4

Yang–Mills theories

4.1 The Yang–Mills field

The successful and simple theory which unifies the weak and electromagnetic interactions is based on the group SU(2) × U(1). We develop the theory in several steps. First we describe, in this chapter, the main features of a gauge theory. Then we will describe a theory containing only electrons and the corresponding neutrinos. Finally, the theory is extended to incorporate hadrons.

The structure of a Yang–Mills theory is almost completely determined by the requirement that the internal symmetry transformations of the fields can be carried out independently at different space-time points. In other words, the theory is invariant under local transformations. Let Ψ be a multiplet of n Dirac fields. The multiplets belong to representations of the group SU(N). We define a transformation of the fermion fields by

$$\begin{aligned} \Psi &\longrightarrow \Psi' = U\Psi, \\ \bar{\Psi} &\longrightarrow \bar{\Psi}' = \bar{\Psi}U^{\dagger}, \end{aligned} \tag{4.1}$$

with U a unitary matrix. We represent U by

$$U = \mathrm{e}^{\mathrm{i}\alpha_j \lambda_j/2}, \tag{4.2}$$

with $j = 1, 2, \ldots,\ N^2 - 1$, where λ_j are the generators of the group with $\lambda_j = \lambda_j^{\dagger}$ and the α_j are real. The generators are familiar in simple cases. When the fermions belong to the fundamental representation of SU(2), the λ_j are the Pauli matrices; for SU(3) they are the Gell-Mann matrices. We distinguish two cases:

(i) when all α_j are constant, we call it a global transformation;
(ii) when the $\alpha_j = \alpha_j(x)$ are functions of x_μ, we call it a local or gauge transformation.

https://doi.org/10.1017/9781009402378.005 Published online by Cambridge University Press

The free Dirac Lagrangian

$$\mathcal{L} = \mathrm{i}\bar{\Psi}(x)\gamma_\mu \partial^\mu \Psi(x) = \mathrm{i}\bar{\Psi}(x)\not{\partial}\Psi(x) \tag{4.3}$$

is invariant under global transformations.

If we allow α_j to be a function of x, then (4.3) is no longer invariant. In fact, the Lagrangian transforms into

$$\mathcal{L} \longrightarrow \mathcal{L}' = \bar{\Psi}\mathrm{i}\gamma_\mu \left[\partial^\mu + \frac{\mathrm{i}}{2}\alpha_j^\mu(x)\lambda_j\right]\Psi, \tag{4.4}$$

with $\alpha_j^\mu(x) = \partial\alpha_j(x)/\partial x_\mu$. Following Yang and Mills, we introduce a set of vector fields $B_i^\mu(x)$ and couple them to the currents as follows:

$$\mathcal{L} = \bar{\Psi}\mathrm{i}\gamma_\mu\left[\partial^\mu + \Lambda_j B_j^\mu\right]\Psi + \mathcal{L}_B, \tag{4.5}$$

where Λ_j is a set of matrices still to be determined and $\mathcal{L}_B$ is a function of the B_j^μ terms only. Each vector field is characterized by a Lorentz index μ and an internal symmetry index j. We now demand that $\mathcal{L}$ remains invariant under the transformations (4.2) with α_j a function of x; this will require that B_j^μ transforms in such a way as to cancel out the additional term in (4.4). Let $\hat{B}_j^\mu$ be the transformed vector field. Then, for $\mathcal{L}$ to remain invariant,

$$U^+ \frac{\partial}{\partial x_\mu} U + U^+ \Lambda_i U B_i^\mu = \Lambda_i \hat{B}_i^\mu \tag{4.6}$$

must hold (see Problem 1). Since the λ_j terms form a complete set of $N \times N$ traceless matrices, we can attempt to write

$$\Lambda_k = \frac{\mathrm{i}}{2} e \lambda_k; \tag{4.7}$$

the imaginary i is there because the λ_k terms and the B terms are Hermitian. Considering infinitesimal transformations,

$$U \simeq 1 + \frac{\mathrm{i}}{2}\alpha_j\lambda_j, \tag{4.8}$$

with

$$\begin{aligned} \left[\frac{1}{2}\lambda_i, \frac{1}{2}\lambda_j\right] &= \mathrm{i} f_{ijk}\frac{1}{2}\lambda_k, \\ \mathrm{Tr}[\lambda_i\lambda_j] &= 2\delta_{ij}, \end{aligned} \tag{4.9}$$

https://doi.org/10.1017/9781009402378.005 Published online by Cambridge University Press

