САНКТ-ПЕТЕРБУРГСКИЙ
ГОСУДАРСТВЕННЫЙ
УНИВЕРСИТЕТ

ПРИОРИТЕТНЫЙ
НАЦИОНАЛЬНЫЙ ПРОЕКТ
"ОБРАЗОВАНИЕ"

Проект «Инновационная образовательная среда в классическом университете»

Пилотный проект № 22 «Разработка и внедрение
инновационной образовательной программы «Прикладные математика и физика»»

Физический факультет

Кафедра физики высоких энергий и элементарных частиц

Ю.В.Новожилов, В.Ю.Новожилов

ELEMENTARY PARTICLE PHYSICS I.
General properties

Учебно-методическое пособие

.

.

Санкт Петербург
2007 г.

- Рецензент: профессор кафедры физики высоких энергий и элементарных частиц физического факультета СПбГУ, д.ф.-м.н. В.А.Франке

- Печатается по решению методической комиссии физического факультета СПбГУ.

- Рекомендовано Ученым советом физического факультета СПбГУ.

Ю.В.Новожилов, В.Ю.Новожилов
Elementary particle physics I. General properties. – СПб., 2007

В учебно-методическом пособии обсуждаются основы описания состояний в релятивистской квантовой теории (учет пространственно-временных симметрий, свойства пространства состояний Фока). В работе дано последовательное изложение современной концепции построения калибровочных теорий взаимодействия элементарных частиц. Рассмотрена калибровочная теория сильных взаимодействий (Квантовая хромодинамика) на примере одного поколения кварков, проделан анализ симметрии основного состояния теории. В пособии рассмотрена теория электрослабого взаимодействия Вайнберга-Салама и проанализированы процессы, определяемые вершинами взаимодействия в этой теории.
Пособие предназначено для студентов старших курсов и аспирантов физических специальностей.

Contents

1 Introduction

The description of states of elementary particles relies on quantum mechanical regularities and concepts of symmetry. Symmetry principles are more general than laws of motion; thus finding them is the first step in the formation of laws. Knowledge of the symmetry group allows one to find immediately a natural set of basic quantities (the generators of the group and their invariants with which to describe particle states. The symmetry method is a general approach to the physical description of any phenomena. It allows one to introduce a logically closed physical description in which every stage of the approximation to reality is endowed with precise symmetry properties. Thus at every stage of the approximation there exists an exact physical interpretation of the theory in terms of an idealized world. Symmetries in fundamental physics are divided into:

• space-time (invariance, conservation laws, quantum numbers, selection rules);

• internal, reflecting interaction properties (symmetry of Lagrangians, gauge symmetry);

• symmetry of solutions (automodelity, renormalization group).

Elementary (or fundamental) particles are defined as particles that are not composed, at the current level of knowledge, of other elementary particles. Experiments have revealed twelve elementary fermions (with spin $s = 1/2$) and four gauge bosons (with spin $s = 1$), not counting the respective antiparticles.

1.1 Quarks and leptons

The fundamental building blocks of strongly interacting particles (hadrons) , the *quarks*, and the fundamental fermions lacking strong interactions, the *leptons*, are summarized in Table 1. Masses are as quoted by the Particle Data Group. Table 1 shows that the twelve fermions form three generations, each composed of two leptons and two quarks. Each charged fermion has an antiparticle counterpart: $\bar{u}, \bar{d}, \bar{c}, \bar{s}, \bar{t}, \bar{b}, e^+, \mu^+, \tau^+$. Neutrinos are believed to be truly neutral.

The quark masses quoted in Table 1 are those which emerge when quarks are probed at distances short compared with 1 fm, the characteristic size of strongly interacting particles and the scale at which QCD becomes too strong to utilize perturbation theory. When regarded as constituents of strongly interacting particles, however, the u and d quarks act

Table 1: The known quarks and leptons. Masses in GeV except where indicated otherwise. Here and elsewhere we take $c = 1$.

Quarks				Leptons			
Charge 2/3		Charge −1/3		Charge −1		Charge 0	
Mass		Mass		Mass		Mass	
u	0.001–0.005	d	0.003–0.009	e	0.000511	ν_e	< 3 eV
c	1.15–1.35	s	0.075–0.175	μ	0.106	ν_μ	< 190 keV
t	172.5 ± 2.3	b	4.0–4.4	τ	1.777	ν_τ	< 18.2 MeV

as quasi-particles with masses of about 0.3 GeV. The corresponding "constituent-quark" masses of s, c, and b are about 0.5, 1.5, and 4.9 GeV, respectively.

The t (top) quark is the heaviest quark found and is still believed to be a fundamental particle. It completes the third-generation structure of the Standard Model (SM), as the isospin partner of the b (bottom) quark. Although mainly produced via the strong interaction at particle colliders (double production via gluon-gluon fusion o $q\bar{q}$ annihilation), the t quark decays through the weak force to a b quark and a W boson with a branching ratio of almost 100 %. Because of their very large mass and decay rate top quarks, unlike any other quarks, are produced and decay as free particles. With very short lifetime (10^{-25}s), the top quark decays before hadronization can take place. For the same reason no toponium bound state with sharp binding energy are expected in the Standard Model; any evidence of a $t\bar{t}$ bound state would be a sign of physics beyond the SM.

1.2 Color

The quarks are distinguished from the leptons by possessing a three-fold charge known as "color" which enables them to interact strongly with one another. The experimental evidence for color comes from several quarters.

1. Quark statistics. One of the lowest-lying hadrons is a particle known as the Δ^{++}, an excited state of the nucleon first produced in $\pi^+ p$ collisions in the mid-1950s at the University of Chicago cyclotron. It can be represented in the quark model as uuu, so it is totally symmetric in flavor. It has spin $J = 3/2$, which is a totally symmetric combination of the three quark spins (each taken to be 1/2). Moreover, as a ground state, it is expected to contain no relative orbital angular momenta among the quarks. This leads to a paradox if there are no additional degrees of freedom. A state composed of fermions should be totally antisymmetric under the interchange of any two fermions, but what we have described so far is totally symmetric under flavor, spin, and space interchanges, hence totally symmetric under their product. Color introduces an additional degree of freedom under which the interchange of two quarks can produce a minus sign, through the representation $\Delta^{++} \sim \epsilon_{abc} u^a u^b u^c$. The totally antisymmetric product of three color triplets is a color singlet.

2. Electron-positron annihilation to hadrons. The charges of all quarks which can be produced in pairs below a given center-of-mass energy is measured by the ratio

$$R \equiv \frac{\sigma(e^+e^- \to \text{hadrons})}{\sigma(e^+e^- \to \mu^+\mu^-)} = \sum_i Q_i^2 \quad . \tag{1.1}$$

For energies at which only $u\bar{u}$, $d\bar{d}$, and $s\bar{s}$ can be produced, i.e., below the charmed-pair threshold of about 3.7 GeV, one expects

$$R = N_c \left[\left(\frac{2}{3}\right)^2 + \left(\frac{-1}{3}\right)^2 + \left(\frac{-1}{3}\right)^2 \right] = \frac{2}{3} N_c \tag{1.2}$$

for N_c "colors" of quarks. Experiment indicate $R = 2$ in this energy range (with a small positive correction associated with the strong interactions of the quarks), indicating $N_c = 3$.

3. Neutral pion decay. The π^0 decay rate is governed by a quark loop diagram in which two photons are radiated by the quarks in $\pi^0 = (u\bar{u} - d\bar{d})/\sqrt{2}$. The predicted rate is

$$\Gamma(\pi^0 \to \gamma\gamma) = \frac{S^2 m_\pi^3}{8\pi f_\pi^2}\left(\frac{\alpha}{2\pi}\right)^2 \quad , \tag{1.3}$$

where $f_\pi = 131$ MeV and $S = N_c(Q_u^2 - Q_d^2) = N_c/3$. The experimental rate is 7.8 ± 0.6 eV, while Eq. (1.3) gives $7.6S^2$ eV, in accord with experiment if $S = 1$ and $N_c = 3$.

4. Triality. Quark composites appear only in multiples of three. Baryons are composed of qqq, while mesons are $q\bar{q}$ (with total quark number zero). This is compatible with our current understanding of QCD, in which only color-singlet states can appear in the spectrum. Thus, mesons M and baryons B are represented by

$$M = \frac{1}{\sqrt{3}}(q^a \bar{q}'_a) \quad , \quad B = \frac{1}{\sqrt{6}}(\epsilon_{abc} q^a q'^b q''^c) \quad . \tag{1.4}$$

Direct evidence for the quanta of QCD, the gluons, was first presented in 1979 on the basis of extra "jets" of particles produced in electron-positron annihilations to hadrons. Normally one sees two clusters of energy associated with the fragmentation of each quark in $e^+e^- \to q\bar{q}$ into hadrons. However, in some fraction of events an extra jet was seen, corresponding to the radiation of a gluon by one of the quarks. The transformations which take one color of quark into another are those of the group SU(3). We shall often refer to this group as $\mathrm{SU}(3)_{\mathrm{color}}$ to distinguish it from the $\mathrm{SU}(3)_{\mathrm{flavor}}$ associated with the quarks u, d, and s.

1.3 Fundamental vector bosons

Experiments have established four fundamental vector bosons: the photon γ, the gluon g, the neutral weak boson Z^0 and the charged weak bosons $W^\pm$, which are antiparticles to each other.

The photon has been studied better than the rest. The photon mass is zero. The electric charge e is the source of photons. All electromagnetic interactions are due to photon exchange. The theory describing electromagnetic interactions (quantum electrodynamics) is very well developed.

Gluons are carriers of the strong interactions. A gluon has two color indices. The number of different gluons is eight: $3 \otimes \bar{3} = 8 + 1$. The singlet combination carries no color charge. In contrast to photon in quantum electrodynamics, which is electrically neutral, gluons in quantum chromodynamics (QCD) are carriers of color charges and thus have to emit and absorb gluons. The result is completely unfamiliar behavior for the strong interactions of quarks and gluons: their interaction energy increases in response to attempts to separate them. As a result, free gluons and quarks cannot exist they are "self-confined" in colorless hadrons.

We shall discuss QCD and theory of weak interactions in detail later.

1.4 Low energy limit

At low energies $E \ll M_W$, the interactions of leptons and hadrons are described by QCD + QED up to corrections of order (E/M_W). If we disregard the electromagnetic interactions, we are left with QCD that contains only a few parameters: the renormalization group invariant scale Λ and the running quark masses $m_u, m_d, m_s \ldots$. The quark masses m_u, m_d and m_s are small on a typical hadronic scale like the mass of the rho or of the proton. It makes therefore sense to consider the limit where these masses are set equal to zero (chiral limit). The remaining quarks $c, b, \ldots$ are not light: although one may of course study the theoretical limit in which these masses also vanish, it does not seem to be possible to recover the actual mass values by an expansion around that limiting case. At low energies, a better approximation is obtained if the quarks $c, b, \ldots$ are instead treated as infinitely heavy. In this limit, the degrees of freedom associated with these quarks freeze and may be ignored in the effective low energy theory. In the chiral limit, QCD contains therefore only one parameter, the scale Λ. The mass of the proton is a pure number multiplying Λ, and likewise for all the other states of the theory – the numbers $M_\rho/M_p, M_\Delta/M_p, \ldots$ are determined in a parameter free manner. In this sense, the chiral limit of QCD may be called a theory without any adjustable parameters: QCD is of course unable to predict the value of M_p in GeV units, but it determines all dimensionless hadronic quantities in a parameter free manner. The elastic cross section for pp scattering e.g. is some fixed function of the variables s/M_p^2 and t/M_p^2 , multiplying the square of the Compton wavelength of the proton. It is unfortunately very difficult to really *calculate* masses, cross sections and decay amplitudes in this theory, because the lagrangian of QCD is formulated in terms of quark and gluon fields which do not create asymptotically observed particles. Several methods have therefore been devised in the past to cope with this problem in different regimes of the energy scale:

i) Processes at high energies. At high energies, the effective coupling constant $\alpha_{\rm QCD}$ becomes small, and conventional perturbation theory in $\alpha_{\rm QCD}$ is applicable.

ii) Lattice calculations. This is the only method known today which leads directly from the QCD lagrangian to the mass spectrum, decay matrix elements, scattering lengths etc. On the other hand, the CPU time needed for full fledged QCD calculations is enormous, and one may still have to wait a long time before this program achieves the accuracy one is aiming at in the framework of effective field theory.

iii) Chiral perturbation theory (ChPT). This method exploits the symmetry of the QCD lagrangian and its ground state: one solves in a perturbative manner the constraints imposed by chiral symmetry and unitarity by expanding the Green functions in powers of the external momenta and of the quark masses m_u, m_d and m_s. To illustrate the idea, consider the process $\pi^+(p_1)\pi^-(p_2) \to \pi^0(p_3)\pi^0(p_4)$. Chiral symmetry implies that the corresponding scattering amplitude has the following form near threshold,

$$T = \frac{M_\pi^2 - s}{F_\pi^2} + O(p^4) \;\; ; \;\; s = (p_1 + p_2)^2 \; , \tag{1.5}$$

where $F_\pi = 92.4$ MeV is the pion decay constant, and M_π denotes the pion mass. This result is due to Weinberg, who used current algebra and PCAC to analyse the Ward

identities for the four-point functions of the axial currents. It displays the first order term in a systematic expansion of the scattering amplitude in powers of momenta and of quark masses. This term algebraically dominates the remainder, denoted by the symbol $O(p^4)$, for sufficiently small energies and thus provides an accurate parameterization of the full amplitude near threshold. As one goes away from threshold, the higher order terms come into play.

2 Elements of Relativistic Quantum Theory

2.1 Space-time symmetry

The theory of elementary particles is a relativistic theory. The basis of its apparatus are quantum mechanics and the Einstein relativity principle. The Einstein relativity principle determines the group of space-time symmetry in classical physics as the inhomogeneous Lorentz group, or the *Poincaré* group $\mathcal{P}_+^\uparrow$.

A quantum state corresponds to not one but a number of state vectors ("unit ray") in Hilbert space. Because of this fact, and because the parameter space of the group $\mathcal{P}_+^\uparrow$ is doubly connected, the space-time symmetry group in quantum mechanics is not the Poincare group $\mathcal{P}_+^\uparrow$ but its universal covering group $\overline{\mathcal{P}_+^\uparrow}$. The irreducible unitary representations of $\overline{\mathcal{P}_+^\uparrow}$ describe elementary particle states. The invariants of *Poincaré* group - mass and spin -characterize the invariant properties of particles, and the generators of the *Poincaré* group - the possible observables, i.e.the variables whose independent measurements fix the state of a relativistic particle.

Infinitesimal transformation *Poincaré* group may be written in the form

$$x^{'\mu} = x^\mu + \varepsilon^\mu + \omega^\mu_\nu x^\nu \ , \tag{2.6}$$

where the infinitesimal parameters ε^μ and ω^μ_ν are real and $\omega_{\mu\nu} = -\omega_{\nu\mu}$. The infinitesimal coordinate transformation (2.6) entails a change of state vectors $|\gamma> \rightarrow (1+\delta U)|\gamma>$ which, as a consequence of the continuity of the transformation, will also be infinitesimal:

$$\delta U = iP_\mu \varepsilon^\mu - i\frac{1}{2}M_{\mu\nu}\omega^{\mu\nu}. \tag{2.7}$$

In formula (2.6) we introduce the operators P_μ and $M_{\mu\nu} = -M_{\nu\mu}$, which are Hermitean since $1 + \delta U$ is unitary.

According to (2.6) and (2.7) the operator

$$1 + iP_\mu \varepsilon^\mu \tag{2.8}$$

describes an infinitesimal translation generated by motion of the reference frame:

$$x^\mu \rightarrow x^{'\mu} = x^\mu + \varepsilon^\mu. \tag{2.9}$$

Infinitesimal three-dimensional rotations and Lorentz transformations are decribed by the operator

$$1 - i\frac{1}{2}M_{\mu\nu}\omega^{\mu\nu}. \tag{2.10}$$

The quantity iP_μ is a generator of translations along the axis x^μ, $-iM_{jk}$ is a generator of rotations in the (jk) plane, and the quantity $-iM_{0j}$ is a generator of a pure Lorentz transformation. Ten generators have been enumerated, corresponding to the 10 single parameter transformations of the *Poincaré* group, and these are the basis quantities in relativistic quantum mechanics. The quantity P_μ is called the energy-momentum or four-momentum vector. The three-vector

$$\mathbf{M} = (M_{23}, M_{31}, M_{12})$$

is the orbital angular momentum. Instead of the components M_{0j} one can introduce the three-vector

$$\mathbf{N} = (M_{01}, M_{02}, M_{03}).$$

If $\Phi(x)$ is an operator of some physical quantity, depending on the coordinate x, then from (2.8) it follows that

$$i[P_\mu, \Phi(x)] = \frac{\partial\Phi(x)}{\partial x_\mu}.$$

The development in time of the dynamical variables of the system is determined by the operator P_0: the Hamiltonian of the system.

