(2s+1)$-dimensional representation of the rotation group $SO(3)$. In other words, there are $2s+1$ polarization states, which transform among themselves under rotations, and which can be distinguished in the standard way by specifying the value of the spin projected along a certain axis. However, for massless particles it is not possible to go to the rest frame and one is forced to restrict oneself to two-dimensional rotations around the direction of motion of the particle. These rotations constitute the group $SO(2)$ (actually, the group of transformations that leave the particle momentum $k_\mu = (\omega(\mathbf{k}),\mathbf{k})$ invariant is somewhat larger, but the extra (noncompact) symmetries must act trivially on the particle states in order to avoid infinite-dimensional representations). The group $SO(2)$ has only one-dimensional complex representations. For spin s these representations just involve the states with spin (i.e. helicity) $\pm s$ in the direction of motion of the particle. Consequently massless particles have only two polarization states, irrespective of the value of their spin.

There is a further difficulty when one attempts to calculate Feynman diagrams for massless spin-1 particles, which is again related to the invariance under gauge transformations. To show this we rewrite (3.1) in the momentum representation,

$$S[A_\mu] = -\tfrac{1}{2}(2\pi)^4 \int d^4k\; A^*_\mu(k)\left[k^2\eta^{\mu\nu} - k^\mu k^\nu\right] A_\nu(k). \tag{3.12}$$

According to the general prescription given in section 2, the propagator is proportional to the inverse of $(k^2\eta_{\mu\nu} - k_\mu k_\nu)$. In this case, however, the inverse does not exist because this matrix has a zero eigenvalue, as we see from

$$(k^2\eta_{\mu\nu} - k_\mu k_\nu)k^\nu = 0. \tag{3.13}$$

The presence of the zero eigenvalue is a direct consequence of the gauge invariance of the theory. Gauge invariance implies that the theory contains fewer degrees of freedom: this fact reflects itself in the presence of zero eigenvalues in the quadratic part of the Lagrangian. Indeed, the null vector associated with the zero eigenvalue is proportional to k_μ, which according to (3.6) characterizes gauge transformations in momentum space. Obviously, the degree of freedom that is absent in (3.12) should not reappear through the interactions. One can show that this is ensured provided that the photon couples to a conserved current.

The standard way to circumvent the singular propagator problem is to make use of a so-called gauge condition. A convenient procedure amounts to explicitly introducing the missing (gauge) degrees of freedom, which formally spoils the gauge invariance. However, the degrees of freedom are introduced only in order to make the propagator well-defined and they will not affect the interactions of the theory. Therefore, the effect of this procedure can still be separated from the true gauge invariant part of the theory, and the physical consequences remain unchanged. It is a rather subtle matter to prove that this is indeed the case. In this section we only present the prescription for defining the propagator. To do that one introduces a so-called "gauge-fixing" term to the Lagrangian. The most convenient choice is to add (3.1)

$$\mathcal{L}_{\text{g.f.}} = -\tfrac{1}{2}(\lambda\,\partial_\mu A^\mu)^2, \tag{3.14}$$

where λ is an arbitrary parameter. Because of this term the Fourier transform of the action corresponding to the combined Lagrangian becomes

$$S[A_\mu] = -\tfrac{1}{2}(2\pi)^4 \int \mathrm{d}^4k\; A^*_\mu(k)\left[k^2\eta^{\mu\nu} - k^\mu k^\nu + \lambda^2 k^\mu k^\nu\right] A_\nu(k), \tag{3.15}$$

so that for $\lambda \neq 0$ the propagator is equal to

$$\begin{aligned}\Delta_{\mu\nu}(k) &= \frac{1}{i(2\pi)^4}\left[k^2\eta_{\mu\nu} - (1-\lambda^2)k_\mu k_\nu\right]^{-1} \\ &= \frac{1}{i(2\pi)^4}\frac{1}{k^2}\left(\eta_{\mu\nu} - (1-\lambda^{-2})\frac{k_\mu k_\nu}{k^2}\right).\end{aligned} \tag{3.16}$$

Clearly the propagator has more poles at $k^2 = 0$ than there are physical photons (characterized by transversal polarizations). However, one must realize that by making the above modification we have somewhat obscured the relation between propagator poles and physical particles. In order to extract the physical content of the theory one should only consider transversal polarizations. This requirement forms an essential ingredient of the proof that physical results do not depend on the parameter λ.

Using the propagator (3.16) one can now construct Feynman diagrams and corresponding scattering and decay amplitudes for photons in the standard fashion. The λ-dependent $k_\mu k_\nu$-term of the propagator residue vanishes when contracting the invariant amplitude with transversal polarization vectors. In order to sum over photon polarizations one may use (for orthonormal polarization vectors)

$$\sum_{\lambda=1,2} \varepsilon_\mu(\mathbf{k},\lambda)\varepsilon^*_\nu(\mathbf{k},\lambda) = \begin{cases}\delta_{\mu\nu} - \dfrac{k_\mu k_\nu}{|\mathbf{k}|^2} & \text{for } \mu,\nu = 1,2,3 \\ 0 & \text{for } \mu \text{ and/or } \nu = 0\end{cases} \tag{3.17}$$

An alternative form is

$$\sum_{\lambda=1,2} \varepsilon_\mu(\mathbf{k},\lambda)\,\varepsilon_\nu^*(\mathbf{k},\lambda) = \eta_{\mu\nu} - \frac{k_\mu \tilde{k}_\nu + \tilde{k}_\mu k_\nu}{k\cdot\tilde{k}}. \tag{3.18}$$

Obviously (3.17) and (3.18) are not manifestly Lorentz covariant, which is related to the fact that the transversality condition $\mathbf{k}\cdot\varepsilon(\mathbf{k},\lambda) = 0$ is not Lorentz invariant. However, if the photons couple to conserved currents, such that the amplitude vanishes when contracted with the photon momentum,

$$k^\mu \mathcal{M}_\mu = 0, \tag{3.19}$$

the noncovariant terms in (3.18) may be dropped. Consequently, when summing $|\mathcal{M}_\mu \varepsilon^\mu(\mathbf{k},\lambda)|^2$ over the transverse polarizations, one has

$$\sum_{\lambda=1,2} |\mathcal{M}_\mu \varepsilon^\mu(\mathbf{k},\lambda)|^2 = \mathcal{M}^\mu \mathcal{M}_\mu, \tag{3.19}$$

which is manifestly Lorentz invariant. We will return to this aspect in section 6.

Problem 6:

To demonstrate that the gauge-fixing term only introduces an extra degree of freedom into the theory that does not interfere with the interactions, consider the Maxwell theory coupled to some conserved current. After addition of the gauge-fixing term (3.14), show that $\partial\cdot A$ satisfies the free massless Klein-Gordon equation, so that the effect of the gauge-fixing term decouples from the rest of the theory.

Problem 7:

To prove the result (3.16) parametrize the propagator as $\Delta_{\mu\nu}(k) = A(k)\,\eta_{\mu\nu} + B(k)\,k_\mu k_\nu$, and solve the equation

$$\left[k^2\eta^{\mu\nu} - k^\mu k^\nu + \lambda^2 k^\mu k^\nu\right]\Delta^{\nu\rho}(k) = \delta^\rho_\mu. \tag{3.20}$$

4. Annihilation of spinless particles by electromagnetic interaction

To get acquainted with the use of Feynman diagrams we will consider the annihilation reaction

$$P^+ + P^- \to S^+ + S^- \tag{4.1}$$

mediated by a virtual photon in tree approximation. The particle $P^\pm$ and $S^\pm$ are hypothetical pointlike particles with no spin. A more relevant reaction is $e^+e^- \to \mu^+\mu^-$, but considering the reaction (4.1) enables us to discuss the characteristic features of such process without having to discuss some of the technical complications related to spin-$\frac{1}{2}$ particles.

As a first step we consider the coupling of a complex spinless field ϕ to photons. This is done by performing the so-called minimal substitution $\partial_\mu\phi \to \partial_\mu\phi - ieA_\mu\,\phi$ in the free Klein-Gordon Lagrangian (cf. (1.14)), where ϕ is a complex scalar field and $\pm e$ is the electric charge of the particle associated with ϕ. Combining this with Maxwell's Lagrangian gives

$$\begin{aligned}\mathcal{L} &= -\tfrac{1}{4}F_{\mu\nu}F^{\mu\nu} - |(\partial_\mu - ieA_\mu)\phi|^2 - m^2|\phi|^2 \\ &= -\tfrac{1}{4}F_{\mu\nu}F^{\mu\nu} - \partial_\mu\phi^*\partial^\mu\phi - m^2\phi^*\phi \\ &\quad - ieA^\mu\left[\phi^*(\partial_\mu\phi) - (\partial_\mu\phi^*)\phi\right] - e^2A_\mu^2\phi^*\phi.\end{aligned} \tag{4.2}$$

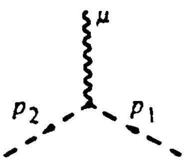
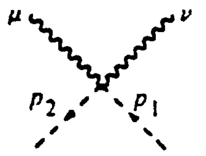
	$\frac{1}{i(2\pi)^4}\frac{1}{p^2+m^2}$
	$\frac{1}{i(2\pi)^4}\frac{1}{k^2}\left(\eta_{\mu\nu}-(1-\lambda^{-2})\frac{k_\mu k_\nu}{k^2}\right)$
	$i(2\pi)^4\,\delta^4(p_1+k-p_2)(-ie)(ip_1+ip_2)_\mu$
	$i(2\pi)^4\,\delta^4(p_1-p_2+k_1+k_2)(-e^2)\eta_{\mu\nu}$

Table 2. Feynman rules for the Lagrangian (4.2)

An important property of the Lagrangian (4.2) is its invariance under the combined gauge transformations

$$A_\mu(x)\to A_\mu(x)+\partial_\mu\xi(x),\qquad \phi(x)\to e^{ie\xi(x)}\phi(x). \tag{4.3}$$

This aspect will be extensively discussed in subsequent sections.

The propagators and vertices implied by (4.2) are shown in Table 2. The arrow on the pion line indicates the flow of (positive) charge rather than the momentum. We choose conventions such that an outgoing arrow on an external line indicates the emission of a positively charged particle absorption of its negatively-charged antiparticle. Combinatorial factors have not been included in the expressions for the vertices.

Assume that the particles $P^\pm$ are associated with the field ϕ introduced above. For the particle $S^\pm$ we introduce a separate field χ, which has the same interactions with the photon (and thus the same electric charge) as ϕ, but a different mass denoted by M. Observe that χ is now subject to the same gauge transformation as ϕ.

Consider now the diagram shown in Fig. 2, which describes the reaction (4.1) in lowest order. Observe that p_1 and p_2 refer to the momenta of the incoming particles P^+ and P^-, while q_1 and q_2 refer to the outgoing particles S^+ and S^-, respectively. Extracting overall factor of $i(2\pi)^4$ and a momentum-conserving δ-function, the invariant amplitude is given by

$$\mathcal{M}=e^2(p_1-p_2)^\mu\frac{1}{(p_1+p_2)^2}\left(\delta_\mu^\nu-(1-\lambda^{-2})\frac{(p_1+p_2)_\mu(q_1+q_2)^\nu}{(p_1+p_2)^2}\right)(q_1-q_2)_\nu, \tag{4.4}$$

where we have set the photon momentum equal to $k_\mu=(p_1+p_2)_\mu=(q_1+q_2)_\mu$. A first important observation is that the gauge-dependent part of the photon propagator vanishes when the pions are taken on the mass shell, because $(p_1+p_2)\cdot(p_1-p_2)=p_1^2-p_2^2=0$ and $(p_1-p_2)\cdot(p_1+p_2)=$

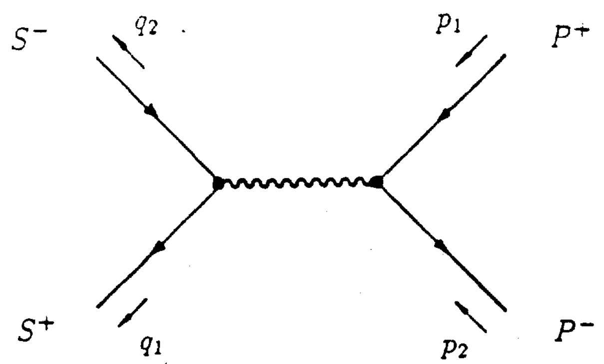


Fig. 2. Lowest-order Feynman diagram for the reaction (4.1)

$p_1^2 - p_2^2 = 0$. This confirms that the physical consequences of the theory have not been affected by introducing the gauge-fixing term into the Lagrangian.