and f_{ijk} are the structure constants of the group. For the infinitesimal transformation we can solve for $\hat{B}_j^\mu$ in (4.6). A convenient method is to rewrite (4.6) as

$$e\lambda_i \hat{B}_i^\mu = U^+ e\lambda_i U B_i^\mu + \lambda_i \frac{\partial \alpha_i}{\partial x_\mu} \tag{4.10}$$

and then expand the unitary matrices to first order in α^i and use the relation in (4.9) to obtain

$$\hat{B}_k^\mu = B_k^\mu + f_{ijk}\alpha_i B_j^\mu + \frac{1}{e}\frac{\partial \alpha_k}{\partial x^\mu}. \tag{4.11}$$

It is convenient to introduce a covariant derivative,

$$D^\mu = \partial^\mu + \frac{\mathrm{i}}{2} e\lambda_j B_j^\mu, \tag{4.12}$$

and rewrite the fermion part in (4.5) as

$$\mathcal{L}_\mathrm{F} = \mathrm{i}\bar{\Psi} \not{D} \Psi = \mathrm{i}\bar{\Psi}\gamma_\mu \left(\partial^\mu + \frac{\mathrm{i}}{2} e\lambda_j B_j^\mu\right)\Psi. \tag{4.13}$$

Covariant derivatives are useful in generating gauge-invariant Lagrangians. A Lagrangian invariant under global transformations becomes locally gauge-invariant when all ordinary derivatives are replaced by covariant derivatives. In quantum electrodynamics this replacement is the well-known minimal-substitution law.

Next we must construct $\mathcal{L}_\mathrm{B}$. It must be Lorentz-invariant and invariant under $B \to \hat{B}$. It must also contain the kinetic term of the B_μ fields. In analogy to the procedure of obtaining gauge-invariant field strengths in electrodynamics, we define

$$F_i^{\mu\nu} = \partial^\nu B_i^\mu - \partial^\mu B_i^\nu + e f_{ijk} B_j^\mu B_k^\nu. \tag{4.14}$$

If we introduce the vector notation

$$\vec{B}^\mu = \left(B_1^\mu, B_2^\mu, \ldots, B_k^\mu\right), \tag{4.15}$$

where $k = N^2 - 1$ and

$$(\vec{A} \times \vec{B})_i = f_{ijk} A_j B_k, \tag{4.16}$$

we can write (4.11) and (4.14) as

$$\vec{F}^{\mu\nu} = \partial^\nu \vec{B}^\mu - \partial^\mu \vec{B}^\nu + e\vec{B}^\mu \times \vec{B}^\nu. \tag{4.17}$$

The last term in (4.17) does not occur in electrodynamics and is introduced to assure that $\vec{F}^{\mu\nu}$ transforms as a vector under gauge transformations. A reason for introducing a generalized $\vec{F}_{\mu\nu}$ is given in Problem 1; in the same problem we

https://doi.org/10.1017/9781009402378.005 Published online by Cambridge University Press

discuss the gauge invariance of $\mathcal{L}_\mathrm{B}$. We can now build a scalar Lagrange function for the $\vec{B}^\mu$ fields,

$$\mathcal{L}_\mathrm{B} = -\frac{1}{4}\vec{F}^{\mu\nu}\vec{F}_{\mu\nu} = -\frac{1}{4}F_i^{\mu\nu}F_{i,\mu\nu}. \tag{4.18}$$

A theory with $\mathcal{L}_\mathrm{B}$ alone is called a pure Yang–Mills theory.

For Lagrangians invariant under symmetries, we can also define currents of the original Lagrangian, which are given by

$$J_\alpha^\mu(x) = \frac{\partial\mathcal{L}}{\partial(\partial_\mu\Psi)}\frac{\lambda_\alpha}{2}\Psi = \mathrm{i}\bar{\Psi}(x)\gamma^\mu\frac{\lambda_\alpha}{2}\Psi(x). \tag{4.19}$$

The invariance of the theory implies that the currents are conserved. We note that these are the same currents as those we introduced in Chapter 3.

To sum up, we constructed a theory that is invariant under gauge transformations. The complete Lagrangian is

$$\mathcal{L} = \mathcal{L}_\mathrm{F} + \mathcal{L}_\mathrm{B}.$$

We found that the invariance requirements are fulfilled by introducing vector fields coupled to conserved currents.