The commutation relations between the operators P_μ and $M_{\mu\nu}$ have the form

$$[M_{\mu\nu}, M_{\lambda\sigma}] = i[g_{\mu\sigma}M_{\nu\lambda} + g_{\nu\lambda}M_{\mu\sigma} - g_{\mu\lambda}M_{\nu\sigma} - g_{\nu\sigma}M_{\mu\lambda}], \tag{2.11}$$

$$[M_{\mu\nu}, P_\lambda] = i[g_{\nu\lambda}P_\mu - g_{\mu\lambda}P_\nu], [P_\mu, P_\nu] = 0. \tag{2.12}$$

One can construct quantities commuting with all generators using $M_{\mu\nu}$ and P_μ. Such quantities do not change under transformations of the group and will be proportional to the unit matrix in each irreducible representation. One may classify the irreducible representations by the values of these quantities. One can form only one invariant quantity

$$m^2 = P_\mu P^\mu \tag{2.13}$$

from the components of momentum. Being a scalar, m^2 commutes with $M_{\mu\nu}$; as a function of momentum, m^2 commutes with P_μ. For $m^2 \geq 0$ m^2 has the meaning of the square of the rest mass. In this case there exists a second invariant quantity constructed from the momenta: the operator of the sign of the energy

$$\varepsilon = \frac{P_0}{|P_0|} \tag{2.14}$$

with eigenvalues $\varepsilon = \pm 1$. Since proper orthochronous transformations $\mathcal{P}_+^\uparrow$ do not change the sign of the components of a time-like vector, ε is invariant.

Using only one of the quantities $M_{\mu\nu}$ it is impossible to form an invariant operator. There is only one combination of generators $M_{\mu\nu}$ and P_σ which commutes with the momentum P_λ, namely the pseudo-vector

$$\omega_\mu = \frac{1}{2}\varepsilon_{\mu\nu\lambda\sigma}M^{\nu\lambda}P^\sigma \tag{2.15}$$

with the property

$$\omega_\mu P^\mu = 0 \tag{2.16}$$

The square of the pseudo-vector ω_μ is a scalar, and, consequently, commutes with $M_{\mu\nu}$. Thus $\omega^2 = \omega_\mu\omega^\mu$ commutes with all 10 generators.

$$\omega^2 = \frac{1}{2}M_{\mu\nu}M^{\mu\nu}P_\lambda P^\lambda - M_{\mu\sigma}M_{\nu\sigma}P^\mu P^\nu \tag{2.17}$$

The physical meaning of the invariant ω^2 may be easily seen when $m^2 \geq 0$. Let us take the four-momenta P_μ diagonal. Then in the rest system, when acting on the state vector $|\mathbf{p} = 0, m >$ the invariant ω^2 is equal to

$$\omega^2|\mathbf{p} = 0, m >= -(M_{12}^2 + M_{23}^2 + M_{31}^2)|\mathbf{p} = 0, m >= -m^2\mathbf{M}^2|\mathbf{p} = 0, m > \tag{2.18}$$

The eigenvalues of the square of the angular momentum (i.e the square of spin), as is well known in quantum mechanics, have the form $J(J+1)$, where $J = 0, \frac{1}{2}, 1, \frac{1}{2}, ...$ Since ω^2 is an invariant operator in any reference frame, applying it to a state $|\mathbf{p}, m >$ and transforming according to an irreducible representation $\mathcal{P}_+^\uparrow$, gives

$$\omega^2|\mathbf{p}; m, J >= -m^2 J(J+1)|\mathbf{p}; m, J > \tag{2.19}$$

To establish a covariant relation between ω_μ and the spin operator J_k we introduce the space-like normals $n_\mu^{(a)}(a = 1, 2, 3) : n_\mu^{(a)}n^{(b)\mu} = -\delta_{ab}, n_\mu^{(a)}p^\mu = 0$. Together with the velocity $v_\mu = p_\mu/m \equiv n_\mu^{(0)}$, they form a complete set of four-normals: $n_\mu^{(\alpha)}g^{\mu\nu}n_\nu^{(\beta)} = g^{\alpha\beta}$. Since $\omega^\mu n_\mu^{(0)} = 0$, there are only three independent space-like components of ω_μ; in covariant form:

$$\omega_\mu = m\sum_{k=1}^{3} J^k n_\mu^{(k)}, J^k = -\frac{1}{m}\omega^\mu n_\mu^{(k)} \tag{2.20}$$

Introducing the vector $\mathbf{J} = (J^1, J^2, J^3)$, one can find by direct calculation:

$$\omega^2 = -m\mathbf{J}^2, [J^i, J^k] = i\varepsilon^{ijk}J^k \tag{2.21}$$

i.e. the vector $\mathbf{J}$ indeed has the properties of a spin operator. The values of J_3 take $2J+1$ values $J_3' \equiv \sigma = -J, -J+1, ...J-1, J$.

Thus an irreducible unitary representation of the *Poincaré* group is distinguished by the values of mass m, spin J and the sign of the energy (for $m^2 \geq 0$. The energy of physical states is always positive, i.e $\varepsilon \geq 0$. Particles with integral spin $J = 0, 1, 2, ...$ are called bosons; particles with half-integer spin $J = \frac{1}{2}, J = \frac{3}{2}, ...$ are fermions.

In the *canonical basis* (for $m^2 \geq 0$) the components of the momentum p_j and spin projection J_3 on a given fixed axis orthogonal to the four-momentum are diagonal. The complete set of mutually commuting quantities contains the operators

$$m^2, J^2, p_j, J_3, (n^{(3)})^2 = -1, p^\mu n_\mu^{(3)} = 0 \tag{2.22}$$

2.2 Fock space

At high energies the colliding particles can produce new particles. In every experiment we observe a definite number of particles. A state with n particles may be characterized by a way function $\phi(k_1, ..., k_n)$ depending on the variables $k_1, ..., k_n$ describing these particles. Since the number of particles is arbitrary (but finite), the state vector arising in collisions may be represented in the form of a superposition of states with a definite number of particles in the form of a Fock column vector:

$$|\Psi >= \begin{pmatrix} \phi_0 \\ \phi_1(k_1) \\ ..... \\ \phi_n(k_1, ..., k_n) \\ ..... \end{pmatrix}$$

Let us consider for simplicity that the variables $k_1, ..., k_n$ take on discrete values, then the Fock column vector with 1 in the nth place (n particles with variables $k_1, ..., k_n$) and the remaining $\phi_i = 0$ is a state vector $|k_1...k_n >$, so that

$$|\Psi >= \sum_{n=0}^{\infty} \sum_{k_1...k_n} \phi_n(k_1...k_n)|k_1...k_n > . \tag{2.23}$$

The space of states $\mathcal{H}$ is thus the direct sum of spaces $\mathcal{H}_n$ formed by the state vectors with a definite number of particles n.

Scalar product of the vectors Ψ and Ψ' has the form

$$< \Psi'|\Psi >= \sum_{n=0}^{\infty} \sum_{k_1...k_n} \phi_n'^*(k_1...k_n)\phi_n(k_1...k_n)\mu_n, \tag{2.24}$$

where μ_n is a relativistically invariant measure.

As a consequence of its asymptotic character, the function $\phi_n(k_1...k_n)$ has the following properties.

1. In the case of identical particles, $\phi_n(k_1...k_n)$ is a symmetrized or antisymmetrized product of single-particle wave functions:

$$\phi_n(k_1...k_n) = \{\phi_1(k_1)...\phi_n(k_n)\}_{\pm},$$

or

$$|k_1...k_n >= \{|k_1 > ...|k_n >\}_{\pm} \tag{2.25}$$

where the symbol $\{.\}_{\pm}$ denotes symmetrization or antisymmetrization together with normalization factors. This symmetry property of the wave function for identical particles is well known in quantum mechanics. The symmetry character of a wave function of a system of identical particles is related to spin: particles with half-integral spin (fermions) are described by antisymmetric functions, and particles with integral spin (bosons) by symmetric functions. This statement ("spin-statistics-theorem") has been proven using various approaches.

2. Under the symmetry transformation $|\alpha >\rightarrow |\alpha_g >$ (where g may refer to transformations in addition to those of the *Poincaré* group) every single-particle state transforms separately:

$$|(k_1...k_n)_g >= U(g)|k_1...k_n >= \{U(g)|k_1 > ...U(g)|k_n >\}_{\pm} \tag{2.26}$$

This property of asymptotic states depends on the possibility of representing each state in the form of a sum of products of single-particle states.

From the property of factorizability of transformations it follows that the single-particle states $|k >$ transform according to unitary representations of the quantum mechanical *Poincaré* group. A mathematical expression of the simplicity of representation is its irreducibility. For this reason, elementary particles are described by the irreducible representations of $\mathcal{P}_+^{\uparrow}$.

Consequently the classification of elementary particles arising from the theory of relativistic invariance consists of the classification of unitary representations of the quantum mechanical *Poincaré* group. The quantities characterizing irreducible representations (mass and spin) describe invariant properties of particles. The quantum numbers characterizing the basis of a irreducible representation are variables describing particle states (e.g. momentum and spin projection in the canonical basis).

2.3 Klein-Fock equation for scalar particle

V.A.Fock presented relativistic generalization of Schroedinger equation for a scalar particle in electromagnetic field on curved space. V.A.Fock introduces five dimensional space with the metric depending on electromagnetic potential A_μ

$$d\sigma^2 = g_{\mu\nu}dx_\mu dx_\nu - \frac{e^2}{m^2c^4}\left(A_\nu dx_\nu + du\right)^2 \tag{2.27}$$

where u is an additional coordinate. In classical physics null geodesic line $d\sigma = 0$ describes a trajectory of charged particle in this space. Corresponding action S will have five dimensional gradient squared equal to zero. Four dimensional action W is related to S

$$S = \frac{e}{c}u + W \tag{2.28}$$

Five dimensional equations are invariant under transformation

$$A_\nu = A'_\nu + \partial_\nu f u = u' - f \tag{2.29}$$

which later were called as **gradient transformation**.

Both classical and quantum eqations act in the same space.Therefore, the corresponding quantum wave equation for the wave function Ψ is the d'Alambert equation in five dimensional space

$$\frac{1}{\sqrt{-g}}\frac{\partial}{\partial x_\mu}\left(\sqrt{-g}g^{\mu\nu}\frac{\partial\Psi}{\partial x_\nu}\right) - 2A^\nu\frac{\partial^2\Psi}{\partial u\partial x_\nu} + \left(A_\nu A^\nu - \frac{m^2c^4}{e^2}\right)\frac{\partial^2\Psi}{\partial u^2} = 0 \tag{2.30}$$

Four dimensional wave function ψ can be found from Ψ by a phase transformation

$$\Psi = e^{i\frac{e}{\hbar c}u}\psi$$

and five dimensional eq. for Ψ produces immediately the Klein-Fock equation for ψ in electromagnetic field and on the curved space.

Gradient transformation of A_ν and u induces transformation of the wave function

$$\psi' = e^{-i\frac{ef}{\hbar c}}\psi \tag{2.31}$$

This transformation rule of wave function under gauge transformation was first written in Fock's paper on the Klein-Fock equation.

3 Gauge theories

3.1 Abelian gauge theories

The Lagrangian describing a free fermion of mass m is $\mathcal{L}_{\text{free}} = \overline{\psi}(i\ \partial\!\!\!/ - m)\psi$. It is invariant under the global phase change $\psi \to \exp(i\alpha)\psi$. (We shall always consider the fermion fields to depend on x.) Now consider independent phase changes at each point:

$$\psi \to \psi' \equiv \exp[i\alpha(x)]\psi. \tag{3.32}$$

Because of the derivative, the Lagrangian then acquires an additional phase change at each point: $\delta\mathcal{L}_{\text{free}} = \overline{\psi}i\gamma^\mu[i\partial_\mu\alpha(x)]\psi$. The free Lagrangian is not invariant under such changes of phase, known as *local gauge transformations*. Local gauge invariance can be restored if we make the replacement $\partial_\mu \to D_\mu \equiv \partial_\mu + ieA_\mu$ in the free-fermion Lagrangian, which now is

$$\mathcal{L} = \overline{\psi}(i\ D\!\!\!/ - m)\psi = \overline{\psi}(i\ \partial\!\!\!/ - m)\psi - e\overline{\psi}\ A\!\!\!/(x)\psi \quad . \tag{3.33}$$

The effect of a local phase in ψ can be compensated if we allow the *vector potential* A_μ to change by a total divergence, which does not change the electromagnetic field strength

$$F_{\mu\nu} \equiv \partial_\mu A_\nu - \partial_\nu A_\nu \quad . \tag{3.34}$$

Indeed, under the transformation $\psi \to \psi'$ and with $A \to A'$ with A' yet to be determined, we have

$$\mathcal{L}' = \overline{\psi}'(i\ \partial\!\!\!/ - m)\psi' - e\overline{\psi}'A\!\!\!/'\psi' = \overline{\psi}(i\ \partial\!\!\!/ - m)\psi - \overline{\psi}[\partial\!\!\!/\alpha(x)]\psi - e\overline{\psi}A\!\!\!/'\psi \quad . \tag{3.35}$$

This will be the same as $\mathcal{L}$ if

$$A'_\mu(x) = A_\mu(x) - \frac{1}{e}\partial_\mu\alpha(x) \quad . \tag{3.36}$$

The derivative D_μ is known as the *covariant derivative*. One can check that under a local gauge transformation, $D_\mu\psi \to e^{i\alpha(x)}D_\mu\psi$. Another way to see the consequences of local gauge invariance suggested by Yang (1974) is to define $-eA_\mu(x)$ as the local change in phase undergone by a particle of charge e as it passes along an infinitesimal space-time

increment between x^μ and $x^\mu + dx^\mu$. For a space-time trip from point A to point B, the phase change is then

$$\Phi_{AB} = \exp\left(-ie\int_A^B A_\mu(x)dx^\mu\right) \quad . \tag{3.37}$$

The phase in general will depend on the path in space-time taken from point A to point B. As a consequence, the phase Φ_{AB} is not uniquely defined. However, one can compare the result of a space-time trip along one path, leading to a phase $\Phi_{AB}^{(1)}$, with that along another, leading to a phase $\Phi_{AB}^{(2)}$. The two-slit experiment in quantum mechanics involves such a comparison; so does the Bohm-Aharonov effect in which a particle beam travelling past a solenoid on one side interferes with a beam travelling on the other side. Thus, phase differences

$$\Phi_{AB}^{(1)}\Phi_{AB}^{(2)*} = \Phi_C = \exp\left(-ie\oint A_\mu(x)dx^\mu\right) \quad , \tag{3.38}$$

associated with *closed paths* in space-time (represented by the circle around the integral sign), are the ones which correspond to physical experiments. The phase Φ_C for a closed path C is independent of the phase convention for a charged particle at any space-time point x_0, since any change in the contribution to Φ_C from the integral up to x_0 will be compensated by an equal and opposite contribution from the integral departing from x_0. The closed path integral (3.38) can be expressed as a surface integral using Stokes' theorem:

$$\oint A_\mu(x)dx^\mu = \int F_{\mu\nu}(x)d\sigma^{\mu\nu} \quad , \tag{3.39}$$

where the electromagnetic field strength $F_{\mu\nu}$ was defined previously and $d\sigma^{\mu\nu}$ is an element of surface area. It is also clear that the closed path integral is invariant under changes (3.36) of $A_\mu(x)$ by a total divergence. Thus $F_{\mu\nu}$ suffices to describe all physical experiments as long as one integrates over a suitable domain. In the Bohm-Aharonov effect, in which a charged particle passes on either side of a solenoid, the surface integral will include the solenoid (in which the magnetic field is non-zero). If one wishes to describe the energy and momentum of free electromagnetic fields, one must include a kinetic term $\mathcal{L}_K = -(1/4)F_{\mu\nu}F^{\mu\nu}$ in the Lagrangian, which now reads

