

Gauge Theories of the Strong and Electroweak Interactions

G. Münster, G. Bergner

Summer term 2011

Notes by B. Echtermeyer

Is nature obeying fundamental laws? Does a comprehensive description of the laws of nature, a kind of theory of everything, exist?

Gauge theories and symmetry principles provide us with a comprehensive description of the presently known fundamental particles and interactions. The Standard Model of elementary particle physics is based on gauge theories, and the interactions between the elementary particles are governed by a symmetry principle, namely local gauge invariance, which represents an infinite dimensional symmetry group.

These notes are not free of errors and typos. Please notify us if you find some.

Contents

1	Introduction	4
1.1	Particles and Interactions	4
1.2	Relativistic Field Equations	13
1.2.1	Klein-Gordon equation	13
1.2.2	Dirac equation	16
1.2.3	Maxwell's equations	19
1.2.4	Lagrangian formalism for fields	22
1.3	Symmetries	30
1.3.1	Symmetries and conservation laws	30
1.3.2	U(1) symmetry, electric charge	32
1.3.3	SU(2) symmetry, isospin	34
1.3.4	SU(3) flavour symmetry	42
1.3.5	Some comments about symmetry	44

1.4	Field Quantisation	46
1.4.1	Quantisation of the real scalar field	47
1.4.2	Quantisation of the complex scalar field	52
1.4.3	Quantisation of the Dirac field	54
1.4.4	Quantisation of the Maxwell field	55
1.4.5	Symmetries and Noether charges	57
1.5	Interacting Fields	58
1.5.1	Interaction picture	58
1.5.2	The S-matrix	61
1.5.3	Wick's theorem	62
1.5.4	Feynman diagrams	63
1.5.5	Fermions	66
1.5.6	Limitations of the perturbative approach	68
2	Quantum Electrodynamics (QED)	69
2.1	Local U(1) Gauge Symmetry	69
2.2	Quantum Electrodynamics	71
3	Non-abelian Gauge Theory	74
3.1	Local Gauge Invariance	74
3.2	Geometry of Gauge Fields	80
3.2.1	Differential geometry	80
3.2.2	Gauge Theory	82
4	Quantum Chromodynamics (QCD)	87
4.1	Lagrangian Density and Symmetries	87
4.1.1	Local SU(3) colour symmetry	88
4.1.2	Global flavour symmetry	90
4.1.3	Chiral symmetry	91
4.1.4	Broken chiral symmetry	95
4.2	Running Coupling	97
4.2.1	Quark-quark scattering	98
4.2.2	Renormalisation	100
4.2.3	Running coupling	101
4.2.4	Discussion	102
4.3	Confinement of Quarks and Gluons	105
4.4	Experimental Evidence for QCD	108
5	Electroweak Theory	111
5.1	Weak Interactions	111
5.1.1	Fermi theory of weak interaction	111

5.1.2	Parity violation	111
5.1.3	V-A theory	113
5.2	Higgs Mechanism	117
5.2.1	Spontaneous breakdown of a global symmetry	117
5.2.2	Higgs mechanism	119
5.3	Glashow-Weinberg-Salam Model	121

1 Introduction

1.1 Particles and Interactions

When reflecting on the constituents of matter, one is lead to the physics of elementary particles. A classification of elementary particles is done by regarding their properties, which are

mass, spin,
 (according to the representations of the inhomogeneous Lorentz group)
 lifetime,
 additional quantum numbers,
 (obtained from conservation laws)
 participation in interactions.

From these properties the following classification arose.

Leptons

e^- , ν_e electron number
 μ^- , ν_μ muon number
 τ^- , ν_τ tauon number

Hadrons strongly interacting particles

Mesons integer spin, baryon number = 0

π^+ , π^- , π^0 , K^+ , K^- , K^0 , η , ρ^+ , ρ^- , ρ^0 , J/ψ etc.

Baryons half integer spin, baryon number = ± 1

n , p , Λ^0 , Σ , Ξ , Δ , Ω^- , Y etc.

Quark model of hadrons (Gell-Mann, SU(3), eightfold way)

The hadrons are build out of two or three quarks.

Mesons

$q \bar{q}$ (quark, antiquark)

Baryons

$q q q$

There are six quarks and their antiparticles. They all have spin 1/2.