Such a theory is a candidate for particle physics. It describes the interaction of massless fermions with massless gauge bosons. It possesses a symmetry that can be SU(2), SU(3), or a larger unitary group. The case of SU(3) is, in fact, realized in Nature as the theory of strong interactions. There the vector bosons are the gluons coupled to quarks and the symmetry is the SU(3)-color. The color symmetry remains unbroken. The electroweak theory is more complicated because it contains masses for the quarks and the gauge bosons. It is a broken symmetry, to be developed in Chapters 5–7.

4.2 Gauge invariance in scalar electrodynamics

The electrodynamic field is described by the four-vector

$$A_\mu(x) = \big(\Phi(x), \vec{A}(x)\big), \tag{4.20}$$

whose components are the standard scalar and vector potentials. The electric and magnetic fields are now determined by

$$\vec{E} = -\vec{\nabla}A_0 - \frac{\partial\vec{A}}{\partial t}, \quad \vec{B} = \vec{\nabla}\times\vec{A}. \tag{4.21}$$

https://doi.org/10.1017/9781009402378.005 Published online by Cambridge University Press

However, to one set of fields $(\vec{E}, \vec{B})$ there correspond many potentials A_μ. The primed potentials obtained by the gauge transformation

$$A'_\mu(x) = A_\mu(x) + \frac{\partial \Lambda(x)}{\partial x^\mu}, \tag{4.22}$$

with $\Lambda(x)$ an arbitrary scalar function, give the same $\vec{E}$ and $\vec{B}$. If someone solves a problem with $A_\mu(x)$ and somebody else does it with $A'_\mu(x)$, both should get the same physical result. In general, only those quantities which are invariant under gauge transformations have physical meaning. Gauge invariance has far-reaching implications for the theories, as we discuss in the following chapters, and clever choices of gauge lead to substantial simplifications of problems.

Here we use gauge invariance to discuss the degrees of freedom for the electromagnetic field. For the pure electromagnetic case

$$\mathcal{L} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu}, \tag{4.23}$$

with

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu. \tag{4.24}$$

The equation of motion

$$\partial_\mu F^{\mu\nu} = 0, \tag{4.25}$$

when written in terms of A_μ, becomes

$$\partial^\mu \partial_\mu A_\nu - \partial_\nu \partial^\mu A_\mu = 0. \tag{4.26}$$

It is well known that the $\vec{E}$ and the $\vec{B}$ fields satisfy a wave equation, but A_μ does not. In order to recover a wave equation from (4.26), we impose the condition

$$\partial_\mu A^\mu = 0 \quad \text{(Lorentz gauge).} \tag{4.27}$$

We used the gauge freedom to obtain this result, but still we did not exhaust all possible gauge transformations, because any gauge function $\chi(x)$ that satisfies

$$\partial^\mu \partial_\mu \chi(x) = 0 \tag{4.28}$$

is still consistent with (4.27). We take advantage of this freedom in order to show that a photon has only two degrees of freedom.

A free photon is represented by a plane wave

$$A_\mu(x) = \varepsilon_\mu e^{-ikx}, \tag{4.29}$$

https://doi.org/10.1017/9781009402378.005 Published online by Cambridge University Press

where ε_μ is called the polarization vector. By substituting (4.29) into (4.26), we find $k^2 = 0$, or $m = 0$, and from (4.27) we find

$$k_\mu \varepsilon^\mu = 0 \quad \text{(Lorentz gauge).} \tag{4.30}$$

We choose a coordinate system with the z-axis along $\vec{k}$ and decompose ε_μ into longitudinal and transverse parts:

$$A_\mu(x) = \left(\varepsilon_\mu^\parallel + \varepsilon_\mu^\perp\right)\mathrm{e}^{-ikx}. \tag{4.31}$$

From (4.30) we conclude that $\varepsilon_\mu^\parallel$ is proportional to k_μ. Instead of the electromagnetic field $A_\mu(x)$, we can choose another one given through a gauge transformation with

$$\chi(x) = \mathrm{i}c\mathrm{e}^{-\mathrm{i}kx}, \qquad c = \text{constant}. \tag{4.32}$$

The new field is

$$A'_\mu(x) = \left(\varepsilon_\mu^\parallel + \varepsilon_\mu^\perp\right)\mathrm{e}^{-ikx} + ck_\mu \mathrm{e}^{-ikx}. \tag{4.33}$$

By an appropriate choice of the constant c, we can eliminate k_μ, i.e. we can gauge away the longitudinal degrees of freedom. Therefore a free photon has only two degrees of freedom.