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \overline{\psi}(i\,\partial\!\!\!/ - m)\psi - e\overline{\psi}\,A\!\!\!/\psi \quad . \tag{3.40}$$

If the electromagnetic current is defined as $J_\mu^{\rm em} \equiv \overline{\psi}\gamma_\mu\psi$, this Lagrangian leads to Maxwell's equations. The local phase changes (3.32) form a U(1) group of transformations. Since such transformations commute with one another, the group is said to be *Abelian*. Electrodynamics, just constructed here, is an example of an *Abelian gauge theory.*

3.2 Non-Abelian gauge theories

One can imagine that a particle travelling in space-time undergoes not only phase changes, but also changes of identity. Such transformations were first considered by Yang and Mills (1954). For example, a quark can change in color (red to blue) or flavor (u to d). In that case we replace the coefficient eA_μ of the infinitesimal displacement dx_μ by an $n \times n$ matrix $-g\mathbf{A}_\mu(x) \equiv -gA_\mu^i(x)\mathbf{T}_i$ acting in the n-dimensional space of the particle's degrees

of freedom. For colors, $n = 3$. The $\mathbf{T}_i$ form a linearly independent basis set of matrices for such transformations, while the A^i_μ are their coefficients. The phase transformation then must take account of the fact that the matrices $\mathbf{A}_\mu(x)$ in general do not commute with one another for different space-time points, so that a *path-ordering* is needed:

$$\Phi_{AB} = \mathcal{P}\left[\exp\left(ig\int_A^B \mathbf{A}_\mu(x)dx^\mu\right)\right] \quad . \tag{3.41}$$

When the basis matrices $\mathbf{T}_i$ do not commute with one another, the theory is *non-Abelian*. We demand that changes in phase or identity conserve probability, i.e., that Φ_{AB} be *unitary*: $\Phi^\dagger_{AB}\Phi_{AB} = 1$. When Φ_{AB} is a matrix, the corresponding matrices $\mathbf{A}_\mu(x)$ in (3.41) must be Hermitian. If we wish to separate out pure phase changes, in which $\mathbf{A}_\mu(x)$ is a multiple of the unit matrix, from the remaining transformations, one may consider only transformations such that $\det(\Phi_{AB}) = 1$, corresponding to *traceless* $\mathbf{A}_\mu(x)$. The $n\times n$ basis matrices $\mathbf{T}_i$ must then be Hermitian and traceless. There will be $n^2 - 1$ of them, corresponding to the number of independent SU(N) generators. (One can generalize this approach to other invariance groups.) The matrices will satisfy the commutation relations

$$[\mathbf{T}_i, \mathbf{T}_j] = ic_{ijk}\mathbf{T}_k \quad , \tag{3.42}$$

where the c_{ijk} are *structure constants* characterizing the group. For SU(2), $c_{ijk} = \epsilon_{ijk}$, while for SU(3), $c_{ijk} = f_{ijk}$, where the f_{ijk} are defined by Gell-Mann and Ne'eman (1964). A 3×3 representation in SU(3) is $\mathbf{T}_i = \lambda_i/2$, where $\lambda_i/2$ are the Gell-Mann matrices normalized such that Tr $\lambda_i\lambda_j = 2\delta_{ij}$. For this representation, then, Tr $\mathbf{T}_i\mathbf{T}_j = \delta_{ij}/2$. In order to define the field-strength tensor $\mathbf{F}_{\mu\nu} = F^i_{\mu\nu}\mathbf{T}_i$ for a non-Abelian transformation, we may consider an infinitesimal closed-path transformation analogous to Eq. (3.38) for the case in which the matrices $\mathbf{A}_\mu(x)$ do not commute with one another. The result is

$$\mathbf{F}_{\mu\nu} = \partial_\mu\mathbf{A}_\nu - \partial_\nu\mathbf{A}_\nu - ig[\mathbf{A}_\mu, \mathbf{A}_\nu] \quad , \quad F^i_{\mu\nu} = \partial_\mu A^i_\nu - \partial_\nu A^i_\mu + gc_{ijk}A^j_\mu A^k_\nu \quad . \tag{3.43}$$

An alternative way to introduce non-Abelian gauge fields is to demand that, by analogy with Eq. (3.32), a theory involving fermions ψ be invariant under local transformations

$$\psi(x) \to \psi'(x) = U(x)\psi(x) \quad , \quad U^\dagger U = 1 \quad , \tag{3.44}$$

where for simplicity we consider unitary transformations. Under this replacement, $\mathcal{L} \to \mathcal{L}'$, where

$$\begin{aligned}\mathcal{L}' &\equiv \overline{\psi}'(i\,\partial\!\!\!/ - m)\psi' = \overline{\psi}U^{-1}(i\,\partial\!\!\!/ - m)U\psi \\ &= \overline{\psi}(i\,\partial\!\!\!/ - m)\psi + i\psi U^{-1}\gamma^\mu(\partial_\mu U)\psi \quad .\end{aligned} \tag{3.45}$$

As in the Abelian case, an extra term is generated by the local transformation. It can be compensated by replacing ∂_μ by

$$\partial_\mu \to \mathbf{D}_\mu \equiv \partial_\mu - ig\mathbf{A}_\mu(x) \quad . \tag{3.46}$$

In this case $\mathcal{L} = \overline{\psi}(i\,\mathbf{D}\!\!\!/ - m)\psi$ and under the change (3.44) we find

$$\mathcal{L}' \equiv \overline{\psi}'(i\,\mathbf{D}\!\!\!/' - m)\psi' = \overline{\psi}U^{-1}(i\,\partial\!\!\!/ + g\,\mathbf{A}\!\!\!/' - m)U\psi$$

$$= \mathcal{L} + \overline{\psi}[g(U^{-1}\not{A}'U - \not{A}) + iU^{-1}(\not{\partial}U)]\psi \quad . \tag{3.47}$$

This is equal to $\mathcal{L}$ if we take

$$\mathbf{A}'_\mu = U\mathbf{A}_\mu U^{-1} - \frac{i}{g}(\partial_\mu U)U^{-1} \quad . \tag{3.48}$$

This reduces to our previous expressions if $g = -e$ and $U = e^{i\alpha(x)}$. The covariant derivative acting on ψ transforms in the same way as ψ itself under a gauge transformation: $\mathbf{D}_\mu\psi \to \mathbf{D}'_\mu\psi' = U\mathbf{D}_\mu\psi$. The field strength $\mathbf{F}_{\mu\nu}$ transforms as $\mathbf{F}_{\mu\nu} \to \mathbf{F}'_{\mu\nu} = U\mathbf{F}_{\mu\nu}U^{-1}$. It may be computed via $[\mathbf{D}_\mu, \mathbf{D}_\nu] = -ig\mathbf{F}_{\mu\nu}$; both sides transform as $U(\quad)U^{-1}$ under a local gauge transformation. In order to obtain propagating gauge fields, as in electrodynamics, one must add a kinetic term $\mathcal{L}_K = -(1/4)F^i_{\mu\nu}F^{i\mu\nu}$ to the Lagrangian. Recalling the representation $\mathbf{F}_{\mu\nu} = F^i_{\mu\nu}\mathbf{T}_i$ in terms of gauge group generators normalized such that $\mathrm{Tr}(\mathbf{T}_i\mathbf{T}_j) = \delta_{ij}/2$, we can write the full Yang-Mills Lagrangian for gauge fields interacting with matter fields as

$$\mathcal{L} = -\frac{1}{2}\mathrm{Tr}(\mathbf{F}_{\mu\nu}\mathbf{F}^{\mu\nu}) + \overline{\psi}(i\not{\mathbf{D}} - m)\psi \quad . \tag{3.49}$$

We shall use Lagrangians of this type to derive the strong, weak, and electromagnetic interactions of the "Standard Model." The interaction of a gauge field with fermions then corresponds to a term in the interaction Lagrangian $\Delta\mathcal{L} = g\overline{\psi}(x)\gamma^\mu\mathbf{A}_\mu(x)\psi(x)$. The $[\mathbf{A}_\mu, \mathbf{A}_\nu]$ term in $\mathbf{F}_{\mu\nu}$ leads to self-interactions of non-Abelian gauge fields, arising solely from the kinetic term. Thus, one has three- and four-field vertices arising from

$$\Delta\mathcal{L}^{(3)}_K = (\partial_\mu A^i_\nu)gc_{ijk}A^{\mu j}A^{\nu k} \quad , \quad \Delta\mathcal{L}^{(4)}_K = -\frac{g^2}{4}c_{ijk}c_{imn}A^{\mu j}A^{\nu k}A^m_\mu A^n_\nu \quad . \tag{3.50}$$

These self-interactions are an important aspect of non-Abelian gauge theories and are responsible in particular for the remarkable *asymptotic freedom* of QCD which leads to its becoming weaker at short distances, permitting the application of perturbation theory.

3.3 Elementary divergent quantities

In most quantum field theories, including quantum electrodynamics, divergences occurring in higher orders of perturbation theory must be removed using charge, mass, and wave function renormalization. This is conventionally done at intermediate calculational stages by introducing a cutoff momentum scale Λ or analytically continuing the number of space-time dimensions away from four. Thus, a vacuum polarization graph in QED associated with external photon momentum k and a fermion loop will involve an integral

$$\Pi_{\mu\nu}(k) \sim \int \frac{d^4p}{(2\pi)^4}\mathrm{Tr}\left(\frac{1}{\not{p} - m}\gamma_\mu\frac{1}{\not{p} + \not{k} - m}\gamma_\nu\right) \quad ; \tag{3.51}$$

a self-energy of a fermion with external momentum p will involve

$$\Sigma(p) \sim \int \frac{d^4q}{(2\pi)^4}\frac{1}{q^2}\gamma_\mu\frac{1}{\not{p} + \not{q} - m}\gamma^\mu \quad , \tag{3.52}$$

and a fermion-photon vertex function with external fermion momenta p, p' will involve

$$\Lambda_\mu(p', p) \sim \int \frac{d^4k}{(2\pi)^4} \frac{1}{k^2} \gamma_\nu \frac{1}{\not{p}' + \not{k} - m} \gamma_\mu \frac{1}{\not{p} + \not{k} - m} \gamma^\nu \quad . \tag{3.53}$$

The integral (3.51) appears to be quadratically divergent. However, the gauge invariance of the theory translates into the requirement $k^\mu \Pi_{\mu\nu} = 0$, which requires $\Pi_{\mu\nu}$ to have the form

$$\Pi_{\mu\nu}(k) = (k^2 g_{\mu\nu} - k_\mu k_\nu)\Pi(k^2) \quad . \tag{3.54}$$

The corresponding integral for $\Pi(k^2)$ then will be only logarithmically divergent. The integral in (3.52) is superficially linearly divergent but in fact its divergence is only logarithmic, as is the integral in (3.53). Unrenormalized functions describing vertices and self-energies involving n_B external boson lines and n_F external fermion lines may be defined in terms of a momentum cutoff Λ and a bare coupling constant g_0 :

$$\Gamma^U_{n_B, n_F} \equiv \Gamma^U_{n_B, n_F}(p_i, g_0, \Lambda) \quad , \tag{3.55}$$

where p_i denote external momenta. *Renormalized* functions Γ^R may be defined in terms of a scale parameter μ, a renormalized coupling constant $g = g(g_0, \Lambda/\mu)$, and renormalization constants $Z_B(\Lambda)$ and $Z_F(\Lambda)$ for the external boson and fermion wave functions:

$$\Gamma^R(p_i, g, \mu) \equiv \lim_{\Lambda\to\infty} [Z_B(\Lambda)]^{n_B} [Z_F(\Lambda)]^{n_F} \Gamma^U_{n_B, n_F}(p_i, g_0, \Lambda) \quad . \tag{3.56}$$

The scale μ is typically utilized by demanding that Γ^R be equal to some predetermined function at a Euclidean momentum $p^2 = -\mu^2$. Thus, for the one-boson, two-fermion vertex, we take

$$\Gamma^R_{1,2}(0, p, -p)|_{p^2=-\mu^2} = \lim_{\Lambda\to\infty} Z_F^2 Z_B \Gamma^U_{1,2}(0, p, -p)|_{p^2=-\mu^2} \equiv g \quad . \tag{3.57}$$

The unrenormalized function Γ^U is *independent* of μ, while Γ^R and the renormalization constants $Z_B(\Lambda)$, $Z_F(\Lambda)$ will depend on μ. For example, in QED, the photon wave function renormalization constant (known as Z_3) behaves as

$$Z_3 = 1 - \frac{\alpha_0}{3\pi} \ln \frac{\Lambda^2}{\mu^2} \quad . \tag{3.58}$$

The bare charge e_0 and renormalized charge e are related by $e = e_0 Z_3^{1/2}$. To lowest order in perturbation theory, $e < e_0$. The vacuum behaves as a normal dielectric; charge is screened. It is the exception rather than the rule that in QED one can define the renormalized charge for $q^2 = 0$; in QCD we shall see that this is not possible.

3.4 Scale changes and the beta function

We differentiate both sides of (3.56) with respect to μ and multiply by μ. Since the functions Γ^U are independent of μ, we find

$$\left(\mu \frac{\partial}{\partial \mu} + \mu \frac{\partial g}{\partial \mu} \frac{\partial}{\partial g}\right) \Gamma^R(p_i, g, \mu)$$

$$= \lim_{\Lambda\to\infty} \left(\frac{n_B}{Z_B}\mu\frac{\partial Z_B}{\partial\mu} + \frac{n_F}{Z_F}\mu\frac{\partial Z_F}{\partial\mu} \right) Z_B^{n_B} Z_F^{n_F} \Gamma^U \quad , \tag{3.59}$$

or

$$\left[\mu\frac{\partial}{\partial\mu} + \beta(g)\frac{\partial}{\partial g} + n_B\gamma_B(g) + n_F\gamma_F(g) \right] \Gamma^R(p_i, g, \mu) = 0 \quad , \tag{3.60}$$

where

$$\beta(g) \equiv \mu\frac{\partial g}{\partial\mu} \quad , \quad \gamma_B(g) \equiv -\frac{\mu}{Z_B}\frac{\partial Z_B}{\partial\mu} \quad , \quad \gamma_F(g) \equiv -\frac{\mu}{Z_F}\frac{\partial Z_F}{\partial\mu} \quad . \tag{3.61}$$

The behavior of any generalized vertex function Γ^R under a change of scale μ is then governed by the universal functions (3.61). Here we shall be particularly concerned with the function $\beta(g)$. Let us imagine $\mu \to \lambda\mu$ and introduce the variables $t \equiv \ln\lambda$, $\bar{g}(g,t) \equiv g(g_0, \Lambda/\lambda\mu)$, Then the relation for the beta-function may be written

$$\frac{d\bar{g}(g,t)}{dt} = \beta(\bar{g}) \quad , \quad \bar{g}(g,0) = g(g_0, \Lambda/\mu) = g \quad . \tag{3.62}$$

Let us compare the behavior of $\bar{g}$ with increasing t (larger momentum scales or shorter distance scales) depending on the sign of $\beta(\bar{g})$. In general we will find $\beta(0) = 0$. We take $\beta(\bar{g})$ to have zeroes at $\bar{g} = 0,\ g_1,\ g_2,\ \ldots$. Then:

1. Suppose $\beta(\bar{g}) > 0$. Then $\bar{g}$ *increases* from its $t = 0$ value $\bar{g} = g$ until a zero g_i of $\beta(\bar{g})$ is encountered. Then $\bar{g} \to g_i$ as $t \to \infty$.

2. Suppose $\beta(\bar{g}) < 0$. Then $\bar{g}$ *decreases* from its $t = 0$ value $\bar{g} = g$ until a zero g_i of $\beta(\bar{g})$ is encountered.