The six quark types are called “flavours”, which are denoted by

u, c, t
 d, s, b

Some baryons	some mesons
$p = uud$	$\pi^+ = u\bar{d}$
$n = udd$	$K^+ = u\bar{s}$
$\Lambda = sud$	$\rho^+ = u\bar{d}$
$\Omega^- = sss$	$D^+ = c\bar{d}$
	$\eta_c = c\bar{c}$

3 Generations of constituents

	mass [MeV]	Q	B
ν_e	≈ 0	0	0
e^-	0.511	-1	0
u	≈ 4	2/3	1/3
d	≈ 7	-1/3	1/3
ν_μ	≈ 0	0	0
μ^-	105.66	-1	0
c	≈ 1300	2/3	1/3
s	≈ 150	-1/3	1/3
ν_τ	< 18.2	0	0
τ^-	1777.0	-1	0
t	$\approx 174\,000$	2/3	1/3
b	$\approx 4\,200$	-1/3	1/3

Table 1: The 3 generations of elementary particles

Fig. 1 and Fig. 2 show multiplets of mesons and baryons arranged in 3-dimensional multiplets¹. The coordinates are

$$(x, y, z) = (\text{Isospin } I, \text{Hypercharge } Y, \text{Charm } C)$$

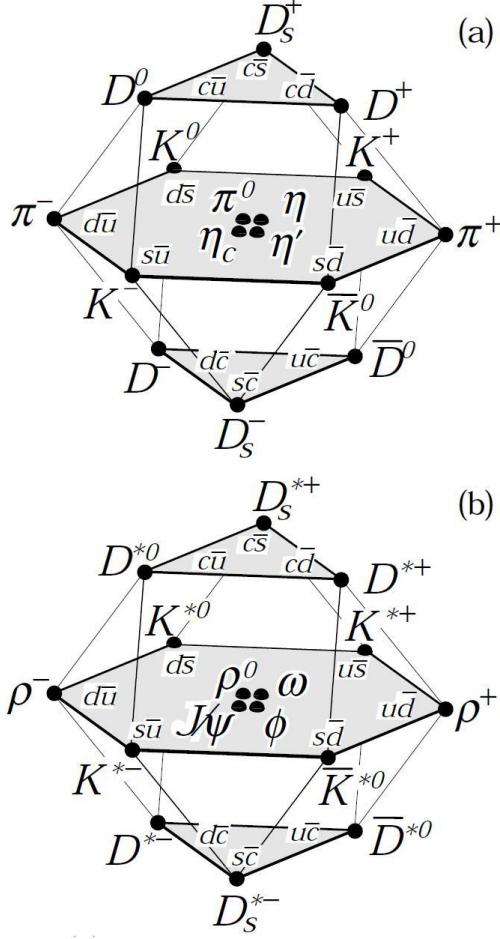
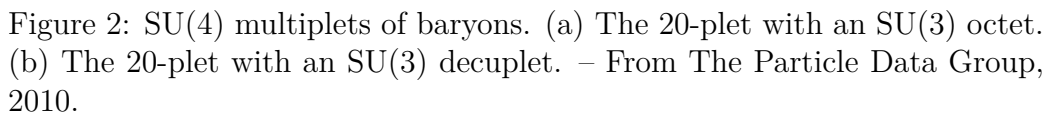


Figure 1: SU(4) multiplets of mesons; 16-plets of pseudoscalar (a) and vector mesons (b). In the central planes the $c\bar{c}$ states have been added. – From The Particle Data Group, 2010.

¹The meson multiplets form an Archimedean solid called cubooctahedron

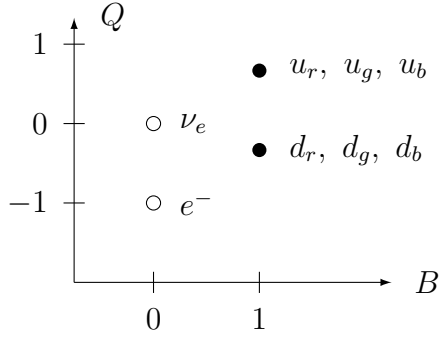


Quarks do not exist as single free particles. There is an additional quantum number, called “colour”. E.g., $\Omega^- = sss$ has spin $3/2$; therefore the wavefunction has to be antisymmetric in the spin-coordinates. It is also symmetric in space coordinates, so the Pauli-principle can only be fulfilled, if the three charmed quarks are different in some additional quantum number.