Introducing Mandelstam variables

$$s = -(p_1 + p_2)^2, \qquad t = -(p_1 - q_1)^2, \qquad u = -(p_1 - q_2)^2, \tag{4.5}$$

which satisfy

$$s + t + u = 2m^2 + 2M^2, \tag{4.6}$$

the amplitude can be written in a simple form

$$\mathcal{M} = e^2 \frac{u - t}{s}. \tag{4.7}$$

In the centre-of-mass frame t and u are expressed in terms of s and the scattering angle θ between $\mathbf{p}_1$ and $\mathbf{q}_1$:

$$\begin{aligned} t &= -\tfrac{1}{2}s + m^2 + M^2 + \tfrac{1}{2}\sqrt{(s - 4m^2)(s - 4M^2)}\cos\theta, \\ u &= -\tfrac{1}{2}s + m^2 + M^2 - \tfrac{1}{2}\sqrt{(s - 4m^2)(s - 4M^2)}\cos\theta, \end{aligned} \tag{4.8}$$

We now use the general formula for the differential cross section for a quasi-elastic scattering reaction $1 + 2 \to 3 + 4$,

$$\frac{d\sigma}{d\Omega_{\mathrm{CM}}} = \frac{1}{64\pi^2}\frac{1}{s}\sqrt{\frac{\lambda(s, m_3^2, m_4^2)}{\lambda(s, m_1^2, m_2^2)}}\,|\mathcal{M}|^2, \tag{4.9}$$

where m_1-m_4 denote the masses of the particles 1-4, and the function λ is defined by

$$\lambda(x, y, z) = x^2 + y^2 + z^2 - 2xy - 2xz - 2yz. \tag{4.10}$$

Application of the above formulae gives rise to

$$\frac{d\sigma}{d\Omega_{\mathrm{CM}}} = \frac{\alpha^2}{4s}\sqrt{\frac{s - 4M^2}{s - 4m^2}}\left(1 - \frac{4m^2}{s}\right)\left(1 - \frac{4M^2}{s}\right)\cos^2\theta. \tag{4.11}$$

Here, α denotes the fine structure constant $\alpha = e^2/4\pi$. Integration over the angles gives the total cross section

$$\sigma = \frac{\pi\alpha^2}{3s}\sqrt{\frac{s - 4M^2}{s - 4m^2}}\left\{1 - \frac{4(m^2 + M^2)}{s} + \frac{16m^2M^2}{s^2}\right\}. \tag{4.12}$$

For $s \gg m^2, M^2$, these results become

$$\frac{d\sigma}{d\Omega_{CM}} = \frac{\alpha^2}{4s}\cos^2\theta. \tag{4.13}$$

and

$$\sigma = \frac{\pi\alpha^2}{3s}. \tag{4.14}$$

For comparison we give the corresponding expression for $e^+e^- \to \mu^+\mu^-$,

$$\frac{d\sigma}{d\Omega_{CM}} = \frac{\alpha^2}{4s}\sqrt{\frac{s-4M^2}{s-4m^2}}\left\{1 + \frac{4(m^2+M^2)}{s} + \left(1-\frac{4m^2}{s}\right)\left(1-\frac{4M^2}{s}\right)\cos^2\theta\right\}, \tag{4.15}$$

where we have averaged over the (incoming) electron and summed over the (outgoing) muon spins. Observe that the $\cos\theta$ dependent terms coincide with (4.11). After integration over the angles, one obtains the total cross-section

$$\sigma = \frac{4\pi\alpha^2}{3s}\sqrt{\frac{s-4M^2}{s-4m^2}}\left\{1 + \frac{2(m^2+M^2)}{s} + \frac{4m^2M^2}{s^2}\right\}, \tag{4.16}$$

where m is the electron mass and M the muon mass. When $E \gg M, m$, as is usually the case, one finds the well-known results

$$\frac{d\sigma}{d\Omega_{CM}} = \frac{\alpha^2}{4s}(1+\cos^2\theta), \tag{4.17}$$

and

$$\sigma = \frac{4\pi\alpha^2}{3s}. \tag{4.18}$$

5. Gauge theory of $U(1)$

In the previous section we have considered scalar electrodynamics, the theory of photons coupled to charged spinless fields. This is one of the simplest examples of an interacting field theory based on gauge invariance. Such theories are called gauge theories. In this section we will present the main ingredients of these theories. For simplicity we will first consider the case of abelian gauge transformations, i.e. gauge transformations that commute. In later sections we shall also discuss theories based on nonabelian gauge groups.

As a first example let us construct a field theory which is invariant under local phase transformations. Our starting point is the free Dirac Lagrangian

$$\mathcal{L}_\psi = -\overline{\psi}\partial\!\!\!/\psi - m\,\overline{\psi}\psi, \tag{5.1}$$

which is obviously invariant under rigid phase transformations, i.e. phase transformations which are the same at each point in space-time. This is so because

$$\psi \to \psi' = e^{iq\xi}\psi, \tag{5.2}$$

implies

$$\overline{\psi} \to \overline{\psi}' = e^{-iq\xi}\overline{\psi}. \tag{5.3}$$

Here we have introduced a parameter q that measures the strength of the phase transformations, because eventually we want to simultaneously consider fields transforming with different strengths. Phase transformations generate the group of 1×1 unitary matrices called $U(1)$.

Let us now consider local phase transformations, and verify whether (5.1) remains invariant. Of course, the local aspect of the transformation is not important for the invariance of the mass term, since the variation of that term only involves the transformation of fields taken at the same point in space-time. But as soon as we compare fields at different points in space-time the local character of the transformation is crucial. A derivative, which depends on the variation of the fields in an infinitesimally small neighbourhood, will be subject to transformations at neighbouring space-time points. To see the effect of this, let us evaluate the effect of a local transformation on $\partial_\mu\psi$:

$$\begin{aligned}\partial_\mu\psi(x) \to (\partial_\mu\psi(x))' &= \partial_\mu(e^{iq\xi(x)}\psi(x)) \\ &= e^{iq\xi(x)}(\partial_\mu\psi(x) + iq\partial_\mu\xi(x)\psi(x)).\end{aligned} \tag{5.4}$$

Clearly $\partial_\mu\psi$ does not have the same transformation rule as ψ itself. There is an extra term induced by the transformations at neighbouring space-time points which is proportional to the derivative of the transformation parameter. This term is responsible for the lack of invariance of the Lagrangian (5.1).

In order to make (5.1) invariant under local phase transformations, one may consider the addition of new terms whose variation will compensate for the $\partial_\mu\xi$ term in (5.4). As a first step one could attempt to construct a modified derivative D_μ transforming according to

$$D_\mu\psi(x) \to (D_\mu\psi(x))' = e^{iq\xi(x)}(D_\mu\psi(x)). \tag{5.5}$$

If such a derivative exists we can then simply replace the ordinary derivative ∂_μ in the Lagrangian (5.1) by D_μ and preserve invariance under local phase transformations.

Since the transformation of $D_\mu\psi$ is entirely determined by the transformation parameter at the same space-time coordinate as ψ, D_μ is called a *covariant derivative*. To appreciate this definition one should realize that a local phase transformation may be regarded as a product of independent phase transformations each acting at a separate space-time point. It is possible that local quantities transform only under the gauge transformation taken at the same space-time point. Such quantities are then said to transform *covariantly*. For instance, according to this nomenclature, the field ψ transforms in a covariant fashion under local phase transformations, whereas the transformation behaviour of ordinary derivatives (cf.(5.4)), although correctly representing the action of the full local group, is clearly noncovariant. It is obviously convenient to have local quantities that transform covariantly, and this is an extra motivation for introducing the covariant derivative.

Let us now turn to an explicit construction of the covariant derivative. Comparing (5.4) and (5.5) we note that the modified derivative D_μ must contain a quantity whose transformation can compensate for the second term in (5.4). If we define

$$D_\mu\psi(x) = (\partial_\mu - iqA_\mu(x))\psi(x), \tag{5.6}$$

we obtain

$$\begin{aligned} D_\mu\psi \to (D_\mu\psi)' &= (\partial_\mu\psi)' - iq(A_\mu\psi)' \\ &= e^{iq\xi}(\partial_\mu\psi + iq\partial_\mu\xi\psi - iqA'_\mu\psi). \end{aligned} \tag{5.7}$$

Comparing this to (5.5) shows that the new quantity A_μ must have the following transformation rule

$$A_\mu \to A'_\mu = A_\mu + \partial_\mu\xi. \tag{5.8}$$

Hence the requirement of local gauge invariance has led us to introduce a new field A_μ, whose transformation is given by (5.8). This new field is called a *gauge field.* Note that the gauge field does *not* transform in a covariant fashion.

Introducing the covariant derivative (5.6) into the Lagrangian (5.1) shows that the theory is no longer free, but describes interactions of the fermions with the gauge field

$$\begin{aligned} \mathcal{L}_\psi &= -\bar\psi \not{D}\psi - m\bar\psi\psi, \\ &= -\bar\psi\not\partial\psi - m\bar\psi\psi + iqA_\mu\bar\psi\gamma^\mu\psi. \end{aligned} \tag{5.9}$$

Usually one assumes that A_μ describes some new and independent degrees of freedom of the system, although this can sometimes be avoided. But in any case it is clear that the requirement of local gauge invariance leads to interacting field theories of a particular structure.

Covariant derivatives play an important role in theories with local gauge invariance, so we discuss them here in more detail. First we note that D_μ consists of two terms which are both related to an infinitesimal transformation. The derivative ∂_μ generates an infinitesimal displacement of the coordinates, $x^\mu \to x^\mu + a^\mu$, whereas the second term $-iqA_\mu\psi$ represents the variation under an infinitesimal gauge transformation $\delta\psi = iq\xi\psi$, with parameter $\xi = -A_\mu$. The combination of an infinitesimal transformation over a distance a_μ and a field-dependent gauge transformation with parameter $\xi = -a_\mu A_\mu$ is sometimes called a *covariant translation.* Under such a translation a field transforms as

$$\delta\psi = a^\mu D_\mu \psi. \tag{5.10}$$

The observation that D_μ corresponds to an infinitesimal variation shows that covariant derivatives must satisfy the Leibnitz rule, just as ordinary derivatives do

$$D_\mu(\psi_1\psi_2) = (D_\mu\psi_1)\psi_2 + \psi_1(D_\mu\psi_2). \tag{5.11}$$

To appreciate this result one should realize that the precise form of the covariant derivative is tied to the transformation character of the quantity on which it acts. For instance, if ψ_1 and ψ_2 transform under local phase transformations with strength q_1 and q_2, respectively, then we have

$$\begin{aligned} D_\mu(\psi_1\psi_2) &= (\partial_\mu - i(q_1+q_2)A_\mu)(\psi_1\psi_2), \\ D_\mu\psi_1 &= (\partial_\mu - iq_1A_\mu)\psi_1, \\ D_\mu\psi_2 &= (\partial_\mu - iq_2A_\mu)\psi_2. \end{aligned} \tag{5.12}$$

With these definitions it is straightforward to verify the validity of (5.11).

Of course, repeated application of covariant derivatives will always yield covariant quantities. This fact may be used to construct a new covariant object which depends only on the gauge fields. Namely, we apply the antisymmetric product of two derivatives on ψ

$$[D_\mu, D_\nu]\,\psi = D_\mu(D_\nu\psi) - D_\nu(D_\mu\psi). \tag{5.13}$$

Writing explicitly

$$D_\mu(D_\nu\psi) = \partial_\mu\partial_\nu\psi - iqA_\mu\partial_\nu\psi - iq(\partial_\mu A_\nu)\psi - iqA_\nu\partial_\mu\psi - q^2A_\mu A_\nu\psi, \tag{5.14}$$

one easily establishes

$$[D_\mu, D_\nu]\,\psi = -iqF_{\mu\nu}\psi, \tag{5.15}$$

where

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu. \tag{5.16}$$

However, since ψ transforms covariantly and the left-hand side of (5.15) is covariant, we may conclude that $F_{\mu\nu}$ is itself a covariant object. In fact, application of the gauge transformation (5.8) shows that $F_{\mu\nu}$ is even gauge *invariant*,

$$\delta F_{\mu\nu} = \partial_\mu\partial_\nu\xi - \partial_\nu\partial_\mu\xi = 0. \tag{5.17}$$

but, as we will see later, this is a coincidence related to the fact that $U(1)$ transformations are abelian.

The result (5.15) is called the *Ricci identity*. It specifies that the comutator of two covariant derivatives is an infinitesimal gauge transformation with parameter $\xi = -F_{\mu\nu}$, where $F_{\mu\nu}$ is called the field strength. This field strength $F_{\mu\nu}$ is sometimes called the *curvature* tensor. The reason for this nomenclature is not difficult to see: if the left-hand side of (5.15) were zero then two successive infinitesimal covariant translations, one in the $\hat{\mu}$ and the other in the $\hat{\nu}$ direction, would lead to the same result when applied in the opposite order. According to the Ricci identity this is not the case for finite $F_{\mu\nu}$. One encounters the same situation when considering translations on a curved surface, which do not commute for finite curvature. As the tensor $F_{\mu\nu}$ on the right-hand side of (5.15) measures the lack of commutativity, its effect is analogous to that of curvature.

We can use covariant derivatives to obtain yet another important identity. Consider the double commutators of covariant derivatives $[D_\mu, [D_\nu, D_\rho]]$. According to the Jacobi identity the cyclic combination vanishes identically,

$$[D_\mu, [D_\nu, D_\rho]] + [D_\nu, [D_\rho, D_\mu]] + [D_\rho, [D_\mu, D_\nu]] = 0, \tag{5.18}$$

as can be verified by writing all the terms. To see the consequence of (5.18) let us write the first term acting explicitly on $\psi(x)$,

$$\begin{aligned}[D_\mu, [D_\nu, D_\rho]]\psi &= D_\mu([D_\nu, D_\rho]\psi) - [D_\nu, D_\rho]D_\mu\psi \\ &= -iq(D_\mu F_{\nu\rho})\psi,\end{aligned} \tag{5.19}$$

where we used (5.11) and (5.15). Therefore the Jacobi identity implies

$$D_\mu F_{\nu\rho} + D_\nu F_{\rho\mu} + D_\rho F_{\mu\nu} = 0. \tag{5.20}$$

In this case the field strength is invariant under gauge transformations so we may replace covariant by ordinary derivatives and obtain

$$\partial_\mu F_{\nu\rho} + \partial_\nu F_{\rho\mu} + \partial_\rho F_{\mu\nu} = 0. \tag{5.21}$$

This result is called the *Bianchi identity*; it implies that $F_{\mu\nu}$ can be expressed in terms of a vector field, precisely in accord with (5.16). The identity (5.21) is well known in electrodynamics as the homogeneous Maxwell equations.

The field strength tensor can now be used to write a gauge invariant Lagrangian for the gauge field itself

$$\mathcal{L} = -\tfrac{1}{4}F_{\mu\nu}^2. \tag{5.22}$$

This Lagrangian can now be combined with (5.9),

$$\begin{aligned}\mathcal{L} &= \mathcal{L}_A + \mathcal{L}_\psi \\ &= -\tfrac{1}{4}(\partial_\mu A_\nu - \partial_\nu A_\mu)^2 - \bar{\psi}\partial\!\!\!/\psi - m\bar{\psi}\psi + iqA_\mu\bar{\psi}\gamma^\mu\psi,\end{aligned} \tag{5.23}$$

so that we have obtained an interacting theory of a vector field and a fermion field invariant under the combined local gauge transformations (5.2) and (5.8). It is not difficult to see that this theory coincides with electrodynamics: the gauge field A_μ is just the vector potential (subject to its familiar gauge transformation), which couples to the fermion field via the minimal substitution, and $F_{\mu\nu}$ is the electromagnetic field strength.

To derive the field equations corresponding to (5.23) is straightforward. They read

$$\partial^\nu F_{\mu\nu} = J_\mu, \tag{5.24}$$

$$(\not\partial + m)\psi = iq\not A\,\psi, \tag{5.25}$$

$$\bar\psi(\overleftarrow{\not\partial} - m) = -iq\bar\psi\not A, \tag{5.26}$$

where the right-hand side of (5.24) is equal to

$$J_\mu = iq\bar\psi\gamma_\mu\psi. \tag{5.27}$$

Clearly (5.24) corresponds to the inhomogeneous Maxwell equation (1.29), while, as was mentioned above, the Bianchi identity (5.21) coincides with the homogeneous Maxwell equations. The field equations (5.25) and (5.26) describe the dynamics of the charged fermion, and are strictly speaking not part of Maxwell's equations.