In either case $\bar{g}$ approaches a point at which $\beta(\bar{g}) = 0$, $\beta'(\bar{g}) < 0$ as $t \to \infty$. Such points are called *ultraviolet fixed points*. Similarly, points for which $\beta(\bar{g}) = 0$, $\beta'(\bar{g}) > 0$ are *infrared fixed points*, and $\bar{g}$ will tend to them for $t \to -\infty$ (small momenta or large distances). The point $e = 0$ is an infrared fixed point for quantum electrodynamics, since $\beta'(e) > 0$ at $e = 0$. It may happen that $\beta'(0) < 0$ for specific theories. In that case $\bar{g} = 0$ is an ultraviolet fixed point, and the theory is said to be *asymptotically free*. We shall see that this property is particular to non-Abelian gauge theories (Gross and Wilczek 1973, Politzer 1974).

3.5 Beta function calculation

In quantum electrodynamics a loop diagram involving a fermion of unit charge contributes the following expression to the relation between the bare charge e_0 and the renormalized charge e:

$$e = e_0 \left(1 - \frac{\alpha_0}{3\pi} \ln\frac{\Lambda}{\mu} \right) \quad , \tag{3.63}$$

as implied by (3.57) and (3.58), where $\alpha_0 \equiv e_0^2/4\pi$. We find

$$\beta(e) = \frac{e_0^3}{12\pi^2} \simeq \frac{e^3}{12\pi^2} \quad , \tag{3.64}$$

where differences between e_0 and e correspond to higher-order terms in e. (Here $\alpha \equiv e^2/4\pi$.) Thus $\beta(e) > 0$ for small e and the coupling constant becomes stronger at larger momentum scales (shorter distances). We shall show an extremely simple way to calculate (3.64) and the corresponding result for a charged scalar particle in a loop. From this we shall be able to first calculate the effect of a charged vector particle in a loop (a calculation first performed by Khriplovich 1969) and then generalize the result to Yang-Mills fields. When one takes account of vacuum polarization, the electromagnetic interaction in momentum space may be written

$$\frac{e^2}{q^2} \to \frac{e^2}{q^2[1+\Pi(q^2)]} \tag{3.65}$$

Here the long-distance ($q^2 \to 0$) behavior has been defined such that e is the charge measured at macroscopic distances, so $\Pi(0) = 0$. We shall reconstruct $\Pi_i(q^2)$ for a loop involving the fermion species i from its imaginary part, which is measurable through the cross section for $e^+e^- \to i\bar{i}$:

$$\text{Im } \Pi_i(s) = \frac{s}{4\pi\alpha}\sigma(e^+e^- \to i\bar{i}) \quad , \tag{3.66}$$

where s is the square of the center-of-mass energy. For fermions f of charge e_f and mass m_f,

$$\text{Im } \Pi_f(s) = \frac{\alpha e_f^2}{3}\left(1+\frac{2m_f^2}{s}\right)\left(1-\frac{4m_f^2}{s}\right)^{1/2}\theta(s-4m_f^2) \quad , \tag{3.67}$$

while for scalar particles of charge e_s and mass m_s,

$$\text{Im } \Pi_s(s) = \frac{\alpha e_s^2}{12}\left(1-\frac{4m_s^2}{s}\right)^{3/2}\theta(s-4m_s^2) \quad . \tag{3.68}$$

The corresponding cross section for $e^+e^- \to \mu^+\mu^-$, neglecting the muon mass, is $\sigma(e^+e^- \to \mu^+\mu^-) = 4\pi\alpha^2/3s$, so one can define

$$R_i \equiv \sigma(e^+e^- \to i\bar{i})/\sigma(e^+e^- \to \mu^+\mu^-) \quad , \tag{3.69}$$

in terms of which $\text{Im } \Pi_i(s) = \alpha R_i(s)/3$. For $s \to \infty$ one has $R_f(s) \to e_f^2$ for a fermion and $R_s(s) \to e_s^2/4$ for a scalar. The full vacuum polarization function $\Pi_i(s)$ cannot directly be reconstructed in terms of its imaginary part via the dispersion relation

$$\Pi_i(s) = \frac{1}{\pi}\int_{4m^2}^{\infty}\frac{ds'}{s'-s}\text{Im } \Pi_i(s') \quad , \tag{3.70}$$

since the integral is logarithmically divergent. This divergence is exactly that encountered earlier in the discussion of renormalization. For quantum electrodynamics we could deal with it by defining the charge at $q^2 = 0$ and hence taking $\Pi_i(0) = 0$. The *once-subtracted* dispersion relation for $\Pi_i(s) - \Pi_i(0)$ would then converge:

$$\Pi_i(s) = \frac{1}{\pi}\int_{4m^2}^{\infty}\frac{ds'}{s'(s'-s)}\text{Im } \Pi_i(s') \quad . \tag{3.71}$$

However, in order to be able to consider cases such as Yang-Mills fields in which the theory is not well-behaved at $q^2 = 0$, let us instead define $\Pi_i(-\mu^2) = 0$ at some spacelike scale $q^2 = -\mu^2$. The dispersion relation is then

$$\Pi_i(s) = \frac{1}{\pi} \int_{4m^2}^{\infty} ds' \left[\frac{1}{s' - s} - \frac{1}{s' + \mu^2} \right] \text{Im } \Pi_i(s') \quad . \tag{3.72}$$

For $|q^2| \gg \mu^2 \gg m^2$, we find

$$\Pi_i(q^2) \to -\frac{\alpha}{3\pi} R_i(\infty) \left[\ln \frac{-q^2}{\mu^2} + \text{const.} \right] \quad , \tag{3.73}$$

and so, from (3.65), the "charge at scale q" may be written as

$$e_q^2 \equiv \frac{e^2}{1 + \Pi_i(q^2)} \simeq e^2 \left[1 + \frac{\alpha}{3\pi} R_i(\infty) \ln \frac{-q^2}{\mu^2} \right] \quad . \tag{3.74}$$

The beta-function here is defined by $\beta(e) = \mu(\partial e/\partial \mu)|_{\text{fixed } e_q}$. Thus, expressing $\beta(e) = -\beta_0 e^3/(16\pi^2) + \mathcal{O}(e^5)$, one finds $\beta_0 = -(4/3)e_f^2$ for spin-1/2 fermions and $\beta_0 = -(1/3)e_s^2$ for scalars. These results will now be used to find the value of β_0 for a single charged massless vector field. We generalize the results for spin 0 and 1/2 to higher spins by splitting contributions to vacuum polarization into "convective" and "magnetic" ones. Furthermore, we take into account the fact that a closed fermion loop corresponds to an extra minus sign in $\Pi_f(s)$ (which is already included in our result for spin 1/2). The "magnetic" contribution of a particle with spin projection S_z must be proportional to S_z^2. For a massless spin-S particle, $S_z^2 = S^2$. We may then write

$$\beta_0 = \begin{cases} (-1)^{n_F}(aS^2 + b)(S = 0) & , \\ (-1)^{n_F}(aS^2 + 2b)(S \neq 0) & , \end{cases} \tag{3.75}$$

where $n_F = 1$ for a fermion, 0 for a boson. The factor of $2b$ for $S \neq 0$ comes from the contribution of each polarization state ($S_z = \pm S$) to the convective term. Matching the results for spins 0 and 1/2,

$$-\frac{1}{3} = b \quad , \quad -\frac{4}{3} = -\left(\frac{a}{4} + 2b \right) \quad , \tag{3.76}$$

we find $a = 8$ and hence for $S = 1$

$$\beta_0 = 8 - \frac{2}{3} = \frac{22}{3} \quad . \tag{3.77}$$

The magnetic contribution is by far the dominant one (by a factor of 12), and is of opposite sign to the convective one. The reversal of sign with respect to the scalar and spin-1/2 results is notable.

3.6 Group-theoretic techniques

The result (3.77) for a charged, massless vector field interacting with the photon is also the value of β_0 for the Yang-Mills group SO(3) $\sim$ SU(2) if we identify the photon with A^3_μ and the charged vector particles with $A^\pm_\mu \equiv (A^1_\mu \mp iA^2_\mu)/\sqrt{2}$. We now generalize it to the contribution of gauge fields in an arbitrary group G. The value of $\beta_{0\ \text{gauge fields}}$ depends on a sum over all possible self-interacting gauge fields that can contribute to the loop with external gauge field labels i and m:

$$\frac{\beta_0[G]}{\beta_0[SU(2)]} = \frac{c^G_{ijk}c^G_{mjk}}{c^{\mathrm{SU}(2)}_{ijk}c^{\mathrm{SU}(2)}_{mjk}} \quad , \tag{3.78}$$

where c^G_{ijk} is the structure constant for G, introduced in Eq. (3.42). The sums in (3.78) are proportional to δ_{im}:

$$c_{ijk}c_{mjk} = \delta_{im}C_2(A) \quad . \tag{3.79}$$

The quantity $C_2(A)$ is the *quadratic Casimir operator* for the adjoint representation of the group G. Since the structure constants for SO(3) $\sim$ SU(2) are just $c^{\mathrm{SU}(2)}_{ijk} = \epsilon_{ijk}$, one finds $C_2(A) = 2$ for SU(2), so the generalization of (3.77) is that $\beta_{0\ \text{gauge fields}} = (11/3)C_2(A)$. The contributions of arbitrary scalars and spin-1/2 fermions in representations R are proportional to $T(R)$, where

$$\mathrm{Tr}\ (T_iT_j) \equiv \delta_{ij}T(R) \tag{3.80}$$

for matrices T_i in the representation R. For a single charged scalar particle (e.g., a pion) or fermion (e.g., an electron), $T(R) = 1$. Thus $\beta_{0\ \text{spin}\ 0} = -(1/3)T_0(R)$, while $\beta_{0\ \text{spin}\ 1/2} = -(4/3)T_{1/2}(R)$, where the subscript on $T(R)$ denotes the spin. Summarizing the contributions of gauge bosons, spin 1/2 fermions, and scalars, we find

$$\beta_0 = \frac{11}{3}C_2(A) - \frac{4}{3}\sum_f T_{1/2}(R_f) - \sum_s \frac{1}{3}T_0(R_s) \quad . \tag{3.81}$$

One often needs the beta-function to higher orders, notably in QCD where the perturbative expansion coefficient is not particularly small. It is

$$\beta(\bar{g}) = -\beta_0\frac{\bar{g}^3}{16\pi^2} - \beta_1\frac{\bar{g}^5}{(16\pi^2)^2} + \dots \quad , \tag{3.82}$$

where the result for gauge bosons and spin 1/2 fermions is

$$\beta_1 = \frac{2}{3}\left\{17[C_2(A)]^2 - 10T(R)C_2(A) - 6T(R)C_2(R)\right\} \quad . \tag{3.83}$$

The first term involves loops exclusively of gauge bosons. The second involves single-gauge-boson loops with a fermion loop on one of the gauge boson lines. The third involves fermion loops with a fermion self-energy due to a gauge boson. The quantity $C_2(R)$ is defined such that

$$[T^i(R)T^i(R)]_{\alpha\beta} = C_2(R)\delta_{\alpha\beta} \quad , \tag{3.84}$$

where α and β are indices in the fermion representation. We now illustrate the calculation of $C_2(A)$, $T(R)$, and $C_2(R)$ for SU(N). More general techniques are given by Slansky

(1981). Any SU(N) group contains an SU(2) subgroup, which we may take to be generated by T_1, T_2, and T_3. The isospin projection I_3 may be identified with T_3. Then the I_3 value carried by each generator T_i (written for convenience in the fundamental N-dimensional representation) may be identified as shown below:

$\leftarrow 2 \rightarrow$		$\leftarrow N-2 \rightarrow$		
0	1	1/2	$\cdots$	1/2
-1	0	$-1/2$	$\cdots$	$-1/2$
$-1/2$	1/2	0	$\cdots$	0
$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$
$-1/2$	1/2	0	$\cdots$	0

Since $C_2(A)$ may be calculated for any convenient value of the index $i = m$ in (3.79), we chose $i = m = 3$. Then

$$C_2(A) = \sum_{\text{adjoint}} (I_3)^2 = 1 + 1 + 4(N-2)\left(\frac{1}{2}\right)^2 = N \quad . \tag{3.85}$$

As an example, the octet (adjoint) representation of SU(3) has two members with $|I_3| = 1$ (e.g., the charged pions) and four with $|I_3| = 1/2$ (e.g., the kaons). For members of the fundamental representation of SU(N), there will be one member with $I_3 = +1/2$, another with $I_3 = -1/2$, and all the rest with $I_3 = 0$. Then again choosing $i = m = 3$ in Eq. (3.80), we find $T(R)|_{\text{fundamental}} = 1/2$. The SU(N) result for β_0 in the presence of n_f spin 1/2 fermions and n_s scalars in the fundamental representation then may be written

$$\beta_0 = \frac{11}{3}N - \frac{2}{3}n_f - \frac{1}{6}n_s \quad . \tag{3.86}$$

The quantity $C_2(A)$ in (3.85) is most easily calculated by averaging over all indices $\alpha = \beta$. If all generators T^i are normalized in the same way, one may calculate the result for an *individual* generator (say, T_3) and then multiply by the number of generators [$N^2 - 1$ for SU(N)]. For the fundamental representation one then finds

$$C_2(R) = \frac{1}{N}(N^2-1)\left[\left(\frac{1}{2}\right)^2 + \left(-\frac{1}{2}\right)^2\right] = \frac{N^2-1}{2N} \quad . \tag{3.87}$$

3.7 The running coupling constant

One may integrate Eq. (3.82) to obtain the coupling constant as a function of momentum scale M and a scale-setting parameter Λ. In terms of $\bar{\alpha} \equiv \bar{g}^2/4\pi$, one has

$$\frac{d\bar{\alpha}}{dt'} = -\beta_0 \frac{\bar{\alpha}^2}{4\pi} - \beta_1 \frac{\bar{\alpha}^3}{(4\pi)^2} \quad , \quad t' \equiv 2t = \ln\left(\frac{M^2}{\Lambda^2}\right) \quad . \tag{3.88}$$

For large t' the result can be written as

$$\bar{\alpha}(M^2) = \frac{4\pi}{\beta_0 t'}\left[1 - \frac{\beta_1}{\beta_0^2}\frac{\ln t'}{t'}\right] + \mathcal{O}(t'^{-2}) \quad . \tag{3.89}$$

Suppose a process involves p powers of $\bar{\alpha}$ to leading order and a correction of order $\bar{\alpha}^{p+1}$:

$$\Gamma = A\bar{\alpha}^p[1 + B\bar{\alpha} + \mathcal{O}(\bar{\alpha}^2)] \quad . \tag{3.90}$$

If Λ is rescaled to $\lambda\Lambda$, then $t' \to t' - 2\ln\lambda = t'(1 - 2\ln\lambda/t')$, so

$$\bar{\alpha}^p \to \bar{\alpha}^p \left(1 + \frac{p\beta_0}{2\pi}\bar{\alpha}\ln\lambda\right) \quad . \tag{3.91}$$

The coefficient B thus depends on the scale parameter used to define $\bar{\alpha}$. Many prescriptions have been adopted for defining Λ. In one ('t Hooft 1973), the "minimal subtraction" or MS scheme, ultraviolet logarithmic divergences are parametrized by continuing the space-time dimension $d = 4$ to $d = 4 - \epsilon$ and subtracting pole terms $\int d^{4-\epsilon}/p^4 \sim 1/\epsilon$. In another (Bardeen *et al.* 1978) (the "modified minimal subtraction or $\overline{\text{MS}}$ scheme) a term

$$\frac{1}{\hat{\epsilon}} = \frac{1}{\epsilon} + \frac{\ln 4\pi - \gamma_E}{2} \tag{3.92}$$

containing additional finite pieces is subtracted. Here $\gamma_E = 0.5772$ is Euler's constant, and one can show that $\Lambda_{\overline{\text{MS}}} = \Lambda_{\text{MS}} \exp[(\ln 4\pi - \gamma_E)/2]$. Many $\mathcal{O}(\bar{\alpha})$ corrections are quoted in the $\overline{\text{MS}}$ scheme. Specification of Λ in any scheme is equivalent to specification of $\bar{\alpha}(M^2)$.

4 QCD with two flavours

In this section, we discuss the flavour symmetries of QCD.