All hadrons are *colourless* combinations of quarks. This phenomenon is called confinement.

There is a characteristic feature for each single generation of leptons and quarks:

$$\sum Q_i = 0.$$



The reason, why (ν, e^-) and (u, d) belong to this same generation and not, for instance (ν, e^-) and (c, s) will be given later in the chapter on weak interactions.

Interactions

An important guiding principle in the history of understanding interactions has been unification. When Newton postulated that the gravitational force which pulls us down to earth and the force between moon and earth are essentially the same, this was a step towards unification of fundamental forces, as was the unification of magnetism and electricity by Faraday and Maxwell, which led to a new understanding of light, or – about a century later – the unification of electromagnetism and weak interactions.

Nowadays one distinguishes four fundamental interactions:

- a) Electromagnetic interactions. They apply to electrically charged particles only (no to neutrinos, for instance).

Since the electrostatic force is proportional to $1/r^2$, one says that the *range* of the electromagnetic interactions is infinite. A further characteristic of interactions is their *relative strength*, when compared with the strength of other interactions. For electromagnetism it is given by Sommerfeld's “Feinstrukturkonstante” α .

$$\text{range} = \infty \tag{1.1}$$

$$\text{relative strength} = \frac{e^2}{4\pi\epsilon_0\hbar c} \approx \frac{1}{137} \tag{1.2}$$

- b) Weak interactions. They are responsible for the β - decay and other processes.

$$\text{range} \approx 10^{-18} \text{ m} \quad (1.3)$$

$$\text{relative strength} \approx 10^{-5} \quad (1.4)$$

- c) Strong interactions. They are responsible for the binding of quarks and for the hadronic interactions. Nuclear forces are also remnants of the strong interactions.

$$\text{range} \approx 10^{-15} \text{ m} \quad (1.5)$$

$$\text{relative strength} \approx 1 \quad (1.6)$$

- d) Gravitation acts on every sort of matter. E.g., it has been shown experimentally that a neutron falls down through a vacuum tube just like any other object on earth. The gravitational force is always attractive. Whereas positive and negative electric charges exist, there are no negative masses and thus the gravitational force cannot be screened. The range of this force is infinite like that of electromagnetism. Comparing the gravitational force between proton and electron in an H-atom with their electrostatic attraction, one finds that the gravitational force is extremely weak.

$$\text{range} = \infty \quad (1.7)$$

$$\text{relative strength} \approx 10^{-39} \quad (1.8)$$

Forces are mediated by the exchange of bosons.

The range is given by the Compton wavelength of the exchange boson. (But there is an exception to this law in QCD due to confinement.)

$$\text{range } R \approx \frac{\hbar}{m c} \quad (1.9)$$

Interaction	bosons	spin	mass, range
electromagnetic	photon γ	1	$m = 0, R = \infty$
weak	W^+, W^-, Z^0	1	$m_W = 80.4 \text{ GeV}$ $m_Z = 91.2 \text{ GeV}$
strong	gluons G	1	$m = 0, R \neq 0$
gravitation	graviton	2	$m = 0$

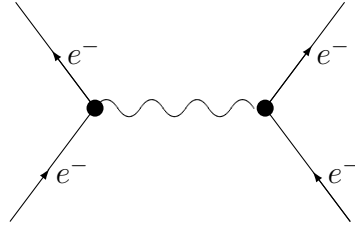
For gluons the spin 1 is a consequence of gauge theory, and the finite range R arises from confinement, which holds for gluons as for quarks. The spin of the exchange boson is related to the possibility of a force being only attractive or both attractive and repulsive. Spin 2 implies that there is only

attraction. The existence of the graviton with zero mass is predicted theoretically and may never be verified by experiment. Measuring gravitational waves is already very challenging, and to identify the quanta of these waves would be extremely difficult.

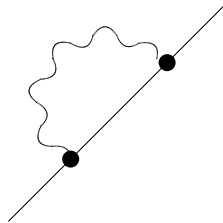
Theories

- a) Quantum Electrodynamics originated in 1927, when in an appendix to the article of Born, Heisenberg and Jordan about matrix mechanics Jordan quantised the free electromagnetic field. It was developed further by Dirac, Jordan, Pauli, Heisenberg and others and culminated before 1950 in the work of Tomonaga, Schwinger, Feynman and Dyson. The calculation of the Lamb shift and the exact value of the gyromagnetic ratio g of the electron are highlights of QED.