It is easy to repeat the above construction for other fields. For example, a complex scalar field ϕ may transform under local phase transformations according to

$$\phi(x) \to \phi'(x) = e^{iq\xi(x)}\phi(x). \tag{5.28}$$

As before, the requirement of local gauge invariance forces one to replace the ordinary derivative by a covariant derivative,

$$D_\mu\phi = (\partial_\mu - iqA_\mu)\phi, \tag{5.29}$$

so that one obtains a gauge invariant version of the Klein-Gordon Lagrangian

$$\begin{aligned}\mathcal{L}_\phi &= -|D_\mu\phi|^2 - m^2|\phi|^2 \\ &= -|\partial_\mu\phi|^2 - m^2|\phi|^2 - iqA_\mu\phi^*\overleftrightarrow{\partial_\mu}\,\phi - q^2A_\mu^2|\phi|^2.\end{aligned} \tag{5.30}$$

This is the Lagrangian for *scalar electrodynamics*, which we have been using in section 4. Observe that the effect of the covariant derivative coincides with that of the minimal substitution procedure. Unlike in *spinor electrodynamics*, which is defined by the Lagrangian (5.9), there are interaction terms in (5.30) that are quadratic in A_μ. Therefore the corresponding expression for the current (5.28) now depends also on the gauge field A_μ; it reads

$$J_\mu = iq((D_\mu\phi^*)\phi - \phi^*(D_\mu\phi)), \tag{5.31}$$

and appears on the right-hand side of the Maxwell equation (5.24). The field equation of the scalar field can be written as

$$D^\mu D_\mu\phi - m^2\phi = 0. \tag{5.32}$$

The parameter q that we have been using to indicate the relative magnitude of the change of phase caused by the gauge transformation, also determines the strength of the interaction with the gauge field A_μ. Hence in electromagnetism the particles described by the various fields carry electric charges $\pm q$. We shall give a more precise definition of the electric charge in section 7.

6. Current conservation

The four-vector current J_μ that appears on the right-hand side of the inhomogeneous Maxwell equation must satisfy an obvious restriction. To see this contract (5.24) with ∂^μ and use the fact that $F_{\mu\nu}$ is antisymmetric in μ and ν. It then follows that J_μ must be conserved. i.e.

$$\partial_\mu J^\mu = 0. \tag{6.1}$$

Another way to derive the same result is to make use of the matter field equations. For instance, for a fermion field one has

$$\begin{aligned}\partial_\mu J^\mu &= \partial_\mu(iq\overline{\psi}\gamma^\mu\psi)\\ &= iq\overline{\psi}(\not{D}\psi) + iq(\overline{\psi}\overleftarrow{\not{D}})\psi\\ &= 0,\end{aligned} \tag{6.2}$$

by virtue of (5.25) and (5.26). Similarly on can show that the current (5.31) is conserved by virtue of the field equation (5.32) and its complex conjugate.

The fact that photons must couple to a conserved current has direct consequences for invariant amplitudes that involve external photon lines. Such an amplitude with one photon and several other incoming and outgoing lines takes the form

$$\mathcal{M}(k,\cdots) = \varepsilon_\mu(\mathbf{k})\mathcal{M}^\mu(k,\cdots), \tag{6.3}$$

where k and $\varepsilon(\mathbf{k})$ denote the momentum and polarization vector of the photon. Current conservation now implies that

$$k_\mu\mathcal{M}^\mu(k,\cdots) = 0, \tag{6.4}$$

provided that all external lines other than that of the photon are on the mass shell. The latter condition can be understood as a consequence of the fact that we had to impose the matter field equation in (6.2).

Actually, the mass-shell condition for the external lines can be somewhat relaxed: it is sufficient to require that only the external lines associated with charged particles are on the mass shell. Therefore it is possible to exploit (6.4) for amplitudes with several off-shell photons, but we should already caution the reader that this result does not hold for nonabelian gauge theories. If all external lines are taken off shell one obtains relations between Green's functions, which have a more complicated structure than (6.4). Such identities are called *Ward identities*. In the context of quantum electrodynamics these identities are usually called Ward-Takahashi identities.

The fact that A_μ couples to a conserved current is essential in order to establish that interactions of the massless spin-1 particle associated with A_μ are Lorentz invariant. We have already observed this in one particular example in section 3. To explain this aspect in more detail, we recall that massless particles have fewer polarization states than massive ones. The spin of massless spin-s particles must be parallel or anti-parallel to the direction of motion of the particle, so that the helicity is equal to $\pm s$. Consequently massless particles have only two polarization states, irrespective of the value of their spin. Physical photons have therefore helicity ± 1, and are described by transverse polarization vectors $\varepsilon(\mathbf{k})$ that satisfy the condition

$$k\cdot\varepsilon(\mathbf{k}) = 0 \quad , \quad \varepsilon_0(\mathbf{k}) = 0. \tag{6.5}$$

corresponding to two linearly independent polarization vectors.

The second condition in (6.5) is obviously not Lorentz invariant, and one may question whether the interactions of massless spin-1 particles will be relativistically invariant. To make this more precise, consider a physical process involving a photon, for which the invariant amplitude takes the form

$$\mathcal{M} = \varepsilon_\mu(\mathbf{k})\mathcal{M}^\mu(k, ...), \tag{6.6}$$

where $\varepsilon(\mathbf{k})$ is the photon polarization vector, k_μ the photon momentum ($k^2 = 0$) and the dots indicate the other particle momenta that are relevant. Obviously, (6.6) has a Lorentz invariant form, but since $\varepsilon(\mathbf{k})$ will not remain transverse after a Lorentz transformation, the amplitude will in general no longer coincide with the expression (6.6) when calculated directly in the new frame.

To examine this question in more detail let us first derive how a transverse polarization vector transforms under Lorentz transformations. What we intend to prove is that $\varepsilon_\mu(\mathbf{k})$, satisfying (6.5) with $k^2 = 0$, transforms into a linear combination of a transverse vector $\varepsilon'(\mathbf{k}')$ and the transformed photon momentum k', where $\varepsilon'(\mathbf{k}')$ is transverse with respect to the new momentum k'. More precisely,

$$\varepsilon_\mu(\mathbf{k}) \rightarrow \varepsilon'_\mu(\mathbf{k}') + \alpha\, k'_\mu, \tag{6.7}$$

with α some unknown coefficient which depends on the Lorentz transformation. To derive (6.7) it is sufficient to note that the condition $k\cdot\varepsilon(\mathbf{k}) = 0$ is Lorentz invariant, so that the right-hand side of (6.7) should vanish when contracted with k'; therefore it follows that this vector can be decomposed into a transverse vector satisfying (6.5) (but now in the new frame) and the momentum k'. What remains to be shown is that the transverse vector $\varepsilon'(\mathbf{k}')$ has the same normalization as $\varepsilon(\mathbf{k})$. This is indeed the case, since

$$\begin{aligned} \varepsilon^\mu(\mathbf{k})\,\varepsilon_\mu(\mathbf{k}) &= \left(\varepsilon'^\mu(\mathbf{k}') + \alpha\, k'^\mu\right)\left(\varepsilon'_\mu(\mathbf{k}') + \alpha\, k'_\mu\right) \\ &= \varepsilon'^\mu(\mathbf{k}')\varepsilon'_\mu(\mathbf{k}'), \end{aligned} \tag{6.8}$$

where we have used (6.5) and $k'^2 = 0$. Using (6.7) one easily establishes that the amplitude (6.3) transforms under Lorentz transformations as

$$\varepsilon_\mu(\mathbf{k})\,\mathcal{M}^\mu(k,\cdots) \longrightarrow \varepsilon'_\mu(\mathbf{k}')\,\mathcal{M}'^\mu(k',\cdots) + \alpha\, k'_\mu\,\mathcal{M}'^\mu(k',\cdots). \tag{6.9}$$

The first term on the right-hand side corresponds precisely to the amplitude that one would calculate in the new frame. Therefore relativistic invariance is ensured provided that the photons couple to a conserved amplitude.

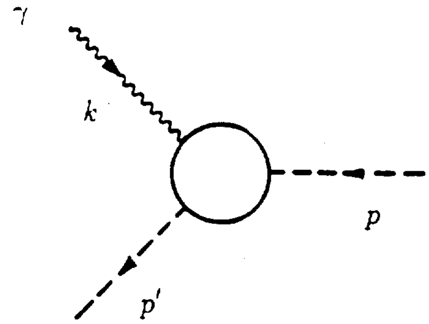


Fig. 3. The absorption of a virtual photon by a spinless particle

7. Conserved charges

In classical field theory current conservation implies that the charge associated with the current is locally conserved. For scattering and decay reactions of elementary particles it seems obvious that charge conservation should imply that the total charge of the incoming particles is equal to the total charge of the outgoing particles. It is the purpose of this section to prove that this is indeed the case and to establish that the charge of a particle can be defined in terms of the invariant amplitude for a particle to emit or absorb a zero-frequency photon. To elucidate this definition consider the amplitude for the absorption of a virtual photon with momentum k by a spinless particle. The corresponding diagram is shown in Fig. 3. The momenta of the incoming and outgoing particles are denoted by p and p', respectively, so that $k = p' - p$. Both p and p' refer to physical particles of the same mass, so that $p^2 = p'^2 = -m^2$. The invariant amplitude can generally be decomposed into two terms

$$\mathcal{M}_\mu(p',p) = F_1(k^2)(p'_\mu + p_\mu) + iF_2(k^2)k_\mu \tag{7.1}$$

where F_1 and F_2 are called form factors. Current conservation implies that $k_\mu \mathcal{M}^\mu(p',p)$ should vanish, so

$$F_1(k^2)(p'^2 - p^2) + iF_2(k^2)k^2 = 0. \tag{7.2}$$

Since the incoming and outgoing particles have the same mass, F_1 drops out from (7.2) and we are left with

$$F_2(k^2) = 0. \tag{7.3}$$

In order to obtain (7.3) it is essential that we assume current conservation for *off-shell* photons. The simplification resulting from this assumption will not imply a loss of generality in what follows, as we will mainly be dealing with physical photons. The function $F_1(k^2)$ is called the *charge form factor*, and as, we have been alluding to above, its value at $k^2 = 0$ *defines* the electric charge of the particle in question. For a pointlike particle this is easily verified, and one finds $F_2(k^2) = 0$ and $F_1(k^2) = e$, where e is the coupling constant in the Lagrangian that measures the strength of the photon coupling. Experimentally it is not possible to measure the probability for absorbing or emitting a zero-frequency photon, so that the charge of a particle is not measured in this way. It is more feasible to use low-energy Compton scattering (also called Thomson scattering) for this purpose. Another process is the elastic scattering of a particle by a Coulomb field.

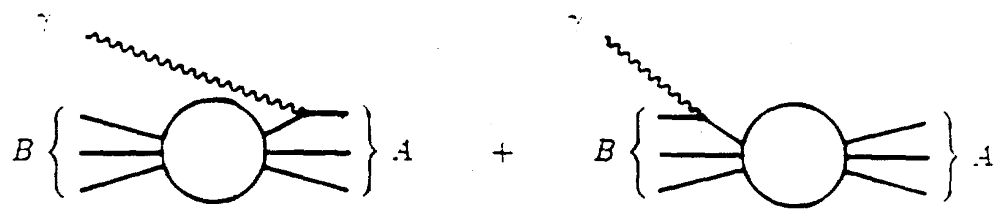


Fig. 4. Born approximation diagrams corresponding to (7.6)

We will now show that, with the above definition of charge, one has charge conservation in any possible elementary particle reaction. The derivation starts from considering a process in which a soft photon is being emitted or absorbed; for instance

$$A \to B + \gamma \tag{7.4}$$

where A and B denote an arbitrary configuration of incoming and outgoing particles. It is possible to divide the amplitude for this process into two terms

$$\mathcal{M}(A \to B + \gamma) = \mathcal{M}^{\mathrm{B}}(A \to B + \gamma) + \mathcal{M}^{\mathrm{R}}(A \to B + \gamma), \tag{7.5}$$

where $\mathcal{M}^{\mathrm{B}}$ consists of all the Born approximation diagrams in which the photon is attached to one of the external lines, as shown in Fig. 4, and $\mathcal{M}^{\mathrm{R}}$ represents the remainder. The Born approximation diagrams have the form

$$\mathcal{M}^{\mathrm{B}} = \varepsilon_\mu(\mathrm{k})\Big\{\sum_i \mathcal{M}(A[i] \to B)\,\frac{(2p_i - k)^\mu}{(p_i - k)^2 + m_i^2}\,F_i(k^2) + \sum_j \frac{(2p_j + k)^\mu}{(p_j + k)^2 + m_j^2}\,F_j(k^2)\,\mathcal{M}(A \to B[j])\Big\}, \tag{7.6}$$

where $\mathcal{M}(A[i] \to B)$ and $\mathcal{M}(A \to B[j])$ denote the invariant amplitude for the process $A \to B$ in which one of the external lines is shifted from its mass shell by an amount k_μ. The index i labels the off-shell incoming line with momentum $p_i - k$ and $p_i^2 = -m_i^2$, whereas j labels the outgoing line with momentum $p_j + k$ and $p_j^2 = -m_j^2$. The reason why we consider the Born approximation diagrams of Fig. 4 separately is that they become singular when the photon momentum k tends to zero because the propagator of the virtual particle diverges in that limit. Therefore we are entitled to restrict ourselves to the charge form factors $F_i(k^2)$ or $F_j(k^2)$ as they are measured for *real* particles, since the deviation from their on-shell value leads to terms in which the propagator pole cancels, and which are therefore regular if k approaches zero. Those terms are thus contained in the second part of (7.5) which is assumed to exhibit no singularities in the soft-photon limit.