4.1 Symmetry of the lagrangian

The lagrangian of QCD is

$$\mathcal{L} = -\frac{1}{2g^2}\langle G_{\mu\nu}G^{\mu\nu}\rangle_c + \mathcal{L}_{ud} \ , \tag{4.1}$$

where

$$\begin{aligned}\mathcal{L}_{ud} &= \bar{u}\not{D}u + \bar{d}\not{D}d - m_u\bar{u}u - m_d\bar{d}d \\ &= (\bar{u}\ \bar{d})\begin{pmatrix} \not{D} - m_u & 0 \\ 0 & \not{D} - m_d \end{pmatrix}\begin{pmatrix} u \\ d \end{pmatrix} \ ; \ \not{D} = i\gamma^u(\partial_\mu - iG_\mu) \ .\end{aligned}$$

$G_{\mu\nu}$ denotes the field strength associated with the gluon field G_μ, and $\langle A\rangle_c$ stands for the color trace of the matrix A. It is useful to introduce left- and right-handed spinors,

$$u_L = \frac{1}{2}(1-\gamma_5)u \ , \ u_R = \frac{1}{2}(1+\gamma_5)u \ ,$$

$$\mathcal{L}_{ud} = (\bar{u}_L\ \bar{d}_L)\begin{pmatrix} \not{D} & 0 \\ 0 & \not{D} \end{pmatrix}\begin{pmatrix} u_L \\ d_L \end{pmatrix} - (\bar{u}_L\ \bar{d}_L)\begin{pmatrix} m_u & 0 \\ 0 & m_d \end{pmatrix}\begin{pmatrix} u_R \\ d_R \end{pmatrix} + L \leftrightarrow R \ .$$

QCD makes sense for any value of the quark masses. For $m_u = m_d = 0$, the lagrangian (4.1) is invariant under $U(2)$ rotations of the left- and right-handed fields,

$$\begin{pmatrix} u_I \\ d_I \end{pmatrix} \Rightarrow V_I\begin{pmatrix} u_I \\ d_I \end{pmatrix} \ ; \ V_I \in U(2) \ , \ I = L, R \ . \tag{4.2}$$

In other words, gluon interactions do not change the helicity of the quarks. On the other hand, the terms proportional to the quark masses are not invariant under the transformations (4.2).

According to the theorem of E. Noether, there is one conserved current for each continuous parameter in the symmetry group. As the group $U(2)$ has four real parameters, one expects eight conserved currents. However, due to quantum effects, one of these currents is not conserved, as a result of which there are only seven conserved currents in the limit of vanishing quark masses,

$$\begin{aligned} L^a_\mu &= \bar{q}_L\gamma_\mu\frac{\tau^a}{2}q_L \ , \ R^a_\mu = \bar{q}_R\gamma_\mu\frac{\tau^a}{2}q_R \ ; \ a = 1,2,3 \\ V_\mu &= \bar{q}\gamma_\mu q \ ; \ q = \begin{pmatrix} u \\ d \end{pmatrix} \ . \end{aligned} \tag{4.3}$$

The world at

$$m_u = m_d = 0$$

is called *the chiral limit of QCD*, and the above statements are summarized as: *In the chiral limit, $\mathcal{L}_{\rm QCD}$ is symmetric under global $SU(2)_L \times SU(2)_R \times U(1)_V$ transformations. The corresponding 7 Noether currents (4.3) are conserved.*

4.2 Symmetry of the ground state

It is useful to introduce in addition the vector and axial currents

$$\begin{aligned} V^{\mu a} &= \bar{q}\gamma^\mu \frac{\tau^a}{2} q = L^{\mu a} + R^{\mu a} \ , \\ A^{\mu a} &= \bar{q}\gamma^\mu \gamma_5 \frac{\tau^a}{2} q = R^{\mu a} - L^{\mu a} \ \ ; \ \ a = 1,2,3 \ . \end{aligned}$$

The corresponding 6 axial and vector charges $Q^a_{A,V}$ are conserved and commute with the hamiltonian $H_0 = H_{\rm QCD}|_{m_u=m_d=0}$,

$$[H_0, Q^a_V] = [H_0, Q^a_A] = 0 \ \ ; \ \ a = 1,2,3 \ . \tag{4.4}$$

Consider now eigenstates of H_0,

$$H_0|\psi\rangle = E|\psi\rangle \ .$$

Then the states $Q^a_A|\psi\rangle$ and $Q^a_V|\psi\rangle$ have the same energy E, but carry opposite parity. On the other hand, there is no trace of such a symmetry in nature. The resolution of the paradox has been provided by Nambu and Lasinio back in 1960 : whereas the vacuum is annihilated by the vector charges, it is not invariant under the action of the axial charges,

$$Q^a_V|0\rangle = 0 \ , \ Q^a_A|0\rangle \neq 0 \ . \tag{4.5}$$

There are two important consequences of this assumption:

i) The spectrum of H_0 contains three massless, pseudoscalar particles (Goldstone bosons (GB)). We will see more of this in the following section.

ii) The axial charges Q^a_A, acting on any state in the Hilbert space, generate Goldstone bosons,

$$Q^a_A|\psi\rangle = |\psi, G_1, \ldots, G_N, \ldots\rangle \ .$$

These are not one-particle states, and are therefore not listed in PDG, and there is therefore no contradiction anymore.

A theory with (4.4), (4.5) is called *spontaneously broken*: the symmetry of the hamiltonian is not the same as the symmetry of the ground state. Where are the three massless, pseudoscalar states? The three pions $\pi^\pm, \pi^0$ are the lightest hadrons. They are not massless, because the quark masses are not zero in the real world :

$$\begin{aligned} m_u &\simeq 5\,\text{MeV} \ , \\ m_d &\simeq 9\,\text{MeV} \ . \end{aligned} \tag{4.6}$$

In the following, we assume that the flavour symmetry of QCD is spontaneously broken to the diagonal subgroup,

$$\boxed{SU(2)_L \times SU(2)_R \to SU(2)_V}$$

and work out the consequences for the *interactions* between the Goldstone bosons.

4.3 A remark on isospin symmetry

Even if the quarks are not massless, the QCD lagrangian has a residual symmetry at $m_u = m_d$: it is invariant under the transformations (4.2) with $V_R = V_L \in SU(2)$. This symmetry is called *isospin symmetry*. We know from experiment that isospin symmetry violations in the strong interactions are small. There are two sources of isospin violations: those due to electromagnetic interactions, and those due to the difference in the up and down quark masses. On the other hand, according to (4.6), one has

$$m_d/m_u \simeq 1.8 \ . \tag{4.7}$$

How can then isospin be a good symmetry if the quark masses differ so much? Consider the neutral and the charged pions: is it so that their masses differ by

$$(M^2_{\pi^+} - M^2_{\pi^0})/M^2_{\pi^0} \simeq \frac{m_d - m_u}{m_d + m_u} \simeq 0.3 \ ?$$

The answer is no: one has

$$\begin{aligned} M^2_{\pi^+} &= (m_u + m_d)B + \cdots \ , \\ M^2_{\pi^0} &= (m_u + m_d)B + \cdots \ , \end{aligned}$$

where the ellipses denote higher order terms in the quark mass expansion. The neutral and the charged pion have the same leading term, the quark mass difference shows up only in the quadratic piece,

$$M^2_{\pi^+} - M^2_{\pi^0} = O[(m_u - m_d)^2] \ .$$

The perturbation due to the quark masses can be written as

$$\begin{aligned} m_u \bar{u}u + m_d \bar{d}d &= \frac{1}{2}(m_u + m_d)(\bar{u}u + \bar{d}d) \\ &+ \frac{1}{2}(m_u - m_d)(\bar{u}u - \bar{d}d) \ . \end{aligned}$$

Isospin is a good symmetry, not because $(m_d - m_u)/(m_d + m_u)$ is small, but because the matrix elements of the operator $\frac{1}{2}(m_d - m_u)(\bar{u}u - \bar{d}d)$ are small with respect to the hadron masses. The bulk part in the pion mass difference is generated by electromagnetic interactions.

5 Goldstone bosons

In this section, we discuss the Goldstone theorem, illustrate it with several examples and consider the interaction of Goldstone bosons at low energy.

5.1 The Goldstone theorem

We consider a quantum field theory which has the following properties:

i) There is a conserved current (i.e., an object that transforms as a four-vector under proper Lorentz transformations),

$$A_\mu(x) \;\; ; \;\; \partial^\mu A_\mu = 0.$$

ii) There is an operator $\Phi(x)$ such that

$$\langle 0|[Q,\Phi]|0\rangle \neq 0 \;\; ; \;\; Q = \int d^3x A_0(x^0,\vec{x}). \tag{5.1}$$

Then the Goldstone theorem applies:

1. There exists a massless particle in the theory,

$$|\pi(\mathbf{p})\rangle \; , \; p^2 = 0 \; .$$

2. The current A_μ couples to the massless state,

$$\langle 0|A_\mu(0)|\pi(\mathbf{p})\rangle = ip_\mu F \neq 0 \; .$$

From the condition (5.1), it is seen that the charge Q does not annihilate the vacuum.

5.2 The free scalar field

We begin with a very simple example, the free, massless scalar field. The lagrangian is given by

$$\mathcal{L}_0 \;=\; \frac{1}{2}\partial_\mu\phi\partial^\mu\phi \; .$$

For the current A_μ, we take

$$A_\mu = \partial_\mu\phi \; .$$

This current is conserved, because ϕ is a free field. Consider now $\Phi = \phi$. From the canonical commutation relations, it follows that the condition (5.1) is satisfied. Therefore, the Goldstone theorem applies. Indeed, we can easily check directly:

- ϕ generates massless states ,

and

- $\langle 0|A_\mu(0)|\pi\rangle = -ip_\mu \neq 0$.

5.3 The linear sigma model

We consider the linear sigma model ($L\sigma M$), because it allows one to illustrate many features of effective field theories. At the same time, it serves as a model with spontaneous symmetry breaking. The lagrangian is

$$\mathcal{L}_\sigma = \frac{1}{2}\partial_\mu \vec{\phi} \cdot \partial^\mu \vec{\phi} - \frac{g}{4}(\vec{\phi}^2 - v^2)^2 \ , \tag{5.2}$$

where $\vec{\phi} = (\phi^0, \phi^1, \phi^2, \phi^3)$ denotes four real fields, and $\vec{\phi}^2 = \phi^k \phi^k$ [repeated indices are summed over in the absence of a summation symbol]. In the following, we assume that

$$v^2 > 0 \ ,$$

and discuss

- the symmetry properties of $\mathcal{L}_\sigma$
- spontaneous symmetry breakdown
- Goldstone bosons
- quantization
- Goldstone boson scattering

Symmetry properties Here, we consider the classical theory and observe that $\mathcal{L}_\sigma$ is invariant under four-dimensional rotations of the vector $\vec{\phi}$,

$$\phi^i \to R^{ik}\phi^k \ , \ R \in O(4) \ .$$

The matrices R can be parametrized in terms of six real parameters. Let us consider infinitesimal rotations

$$R = 1 + \varepsilon + O(\varepsilon^2) \ .$$

Because R is an orthogonal matrix, ε is antisymmetric, $\varepsilon + \varepsilon^T = 0$. Every real and antisymmetric four by four matrix can be expanded in terms of six generators,

$$\varepsilon = \sum_{i=1}^{3} \Big(c_i \varepsilon_V^i + d_i \varepsilon_A^i \Big) \ ,$$

where c_i, d_i are 6 real parameters. The generators satisfy the commutation relations $[\varepsilon_V^a, \varepsilon_V^b] = \varepsilon^{abc}\varepsilon_V^c$, $[\varepsilon_V^a, \varepsilon_A^b] = \varepsilon^{abc}\varepsilon_A^c$, $[\varepsilon_A^a, \varepsilon_A^b] = \varepsilon^{abc}\varepsilon_V^c$, with $\varepsilon^{123} = 1$, cycl. The linear combinations

$$Q_L^a = \frac{1}{2}(\varepsilon_V^a - \varepsilon_A^a) \ , \ Q_R^a = \frac{1}{2}(\varepsilon_V^a + \varepsilon_A^a) \ ,$$

generate two commuting $SU(2)$ Lie-algebras (up to a factor i),

$$\begin{aligned} &[Q_I^a, Q_I^b] = \varepsilon^{abc} Q_I^c \qquad ; \ I = L, R \ , \\ &[Q_L^a, Q_R^b] = 0 \ . \end{aligned}$$

In other words, the lagrangian $\mathcal{L}_\sigma$ has a $SU(2)_L \times SU(2)_R$ symmetry. As a result of this, there are six conserved Noether currents, which we take to be

$$\begin{aligned} V_\mu^a &= \varepsilon^{abc}\phi^b\partial_\mu\phi^c \ , \\ A_\mu^a &= -\phi^0 \overleftrightarrow{\partial}_\mu \phi^a \ \ ; \ \ a = 1,2,3 \ . \end{aligned}$$

Spontaneous symmetry breaking

The potential $V = g(\vec{\phi}^{\,2} - v^2)^2/4$ is extremal at $\phi = 0$ and at $\vec{\phi}^{\,2} = v^2$. The latter configuration corresponds to a global minimum. The vector

$$\vec{\phi}_G = (v, \vec{0}) \ ,$$

which realizes this global minimum, is only invariant under the subgroup $H = O(3)$ (with generators ε_V^a), for the case where the symmetry group is $O(2)$. The number of generators that do not leave invariant $\vec{\phi}_G$ is $n_G - n_H = 3$, where

$$\begin{aligned} n_G : &\quad \text{number of parameters in O(4)} \ , \\ n_H : &\quad \text{number of parameters in O(3)} \ . \end{aligned}$$

Therefore, one expects three Goldstone bosons in the spectrum of the theory.

Goldstone bosons

In order to identify the Goldstone bosons, we consider fluctuations around the configuration $\vec{\phi}_G$, and write

$$\vec{\phi} = (v + \varphi_0, \vec{\pi}) \ \ ; \ \ \vec{\pi} = (\pi^1, \pi^2, \pi^3) \ ,$$

where we have introduced the pion fields $\vec{\pi}$. In terms of the new fields, the lagrangian becomes

$$\begin{aligned} \mathcal{L}_\sigma &= \frac{1}{2}[\partial_\mu\varphi_0\partial^\mu\varphi_0 - 2gv^2\varphi_0^2] + \frac{1}{2}\partial_\mu\vec{\pi}\cdot\partial^\mu\vec{\pi} \\ &\quad -gv\varphi_0\,(\varphi_0^2 + \vec{\pi}^{\,2}) - \frac{g}{4}(\varphi_0^2 + \vec{\pi}^{\,2})^2 \ . \end{aligned} \tag{5.3}$$

The kinetic term shows that there is indeed one massive field φ_0, with mass $m = \sqrt{2g}v$, together with three massless fields $\vec{\pi}$.

5.4 QCD with two flavors

We now go back to the discussion of the Goldstone theorem in the framework of QCD. In section 4, we noted that the three axial currents A_μ^a are conserved in the case of vanishing quark masses. For the field Φ in (5.1), we choose $\Phi = \bar{q}\gamma_5\tau^a q$ (three fields, one for each a). Applying canonical commutation relations, one finds that

$$[Q_A^a, \Phi^b] = -\delta^{ab}(\bar{u}u + \bar{d}d) \ \ ; \ \ a, b = 1,2,3 \ .$$

Provided that the vacuum expectation value of the quark bilinear is different from zero, the Goldstone theorem applies:

$$\langle 0|\bar{u}u|0\rangle \neq 0 \Rightarrow 3 \text{ Goldstone bosons} .$$

(The vacuum expectation value of $\bar{u}u$ is equal to the one of $\bar{d}d$ by isospin symmetry.) Lattice calculations support the conjecture that $\langle 0|\bar{u}u|0\rangle$ is different from zero.

How does one evaluate Goldstone boson (GB) scattering in QCD? This is a very complicated affair: the QCD lagrangian contains quark and gluon fields, not pion fields. On the other hand, if QCD is spontaneously broken in the manner just discussed, the Goldstone theorem guarantees that the axial current can be used as an interpolating field for the pion. Let us consider therefore the matrix element

$$G_\mu(p_4, p_3; p_1) = \langle \pi(p_3)\pi(p_4)\text{out}|A_\mu(0)|\pi(p_1)\rangle \ , \tag{5.4}$$

where I have suppressed all isospin indices. This matrix element has two parts: one, where the axial current generates a pion pole, and a second one, which is free from one-particle singularities .