The argument fails for a massive $A_\mu(x)$ field. The addition of a mass term $\frac{1}{2}\mu^2 A_\mu A^\mu$ to Eq. (4.23) breaks the gauge invariance of the theory. In this case the equation of motion

$$\partial_\mu F^{\mu\nu} + \mu^2 A^\nu = 0 \tag{4.34}$$

implies

$$\mu^2 \partial_\nu A^\nu = 0. \tag{4.35}$$

For $\mu^2 \neq 0$, A_μ satisfies the Lorentz condition, which again eliminates one degree of freedom. But now we cannot repeat the steps between Eqs. (4.28) and (4.33). Therefore a massive field has three degrees of freedom.

Next we study the interaction of a photon with a charged scalar field: scalar electrodynamics. The Lagrangian is

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + (D^\mu\phi)^*(D_\mu\phi) - V(\phi^*\phi). \tag{4.36}$$

Here $D^\mu = (\partial^\mu + \mathrm{i}eA^\mu)$ is the covariant derivative, in agreement with the rule of replacing ordinary derivatives with covariant ones. We represent the field as

$$\phi(x) = \frac{1}{\sqrt{2}}(\phi_1(x) + \mathrm{i}\phi_2(x)) \tag{4.37}$$

https://doi.org/10.1017/9781009402378.005 Published online by Cambridge University Press

and introduce the potential

$$V(\phi^*\phi) = -\mu^2\phi^*\phi + \lambda(\phi^*\phi)^2. \tag{4.38}$$

This Lagrangian is invariant under the local transformation

$$\begin{aligned} A^\mu(x) &\longrightarrow A^\mu(x) + \partial^\mu\omega(x), \\ \phi(x) &\longrightarrow \mathrm{e}^{-\mathrm{i}\omega(x)}\phi(x), \\ \phi^*(x) &\longrightarrow \mathrm{e}^{\mathrm{i}\omega(x)}\phi^*(x), \end{aligned} \tag{4.39}$$

where $\omega(x)$ is an arbitrary real function. The photon is again massless and carries two independent degrees of freedom. This follows from the same arguments as in the free-photon case. First we can go to the Lorentz gauge, which again simplifies the equations of motion both for $A_\mu(x)$ and for the scalar field $\phi(x)$. Then, since the photon is again massless, we can introduce a new gauge transformation satisfying (4.28) and eliminate the longitudinal degrees of freedom. In gauge theories, masses for the gauge bosons are introduced, not through the ad-hoc procedure of the previous paragraph, but through spontaneous breaking of the symmetry. We study this topic in Chapter 6 and return to scalar electrodynamics in Section 5.3.

Problems for Chapter 4

1. Consider the fermion Lagrangian in Eq. (4.13).
   (i) Show that invariance under the local transformation (4.2) requires that the covariant derivative satisfies
   $$D'_\mu = U^\dagger D_\mu U.$$
   (ii) Show that this result, together with the definition of D_μ, implies the transformation property for B^μ given in (4.6).
   (iii) Show that the Hermitian quantity $F_{\mu\nu} = -\mathrm{i}[D_\mu, D_\nu]$ is the field tensor whose transformation under local transformations is
   $$F^{i'}_{\mu\nu} = U^\dagger F^i_{\mu\nu} U.$$

   It is now easy to build an invariant term given by
   $$\mathcal{L}_{\mathrm{YM}} = \frac{1}{4}\,\mathrm{Tr}\big(F^i_{\mu\nu}F_i^{\mu\nu}\big)\,.$$
2. Show that, under local transformations, the field-strength tensor $F^i_{\mu\nu}$ transforms as a vector on the index i. The result holds including terms linear in $\varepsilon_i(x)$, i.e.
   $$\vec{F}'_{\mu\nu} = \vec{F}_{\mu\nu} - \vec{\varepsilon}\times\vec{F}_{\mu\nu} + O(\varepsilon^2).$$

https://doi.org/10.1017/9781009402378.005 Published online by Cambridge University Press

You may need the Jacobi identity

$$f_{ABm} f_{mC\alpha} + f_{BCm} f_{mA\alpha} + f_{CAm} f_{mB\alpha} = 0,$$

which follows from

$$\left[\frac{\lambda_A}{2}, \frac{\lambda_B}{2}\right] = \mathrm{i} f_{ABm} \frac{\lambda_m}{2} \quad \text{and} \quad \mathrm{Tr}(\lambda_A \lambda_B) = 2\delta_{AB}.$$

Select bibliography

Anderson, P. W. (1963), *Phys. Rev*. **130**, 439
Yang, C. N., and Mills, R. (1954), *Phys. Rev*. **96**, 191