We now use current conservation on the full amplitude (7.5), i.e., we require that the amplitude vanishes when the photon polarization vector $\varepsilon_\mu(\mathrm{k})$ is replaced by k_μ. Contracting the momentum factors in the Born approximation amplitude with k_μ leads to the following factors

$$k_\mu\frac{(2p_i - k)^\mu}{(p_i - k)^2 + m_i^2} = \frac{-(p_i - k)^2 + p_i^2}{(p_i - q)^2 + m_i^2} = -1, \tag{7.7}$$

$$k_\mu \frac{(2p_j + k)^\mu}{(p_j + k)^2 + m_j^2} = \frac{(p_j + k)^2 - p_j^2}{(p_j + k)^2 + m_j^2} = +1. \tag{7.8}$$

Using the fact that the remaining diagrams in (7.5) are regular for vanishing k, we thus find

$$-\sum_i \mathcal{M}(A[i] \rightarrow B)\, F_i(k^2) + \sum_j F_j(k^2)\, \mathcal{M}(A \rightarrow B[j]) = O(k), \tag{7.9}$$

or, in the soft-photon limit

$$\Big(\sum_j F_j(0) - \sum_i F_i(0)\Big)\, \mathcal{M}(A \rightarrow B) = 0, \tag{7.10}$$

where $\mathcal{M}(A \rightarrow B)$ is now the full on-shell amplitude for the process $A \rightarrow B$. The implication of (7.10) should be obvious. In every possible process the sum of the charges of the incoming particles should equal the sum of the charges of the outgoing particles. This result justifies the definition of electric charge as the charge form factor taken at zero momentum transfer.

8. Nonabelian gauge fields

In the previous chapter we have introduced theories with local gauge invariance. In order to demonstrate the essential ingredients we stayed primarily within the context of theories such as electrodynamics that are invariant under local phase transformations. However, the same framework can be applied to theories that are invariant under more complicated gauge transformations. One distinctive feature of the latter is that they depend on several parameters and also that they are not always commuting. Groups of noncommuting transformations are called *nonabelian*. This in contradistinction with phase transformations, which depend on a single parameter ξ and are obviously commuting (and therefore called abelian). The fact that the gauge transformations depend on several parameters forces us to introduce several independent gauge fields. Also the matter fields must have a certain multiplicity in order that the gauge transformations can act on them. In more mathematical terms, the fields should transform according to *representations* of the gauge group. Each representation consists of a set of fields which transform among themselves, just as the components of a three-dimensions vector transform among themselves under the group of rotations. The coupling of the gauge fields to matter will involve certain matrices, which will appear in the expressions for the charges. Charges can be defined along the same lines as in section 7. If the gauge transformations do not commute, these matrices will not commute either. This requires that the nonabelian gauge fields exhibit selfinteractions (in other words, they are not neutral as the photon), so that the Lagrangian for the gauge fields will be considerably more complicated than the Lagrangian (5.14). One way to discover the need for selfinteractions of nonabelian gauge fields follows from analyzing Feynman diagrams with several external lines associated with nonabelian gauge fields. One can then establish that the corresponding amplitudes are only conserved if the gauge fields have direct interactions with themselves. However, we will proceed differently and start in the same vein as in section 5, assuming invariance under nonabelian transformations.

The relevant gauge groups consist of transformations, usually represented by matrices, that can be parametrized in an analytic fashion in terms of a finite number of parameters. Such groups are called *Lie groups*. The number of independent parameters defines the *dimension* of the group.* For instance, phase transformations constitute the group $U(1)$, which is clearly of dimension one. As mentioned above, this group is abelian because phase transformations commute. Two important classes of nonabelian groups are the groups $SO(N)$ of real rotations in N dimensions ($N > 2$), and the groups $SU(N)$ of $N \times N$ unitary matrices with unit determinant ($N > 1$). As we shall see in a moment the dimension of these groups is $\frac{1}{2}N(N-1)$ and $N^2 - 1$, respectively.

Let us generally consider fields that transform according to a representation of a certain Lie group G. This means that, for every element of the group G, we have a matrix U; these matrices U satisfy the same multiplication rules as the corresponding elements of G. Under a group transformation the fields rotate as follows

$$\psi(x) \rightarrow \psi'(x) = U\,\psi(x), \tag{8.1}$$

where ψ denotes an array of different fields written as a column vector. More explicitly, we may write

$$\psi_i(x) \rightarrow \psi_i'(x) = U_{ij}\,\psi_j(x). \tag{8.1'}$$

For most groups the matrices U can generally be written in exponential form

$$U = \exp(\xi^a t_a), \tag{8.2}$$

where the matrices t_a are called the *generators* of the group defined in the representation appropriate to ψ, and the ξ^a constitute a set of *real* parameters in terms of which the group elements can be described. The number of generators, which is obviously equal to the number of independent parameters ξ^a and therefore to dimension of the group, is unrelated to the dimension of the matrices U and t_a. It is usually straightforward to determine the generators for a given group. For example, the generators of the $SO(N)$ group must consist of the $N \times N$ real and antisymmetric matrices, in order that (8.2) defines an orthogonal matrix: $U^T = U^{-1}$. As there are $\frac{1}{2}N(N-1)$ independent real and antisymmetric matrices the dimension of the $SO(N)$ group is equal to $\frac{1}{2}N(N-1)$. For the $SU(N)$ group, the defining relation $U^\dagger = U^{-1}$ requires the generators t_a to be antihermitean $N \times N$ matrices. Furthermore, to have a matrix (8.2) with unit determinant it is necessary that these antihermitean matrices t_a are traceless. There are $N^2 - 1$ independent antihermitean traceless matrices so that the dimension of $SU(N)$ is equal to $N^2 - 1$. To verify these properties it is usually sufficient to consider infinitesimal transformations, where the parameters ξ^a are small, so that $U = 1 + \xi^a t_a + O(\xi^2)$.

Because the matrices U defined in (8.2) constitute a representation of the group, products of these matrices must be of the same exponential form. This leads to an important condition on the matrices t_a, which can already be derived by considering a product of two infinitesimal

* Mathematically, a set of transformations forms a group if the product of every two transformations, the identity and the inverse of each transformation is contained in the set, and if the product of transformations is associative. It is usually rather obvious that the complete set of transformations that leave a theory invariant, forms a group.

transformations: the matrices t_a generate a group representation if and only if their commutators can be decomposed into the same set of generators. These commutation relations define the *Lie algebra* g corresponding to the Lie group G,

$$[t_a, t_b] = f_{ab}{}^c\, t_c, \tag{8.3}$$

where the proportionality constants $f_{ab}{}^c$ are called the *structure constants*, because they define the multiplication properties of the Lie group. As we shall see the Lie algebra relation (8.3) plays a central role in what follows.

Let us now follow the same approach as in section 5 and consider the extension of the group G to a group of *local* gauge transformations. This means that the parameters of G will become functions of the space-time coordinates x^μ. As long as one considers variations of the field at a single point in space-time this extension is trivial, but the local character of the transformations becomes important when comparing changes at different space-time points. In particular this is relevant when considering the effect of local transformations on derivatives of the fields, i.e.,

$$\psi(x) \to \psi'(x) = U(x)\,\psi(x), \tag{8.4}$$

$$\partial_\mu\psi(x) \to (\partial_\mu\psi(x))' = U(x)\,\partial_\mu\psi(x) + (\partial_\mu U(x))\,\psi(x). \tag{8.5}$$

Just as in section 5 local quantities such as the vector ψ, which transform according to a representation of the group G at the *same* space-time point, are called *covariant*. Due to the presence of the second term on the right-hand side of (8.5), $\partial_\mu\psi$ does not transform covariantly. Although the action of the space-time dependent extension of G is still correctly realized by (8.5) this type of behaviour under symmetry variations is difficult to work with. Therefore one attempts to replace ∂_μ by a so-called *covariant derivative* D_μ, which constitutes a covariant quantity when applied on ψ,

$$D_\mu\psi(x) \to (D_\mu\psi(x))' = U(x)\,D_\mu\psi(x). \tag{8.6}$$

The construction of a covariant derivative has been discussed in the previous chapter for abelian transformations, where it was noted that a covariant derivative can be viewed as the result of a particular combination of an infinitesimal displacement generated by the ordinary derivative and a field-dependent infinitesimal gauge transformation. Such an infinitesimal displacement was called a *covariant translation*. Its form suggests an immediate generalization to the covariant derivative for an arbitrary group. Namely, we take the linear combination of an ordinary derivative and an infinitesimal gauge transformation, where the parameters of the latter *define* the nonabelian gauge fields. Hence

$$D_\mu\psi \equiv \partial_\mu\psi - W_\mu\,\psi, \tag{8.7}$$

where W_μ is a matrix of the type generated by an infinitesimal gauge transformation. This means that W_μ takes values in the Lie-algebra corresponding to the group G, i.e., W_μ can be decomposed into the generators t_a,

$$W_\mu = W_\mu^a t_a. \tag{8.8}$$

Indeed W_μ has the characteristic feature of a gauge field, as it can carry information regarding the group from one space-time point to another.

Let us now examine the consequences of (8.7). Combining (8.5) and (8.6) shows that $W_\mu\psi$ must transform under gauge transformations as

$$\begin{aligned}(W_\mu\psi)' &= \partial_\mu\psi' - (D_\mu\psi)' \\ &= U(\partial_\mu\psi) + (\partial_\mu U)\psi - U(D_\mu\psi) \\ &= \{UW_\mu U^{-1} + (\partial_\mu U)U^{-1}\}\,\psi'. \end{aligned} \tag{8.9}$$

This implies the following transformation rule for W_μ,

$$W_\mu \rightarrow W'_\mu = UW_\mu U^{-1} + (\partial_\mu U)U^{-1}. \tag{8.10}$$

Clearly the gauge fields do not transform covariantly. The first term in (8.10) indicates that the gauge fields W^a_μ transform according to the so-called adjoint representation of the group; the second noncovariant term is a modification that is characteristic for gauge fields. It is easy to evaluate (8.10) for infinitesimal transformations by using (8.3) and we find

$$W^a_\mu \rightarrow (W^a_\mu)' = W^a_\mu + f_{bc}{}^a\xi^b W^c_\mu + \partial_\mu\xi^a + O(\xi^2). \tag{8.11}$$

This result differs from the transformation law of abelian gauge fields by the presence of the term $f_{bc}{}^a\xi^b W^c_\mu$.

We have already made use of the observation that the D_μ can be viewed as the generators of covariant translations, which consist of infinitesimal space-time translations combined with infinitesimal field-dependent gauge transformations in order to restore the covariant character of the translated quantity. Since both these infinitesimal transformations satisfy $\delta(\phi\psi) = (\delta\phi)\psi + \phi(\delta\psi)$, we have Leibnitz' rule for covariant derivatives,

$$D_\mu(\phi\psi) = (D_\mu\phi)\psi + \phi(D_\mu\phi). \tag{8.12}$$

Note that the covariant derivative always depends on the representation of the fields on which it acts through the choice of the generators t_a. Hence each of the three terms in (8.12) may contain a different representation for the generators (see the simple abelian example in (5.12)).

Unlike ordinary differentiations, two covariant differentiations do not necessarily commute. It is easy to see that the commutator of two covariant derivatives D_μ and D_ν, which is obviously a covariant quantity, is given by

$$\begin{aligned}[D_\mu, D_\nu]\psi &= D_\mu(D_\nu\psi) - D_\nu(D_\mu\psi) \\ &= -(\partial_\mu W_\nu - \partial_\nu W_\mu - [W_\mu, W_\nu])\psi. \end{aligned} \tag{8.13}$$

This result leads to the definition of a covariant antisymmetric tensor $G_{\mu\nu}$,

$$G_{\mu\nu} = \partial_\mu W_\nu - \partial_\nu W_\mu - [W_\mu, W_\nu], \tag{8.14}$$

which is called the *field strength.* As ψ and $D_\mu D_\nu\psi$ transform identically under the gauge transformations the field strength must transform covariantly according to

$$G_{\mu\nu} \rightarrow G'_{\mu\nu} = UG_{\mu\nu}U^{-1}. \tag{8.15}$$

Because W_μ is Lie-algebra valued and the quadratic term in (8.14) is a commutator, the field strength is also Lie-algebra valued, i.e., $G_{\mu\nu}$ can also be decomposed in terms of the group generators t_a,

$$G_{\mu\nu} = G^a_{\mu\nu}t_a, \tag{8.16}$$

with

$$G^a_{\mu\nu} = \partial_\mu W^a_\nu - \partial_\nu W^a_\mu - f_{bc}{}^a W^b_\mu W^c_\nu. \tag{8.17}$$

Note that(8.17) differs from the abelian field strength derived in section 5 by the presence of the term quadratic in the gauge fields.

Under an infinitesimal transformation $G_{\mu\nu}$ transforms as

$$G_{\mu\nu} \rightarrow G'_{\mu\nu} = G_{\mu\nu} + [\xi^a t_a, G_{\mu\nu}], \tag{8.18}$$

or, equivalently, as

$$G^a_{\mu\nu} \rightarrow (G^a_{\mu\nu})' = G^a_{\mu\nu} + f_{bc}{}^a\xi^b G^c_{\mu\nu}. \tag{8.19}$$

For abelian groups the structure constants vanish (i.e., the adjoint representation is trivial for an abelian group), so that the field strengths are invariant in that case.

The result (8.13) can now be expressed in a representation independent form

$$[D_\mu, D_\nu] = -G_{\mu\nu}, \tag{8.20}$$

implying that the commutator of two covariant derivatives is equal to an infinitesimal gauge transformation with $-G^a_{\mu\nu}$ as parameters. This is the *Ricci identity.* Precisely as for the abelian case we may apply further covariant derivatives to (8.20). In particular we consider

$$[D_\mu[D_\nu, D_\rho]] + [D_\nu[D_\rho, D_\mu]] + [D_\rho[D_\mu, D_\nu]]$$

which vanishes identically because of the Jacobi identity. Inserting (8.20) we obtain the result

$$D_\mu G_{\nu\rho} + D_\nu G_{\rho\mu} + D_\rho G_{\mu\nu} = 0, \tag{8.21}$$

where, acording to (8.18), the covariant derivative of $G_{\mu\nu}$ equals

$$D_\mu G_{\nu\rho} = \partial_\mu G_{\nu\rho} - [W_\mu, G_{\nu\rho}], \tag{8.22}$$

or, in components,

$$D_\mu G^a_{\nu\rho} = \partial_\mu G^a_{\nu\rho} - f_{bc}{}^a W^b_\mu G^c_{\nu\rho}. \tag{8.23}$$

The relation (8.21) is called the *Bianchi identity*; in the abelian case the Bianchi identity corresponds to the homogeneous Maxwell equations.