According to the LSZ reduction formula, the quantity G_μ has the structure

$$G_\mu = \frac{Fq_\mu}{q^2} T(p_3, p_4; p_1) + R_\mu \ , \tag{5.5}$$

where T denotes the elastic $\pi\pi$ scattering matrix element, and where F is the pion decay constant,

$$\langle 0|A_\mu(0)|\pi(p)\rangle = ip_\mu F \ .$$

The remainder R_μ is non singular when $q^\mu = (p_3 + p_4 - p_1)^\mu$ is sent to zero. We contract both sides in (5.5) with q^μ. Because the axial current is conserved, the left-hand side vanishes, and therefore

$$FT(p_3, p_4; p_1) + q^\mu R_\mu = 0 \ .$$

One concludes that the scattering matrix element vanishes at $q^\mu = 0$. We find again that the GB do not interact at vanishing momenta. One can even go further: as already mentioned in the introduction, Weinberg determined in 1966 - using current algebra – the leading term of the $\pi\pi$ scattering amplitude in a systematic expansion of the momenta and of the pion mass.

6 Effective field theories

As we have just seen, it is possible to get a great deal of information about the interactions of GB in QCD without actually solving the theory. *Effective field theories* (EFT) provide the proper framework to perform detailed calculations in a systematic manner. EFT are

valid in a restricted energy region, describing there an underlying theory that is valid on a wider energy scale.

Effective field theories are in particular useful, when a full calculation is not yet possible, as is the case with the Standard Model at low energies. One sets up an EFT with the same symmetry properties – in case of the Standard Model, this effective theory is called *Chiral perturbation theory.* A second possibility for the use of EFT occurs in case that the calculations can be done in the underlying theory, but are very complicated. Further applications concern the case where the underlying theory is not known – one builds the effective theory in terms of the light fields and attempts to determine from experiment the unknown coupling constants that occur in there.

6.1 Linear sigma model at low energy

We have seen in the last section that the $O(4)$ version of the $L\sigma M$ in its broken phase develops three Goldstone bosons, which interact weakly at low energy (we have proven this at tree level only).

In the following, we can construct an effective theory that contains only pions and their interactions - the sigma particle is removed from the theory. This EFT is constructed in such a manner that the Green functions with pion fields, evaluated in the $L\sigma M$ at low energies, are recovered by the EFT.

As a first step, one constructs an EFT which reproduces all *tree graphs* at low energies, to any order in the low-energy expansion: as is seen from the explicit expression, the scattering matrix element contains arbitrarily high powers in the momenta even at tree level. The result can be written in various equivalent forms. Here, we use

$$\mathcal{L}^{\sigma}_{\text{eff}} \;=\; \mathcal{L}^{\sigma}_{2} + \mathcal{L}^{\sigma}_{4} + \cdots \; .$$

The lagrangians $\mathcal{L}^{\sigma}_{n}$ contain n derivatives of the pion fields. These derivatives become momenta in Fourier space: the effective lagrangian provides a momentum expansion of the amplitudes. Explicitly, one has for the leading term

$$\mathcal{L}^{\sigma}_{2} \;=\; \frac{v^2}{4}\langle \partial_\mu U \partial^\mu U^\dagger \rangle \; , \tag{6.1}$$

where U is a 2×2 unitary matrix,

$$U \;=\; \sigma \cdot \mathbf{1}_{2\times 2} + \frac{i}{v}\tau^k \pi^k \quad ; \quad \sigma^2 + \frac{\vec{\pi}^2}{v^2} = 1 \; , \vec{\pi} = (\pi^1, \pi^2, \pi^3) \; .$$

The symbol $\langle A \rangle$ denotes the trace of the matrix A, and τ^k are the Pauli matrices. In order to illustrate the structure of this lagrangian, we expand it in terms of the pion fields:

$$\mathcal{L}^{\sigma}_{2} = \frac{1}{2}\partial_\mu \vec{\pi} \cdot \partial^\mu \vec{\pi} + \frac{1}{8v^2}\partial_\mu \vec{\pi}^2 \partial^\mu \vec{\pi}^2 + O(\vec{\pi}^6) \; .$$

The lagrangians $\mathcal{L}^{\sigma}_{4,6,8\ldots}$ have a similar structure.

Comments

- Only pions occur in the effective theory, the heavy particle has disappeared.
- It is easy to calculate the $\pi\pi$ scattering amplitude at tree level with $\mathcal{L}_2^\sigma$ - only one diagram remains to be calculated.
- The mass of the sigma particle is given by $m = \sqrt{2g}v$. As a result of this, the interactions disappear formally in the large-mass limit.

5. Where has the $O(4)$ symmetry of the $L\sigma M$ gone? $\mathcal{L}_2^\sigma$ is invariant under

$$U \to V_R U V_L^\dagger;\ V_{R,L} \in SU(2)\ .$$

Therefore, the effective theory has an $SU(2)_R \times SU(2)_L$ symmetry, as the original theory.

6. How is the spontaneously broken symmetry realized?

$U_G = \mathbf{1}_{2\times 2}$ is the ground state. It is only invariant under

$$U_G \to V U_G V^\dagger\ .$$

Therefore, the theory is spontaneously broken,

$$SU(2)_R \times SU(2)_L \Rightarrow SU(2)_V\ ,$$

and generates three Goldstone bosons.

The lagrangian $\mathcal{L}_{\text{eff}}^\sigma$ reproduces the tree graphs of the $L\sigma M$ - how about loops? Indeed, one may calculate the scattering amplitudes in the framework of the $L\sigma M$ to any order in the loop expansion, perform a low-energy expansion of the result, and finally construct an effective theory that reproduces the result of this calculation, order by order in the low-energy expansion.

6.2 QCD at low energy

The effective theory of QCD is formulated in terms of asymptotic pion fields. As is the case for the $L\sigma M$, the effective theory consists of an infinite number of terms, with more and more derivatives:

$$\mathcal{L}_{\text{eff}} = \mathcal{L}_2 + \mathcal{L}_4 + \mathcal{L}_6 + \cdots\ . \tag{6.2}$$

All that goes into the construction of $\mathcal{L}_{\text{eff}}$ are symmetry properties of QCD. The result for the leading term is

$$\mathcal{L}_2 \quad = \quad \frac{F^2}{4}\langle \partial_\mu U \partial^\mu U^\dagger + 2B\mathcal{M}(U + U^\dagger)\rangle\ , \tag{6.3}$$

where the field U is the same as above, and where

$$\mathcal{M} = \begin{pmatrix} m_u & 0 \\ 0 & m_d \end{pmatrix}$$

contains the quark masses m_u, m_d. A glance at (6.1) shows that, at $m_u = m_d = 0$, the leading order lagrangians in the $L\sigma M$ and in QCD agree, provided that one sets $v = F$. The leading term (6.3) contains the two constants F and B which are related to the pion decay constant and to the quark condensate, respectively,

$$\begin{aligned} \langle 0|A^a_\mu(0)|\pi^b(p)\rangle &= ip_\mu F\delta^{ab} \ , \\ \langle 0|\bar{u}u|0\rangle_{|m_u=m_d=0} &= -F^2 B \ . \end{aligned}$$

We have made a very big step: we have replaced the QCD lagrangian $\mathcal{L}_{\rm QCD}$ by the effective lagrangian $\mathcal{L}_{\rm eff}$. The crucial point to observe is the fact that the transition

$$\mathcal{L}_{\rm QCD} \Rightarrow \mathcal{L}_{\rm eff}$$

is a non perturbative phenomenon. It is very different from the construction of the effective theory for the linear sigma model, where the loop expansion in the original theory does make sense, and where the low-energy representation can be worked out order by order in the loop expansion.

7 W bosons

7.1 Fermi theory of weak interactions

The effective four-fermion Hamiltonian for the $V - A$ theory of the weak interactions is

$$\mathcal{H}_W = \frac{G_F}{\sqrt{2}}[\overline{\psi}_1\gamma_\mu(1-\gamma_5)\psi_2][\overline{\psi}_3\gamma^\mu(1-\gamma_5)\psi_4] = 4\frac{G_F}{\sqrt{2}}(\overline{\psi}_{1L}\gamma_\mu\psi_{2L})(\overline{\psi}_{3L}\gamma^\mu\psi_{4L}) \quad , \tag{7.4}$$

where G_F is a Fermi constant. We wish to write instead a Lagrangian for interaction of particles with charged W bosons which reproduces (7.4) when taken to second order at low momentum transfer. We shall anticipate a result of Section 3 by introducing the W through an SU(2) symmetry, in the form of a gauge coupling. In the kinetic term in the Lagrangian for fermions,

$$\mathcal{L}_{Kf} = \overline{\psi}(i\ \partial\!\!\!/ - m)\psi = \overline{\psi}_L(i\ \partial\!\!\!/)\psi_L + \overline{\psi}_R(i\ \partial\!\!\!/)\psi_R - m\overline{\psi}\psi \quad , \tag{7.5}$$

the $\partial\!\!\!/$ term does not mix ψ_L and ψ_R, so in the absence of the $\overline{\psi}\psi$ term one would have the freedom to introduce different covariant derivatives $\mathbf{D}\!\!\!/$ acting on left-handed and right-handed fermions. We shall find that the same mechanism which allows us to give masses to the W and Z while keeping the photon massless will permit the generation of fermion masses even though ψ_L and ψ_R will transform differently under our gauge group.

We now let the left-handed spinors be doublets of an SU(2), such as

$$\begin{bmatrix} \nu_e \\ e^- \end{bmatrix}_L \ , \quad \begin{bmatrix} \nu_\mu \\ \mu^- \end{bmatrix}_L \ , \quad \begin{bmatrix} \nu_\tau \\ \tau^- \end{bmatrix}_L \ . \tag{7.6}$$

The W is introduced via the replacement

$$\partial_\mu \to \mathbf{D}_\mu \equiv \partial_\mu - ig\mathbf{T}^i W^i_\mu \ , \quad \mathbf{T}^i \equiv \tau^i/2 \quad , \tag{7.7}$$

where τ^i are the Pauli matrices and W^i_μ are a triplet of massive vector mesons. Here we will be concerned only with the $W^\pm$, defined by $W^\pm_\mu \equiv (W^1_\mu \mp iW^2_\mu)/\sqrt{2}$. The field W^+_μ annihilates a W^+ and creates a W^-, while W^-_μ annihilates a W^- and creates a W^+. Then $W^1_\mu = (W^+_\mu + W^-_\mu)/\sqrt{2}$ and $W^2_\mu = i(W^+_\mu - W^-_\mu)/\sqrt{2}$, so

$$\mathbf{T}^i W^i_\mu = \frac{1}{2}\begin{bmatrix} W^3_\mu & \sqrt{2}W^+_\mu \\ \sqrt{2}W^-_\mu & -W^3_\mu \end{bmatrix} \quad . \tag{7.8}$$

The interaction arising from (7.5) for a lepton $l = e, \mu, \tau$ is then

$$\mathcal{L}^{(W^\pm)}_{\text{int},\, l} = \frac{g}{\sqrt{2}}\left[\bar{\nu}_{lL}\gamma^\mu W^+_\mu l_L + \bar{l}_L\gamma^\mu W^-_\mu \nu_{lL}\right] \quad , \tag{7.9}$$

where we temporarily neglect the W^3_μ terms. Taking this interaction to second order and replacing the W propagator $(M^2_W - q^2)^{-1}$ by its $q^2 = 0$ value, we find an effective interaction of the form (7.4), with

$$\frac{G_F}{\sqrt{2}} = \frac{g^2}{8M^2_W} \quad . \tag{7.10}$$

7.2 Charged-current quark interactions

The experiment produces estimate for the relative strengths of the charge-current weak transitions between the quarks are summarized in Table 2.

Table 2: Relative strengths of charge-changing weak transitions.

Relative amplitude	Transition	Source of information (example)
~ 1	$u \leftrightarrow d$	Nuclear β-decay
~ 1	$c \leftrightarrow s$	Charmed particle decays
~ 0.22	$u \leftrightarrow s$	Strange particle decays
~ 0.22	$c \leftrightarrow d$	Neutrino prod. of charm
~ 0.04	$c \leftrightarrow b$	b decays
~ 0.003–0.004	$u \leftrightarrow b$	Charmless b decays
~ 1	$t \leftrightarrow b$	Dominance of $t \to Wb$
~ 0.04	$t \leftrightarrow s$	Only indirect evidence
~ 0.01	$t \leftrightarrow d$	Only indirect evidence

The left-handed quark doublets may be written

$$\begin{bmatrix} u \\ d' \end{bmatrix}_L \quad , \quad \begin{bmatrix} c \\ s' \end{bmatrix}_L \quad , \quad \begin{bmatrix} t \\ b' \end{bmatrix}_L \quad , \tag{7.11}$$

where d', s', and b' are related to the mass eigenstates d, s, b by a unitary transformation

$$\begin{bmatrix} d' \\ s' \\ b' \end{bmatrix} = V \begin{bmatrix} d \\ s \\ b \end{bmatrix} \quad , \quad V^\dagger V = 1 \quad . \tag{7.12}$$

The rationale for the unitary matrix V of Kobayashi and Maskawa (1973) will be reviewed in the next Section when we discuss the origin of fermion masses in the electroweak theory. The interaction Lagrangian for W's with quarks then is

$$\mathcal{L}^{(W^\pm)}_{\text{int, quarks}} = \frac{g}{\sqrt{2}}(\bar{U}_L \gamma^\mu W^+_\mu V D_L) + \text{h.c.} \quad , \quad U \equiv \begin{bmatrix} u \\ c \\ t \end{bmatrix} \quad , \quad D \equiv \begin{bmatrix} d \\ s \\ b \end{bmatrix} \quad . \tag{7.13}$$

A convenient parametrization of V (conventionally known as the Cabibbo-Kobayashi-Maskawa matrix, or CKM matrix) is

$$V \equiv \begin{bmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{bmatrix} = \begin{bmatrix} 1 - \frac{\lambda^2}{2} & \lambda & A\lambda^3(\rho - i\eta) \\ -\lambda & 1 - \frac{\lambda^2}{2} & A\lambda^2 \\ A\lambda^3(1 - \rho - i\eta) & -A\lambda^2 & 1 \end{bmatrix} \quad . \tag{7.14}$$

Experimentally $\lambda \simeq 0.22$ and $A \simeq 0.85$.

7.3 Decays of the τ lepton

The τ lepton provides a good example of "standard model" charged-current physics. The τ^- decays to a ν_τ and a virtual W^- which can then materialize into any kinematically allowed final state: $e^-\bar{\nu}_e$, $\mu^-\bar{\nu}_\mu$, or three colors of $\bar{u}d'$, where, in accord with (7.14), $d' \simeq 0.975d + 0.22s$. Neglecting strong interaction corrections and final fermion masses, the rate for τ decay is expected to be

$$\Gamma(\tau^- \to \text{all}) = 5G_F^2 \frac{m_\tau^5}{192\pi^3} \simeq 2 \times 10^{-3} \text{ eV} \quad , \tag{7.15}$$

corresponding to a lifetime of $\tau_\tau \simeq 3 \times 10^{-13}$ s as observed. The factor of $5 = 1 + 1 + 3$ corresponds to equal rates into $e^-\bar{\nu}_e$, $\mu^-\bar{\nu}_\mu$, and each of the three colors of $\bar{u}d'$. The branching ratios are predicted to be

$$\mathcal{B}(\tau^- \to \nu_\tau e^- \bar{\nu}_e) = \mathcal{B}(\tau^- \to \nu_\tau \mu^- \bar{\nu}_\mu) = \frac{1}{3}\mathcal{B}(\tau^- \to \nu_\tau \bar{u}d') = 20\% \quad . \tag{7.16}$$

Measured values for the purely leptonic branching ratios are slightly under 18%, as a result of the enhancement of the hadronic channels by a QCD correction whose leading-order behavior is $1 + \alpha_S/\pi$, the same as for R in e^+e^- annihilation. The τ decay is thus further evidence for the existence of three colors of quarks.