9. Gauge invariant Lagrangians for spin-0 and spin-$\frac{1}{2}$ fields

By making use of the covariant derivatives constructed in the previous section it is rather straightforward to construct gauge invariant Lagrangians for spin-0 and spin-$\frac{1}{2}$ fields. To demonstrate this, consider a set of N spinor fields ψ_i transforming under transformations U belonging to a certain group G according to ($i, j = 1, \cdots, N$)

$$\psi_i \to \psi_i' = U_{ij}\,\psi_j, \tag{9.1}$$

or, suppressing indices i, j, and writing ψ as an N-dimensional column vector,

$$\psi \to \psi' = U\psi. \tag{9.2}$$

Conjugate spinors $\bar{\psi}_i$ then transform as

$$\bar{\psi}_i \to \bar{\psi}_i' = U_{ij}^{*}\,\bar{\psi}_j, \tag{9.3}$$

or, regarding $\bar{\psi}$ as a row vector and again suppressing indices, as

$$\bar{\psi} \to \bar{\psi}' = \bar{\psi}\,U^{\dagger}. \tag{9.4}$$

Obviously, if U is unitary, i.e., if $U^{\dagger} = U^{-1}$, the massive Dirac Lagrangian,

$$\mathcal{L} = -\bar{\psi}_i \partial\!\!\!/ \psi_i - m\,\bar{\psi}_i\psi_i, \tag{9.5}$$

is invariant under G. This Lagrangian thus describes N spin-$\frac{1}{2}$ (anti)particles of equal mass m.

We now require that the Lagrangian be invariant under *local* G transformations. To achieve this we simply replace the ordinary derivative in (9.5) by a covariant derivative (here and henceforth we will suppress indices i, j, etc.),

$$\begin{aligned}\mathcal{L} &= -\bar{\psi} D\!\!\!\!/\, \psi - m\,\bar{\psi}\psi \\ &= -\bar{\psi}\partial\!\!\!/ \psi - m\,\bar{\psi}\psi + \bar{\psi}\gamma^{\mu}W_{\mu}\psi,\end{aligned} \tag{9.6}$$

where $W_\mu = W_\mu^a t_a$ is the Lie-algebra valued gauge field introduced in the previous section. Hence the gauge field interactions are given by

$$\mathcal{L}_{\rm int} = W_\mu^a \bar{\psi}\gamma^{\mu} t_a \psi, \tag{9.7}$$

where t_a are the parameters of the gauge group G in the representation appropriate to ψ. Observe that the matrices t_a are antihermitean in order that the gauge transformations be unitary. The reader will have noticed that there is no obvious coupling constant in (9.7), but we shall see in the next section how this coupling constant can be extracted from the fields W_μ^a. The matrices $(t_a)_{ij}$ can be regarded as nonabelian charges (up to a proportionality factor i). The commutation relation for these charges is a consequence of the Lie algebra relation (9.3) which is necessary and sufficient in order that the nonabelian gauge transformations form a group.

To illustrate the above construction, let us explicitly construct a gauge invariant Lagrangian for fermions transforming as doublets under the group $SU(2)$. This group consists of all 2×2 unitary matrices with unit determinant. Such matrices can be written in exponentiated form

$$U(\xi) = \exp(\xi^a t_a), \qquad (a = 1,2,3) \tag{9.8}$$

where the three generators of $SU(2)$ are expressed in terms of the isotopic spin matrices τ_a,

$$t_a = \tfrac{1}{2} i \tau_a, \tag{9.9}$$

which coincide with the Pauli matrices used in the context of ordinary spin.

$$\tau_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \tau_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \tau_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{9.10}$$

As can be explicitly verified the generators t_a satisfy the commutation relations

$$[t_a, t_b] = -\epsilon_{abc}\, t_c, \tag{9.11}$$

ensuring that the matrices (9.8) form a group.

Historically the first construction of a nonabelian gauge field theory was based on $SU(2)$ and was motivated by the existence of the approximate isospin invariance in Nature. According to the notion of isospin (or isobaric spin) invariance the proton and the neutron can be regarded as an isospin doublet. Therefore one introduces a doublet field

$$\psi = \begin{pmatrix} \psi_p \\ \psi_n \end{pmatrix}, \tag{9.12}$$

analogous to the $s = \frac{1}{2}$ doublet of ordinary spin. Conservation of isospin is just the requirement of invariance under isospin rotations

$$\psi \to \psi' = U\psi, \tag{9.13}$$

where U is an $SU(2)$ matrix as defined in (9.8). If isospin invariance would be an exact symmetry then it is a matter of convention which component of ψ would correspond to the proton and which one to the neutron. If one insists on being able to define this convention at any space-time point separately, then one is led to the construction of a gauge field theory based on local isospin transformations (this is the heuristic argument that motivated Yang and Mills to attempt the construction of the gauge theory of $SU(2)$)

Starting from (9.13) it is straightforward to construct the covariant derivative on ψ,

$$\begin{aligned} D_\mu \psi &= \begin{pmatrix} \partial_\mu \psi_p \\ \partial_\mu \psi_n \end{pmatrix} - \frac{1}{2} i W_\mu^a \tau_a \begin{pmatrix} \psi_p \\ \psi_n \end{pmatrix} \\ &= \begin{pmatrix} \partial_\mu \psi_p \\ \partial_\mu \psi_n \end{pmatrix} - \frac{1}{2} i \begin{pmatrix} W_\mu^3 & W_\mu^1 - iW_\mu^2 \\ W_\mu^1 + iW_\mu^2 & -W_\mu^3 \end{pmatrix} \begin{pmatrix} \psi_p \\ \psi_n \end{pmatrix}. \end{aligned} \tag{9.14}$$

A locally $SU(2)$ invariant Lagrangian is then obtained by replacing the ordinary derivative by a covariant one in the Lagrangian of a degenerate doublet of spin-$\frac{1}{2}$ fields.

$$\begin{aligned}\mathcal{L} &= -\bar{\psi}\not{D}\psi - m\,\bar{\psi}\psi \\ &= -\bar{\psi}\not{\partial}\psi - m\,\bar{\psi}\psi + \tfrac{1}{2}i\,W_\mu^a\,\bar{\psi}\gamma^\mu\tau_a\psi.\end{aligned} \tag{9.15}$$

The field strength tensors follow straightforwardly from the $SU(2)$ structure constants exhibited in (9.11)

$$G_{\mu\nu}^a = \partial_\mu W_\nu^a - \partial_\nu W_\mu^a + \epsilon_{abc}W_\mu^b W_\nu^c, \tag{9.16}$$

while under infinitesimal $SU(2)$ transformations the gauge fields transform according to

$$W_\mu^a \to (W_\mu^a)' = W_\mu^a + \epsilon_{abc}W_\mu^b\xi^c + \partial_\mu\xi^a. \tag{9.17}$$

However, contrary to initial expectations, *local* $SU(2)$ transformations have no role to play in the strong interactions. Instead these forces are governed by an $SU(3)$ gauge theory called quantum chromodynamics because one has introduced the term *colour* for the degrees of freedom transforming under $SU(3)$. The corresponding gauge fields are called gluon fields because they are assumed to bind the elementary hadronic constituents, called quarks, into hadrons. The quarks transform as triplets under $SU(3)$ and can thus be viewed as a straightforward extension of the $SU(2)$ doublet (9.12). Theories based on $SU(2)$ gauge transformations are relevant for the weak interactions.

Gauge invariant Lagrangians with spin-0 fields are constructed in the same way. For instance, consider an array of complex scalar fields transforming under transformations U as in (9.1). If we regard ϕ as a column vector and the complex conjugate fields as a row vector ϕ^*, we may write

$$\begin{aligned}\phi &\to \phi' = U\phi \\ \phi^* &\to (\phi^*)' = \phi^* U^\dagger.\end{aligned} \tag{9.18}$$

Covariant derivatives read

$$\begin{aligned}D_\mu\phi &= \partial_\mu\phi - W_\mu^a t_a\phi, \\ D_\mu\phi^* &= \partial_\mu\phi^* - \phi^* t_a^\dagger W_\mu^a.\end{aligned} \tag{9.19}$$

Provided that the transformation matrices U in (9.18) are unitary (so that $t_a^\dagger = -t_a$) the following Lagrangian is gauge invariant

$$\mathcal{L} = -|D_\mu\phi|^2 - m^2|\phi|^2 - \lambda|\phi|^4 \tag{9.20}$$

where we have used a complex inner product $|\phi|^2 \equiv \phi_i^*\phi_i$. Substituting (9.19) leads to

$$\begin{aligned}\mathcal{L} = &-(\partial_\mu\phi^* + \phi^* t_a W_\mu^a)(\partial_\mu\phi - W_\mu^b t_b\phi) - m^2|\phi|^2 - \lambda|\phi|^4 \\ = &-|\partial_\mu\phi|^2 - m^2|\phi|^2 - \lambda|\phi|^4 \\ &- W_\mu^a(\phi^* t_a\partial_\mu\phi - (\partial_\mu\phi^*)t_a\phi) + W_\mu^a W_\mu^b(\phi^* t_a t_b\phi),\end{aligned} \tag{9.21}$$

where in the gauge field interaction terms ϕ^* and ϕ are written as row and column vectors. This result once more exhibits the role played by the generators t_a as matrix generalizations of the charge. Using (9.9) it is easy to give the corresponding Lagrangian invariant under $SU(2)$. In that case (9.21) reads

$$\begin{aligned}\mathcal{L} = &- |\partial_\mu \phi|^2 - m^2|\phi|^2 - \lambda|\phi|^4 \\ &- \tfrac{1}{2} i W_\mu^a \left(\phi^* \tau_a \overleftrightarrow{\partial}_\mu \phi\right) - \tfrac{1}{4}\left(W_\mu^a\right)^2 |\phi|^2,\end{aligned} \qquad (9.22)$$

where we have used $\tau_a\tau_b + \tau_b\tau_a = 1\,\delta_{ab}$.

10. The gauge field Lagrangian

In the preceding section we discussed how to construct locally invariant Lagrangians for matter fields. Starting from a Lagrangian that is invariant under the corresponding rigid transformations, one replaces ordinary derivatives by covariant ones. Until that point the gauge fields are not yet treated as new dynamical degrees of freedom. For that purpose one must also specify a Lagrangian for the gauge fields, which must be separately locally gauge invariant. A transparent construction of such invariants make use of the field strength tensor $G_{\mu\nu}$. Let us recall that $G_{\mu\nu}$ transforms according to (cf. (8.15)),

$$G_{\mu\nu} \to G'_{\mu\nu} = U G_{\mu\nu} U^{-1}, \qquad (10.1)$$

so that for any product of these tensors we have

$$G_{\mu\nu}G_{\rho\sigma}\cdots G_{\lambda\tau} \to G'_{\mu\nu}G'_{\rho\sigma}\cdots G'_{\lambda\tau} = U(G_{\mu\nu}G_{\rho\sigma}\cdots G_{\lambda\tau})U^{-1}. \qquad (10.2)$$

Consequently the trace of arbitrary products of the form (10.2) is gauge invariant, i.e.,

$$\mathrm{Tr}(G_{\mu\nu}\cdots G_{\lambda\tau}) \to \mathrm{Tr}(UG_{\mu\nu}\cdots G_{\lambda\tau}U^{-1}) = \mathrm{Tr}(G_{\mu\nu}\cdots G_{\lambda\tau}), \qquad (10.3)$$

by virtue of the cyclicity of the trace operation. The simplest Lorentz invariant and parity conserving Lagrangian can therefore be expressed as a quadratic form in $G_{\mu\nu}$,

$$\mathcal{L}_W = \frac{1}{4g^2}\mathrm{Tr}(G_{\mu\nu}G^{\mu\nu}), \qquad (10.4)$$

where we have introduced $(4g^2)^{-1}$ as an arbitrary normalization constant. Notice that two alternative forms $G_{\mu\mu}$ and $G_{\mu\nu}G_{\rho\sigma}\epsilon_{\mu\nu\rho\sigma}$ are excluded: the first one vanishes by antisymmetry of $G_{\mu\nu}$, and the second one is not parity conserving (in fact one can show that the second term is equal to a total divergence).

After rescaling W_μ to gW_μ, the gauge field Lagrangian (10.4) acquires the form

$$\begin{aligned}\mathcal{L}_W = \mathrm{Tr}(t_a t_b)\Big\{&\tfrac{1}{4}(\partial_\mu W_\nu^a - \partial_\nu W_\mu^a)(\partial_\mu W_\nu^b - \partial_\nu W_\mu^b) \\ &- g f_{cd}{}^a W_\mu^c W_\nu^d \partial_\mu W_\nu^b + \tfrac{1}{4} g^2 f_{cd}{}^a f_{ef}{}^b W_\mu^c W_\mu^e W_\nu^d W_\nu^f\Big\},\end{aligned} \qquad (10.5)$$

where one may distinguish a kinetic term, which resembles the abelian Lagrangian (5.12), and W^3- and W^4-interaction terms, which depend on the structure constants of the gauge group.

For the Lie groups that we will be interested in, we have $t_a^\dagger = -t_a$ and the generators can be defined such that $\mathrm{Tr}(t_a t_b)$ becomes equal to $-\delta_{ab}$. In that case the Lagrangian reads

$$\begin{aligned}\mathcal{L}_W = &-\tfrac{1}{4}(\partial_\mu W_\nu^a - \partial_\nu W_\mu^a)^2 + g f_{abc} W_\mu^a W_\nu^b \partial_\mu W_\nu^c \\ &- \tfrac{1}{4} g^2 f_{abc} f_{ade} W_\mu^b W_\mu^d W_\nu^c W_\nu^e. \end{aligned} \tag{10.6}$$

In this particular case there is no need to distinguish between upper and lower indices $a, b, \ldots$ and $f_{abc} \equiv f_{ab}{}^c$ is totally antisymmetric.

One may combine the gauge field Lagrangian (10.6) with a gauge invariant Lagrangian for the matter fields, and derive the corresponding Euler-Lagrange equations from Hamilton's principle. To be specific let us choose the gauge invariant Lagrangian (9.6) for fermions and combine it with (10.6),

$$\begin{aligned}\mathcal{L} &= \mathcal{L}_W + \mathcal{L}_\psi \\ &= \frac{1}{4g^2}\mathrm{Tr}(G_{\mu\nu}G_{\mu\nu}) - \bar{\psi}\not{D}\psi - m\bar{\psi}\psi \end{aligned} \tag{10.7}$$

The field equations for the fermions are obviously a covariant version of the Dirac equation, i.e.,

$$(\not{D} + m)\psi = 0, \qquad \bar{\psi}(\overleftarrow{\not{D}} - m) = 0. \tag{10.8}$$

or explicitly (using $t_a^\dagger = -t_a$)

$$(\not\partial + m)\psi = \not{W}^a t_a \psi, \tag{10.9}$$

$$\bar{\psi}(\overleftarrow{\not\partial} + m) = -\bar{\psi} t_a \not{W}^a, \tag{10.10}$$

where we have not yet extracted the gauge coupling constant g from W_μ. These equations are the nonabelian generalizations of (5.12) and (5.13).

To derive the field equations for the gauge fields requires more work and we give the result without proof,

$$D^\nu G_{\mu\nu}^a = J_\mu^a, \tag{10.11}$$

which is obviously a nonabelian extension of (5.24). The explicit form of this equation is rather complicated,

$$\begin{aligned}\partial^\nu(\partial_\mu W_\nu^a - \partial_\nu W_\mu^a) + f_{bc}{}^a(W^{b\nu}\partial_\mu W_\nu^c + W_\mu^b \partial_\nu W^{c\nu} - 2W^{b\nu}\partial_\nu W_\mu^c) \\ + f_{bc}{}^a f_{de}{}^c W_\mu^d W_\nu^b W^{e\nu} = J_\mu^a, \end{aligned} \tag{10.12}$$

where

$$J_\mu^a = \bar{\psi}\gamma_\mu t_a \psi. \tag{10.13}$$

To examine whether J_μ^a is conserved, we apply a covariant derivative D_μ to (10.11). On the left-hand side this leads to

$$\begin{aligned}D_\mu D_\nu G^{a\mu\nu} &= \tfrac{1}{2}[D_\mu, D_\nu]\, G^{a\mu\nu} \\ &= \tfrac{1}{2} G_{\mu\nu}^b\, f_{bc}{}^a\, G^{c\mu\nu}, \end{aligned} \tag{10.14}$$

where we have used the Ricci identity (10.28). Because $f_{bc}{}^{a}$ is antisymmetric in b and c, (10.14) vanishes. Therefore the current satisfies a covariant divergence equation

$$D_\mu J^{a\mu} = 0, \tag{10.15}$$

or, explicitly,

$$\partial_\mu J^{a\mu} - f_{bc}{}^{a} W_\mu^b J^{c\mu} = 0. \tag{10.16}$$

This result implies that gauge fields can only couple consistently to currents that are *covariantly* constant. According to (10.15) the charges associated with the current are not quite conserved. The reason is obviously that the gauge fields are not neutral. Their contributions must be included in order to define charges that are conserved.

11. Spontaneously broken symmetries

It is possible for a theory to be exactly invariant under a symmetry, while its ground-state solution does not exhibit this symmetry. In itself it is not surprising that a symmetric theory can give rise to nonsymmetric states; take, for instance, the hydrogen atom which is described by a hamiltonian that is rotationally invariant, while its eigenstates with nonvanishing angular momentum are not inert under rotations. Nevertheless its ground state has zero angular momentum, and is thus rotationally symmetric.

An example of a rotationally invariant system which is realized in such a way that the ground state is not symmetric, is the ferromagnet. Nonferromagnetic materials have a rotationally symmetric ground state in which the atomic spins are randomly oriented. Therefore the gross magnetization is zero. However, in a ferromagnet the spin-spin interactions are such that in the state of lowest energy all spins are aligned. This gives rise to a finite magnetization which breaks the manifest rotational symmetry; thus the rotational symmetry is realized in a spontaneously broken way. This does not mean that rotational symmetry has no consequences anymore, but the most obvious implications of having a symmetric theory are absent. One important aspect of a spontaneously broken realization is that the ground state must be infinitely degenerate. From a nonsymmetric ground state one may construct an infinite number of states by applying the symmetry transformations on the ground state. All these different states must have the same energy as the original one, because of the symmetry of the theory; the hamiltonian of the system still commutes with all symmetry transformations. Indeed for the ferromagnet with all spins aligned in a given direction, one obtains an infinite set of ground states by rotations of the magnet.

We now discuss these phenomena in the context of a field-theoretic model based on a complex spinless field ϕ :

$$\mathcal{L} = -|\partial_\mu \phi|^2 - V(|\phi|). \tag{11.1}$$

This Lagrangian is invariant under constant phase transformations of ϕ

$$\phi(x) \rightarrow \phi'(x) = e^{iq\xi}\phi(x). \tag{11.2}$$

Such $U(1)$ transformations can also be represented as two-dimensional rotations of the real and imaginary parts of ϕ. Hence the groups $U(1)$ and $O(2)$ are equivalent.

In theories such as (11.1) the fields are usually expanded about some constant value for which the potential (and thus the energy) has an absolute minimum. This value characterizes the ground state of the system, in the same way as the magnetization of the ferromagnet specifies its ground state. Of course, the actual value changes when quantum corrections are included, but such effects will not concern us here. The field ϕ will have fluctuations about this classical value corresponding to dynamical degrees of freedom. Such degrees of freedom can be associated with particles: the constant field value which represents the field configuration with minimal energy is called the vacuum expectation value. The nature of the fluctuations about this value is determined by the Lagrangian. Expanding the field about its vacuum expectation value v, we obtain a Lagrangian of the Klein-Gordon type for the two field components contained in ϕ, which describe particles with a mass determined by the second derivative of $V(|\phi|)$ at $\phi = v$. Rather than working this out in detail, we give a systematic description of the various possibilities.

The first possibility is that the potential acquires its minimum at $\phi = 0$. In that case expanding the Lagrangian about $\phi = 0$ gives rise to

$$\mathcal{L} = -|\partial_\mu \phi|^2 - \mu^2 |\phi|^2 + \ldots . \tag{11.3}$$

This shows that the excitations described by ϕ correspond to particles with mass μ; note that μ^2 must be positive because we have expanded the potential about a minimum. We may thus distinguish two kinds of particles, corresponding to the real and imaginary part of ϕ. Both have the same mass, which can be understood on the basis of the symmetry (11.2) which rotates the real and imaginary parts of the field. Such a symmetric realization of the theory is called the Wigner-Weyl mode.

We now consider the case where the minimum is acquired for a non-zero field value. In that case one immediately realizes that there must be an infinite set of minima because of the symmetry (11.2). In the plane of real and imaginary components of ϕ these minima are located on a circle (see Fig. 5), and each of them represents a possible ground state. This situation describes a spontaneously broken realization of the symmetry (11.2), because the fact that we are forced to consider the theory for nonvanishing vacuum expectation value means that the symmetry is no longer manifest. However, further inspection shows that the symmetry still has an important implication. Because of the degeneracy there is one direction in which the potential remains constant when expanding about the minimum. Consequently, one of the excitations about the ground state value of ϕ is massless. This is in accordance with the Goldstone theorem, which states that to every generator of the symmetry group that is spontaneously broken, there corresponds a massless particle. This particle is called the Goldstone particle, and the spontaneously broken realization is called the Goldstone mode. One recognizes that the massless degrees of freedom are related to the symmetry that is broken, since it is this symmetry that causes the degeneracy of the potential. Because the symmetry in this case is generated by a scalar parameter, which shifts the phase of ϕ , the particle is a scalar particle. But more complicated examples of spontaneously broken symmetries are possible.

We now consider the Goldstone mode in somewhat more detail. Since we are expanding the field about some nonvanishing value, it is convenient to make a decomposition

$$\phi(x) = \frac{1}{\sqrt{2}} \rho(x)\, e^{i\theta(x)}. \tag{11.4}$$

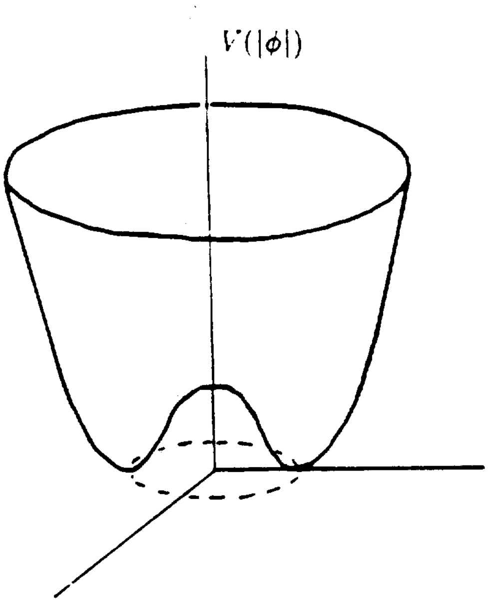


Fig. 5. The potential $V(|\phi|)$

This leads to

$$\partial_\mu \phi = \frac{1}{\sqrt{2}} e^{i\theta(x)} \left(\partial_\mu \rho + i\rho\, \partial_\mu \theta\right), \tag{11.5}$$

which is inserted into the Lagrangian (11.1)

$$\mathcal{L} = -\frac{1}{2}(\partial_\mu \rho)^2 - \frac{1}{2}\rho^2(\partial_\mu \theta)^2 - V(\rho/\sqrt{2}) \tag{11.6}$$

Clearly the radial degrees of freedom describe a particle with a mass given by

$$\mu^2 = \frac{\partial^2}{\partial \rho^2} V(\rho/\sqrt{2})\Big|_{\rho=v} = \tfrac{1}{2} V''(\rho/\sqrt{2}). \tag{11.7}$$

But the angular degrees of freedom related to θ do not have a mass, and we find a standard kinetic term for a massless scalar field with some additional derivative interactions. This confirms the result of our heuristic considerations, and is in agreement with Goldstone's theorem.

12. The Brout-Englert-Higgs mechanism

The existence of two possible realizations of a symmetry naturally raises the question whether a similar phenomenon exists for local gauge symmetries. We will see that this is indeed the case: there exists a second realization of theories with local gauge invariance, which causes the generation of a mass term for the gauge fields. To analyse this in detail we extend the model of the previous section by introducing an abelian gauge field A_μ and by requiring invariance under local $U(1)$ transformations. The combined transformation rules thus read

$$\begin{aligned} \phi(x) &\rightarrow \phi'(x) = e^{iq\xi(x)}\phi(x), \\ A_\mu(x) &\rightarrow A'_\mu(x) = A_\mu(x) + \partial_\mu\xi(x). \end{aligned} \tag{12.1}$$

Following the rules of the previous sections, it is easy to write down a Lagrangian invariant under these transformations. We add a kinetic term for the gauge field to (12.1), and replace the derivatives of ϕ by covariant derivatives

$$\begin{aligned} \mathcal{L} &= -\frac{1}{4}F^2_{\mu\nu}(A) - |D_\mu\phi|^2 - V(|\phi|), \\ F_{\mu\nu}(A) &= \partial_\mu A_\nu - \partial_\nu A_\mu, \\ D_\mu\phi &= \partial_\mu\phi - iqA_\mu\phi. \end{aligned} \tag{12.2}$$

We assume that the potential acquires an absolute minimum for nonvanishing field values; therefore we adopt the decomposition (11.4). In this parametrization the phase transformation is expressed by

$$\theta(x) \rightarrow \theta'(x) = \theta(x) + q\,\xi(x), \tag{12.3}$$

and the covariant derivative takes the form

$$D_\mu\phi = \frac{1}{\sqrt{2}}e^{i\theta}\left(\partial_\mu\rho - iq\rho(A_\mu - q^{-1}\partial_\mu\theta)\right). \tag{12.4}$$

We now define a new field B_μ by

$$B_\mu = A_\mu - q^{-1}\partial_\mu\theta, \tag{12.5}$$

which is inert under the gauge transformations. The covariant derivative can then be written as

$$D_\mu\phi = \frac{1}{\sqrt{2}}\,e^{i\theta}\left(\partial_\mu\rho - iq\rho B_\mu\right). \tag{12.6}$$

Since the relation between B_μ and A_μ takes the form of a (field-dependent) gauge transformation, we can simply replace the field strength $F_{\mu\nu}(A)$ by the corresponding tensor $F_{\mu\nu}(B)$. Therefore the Lagrangian can be expressed entirely in terms of the fields ρ and B_μ, which are both gauge invariant,

$$\mathcal{L} = -\tfrac{1}{4}F^2_{\mu\nu}(B) - \tfrac{1}{2}(\partial_\mu\rho)^2 - \tfrac{1}{2}q^2\rho^2B^2_\mu - V(\rho/\sqrt{2}). \tag{12.7}$$

If we now expand the field ρ about its vacuum expectation value v, we find that the Lagrangian (12.7) describes a *massive* spin-1 field B_μ, with a mass given by

$$M_B = |qv|. \tag{12.8}$$

The massless field that corresponds to the Goldstone particle in the model of the previous section has simply disappeared, while the massive spinless field remains.

At this point we realize that we could have derived (12.7) directly, by exploiting the gauge invariance in order to put $\theta(x) = 0$ from the beginning. This amounts to choosing a gauge condition $\phi = \rho/\sqrt{2}$, which is called the unitary gauge. The advantage of this gauge is that the physical content of the model is immediately clear. However, this gauge is extremely inconvenient for calculating quatum corrections, because it leads to many more ultraviolet divergences than the so-called renormalizable gauge conditions.

It is important to realize that the degeneracy of the ground state that was present in the case without local gauge invariance, has disappeared. The degeneracy is related to the fictitious degrees of freedom that are affected by the gauge transformations; these degrees of freedom have no physical content. Hence the phase of ϕ becomes irrelevant and only the radial degree of freedom, which is gauge invariant, has physical significance. The same remark applies to the vacuum expectation of ϕ. Actually the term vacuum expectation value is somewhat misleading in this context. Since the *physical* states are gauge invariant it is not possible to have nonvanishing expectation values for quantities that are not gauge invariant. Strictly speaking the only relevant quantity is the expectation value of $|\phi|$, while the phase of ϕ is not relevant. This represents a crucial difference with the situation described in the previous section where the phase of ϕ does represent a physical degree of freedom. In that case the physical states are not required to be invariant under the symmetry, which leads to the connection between a nonzero expectation value and the infinite degeneracy of the ground state.

Hence we have discovered that in the spontaneously broken mode the gauge-dependent degrees of freedom, which effectively reside in the phase θ, decouple from the theory. The reason for this decoupling is rather obvious, since a gauge invariant theory does not depend on gauge degrees of freedom. Unlike in the realization where the potential acquires a minimum at $\phi = 0$ and the gauge field remains massless, the decoupling takes place in a purely algebraic manner without the necessity of making nonlocal field redefinitions. The number of field components has not been changed in this way. Previously we had a complex field and a gauge field representing $2 + 3 = 5$ degrees of freedom; in this realization we have only one spinless field and a vector field, but no gauge invariance. Hence we still count $1 + 4 = 5$ degrees of freedom. Also the number of physical degrees of freedom has not changed. Originally we had two scalar particles and a massless spin-1 particle; since the latter has two physical degrees of freedom, the total number of physical degrees of freedom is four. In the spontaneously broken realization we have one scalar and one massive spin-1 particle. Massive spin-1 particles have three polarizations, so that we count again four physical degrees of freedom.