7.4 W decays

We shall calculate the rate for the process $W \to f\bar{f}'$ and then generalize the result to obtain the total W decay rate. The interaction Lagrangian (7.9) implies that the covariant matrix element for the process $W(k) \to f(p)\bar{f}'(p')$ is

$$\mathcal{M}^{(\lambda)} = \frac{g}{2\sqrt{2}}\bar{u}_f(p)\gamma^\mu(1-\gamma_5)v_{f'}(p')\epsilon_\mu^{(\lambda)}(k) \quad . \tag{7.17}$$

Here λ describes the polarization state of the W. The partial width is

$$\Gamma(W^- \to f\bar{f}') = \frac{1}{2M_W}\frac{1}{3}\sum_{\text{pols}}|\mathcal{M}^{(\lambda)}|^2\frac{p^*}{4\pi M_W} \quad , \tag{7.18}$$

where $(2M_W)^{-1}$ is the initial-state normalization, 1/3 corresponds to an average of W polarizations, the sum is over both W and lepton polarizations, and p^* is the final center-of-mass (c.m.) 3-momentum. We use the identity

$$\sum_\lambda \epsilon_\mu^{(\lambda)}(k)\epsilon_\nu^{(\lambda)*}(k) = -g_{\mu\nu} + \frac{k_\mu k_\nu}{M_W^2} \tag{7.19}$$

for sums over W polarization states. The result is that

$$\sum_{\text{pols}}|\mathcal{M}^{(\lambda)}|^2 = g^2\left[M_W^2 - \frac{1}{2}(m^2+m'^2) - \frac{(m^2-m'^2)^2}{2M_W^2}\right] \tag{7.20}$$

for any process $W \to f\bar{f}'$, where m is the mass of f and m' is the mass of f'. Recalling the relation between G_F and g^2, this may be written in the simpler form

$$\Gamma(W \to f\bar{f}') = \frac{G_F}{\sqrt{2}}\frac{M_W^3}{6\pi}\Phi_{ff'} \quad , \quad \Phi_{ff'} \equiv \frac{2p^*}{M_W}\frac{p^{*2}+3EE'}{M_W^2} \quad . \tag{7.21}$$

Here $E = (p^{*2}+m^2)^{1/2}$ and $E' = (p^{*2}+m'^2)^{1/2}$ are the c.m. energies of f and f'. The factor $\Phi_{ff'}$ reduces to 1 as $m, m' \to 0$. The present experimental average for the W mass is $M_W = 80.451 \pm 0.033$ GeV. Using this value, we predict $\Gamma(W \to e^-\bar{\nu}_e) = 227.8 \pm 2.3$ MeV. The widths to various channels are expected to be in the ratios

$$e^-\bar{\nu}_e : \mu^-\bar{\nu}_\mu : \tau^-\bar{\nu}_\tau : \bar{u}d' : \bar{c}s' = 1 : 1 : 1 : 3\left[1+\frac{\alpha_S(M_W^2)}{\pi}\right] : 3\left[1+\frac{\alpha_S(M_W^2)}{\pi}\right] \quad , \tag{7.22}$$

so $\alpha_S(M_W^2) = 0.120 \pm 0.002$ leads to the prediction $\Gamma_{\text{tot}}(W) = 2.10 \pm 0.02$ GeV. This is to be compared with a value obtained at LEP II by direct reconstruction of W's: $\Gamma_{\text{tot}}(W) = 2.150 \pm 0.091$ GeV. Higher-order electroweak corrections are not expected to play a major role here. This agreement means, among other things, that we are not missing a significant channel to which the charged weak current can couple below the mass of the W.

7.5 W pair production

We shall outline a calculation which indicates that the weak interactions cannot possibly be complete if described only by charged-current interactions. We consider the process $\nu_e(q) + \bar{\nu}_e(q') \to W^+(k) + W^-(k')$ due to exchange of an electron e^- with momentum p. The matrix element is

$$\mathcal{M}^{(\lambda,\lambda')} = \frac{G_F M_W^2}{\sqrt{2}} \bar{v}(q') \not{\epsilon}^{(\lambda')}(k')(1-\gamma_5)\frac{\not{p}}{p^2} \not{\epsilon}^{(\lambda)}(k) u(q) \quad . \tag{7.23}$$

For a longitudinally polarized W^+, this matrix element grows in an unacceptable fashion for high energy. In fact, an inelastic amplitude for any given partial wave has to be bounded, whereas $\mathcal{M}^{(\lambda,\lambda')}$ will not be. The polarization vector for a longitudinal W^+ travelling along the z axis is

$$\epsilon_\nu^{(\lambda)}(k) = (|\vec{k}|, 0, 0, M_W) \simeq k_\nu / M_W \quad , \tag{7.24}$$

with a correction which vanishes as $|\vec{k}| \to \infty$. Replacing $\epsilon_\nu^{(\lambda)}(k)$ by k_ν/M_W, using $\not{k} = \not{q} - \not{p}$ and $\not{q} u(q) = 0$, we find

$$\mathcal{M}^{(\lambda,\lambda')} \simeq -\sqrt{2} G_F M_W \bar{v}(q') \not{\epsilon}^{(\lambda')}(k') u(q) \quad , \tag{7.25}$$

$$\sum_{\text{lepton pol.}} |\mathcal{M}^{(\lambda,\lambda')}|^2 = 2G_F^2 M_W^2 [8q' \cdot \epsilon^{(\lambda')} q \cdot \epsilon^{(\lambda')} - 4q \cdot q' \epsilon^{(\lambda')} \cdot \epsilon^{(\lambda')}] \quad . \tag{7.26}$$

This quantity contributes only to the lowest two partial waves, and grows without bound as the energy increases. Such behavior is not only unacceptable on general grounds because of the boundedness of inelastic amplitudes, but it leads to divergences in higher-order perturbation contributions, e.g., to elastic $\bar{\nu}\nu$ scattering. Two possible contenders for a solution of the problem in the early 1970s were (1) a neutral gauge boson Z^0 coupling to $\nu\bar{\nu}$ and W^+W^- (Glashow 1961, Weinberg 1967, Salam 1968), or (2) a left-handed heavy lepton E^+ (Georgi and Glashow 1972a) coupling to $\nu_e W^+$. Either can reduce the unacceptable high-energy behavior to a constant. The Z^0 alternative seems to be the one selected in nature. In what follows we will retrace the steps of the standard electroweak theory, which led to the prediction of the W and Z and all the phenomena associated with them.

8 Electroweak unification

8.1 Guidelines for symmetry

We now return to the question of what to do with the "neutral W" (the particle we called W^3 in the previous Section). The time component of the charged weak current

$$J_\mu^{(+)} = \bar{N}_L \gamma_\mu L_L + \bar{U}_L \gamma_\mu V D_L \quad , \tag{8.27}$$

where N_L and L_L are neutral and charged lepton column vectors defined in analogy with U_L and D_L, may be used to define operators

$$Q^{(+)} \equiv \int d^3x J_0^{(+)} \quad , \quad Q^{(-)} \equiv Q^{(+)\dagger} \tag{8.28}$$

which are charge-raising and -lowering members of an SU(2) triplet. If we define $Q_3 \equiv (1/2)[Q^{(+)}, Q^{(-)}]$, the algebra closes: $[Q_3, Q^{(\pm)}] = \pm Q^{(\pm)}$. The form (8.27) (with unitary V) guarantees that the corresponding neutral current will be

$$J_\mu^{(3)} = \frac{1}{2}\left[\bar{N}_L\gamma_\mu N_L - \bar{L}_L\gamma_\mu L_L + \bar{U}_L\gamma_\mu U_L - \bar{D}_L\gamma_\mu D_L\right] \quad , \tag{8.29}$$

which is *diagonal* in neutral currents. This can only succeed, of course, if there are equal numbers of charged and neutral leptons, and equal numbers of charge 2/3 and charge $-1/3$ quarks. It would have been possible to define an SU(2) algebra making use only of a doublet (Gell-Mann and Lévy 1960)

$$\begin{bmatrix} u \\ d' \end{bmatrix}_L = \begin{bmatrix} u \\ V_{ud}d + V_{us}s \end{bmatrix}_L \tag{8.30}$$

which was the basis of the Cabibbo theory of the charge-changing weak interactions of strange and nonstrange particles. If one takes $V_{ud} = \cos\theta_C$, $V_{us} = \sin\theta_C$, as is assumed in the Cabibbo theory, the u, d, s contribution to the neutral current $J_\mu^{(3)}$ is

$$J_\mu^{(3)}|_{u,d,s} = \frac{1}{2}[\bar{u}_L\gamma_\mu u_L - \cos^2\theta_C \bar{d}_L\gamma_\mu d_L$$

$$- \sin^2\theta_C \bar{s}_L\gamma_\mu s_L - \sin\theta_C\cos\theta_C(\bar{d}_L\gamma_\mu s_L + \bar{s}_L\gamma_\mu d_L)] \quad . \tag{8.31}$$

This expression contains *strangeness-changing neutral currents*, leading to the expectation of many processes like $K^+ \to \pi^+\nu\bar{\nu}$, $K_L^0 \to \mu^+\mu^-$, ..., at levels far above those observed. It was the desire to banish strangeness-changing neutral currents that led Glashow *et al.* (1970) to introduce the charmed quark c (proposed earlier by several authors on the basis of a quark-lepton analogy) and the doublet

$$\begin{bmatrix} c \\ s' \end{bmatrix}_L = \begin{bmatrix} c \\ V_{cd}d + V_{cs}s \end{bmatrix}_L \quad . \tag{8.32}$$

In this four-quark theory, one assumes the corresponding matrix V is unitary. By suitable phase changes of the quarks, all elements can be made real, making V an orthogonal matrix with $V_{ud} = V_{cs} = \cos\theta_C$, $V_{us} = -V_{cd} = \sin\theta_C$. Instead of (8.31) one then has

$$J_\mu^{(3)}|_{u,d,s,c} = \frac{1}{2}[\bar{u}_L\gamma_\mu u_L + \bar{c}_L\gamma_\mu c_L - \bar{d}_L\gamma_\mu d_L - \bar{s}_L\gamma_\mu s_L] \quad , \tag{8.33}$$

which contains no flavor-changing neutral currents. The charmed quark also plays a key role in higher-order charged-current interactions. Let us consider K^0–$\overline{K}^0$ mixing. The shared $\pi\pi$ intermediate state and other low-energy states like π^0, η, and η' are chiefly responsible for CP-conserving K^0–$\overline{K}^0$ mixing.

If the only charge 2/3 quark contributing to this process were the u quark, one would expect a contribution to Δm_K of order

$$\Delta m_K|_u \sim g^4 f_K^2 m_K \sin^2\theta_C \cos^2\theta_C/16\pi^2 M_W^2 \sim G_F f_K^2 m_K (g^2/16\pi^2) \quad , \tag{8.34}$$

where f_K is the amplitude for $d\bar{s}$ to be found in a K^0, and the factor of $16\pi^2$ is characteristic of loop diagrams. This is far too large, since $\Delta m_K \sim \Gamma_{K_S} \sim G_F^2 f_K^2 m_K^3$. However, the introduction of the charmed quark, coupling to $-d\sin\theta_C + s\cos\theta_C$, cancels the leading contribution, leading to an additional factor of $[(m_c^2 - m_u^2)/M_W^2]\ln(M_W^2/m_c^2)$ in the above expression. Using such arguments Glashow *et al.* (1970) and Gaillard and Lee (1974) estimated the mass of the charmed quark to be less than several GeV. The discovery of the J/ψ confirmed this prediction; charmed hadrons produced in neutrino interactions and in e^+e^- annihilations followed soon after. An early motivation for charm relied on an analogy between quarks and leptons. Hara (1964), Maki and Ohnuki (1964), and Bjorken and Glashow (1964) inferred the existence of a charmed quark coupling mainly to the strange quark from the existence of the $\mu - \nu_\mu$ doublet:

$$\begin{pmatrix} \nu_\mu \\ \mu^- \end{pmatrix} : \text{leptons} \Rightarrow \begin{pmatrix} c \\ s \end{pmatrix} : \text{quarks} \quad . \tag{8.35}$$

Further motivation for the quark-lepton analogy was noted by Georgi and Glashow (1972b), and Gross and Jackiw (1972). In a gauge theory of the electroweak interactions, triangle anomalies have to be avoided. This cancellation requires the fermions f in the theory to contribute a total of zero to the sum over f of $Q_f^2 I_{3L}^f$. Such a cancellation can be achieved by requiring quarks and leptons to occur in complete families so that the terms

$$\text{Leptons}: \quad (0)^2\left(\frac{1}{2}\right) + (-1)^2\left(-\frac{1}{2}\right) = -\frac{1}{2} \tag{8.36}$$

$$\text{Quarks}: \quad 3\left[\left(\frac{2}{3}\right)^2\left(\frac{1}{2}\right) + \left(-\frac{1}{3}\right)^2\left(-\frac{1}{2}\right)\right] = \frac{1}{2} \tag{8.37}$$

sum to zero for each family.

We are then left with a flavor-preserving neutral current $J_\mu^{(3)}$, given by (8.33), whose interpretation must still be given. It cannot correspond to the photon, since the photon couples to both left-handed and right-handed fermions. At the same time, the photon is somehow involved in the weak interactions associated with W exchange. In particular, the $W^\pm$ themselves are charged, so any theory in which electromagnetic current is conserved must involve a $\gamma W^+ W^-$ coupling. Moreover, the charge is sensitive to the third component of the SU(2) algebra we have just introduced. We shall refer to this as $\mathrm{SU}(2)_L$, recognizing that only left-handed fermions ψ_L transform non-trivially under it. Then we can define a *weak hypercharge* Y in terms of the difference between the electric charge Q and the third component I_{3L} of $\mathrm{SU}(2)_L$ (*weak isospin*):

$$Q = I_{3L} + \frac{Y}{2} \quad . \tag{8.38}$$

Values of Y for quarks and leptons are summarized in Table 3.

These weak hypercharge assignments follow naturally in unified theories (*grand unified theories*) of the electroweak and strong interactions. A formula for Y, which may have

Table 3: Values of charge, I_{3L}, and weak hypercharge Y for quarks and leptons.

Particle(s)	Q	I_{3L}	Y
ν_{eL}	0	1/2	−1
e_L^-	−1	−1/2	−1
u_L	2/3	1/2	1/3
d_L	−1/3	−1/2	1/3
e_R^-	−1	0	−2
u_R	2/3	0	4/3
d_R	−1/3	0	−2/3

deeper significance (Pati and Salam 1973), is $Y = 2I_{3R} + (B - L)$, where I_{3R} is the third component of "right-handed" isospin, B is baryon number (1/3 for quarks), and L is lepton number (1 for leptons such as e^- and ν_e). The orthogonal component of I_{3R} and $B-L$ may correspond to a higher-mass, as-yet-unseen vector boson, an example of what is called a Z'. The search for Z' bosons with various properties is an ongoing topic of interest; current limits are quoted by the Particle Data Group. The gauge theory of charged and neutral W's thus must involve the photon in some way. It will then be necessary, in order to respect the formula (8.38), to introduce an additional U(1) symmetry associated with weak hypercharge. The combined electroweak gauge group will have the form $\mathrm{SU}(2)_L \otimes \mathrm{U}(1)_Y$.

8.2 Symmetry breaking

Any unified theory of the weak and electromagnetic interactions must be broken, since the photon is massless while the W bosons (at least) are not. An explicit mass term in a gauge theory of the form $m^2 A^i_\mu A^{\mu i}$ violates gauge invariance. It is not invariant under the replacement (3.48). Another means must be found to introduce a mass. The symmetry must be broken in such a way as to preserve gauge invariance. A further manifestation of symmetry breaking is the presence of fermion mass terms. Any product $\overline{\psi}\psi$ may be written as

$$\overline{\psi}\psi = (\overline{\psi}_L + \overline{\psi}_R)(\psi_L + \psi_R) = \overline{\psi}_L\psi_R + \overline{\psi}_R\psi_L \quad , \tag{8.39}$$

using the fact that $\overline{\psi}_L = \overline{\psi}(1+\gamma_5)/2$, $\overline{\psi}_R = \overline{\psi}(1-\gamma_5)/2$. Since ψ_L transforms as an $\mathrm{SU}(2)_L$ doublet but ψ_R as an $\mathrm{SU}(2)_L$ singlet, a mass term proportional to $\overline{\psi}\psi$ transforms as an overall $\mathrm{SU}(2)_L$ doublet. Moreover, the weak hypercharges of left-handed fermions and their right-handed counterparts are different. Hence one cannot even have explicit *fermion* mass terms in the Lagrangian and hope to preserve local gauge invariance. One way to generate a fermion mass without explicitly violating gauge invariance is to assume the existence of a complex scalar $\mathrm{SU}(2)_L$ doublet ϕ coupled to fermions via a Yukawa interaction:

$$\mathcal{L}_Y = -g_Y(\overline{\psi}_L \phi \psi_R + \text{h.c.}) \quad , \quad \phi \equiv \begin{bmatrix} \phi^+ \\ \phi^0 \end{bmatrix} \quad . \tag{8.40}$$

Thus, for example, with $\overline{\psi}_L = (\bar{\nu}_e, e)_L$ and $\psi_R = e_R$, we have

$$\mathcal{L}_{Y,e} = -g_{Ye}(\bar{\nu}_{eL}\phi^+ e_R + \bar{e}_L \phi^0 e_R + \text{h.c.}) \quad . \tag{8.41}$$

If ϕ^0 acquires a vacuum expectation value, $\langle \phi^0 \rangle \neq 0$, this quantity will automatically break $\text{SU}(2)_L$ and $\text{U}(1)_Y$, and will give rise to a non-zero electron mass. A neutrino mass is not generated, simply because no right-handed neutrino has been assumed to exist. The gauge symmetry is not broken in the Lagrangian, but only in the solution. This is similar to the way in which rotational invariance is broken in a ferromagnet, where the fundamental interactions are rotationally invariant but the ground-state solution has a preferred direction along which the spins are aligned. The d quark masses are generated by similar couplings involving $\overline{\psi}_L = (\bar{u}, \bar{d})_L$, $\psi_R = d_R$, so that

$$\mathcal{L}_{Y,d} = -g_{Yd}(\bar{u}_L \phi^+ d_R + \bar{d}_L \phi^0 d_R + \text{h.c.}) \quad . \tag{8.42}$$

To generate u quark masses one must either use the multiplet

$$\tilde{\phi} \equiv \begin{bmatrix} \bar{\phi}^0 \\ -\phi^- \end{bmatrix} = i\tau^2 \phi^* \quad , \tag{8.43}$$

which also transforms as an SU(2) doublet, or a separate doublet of scalar fields

$$\phi' = \begin{bmatrix} \phi'^0 \\ \phi'^- \end{bmatrix} \quad . \tag{8.44}$$

With $\overline{\psi}_L = (\bar{u}, \bar{d})_L$ and $\psi_R = u_R$, we then find

$$\mathcal{L}_{Y,u} = -g_{Yu}(\bar{u}_L \bar{\phi}^0 u_L - \bar{d}_L \phi^- u_L + \text{h.c.}) \tag{8.45}$$

if we make use of $\tilde{\phi}$, or

$$\mathcal{L}_{Y,u} = -g_{Yu}(\bar{u}_L \phi'^0 u_L + \bar{d}_L \phi'^- u_L + \text{h.c.}) \tag{8.46}$$

if we use ϕ'. For present purposes we shall assume the existence of a single complex doublet, though many theories (notably, some grand unified theories or supersymmetry) require more than one.

8.3 Scalar fields and the Higgs mechanism

Suppose a complex scalar field of the form (8.40) is described by a Lagrangian

$$\mathcal{L}_\phi = (\partial_\mu \phi)^\dagger (\partial^\mu \phi) - \frac{\lambda}{4}(\phi^\dagger \phi)^2 + \frac{\mu^2}{2}\phi^\dagger \phi \quad . \tag{8.47}$$

Note the "wrong" sign of the mass term. This Lagrangian is invariant under $\mathrm{SU}(2)_L \otimes \mathrm{U}(1)_Y$. The field ϕ will acquire a constant vacuum expectation value which we calculate by asking for the stationary value of $\mathcal{L}_\phi$:

$$\frac{\partial \mathcal{L}_\phi}{\partial(\phi^\dagger\phi)} = 0 \Rightarrow \langle\phi^\dagger\phi\rangle = \frac{\mu^2}{\lambda} \quad . \tag{8.48}$$

We still have not specified which component of ϕ acquires the vacuum expectation value. At this point only $\phi^\dagger\phi = |\phi^+|^2 + |\phi^0|^2$ is fixed, and (Re ϕ^+, Im ϕ^+, Re ϕ^0, Im ϕ^0) can range over the surface of a four-dimensional sphere. The Lagrangian (8.47) is, in fact, invariant under rotations of this four-dimensional sphere, a group SO(4) isomorphic to $\mathrm{SU}(2) \otimes \mathrm{U}(1)$. A lower-dimensional analogue of this surface would be the bottom of a wine bottle along which a marble rolls freely in an orbit a fixed distance from the center. Let us *define* the vacuum expectation value of ϕ to be a real parameter in the ϕ^0 direction:

$$\langle\phi\rangle = \begin{bmatrix} 0 \\ v/\sqrt{2} \end{bmatrix} \quad . \tag{8.49}$$

The factor of $1/\sqrt{2}$ is introduced for later convenience. We then find, from the discussion in the previous section, that Yukawa couplings of ϕ to fermions ψ_i generate mass terms $m_i = g_{Yi}v/\sqrt{2}$. We must now see what such vacuum expectation values do to gauge boson masses.

In order to introduce gauge interactions with the scalar field ϕ, one must replace ∂_μ by $\mathbf{D}_\mu$ in the kinetic term of the Lagrangian (8.47). Here

$$\mathbf{D}_\mu = \partial_\mu - ig\frac{\tau^i W^i_\mu}{2} - ig'\frac{Y}{2}B_\mu \quad , \tag{8.50}$$

where the $\mathrm{U}(1)_Y$ interaction is characterized by a coupling constant g' and a gauge field B_μ, and we have written g for the $SU(2)$ coupling discussed earlier. It will be convenient to write ϕ in terms of four independent real fields $(\xi^i,\ \eta)$ in a slightly different form:

$$\phi = \exp\left(\frac{i\xi^i\tau^i}{2v}\right)\begin{bmatrix} 0 \\ \frac{v+\eta}{\sqrt{2}} \end{bmatrix} \quad . \tag{8.51}$$

We then perform an $\mathrm{SU}(2)_L$ gauge transformation to remove the ξ dependence of ϕ, and rewrite it as

$$\phi = \begin{bmatrix} 0 \\ \frac{v+\eta}{\sqrt{2}} \end{bmatrix} \quad . \tag{8.52}$$

The fermion and gauge fields are transformed accordingly; we rewrite the Lagrangian for them in the new gauge. The resulting kinetic term for the scalar fields, taking account that $Y = 1$ for the Higgs field (8.40), is

$$\begin{aligned} \mathcal{L}_{K,\phi} &= (\mathbf{D}_\mu\phi)^\dagger(\mathbf{D}^\mu\phi) \\ &= \left|\left\{\partial_\mu - \frac{ig}{2}\begin{bmatrix} W^3_\mu & W^1_\mu - iW^2_\mu \\ W^1_\mu + iW^2_\mu & -W^3_\mu \end{bmatrix} - \frac{ig'}{2}B_\mu\right\}\begin{bmatrix} 0 \\ \frac{v+\eta}{\sqrt{2}} \end{bmatrix}\right|^2 \quad . \end{aligned} \tag{8.53}$$

This term contains several contributions.

1. There is a kinetic term $\frac{1}{2}(\partial_\mu\eta)(\partial^\mu\eta)$ for the scalar field η.
2. A term $v\partial_\mu\eta$ is a total divergence and can be neglected.
3. There are $WW\eta$, $BB\eta$, $WW\eta^2$, and $BB\eta^2$ interactions.
4. The v^2 term leads to a mass term for the Yang-Mills fields:

$$\mathcal{L}_{m,YM} = \frac{v^2}{8}\left\{g^2[(W^1)^2 + (W^2)^2] + (gW^3 - g'B)^2\right\} \quad . \tag{8.54}$$

The spontaneous breaking of the SU(2) $\otimes$ U(1) symmetry thus has led to the appearance of a mass term for the gauge fields. This is an example of the *Higgs mechanism.* An unavoidable consequence is the appearance of the scalar field η, the *Higgs field.* The masses of the charged W bosons may be identified by comparing Eqs. (8.54) and (7.8):

$$(gv)^2/8 = M_W^2/2 \quad , \quad \text{or } M_W = gv/2 \quad . \tag{8.55}$$

Since the Fermi constant is related to g/M_W, one finds

$$\frac{G_F}{\sqrt{2}} = \frac{g^2}{8M_W^2} = \frac{1}{2v^2} \quad , \quad \text{or} \quad v = 2^{-1/4}G_F^{-1/2} = 246 \text{ GeV} \quad . \tag{8.56}$$

The combination $gW_\mu^3 - g'B_\mu$ also acquires a mass. We must normalize this combination suitably so that it contributes properly in the kinetic term for the Yang-Mills fields:

$$\mathcal{L}_{K,YM} = -\frac{1}{4}W_{\mu\nu}^i W^{\mu\nu i} - \frac{1}{4}B_{\mu\nu}B^{\mu\nu} \quad , \tag{8.57}$$

where

$$W_{\mu\nu}^i \equiv \partial_\mu W_\nu^i - \partial_\nu W_\mu^i + g\epsilon_{ijk}W_\mu^j W_\nu^k \quad , \quad B_{\mu\nu} \equiv \partial_\mu B_\nu - \partial_\nu B_\mu \quad . \tag{8.58}$$

Defining

$$\cos\theta \equiv \frac{g}{(g^2+g'^2)^{1/2}} \quad \left[\text{so that} \quad \sin\theta = \frac{g'}{(g^2+g'^2)^{1/2}}\right] \quad , \tag{8.59}$$

we may write the normalized combination $\sim gW_\mu^3 - g'B_\mu$ which acquires a mass as

$$Z_\mu \equiv W_\mu^3\cos\theta - B_\mu\sin\theta \quad . \tag{8.60}$$

The orthogonal combination does not acquire a mass. It may then be identified as the photon:

$$A_\mu = B_\mu\cos\theta + W_\mu^3\sin\theta \quad . \tag{8.61}$$

The mass of the Z is given by

$$\frac{(g^2+g'^2)v^2}{8} = \frac{M_Z^2}{2} \quad , \quad \text{or} \quad M_Z = M_W(g^2+g'^2)^{1/2}/g = M_W/\cos\theta \quad , \tag{8.62}$$

using (8.59) in the last relation. The W's and Z's have acquired masses, but they are not equal unless g' were to vanish. We shall see in the next subsection that both g and g' are

nonzero, so one expects the Z to be heavier than the W. It is interesting to stop for a moment to consider what has taken place. We started with four scalar fields ϕ^+, ϕ^-, ϕ^0, and $\bar{\phi}^0$. Three of them [ϕ^+, ϕ^-, and the combination $(\phi^0 - \bar{\phi}^0)/i\sqrt{2}$] could be absorbed in the gauge transformation in passing from (8.51) to (8.52), which made sense only as long as $(\phi^0 + \bar{\phi})/\sqrt{2}$ had a vacuum expectation value v. The net result was the generation of mass for three gauge bosons W^+, W^-, and Z. If we had not transformed away the three components ξ^i of ϕ in (8.51), the term $\mathcal{L}_{K,\phi}$ in the presence of gauge fields would have contained contributions $W_\mu \partial^\mu \phi$ which mixed gauge fields and derivatives of ϕ. These can be expressed as

$$W_\mu \partial^\mu \phi = \partial^\mu (W_\mu \phi) - (\partial^\mu W_\mu)\phi \tag{8.63}$$

and the total divergence (the first term) discarded. One thus sees that such terms *mix longitudinal components of gauge fields* (proportional to $\partial^\mu W_\mu$) *with scalar fields.* It is necessary to redefine the gauge fields by means of a gauge transformation to get rid of such mixing terms. It is just this transformation that was anticipated in passing from (8.51) to (8.52). The three "unphysical" scalar fields provide the necessary longitudinal degrees of freedom in order to convert the massless $W^\pm$ and Z to massive fields. Each massless field possesses only two polarization states ($J_z = \pm J$), while a massive vector field has three ($J_z = 0$ as well). Such counting rules are extremely useful when more than one Higgs field is present, to keep track of how many scalar fields survive being "eaten" by gauge fields.

8.4 Interactions in the SU(2) $\otimes$ U(1) theory

By introducing gauge boson masses via the Higgs mechanism, and letting the simplest non-trivial representation of scalar fields acquire a vacuum expectation value v, we have related the Fermi coupling constant to v, and the gauge boson masses to gv or $(g^2+g'^2)^{1/2}v$. We still have two arbitrary couplings g and g' in the theory, however. We shall show how to relate the electromagnetic coupling to them, and how to measure them separately. The interaction of fermions with gauge fields is described by the kinetic term $\mathcal{L}_{K,\psi} = \bar{\psi}\, \not{D}\psi$. Here, as usual,

$$\not{D} = \not{\partial} - ig\frac{\tau^i}{2}\not{W}^i - ig'\frac{Y}{2}\not{B} \quad . \tag{8.64}$$

The charged-W interactions have already been discussed. They are described by the terms (7.9) for leptons and (7.13) for quarks. The interactions of W^3 and B may be re-expressed in terms of A and Z via the inverse of (8.60) and (8.61):

$$W^3_\mu = Z_\mu \cos\theta + A_\mu \sin\theta \quad , \quad B_\mu = -Z_\mu \sin\theta + A_\mu \cos\theta \quad . \tag{8.65}$$

Then the covariant derivative for neutral gauge bosons is

$$\not{D}|_{\text{neutral}} = \not{\partial} - igI_{3L}(\not{Z}\cos\theta + \not{A}\sin\theta) - ig'(Q - I_{3L})(-\not{Z}\sin\theta + \not{A}\cos\theta) \quad . \tag{8.66}$$

Here we have substituted $Y/2 = (Q - I_{3L})$. We identify the electromagnetic contribution to the right-hand side of (8.66) with the familiar one $-ieQ\,\not{A}$, so that

$$e = g'\cos\theta = g\sin\theta \quad . \tag{8.67}$$

The second equality, stemming from the demand that $I_{3L}\,\not{A}$ terms cancel one another in (8.66), is automatically satisfied as a result of the definition (8.59). Combining (8.59) and (8.67), we find

$$e = \frac{gg'}{\sqrt{g^2+g'^2}} \quad \text{or} \quad \frac{1}{e^2} = \frac{1}{g^2} + \frac{1}{g'^2} \quad , \tag{8.68}$$

the result advertised in the Introduction. The interaction of the Z with fermions may be determined from Eq. (8.66) with the help of (8.59), noting that $g\cos\theta + g'\sin\theta = (g^2+g'^2)^{1/2}$ and $g'\sin\theta = (g^2+g'^2)^{1/2}\sin^2\theta$. We find

$$\not{D}|_{\text{neutral}} = \not{\partial} - ieQ\,\not{A} - i(g^2+g'^2)^{1/2}(I_{3L} - Q\sin^2\theta)\,\not{Z} \quad . \tag{8.69}$$

Knowledge of the weak mixing angle θ will allow us to predict the W and Z masses. Using $G_F/\sqrt{2} = g^2/8M_W^2$ and $g\sin\theta = e$, we can write

$$M_W = \left[\frac{\pi\alpha}{\sqrt{2}G_F}\right]^{1/2} \frac{1}{\sin\theta} = \frac{37.3 \ \text{GeV}}{\sin\theta} \tag{8.70}$$

if we were to use $\alpha^{-1} = 137.036$. However, we shall see in the next Section that it is more appropriate to use a value of $\alpha^{-1} \simeq 129$ at momentum transfers characteristic of the W mass. With this and other electroweak radiative corrections, the correct estimate is raised to $M_W \simeq 38.6$ GeV$/\sin\theta$, leading to the successful predictions. The Z mass is expressed in terms of the W mass by $M_Z = M_W/\cos\theta$.

9 Recommended Literature

Particle Data Group Collaboration, Phys. Rev. D **66**, 010001 (2002).
Yu.V.Novozhilov, Introduction to Elementary Particle Theory, Pergamon Press, 1975.
S. Weinberg, The Quantum Theory of Fields, vol.1, Cambridge University Press, 1995.
S.Adler, Adventures in Theoretical Physics Selected Papers with Commentaries, World Scientific, 2006.
E.Shuryak, The QCD Vacuum Hadrons and Superdense Matter, World Scientific, 1988.
L.B.Okun, Leptons and Quarks, Nauka, 1990.
M.E.Peskin and D.V.Schroeder, An Introduction to Quantum Field Theory, Addison-Wesley Publishing Company, 1995.
P.Ramon, Field Theory, the Benjamin/Cummings Publishing Company, Inc., 1981.