It is possible to understand the above phenomenon in more physical terms. The gauge field A_μ mediates a force between charged particles which is of long range. However, when this field is generated in a medium, it is not obvious that it will still manifest itself as a long-range force. The medium may polarize under the influence of an electromagnetic field, so that the electromagnetic forces will be screened. The characteristic screening length is then inversely proportional to the mass of the gauge field. In fact this phenomenon is well known in superconductivity.

13. Massive $SU(2)$ gauge fields

We will now apply the Brout-Englert-Higgs mechanism to an $SU(2)$ gauge theory. Consider $SU(2)$ gauge fields W_μ^a coupled to a doublet of spinless fields denoted by ϕ. The relevant Lagrangians were already given in the previous sections (cf. (10.6) and (9.22)) and we find

$$\begin{aligned}\mathcal{L} = &- \tfrac{1}{4}(\partial_\mu W_\nu^a - \partial_\nu W_\mu^a)^2 - g\epsilon_{abc} W_\mu^a W_\nu^b \partial_\mu W_\nu^c \\ &- \tfrac{1}{4} g^2 \epsilon_{abc}\epsilon_{ade} W_\mu^b W_\mu^d W_\nu^c W_\nu^e \\ &- |\partial_\mu \phi|^2 + \mu^2 |\phi|^2 - \lambda |\phi|^4 \\ &- \tfrac{1}{2} i g W_\mu^a \left(\phi^* \tau_a \overleftrightarrow{\partial}_\mu \phi\right) - \tfrac{1}{4} g^2 \left(W_\mu^a\right)^2 |\phi|^2, \end{aligned} \tag{13.1}$$

With $\mu^2, \lambda > 0$ the potential acquires a minimum for a nonzero value of the field ϕ. Following the example of the previous section we decompose ϕ according to

$$\phi(x) = \begin{pmatrix} \phi_1(x) \\ \phi_2(x) \end{pmatrix} = \frac{1}{\sqrt{2}} \Phi(x) \begin{pmatrix} 0 \\ \rho(x) \end{pmatrix}, \tag{13.2}$$

where $\Phi(x)$ is an x-dependent $SU(2)$ matrix, which is a generalization of the phase factor $\exp i\theta$ (which is a $U(1)$ "matrix") used in the previous sections. Here we make use of the fact that the doublet ϕ can be brought into the form $(0, \rho/\sqrt{2})$ by a suitable gauge transformation. The field ρ is gauge invariant and represents the invariant length of the doublet field ϕ.

In principle, we could now substitute the parametrization (13.2) into the Lagrangian, redefine the gauge fields in a way analogous to (12.5), and observe that the generalized phase factor Φ disappears from the Lagrangian. However, it is more convenient to adopt the unitary gauge: owing to the local $SU(2)$ invariance we can simply ignore the matrix Φ and replace ϕ by $(0, \rho/\sqrt{2})$. The Lagrangian then takes the form

$$\begin{aligned}\mathcal{L} = &- \tfrac{1}{4}(\partial_\mu W_\nu^a - \partial_\nu W_\mu^a)^2 - g\epsilon_{abc} W_\mu^a W_\nu^b \partial_\mu W_\nu^c \\ &- \tfrac{1}{4} g^2 \epsilon_{abc}\epsilon_{ade} W_\mu^b W_\mu^d W_\nu^c W_\nu^e \\ &- \tfrac{1}{2}\left(\partial_\mu \rho\right)^2 + \tfrac{1}{2}\mu^2\rho^2 - \tfrac{1}{4}\lambda\rho^4 - \tfrac{1}{8} g^2 \rho^2 \left(W_\mu^a\right)^2 . \end{aligned} \tag{13.3}$$

Let us now determine the values $\rho = \pm v$ for which the potential $V(\rho) = -\frac{1}{2}\mu^2\rho^2 + \frac{1}{4}\lambda\rho^4$ acquires a minimum. The derivative of $V(\rho)$ vanishes whenever $\rho = 0$ or $-\mu^2 + \lambda\rho^2 = 0$. At $\rho = 0$ we have a local maximum, while the minima are at $\rho = \pm v$ with

$$v = \sqrt{\frac{\mu^2}{\lambda}}. \tag{13.4}$$

The mass of the so-called Higgs particle, which is associated with ρ, equals

$$m_\rho^2 = 2\lambda v^2, \tag{13.5}$$

while the W-masses follow from substituting $\rho^2 = v^2$ into the last term in the Lagrangian (13.3),

$$M_W^2 = \tfrac{1}{4} g^2 v^2. \tag{13.6}$$

The field ρ seems to play a minor role. It only interacts with the gauge fields through the $\rho^2 W^2$ interaction. In the limit $\lambda \to \infty$, keeping v fixed, the degrees of freedom associated with ρ are suppressed, and one is left with the standard Lagrangian for $SU(2)$ gauge fields with an extra mass term. At the *classical* level this procedure is harmless and there is no reason why one cannot drop the Higgs field. However, for the ***quantum*** theory, the situation is quite different, at least for the nonabelian case. With an explicit mass term the theory is not renormalizable, so that there is no way to obtain sensible predictions. As it turns out the presence of the extra scalar field has a smoothening effect on the the quantum corrections and makes the theory renormalizable. This aspect, and not so much the gauge invariance of the original theory which is no longer manifest in (13.3) anyway, forms the prime motivation for constructing theories according to the Brout-Englert-Higgs recipe.

14. The prototype model for $SU(2)\otimes U(1)$ electroweak interactions

We will now extend the model of the previous section in two respects. First we introduce an extra $U(1)$ gauge group, so that the resulting gauge group is $SU(2)\otimes U(1)$. Secondly we introduce two fermions denoted by p and n, which will also transform under the combined gauge group in a way that we will specify shortly. Our goal is to exhibit the essential features of the standard model for electroweak interactions, which is based on this gauge group. As the gauge group is of dimension 4, there will be four gauge fields: the three gauge fields of $SU(2)$ will be denoted by W_μ^a and the $U(1)$ gauge field by B_μ. However, the Brout-Englert-Higgs mechanism now introduces a novel feature. Initially the three gauge fields of $SU(2)$ and the gauge field of $U(1)$ are massless and have no direct interactions. After the emergence of a mass term, however, it turns out that there is one nontrivial mixture of the gauge fields which remains massless. This field is associated with a nontrivial subgroup of $SU(2)\otimes U(1)$ and will describe the photon.

Let us first comment on the way in which the fermions transform under the gauge transformations. To that order we decompose the fields in chiral components with the help of the projection operators $\frac{1}{2}(1\pm\gamma_5)$. Their left-handed components are assigned to a doublet representation of $SU(2)$; their right-handed counterparts are singlets:

$$\psi_L = (p_L, n_L);\quad p_R;\quad n_R. \tag{14.1}$$

In the original leptonic version of this model p and n correspond to the neutrino and the electron, respectively. The right-handed neutrino was chosen to decouple in that case, and only occurs as a free field. For hadrons, p and n may for instance correspond to the "up" and the "down" quark, respectively.

Under the additional $U(1)$ group the doublet fields transform as

$$\phi \to \phi' = e^{\frac{1}{2}iq\xi}\,\phi, \tag{14.2}$$

$$\psi_L \to \psi_L' = e^{\frac{1}{2}iq_1\xi}\,\psi_L, \tag{14.3}$$

and the singlets as

$$p_R \to p_R' = e^{\frac{1}{2}iq_2\xi}\,p_R, \tag{14.4}$$

$$n_R \to n_R' = e^{\frac{1}{2}iq_3\xi}\,n_R, \tag{14.5}$$

where ξ is the parameter of the $U(1)$ transformations and, for the moment, q, q_1, q_2 and q_3 are arbitrary numbers.

We now assume that the potential is such that the potential acquires a minimum for $\phi \neq 0$. In that case we can decompose ϕ according to (13.2). Ignoring the matrix Φ, which amounts to choosing the unitary gauge, gives

$$\phi(x) = (0, \rho(x)/\sqrt{2}), \tag{14.6}$$

with $\rho(x)$ a real scalar field. The form of (14.6) is left invariant under a nontrivial $U(1)$ subgroup of $SU(2) \otimes U(1)$. To identify this subgroup, consider first a somewhat larger subgroup of $SU(2) \otimes U(1)$ consisting of the diagonal matrices. They are parametrized as

$$U(\xi^a, \xi) = \begin{pmatrix} e^{\frac{i}{2}(g\xi^3 + q\xi)} & 0 \\ 0 & e^{\frac{i}{2}(-g\xi^3 + q\xi)} \end{pmatrix}, \tag{14.7}$$

where we have rescaled the $SU(2)$ parameter ξ^3 with the $SU(2)$ gauge coupling constant g in accordance with the procedure outlined for the gauge fields in section 10. In order that (14.6) be left invariant under the transformations (14.7), we must obviously have $-g\xi^3 + q\xi = 0$. This motivates the following decomposition of ξ^3 and ξ,

$$\begin{aligned} \xi^3 &= \cos\theta_W\, \xi^Z + \sin\theta_W\, \xi^{EM}, \\ \xi &= \cos\theta_W\, \xi^{EM} - \sin\theta_W\, \xi^Z, \end{aligned} \tag{14.8}$$

where the *weak mixing angle* θ_W satisfies the condition

$$\tan\theta_W = \frac{q}{g}, \tag{14.9}$$

such that the $U(1)$ subgroup generated by the parameter ξ^{EM} leaves (14.6) invariant. To see this, substitute (14.8) into (14.7). Using (14.9) it is then obvious that (14.6) is left invariant under transformations (14.7) with $\xi^Z = 0$. The group generated by ξ^{EM}, which we denote by $U(1)^{EM}$ henceforth, thus remains a manifest local gauge symmetry and is not affected by the nonzero value of the field ϕ. Hence $U(1)^{EM}$ corresponds to the electromagnetic gauge transformations in this model, and the weak mixing angle θ_W characterizes the embedding of $U(1)^{EM}$ into the full gauge group $SU(2) \otimes U(1)$. Observe that, although we have not yet considered a Lagrangian, the symmetry structure of the model is already to a large extent determined by the representation content of the scalar fields. The fact that the model of this section has precisely one massless gauge field is a consequence of choosing a doublet field. For other scalar field configurations one would obtain a different mass spectrum for the gauge fields.

We now redefine the gauge fields W_μ^3 and B_μ in accordance with the decomposition (14.8),

$$\begin{aligned} W_\mu^3 &= \cos\theta_W\, Z_\mu + \sin\theta_W\, A_\mu, \\ B_\mu &= \cos\theta_W\, A_\mu - \sin\theta_W\, Z_\mu. \end{aligned} \tag{14.10}$$

Let us now examine how the various gauge fields transform under the two $U(1)$ transformations parametrized by ξ^{EM} and ξ^Z. Using the infinitesimal gauge transformations of W_μ^a and B_μ in terms of the original parameters of $SU(2) \otimes U(1)$,

$$\begin{aligned} \delta W_\mu^a &= \partial_\mu \xi^a + g\epsilon^a_{bc}\xi^b W_\mu^c, \\ \delta B_\mu &= \partial_\mu \xi, \end{aligned} \tag{14.11}$$

it follows that the fields A_μ and Z_μ transform according to

$$\begin{aligned} \delta A_\mu &= \partial_\mu \xi^{EM}, \\ \delta Z_\mu &= \partial_\mu \xi^{Z}, \end{aligned} \tag{14.12}$$

which identifies A_μ as the physical photon field. The other field Z_μ corresponds to a neutral massive vector boson, whose mass will be different from that of the fields $W_\mu^{1,2}$ because of the electroweak-mixing effects. The fields $W_\mu^{1,2}$ are electrically charged since they transform under electromagnetic gauge transformations. It is convenient to decompose them according to

$$W_\mu^{\pm} = \tfrac{1}{2}\sqrt{2}(W_\mu^1 \mp iW_\mu^2). \tag{14.13}$$

Under infinitesimal electromagnetic gauge transformations $W_\mu^{\pm}$ transform as

$$\delta W_\mu^{\pm} = \pm i(g\sin\theta_W)\,\xi^{EM} W_\mu^{\pm}, \tag{14.14}$$

which shows that the W bosons associated with $W_\mu^{\pm}$ carry an electric charge equal to $\pm(g\sin\theta_W)$. This charge will be denoted by e, so that we have

$$e = g\sin\theta_W = q\cos\theta_W. \tag{14.15}$$

The electric charge Q^{EM}, measured in units of e, is defined by the lowest-order coupling of the photon field A_μ. From (14.10) we can determine the photon coupling in terms of the lowest-order coupling of the fields B_μ and W_μ^3. The former is determined by the constants q_i defined in (14.2-5), and the latter is given by the $SU(2)$ generator associated with ξ^3. We will denote this generator by i times the hermitean matrix T_3. Since we are only dealing with singlets or doublets, T_3 is equal to zero or to the matrix $\frac{1}{2}\tau_3$. The charge Q^{EM} is thus given by

$$\begin{aligned} eQ^{EM} &= \tfrac{1}{2}q_i\cos\theta_W + g\sin\theta_W T_3 \\ &= e(T_3 + \tfrac{1}{2}q_i/q) \\ &= e(T_3 + \tfrac{1}{2}Y). \end{aligned} \tag{14.16}$$

The operator Y, which is often called the "weak hypercharge", measures the $U(1)$ "charges" q_i in units of the coupling constant q of the field ϕ. Hence ϕ has $Y = 1$ by definition. According to (14.16) charge *differences* within $SU(2)$ multiplets are necessarily multiples of e. Using (14.10) we can also derive a similar expression for the lowest-order coupling of the neutral vector boson Z_μ to the other fields, which we denote by Q^Z,

$$\begin{aligned} Q^Z &= -\tfrac{1}{2}q_i\sin\theta_W + g\cos\theta_W T_3, \\ &= -\tfrac{1}{2}q\sin\theta_W Y + g\cos\theta_W T_3, \\ &= \frac{g}{\cos\theta_W}\Big(T_3 - \sin^2\theta_W Q^{EM}\Big). \end{aligned} \tag{14.17}$$

The Lagrangian for the gauge fields can now be presented. We define the Lie algebra valued form

$$gW_\mu = gW_\mu^a(\tfrac{1}{2}i\tau_a) = \tfrac{1}{2}i\begin{pmatrix} g\cos\theta_W Z_\mu + eA_\mu & g\sqrt{2}\,W_\mu^+ \\ g\sqrt{2}\,W_\mu^- & -g\cos\theta_W Z_\mu - eA_\mu \end{pmatrix}, \tag{14.18}$$

and its corresponding field strength

$$G_{\mu\nu} = \partial_\mu W_\nu - \partial_\nu W_\mu - g[W_\mu, W_\nu], \qquad (14.19)$$

or, in component form,

$$G^a_{\mu\nu} = \partial_\mu W^a_\nu - \partial_\nu W^a_\mu + g\,\epsilon_{abc} W^b_\mu W^c_\nu. \qquad (14.20)$$

It is convenient to decompose the $SU(2)$ field strengths according to

$$\begin{aligned} G^\pm_{\mu\nu} &\equiv \tfrac{1}{2}\sqrt{2}(G^1_{\mu\nu} \mp G^2_{\mu\nu}) \\ &= \partial^{EM}_\mu W^\pm_\nu - \partial^{EM}_\nu W^\pm_\mu \pm ig\cos\theta_W(W^\pm_\mu Z_\nu - W^\pm_\nu Z_\mu), \qquad (14.21) \\ G^3_{\mu\nu} &= \cos\theta_W(\partial_\mu Z_\nu - \partial_\nu Z_\mu) + \sin\theta_W(\partial_\mu A_\nu - \partial_\nu A_\mu) - ig(W^+_\mu W^-_\nu - W^+_\nu W^-_\mu), \qquad (14.22) \end{aligned}$$

where $\partial^{EM}_\mu W^\pm_\nu = (\partial_\mu \mp ieA_\mu)W^\pm_\nu$. The $U(1)$ field strength becomes

$$\begin{aligned} G^0_{\mu\nu} &\equiv \partial_\mu B_\nu - \partial_\nu B_\mu \\ &= \cos\theta_W(\partial_\mu A_\nu - \partial_\nu A_\mu) - \sin\theta_W(\partial_\mu Z_\nu - \partial_\nu Z_\mu). \qquad (14.23) \end{aligned}$$

In terms of the above field strengths the gauge field Lagrangian is equal to

$$\mathcal{L}_G = -\tfrac{1}{2}G^+_{\mu\nu}G^-_{\mu\nu} - \tfrac{1}{4}G^3_{\mu\nu}G^3_{\mu\nu} - \tfrac{1}{4}G^0_{\mu\nu}G^0_{\mu\nu}. \qquad (14.24)$$

The quadratic terms in (14.24) read

$$\mathcal{L}_0 = -\tfrac{1}{2}(\partial_\mu W^+_\nu - \partial_\nu W^+_\mu)(\partial_\mu W^-_\nu - \partial_\nu W^-_\mu) - \tfrac{1}{4}(\partial_\mu Z_\nu - \partial_\nu Z_\mu)^2 - \tfrac{1}{4}(\partial_\mu A_\nu - \partial_\nu A_\mu)^2. \qquad (14.25)$$

Note that the normalization factor for the W-fields is different, because $W^\pm_\mu$ is complex. The normalization convention for real and complex fields was already discussed in sections 1 and 2. In addition there are cubic and quartic gauge field interactions. The electromagnetic interactions follow from replacing the derivatives on the charged fields $W^\pm_\mu$ by the covariant derivatives ∂^{EM}_μ defined above. In addition there is an interaction with the photon field wich is separately gauge invariant and follows from (14.22). It gives rise to a magnetic moment for the W, and is equal to

$$\mathcal{L}^{\rm mag} = ie(\partial_\mu A_\nu - \partial_\nu A_\mu)W^{+\mu}W^{-\nu}. \qquad (14.26)$$

To find the masses of the $W^\pm_\mu$ and Z_μ bosons we examine the Lagrangian for the scalar field ϕ,

$$\mathcal{L}_\phi = -|D_\mu\phi|^2, \qquad (14.27)$$

where the explicit form for the covariant derivative on ϕ is equal to

$$\begin{aligned} D_\mu\phi &= \begin{pmatrix} \partial_\mu - ieA_\mu - \tfrac{1}{2}ig\cos^{-1}\theta_W(1-\sin^2\theta_W)Z_\mu & -\tfrac{1}{2}\sqrt{2}\,igW^+_\mu \\ -\tfrac{1}{2}\sqrt{2}\,igW^-_\mu & \partial_\mu + \tfrac{1}{2}ig\cos^{-1}\theta_W Z_\mu \end{pmatrix}\begin{pmatrix}\phi_1 \\ \phi_2\end{pmatrix}, \\ &= \begin{pmatrix} -\tfrac{1}{2}igW^+_\mu\rho \\ \tfrac{1}{2}\sqrt{2}\left(\partial_\mu\rho + \tfrac{1}{2}i\dfrac{g}{\cos\theta_W}Z_\mu\rho\right)\end{pmatrix}. \qquad (14.28) \end{aligned}$$

Substituting this result into (14.27) we find

$$\mathcal{L}_\phi = -\tfrac{1}{2}(\partial_\mu\rho)^2 - \tfrac{1}{2}g^2\cos^{-2}\theta_W\rho^2 Z_\mu^2 - \tfrac{1}{4}g^2\rho^2|W_\mu^+|^2. \tag{14.29}$$

We can now read off the masses of the W and the Z bosons after combining (14.25) with (14.29) and substituting $\rho = v$.

$$M_W = \tfrac{1}{2}gv, \qquad M_Z = \tfrac{1}{2}\frac{gv}{\cos\theta_W}. \tag{14.30}$$

This leads to the well-known relation

$$\frac{M_W}{M_Z} = \cos\theta_W. \tag{14.31}$$

The gauge boson couplings to the fermions follow directly from substituting covariant derivatives into the free massless Dirac Lagrangians,

$$\begin{aligned}\mathcal{L}_f &= -\overline{\psi}_L \not{D}\psi_L - \overline{p}_R\not{D}p_R - \overline{n}_R\not{D}n_R \\ &= -\overline{\psi}_L\hat{\not\partial}\psi_L - \overline{p}_R\hat{\not\partial}p_R - \overline{n}_R\hat{\not\partial}n_R \\ &\quad + \tfrac{1}{2}\sqrt{2}\,ig\Big(W_\mu^-\,\overline{n}_L\gamma^\mu p_L + W_\mu^+\,\overline{p}_L\gamma^\mu n_L\Big) \\ &\quad + \frac{ig}{\cos\theta_W}Z_\mu\Big(\tfrac{1}{2}\overline{p}_L\gamma^\mu p_L - \tfrac{1}{2}\overline{n}_L\gamma^\mu n_L\Big),\end{aligned} \tag{14.32}$$

where we have used (14.17) and the definition

$$\hat{\partial}_\mu = \partial_\mu - i\Big(eA_\mu - gZ_\mu\frac{\sin^2\theta_W}{\cos\theta_W}\Big)Q^{EM}. \tag{14.33}$$

The Fermi coupling constant G_F is defined by the strength of the four-fermion coupling caused by the exchange of a charged intermediate W-boson in the limit of zero momentum transfer. In this model its value is given by

$$\begin{aligned}\frac{G_F}{\sqrt{2}} &= \frac{1}{4}\frac{(\frac{1}{2}\sqrt{2}g)^2}{M_W^2} \\ &= \frac{1}{2v^2} \\ &= \frac{e^2}{8M_W^2\sin^2\theta_W}.\end{aligned} \tag{14.34}$$

We now discuss the generation of fermion masses in this model. Mass terms are constructed from the product of a right-handed and a left-handed fermion field. However, the right- and left-handed fermions belong to different $SU(2)$ representations, so that a direct construction of an invariant mass term is excluded. Therefore the only way for the fermions have to acquire masses is via a Yukawa coupling of the scalar doublet ϕ to products of a right- and a left-handed fermion field. Expanding ϕ about v will then lead to fermionic mass terms. In order to construct the necessary Yukawa couplings we first form two left-handed $SU(2)$ singlets, ψ_1 and ψ_2, by taking the invariant products of ψ_L with ϕ,

$$\begin{aligned}\psi_1 &= \phi^{*a}\psi_{L_a} = \phi_1^*\,p_L + \phi_2^*\,n_L, \\ \psi_2 &= -\epsilon^{ab}\phi_a\psi_{L_b} = -\phi_1\,n_L + \phi_2\,p_L.\end{aligned} \tag{14.35}$$

Using the parametrization (14.6) these singlets take the following form

$$\begin{aligned} \psi_1(x) &= \frac{1}{\sqrt{2}}\rho(x)\, n_L(x), \\ \psi_2(x) &= \frac{1}{\sqrt{2}}\rho(x)\, p_L(x). \end{aligned} \tag{14.36}$$

Under U(1) ψ_1 and ψ_2 transform as

$$\begin{aligned} \psi_1 &\to \psi_1' = e^{\frac{1}{2}i(q_1-q)\xi}\, \psi_1, \\ \psi_2 &\to \psi_2' = e^{\frac{1}{2}i(q_1+q)\xi}\, \psi_2. \end{aligned} \tag{14.37}$$

We can now construct two invariant Yukawa couplings, if we assume the following relations among the $U(1)$ coupling constants,

$$q_2 = q_1 + q, \tag{14.38a}$$

$$q_3 = q_1 - q. \tag{14.39a}$$

The corresponding invariants are, respectively,

$$\mathcal{L}_p = -\sqrt{2}\, G_p\, \overline{p}_R \psi_2 + h.c., \tag{14.38b}$$

$$\mathcal{L}_n = -\sqrt{2}\, G_n\, \overline{n}_R \psi_1 + h.c., \tag{14.39b}$$

which, using (14.6), can be written as

$$\mathcal{L}_p = -G_p\, \rho\, \overline{p}_R p_L + h.c., \tag{14.38c}$$

$$\mathcal{L}_n = -G_n\, \rho\, \overline{n}_R n_L + h.c. \tag{14.39c}$$

The coupling constants G_p and G_n can be chosen real by absorbing possible phases into the definition of p_R and n_R. Expanding $\rho(x)$ about the constant v then gives rise to the following expression for the masses

$$m_p = G_p v, \qquad m_n = G_n v. \tag{14.40}$$

Combining (14.3) and (14.40) it follows that,

$$G_{p,n} = \sqrt{\sqrt{2}\, G_F}\; m_{p,n}, \tag{14.41}$$

which shows that the Higgs field ρ couples weakly to the fermions and also that its coupling is proportional to the fermion mass. Consequently Higgs bosons are expected to couple more strongly to heavy flavours.

Let us now briefly discuss the two versions of this prototype model. In its *leptonic* version p and n correspond to the neutrino and its corresponding lepton, respectively. The right-handed neutrino is assumed to decouple from the other particles as a free massless field. Therefore one ignores the Yukawa coupling $\mathcal{L}_p$ and the corresponding restriction (14.38a), and chooses $q_2 = 0$. The model then depends on three independent gauge coupling constants q_1, θ_W, and g. As follows from (14.17) the requirement that the (left-handed) neutrino is electrically neutral forces one to choose $q_1 = -q$. The condition (14.39a) then leads to $q_3 = -2q$. Consequently, the left-handed leptons have $Y = -1$ and the right-haded lepton has $Y = 2$.

The *hadronic* version of the model requires the presence of both Yukawa couplings in order to generate masses for both the quarks corresponding to p and n. Because of the two restrictions (14.38a) and (14.39a) we have again three independent gauge coupling constants, say, q_1, θ_W, and g. Since the mass terms (14.40) and (14.41) are invariant under electromagnetic gauge transformations the coupling of the photon must be purely vectorlike. This can be verified explicitly by using (14.16), (14.38a) and (14.39a). The charges of the two quarks are equal to $\frac{1}{2}e(q_1/q + 1)$ and $\frac{1}{2}e(q_1/q - 1)$. The choice $q_1 = \frac{1}{3}q$ leads to the desired quark charges $\frac{2}{3}e$ and $-\frac{1}{3}e$, and also implies $q_2 = \frac{4}{3}q, q_3 = -\frac{2}{3}q$. Therefore the left-handed quarks have $Y = \frac{1}{3}$, and the right-handed quarks have $Y = \frac{4}{3}$ and $Y = -\frac{2}{3}$

Let us now consider this model coupled to a full generation, i.e. to a lepton pair and three (because of colour) quark pairs. An important property of this fermion configuration is that, provided we make the weak-hypercharge assignments for leptons and quarks that we found above, this model is anomaly-free. By anomaly-free we mean that certain divergent Feynman diagrams, consisting of fermion loops with external gauge fields, that are known to spoil the renormalizability of the theory, are absent (or, at least, their leading divergence cancels). For the model at hand, it is only the $U(1)$ gauge field coupling that may lead to difficulties. In order to eliminate the anomaly, we must satisfy two conditions. According to the first one, the sum of the hypercharges of the left-handed fermion doublets must equal the sum of the hypercharges of the right-handed doublets. Since there are only left-handed doublets, we must have

$$\sum_{\substack{\text{fermion}\\ \text{doublets}}} Y = 0. \tag{14.42}$$

According to the discussion above, a lepton doublet has $Y = -1$, and a quark doublet has $Y = \frac{1}{3}$. Since a generation contains three quark doublets, the hypercharges add up to zero in agreement with (14.42).

The second condition states that the sum of Y^3 for all the left-handed fermions must equal the sum of Y^3 for all right-handed fermions,

$$\sum_{\substack{\text{left-handed}\\ \text{fermions}}} Y^3 = \sum_{\substack{\text{right-handed}\\ \text{fermions}}} Y^3. \tag{14.43}$$

According to the hypercharge assignments, the leptons contribute $2(-1)^3$ to the left-hand side and $(-2)^3$ to the right-hand side of (14.43). For each quark colour, the corresponding contributions are $2(\frac{1}{3})^3$ and $(\frac{4}{3})^3 + (-\frac{2}{3})^3$, respectively. With these values it is easy to verify that also (14.43) is satisfied,

$$2(-1)^3 + 2(\tfrac{1}{3})^3 = (-2)^3 + 3\Big((\tfrac{4}{3})^3 + (-\tfrac{2}{3})^3\Big). \tag{14.44}$$