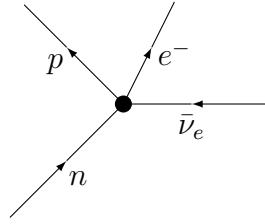
Here is an example of a Feynman diagram for the scattering of two electrons by exchanging a photon.



The vertex stands for a number, in QED this is $\alpha \approx 1/137$. The propagation of electrons is affected by the emission and absorption of virtual photons, as shown in the following Feynman diagram.



- b) The theory of weak interactions begun in 1932 with Fermi's theory for the β^- -decay. The Feynman graph for the decay of neutrons involves a 4-fermion coupling.



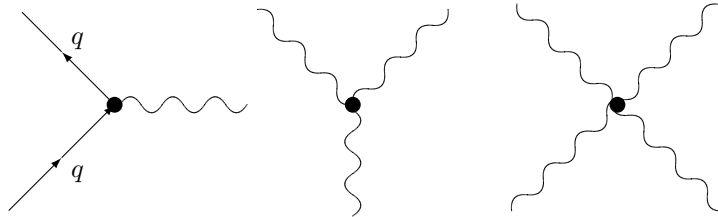
Improvements of the theory of β -decay in nucleons were made by the V-A theory, taking care of parity violation.

Theoretical problems: while in QED perturbation theory in powers of α works extremely well, it leads to infinities in the Fermi theory of weak interactions. The problems were overcome in 1961 – 1968 by Glashow, Weinberg, Salam and others, developing the unified theory of weak and electromagnetic interactions. The bosons mediating the electroweak interactions are

Vector bosons $W^\pm, Z^0$ and photon γ .

- c) Strong interactions between quarks are described by Quantum Chromodynamics (QCD), which was formulated by Fritzsch, Gell-Mann and Leutwyler, and further developed by 't Hooft and others. There are three “strong charges”, sources for the forces, named red, green and blue charge. The gauge bosons which mediate strong interactions are called gluons.

Unlike the electrically neutral photons in QED, gluons carry colour charges themselves and interact with each other. Due to their self-interactions, gluons may form glueballs, and a “theory of pure glue” is a non-trivial theory.



Feynman diagrams with quarks and gluons

- d) Gravitation is described by General Relativity (GRT), a nonlinear theory. A quantum theory of gravitation is not yet known. String theory, Superstring theory or Loop gravity might be candidates.

The Standard Model

This means the theory of Glashow, Weinberg and Salam (G.W.S.) plus QCD. There is no mixing between the Lagrangians for electroweak and strong interactions, therefore, we do not speak of a unification of these interactions. The theoretical predictions of the Standard Model are so far consistent with the experimental results.

Common to all parts of the Standard Model are exchange bosons, which are related to gauge fields showing local gauge symmetries. (Gravitation is also based on a local symmetry.) Gauge theories are based on gauge groups. The groups belonging to the Standard Model are

$$\underbrace{\mathrm{SU}(3)}_{\mathrm{QCD}} \otimes \underbrace{\mathrm{SU}(2) \otimes \mathrm{U}(1)}_{\mathrm{G.W.S.}}. \quad (1.10)$$

The principles of the Standard Model are:

- local gauge symmetry,
- Higgs mechanism giving masses to $W^\pm, Z^0$ and quarks.

The Higgs mechanism is due to P. Anderson, F. Englert, R. Brout, P. Higgs, G. Guralnik, C. R. Hagen and T. Kibble. It uses the Higgs field, associated with a Higgs-boson. This does not fit into a local gauge theory, so the Higgs boson might not be a fundamental particle. There is no other reason for the Higgs field than the mechanism to give the above mentioned masses.

Outlook

A further unification of interactions is attempted in Grand Unified Theories (GUT). The idea is to extend the semisimple² Lie group $\mathrm{SU}(3) \otimes \mathrm{SU}(2) \otimes \mathrm{U}(1)$ to a simple Lie group as for example $\mathrm{SU}(5)$, $\mathrm{SO}(10)$ or the exceptional Lie group E_6 . GUTs predict proton decay and several Higgs particles.

²A group is called semisimple, if it is the direct product of simple groups. A group is simple, if it has no normal subgroups besides the trivial ones.

1.2 Relativistic Field Equations

In classical physics there are two distinct kinds of objects: particles – point particles or continuous distributions of mass – and secondly fields, like gravitational or electromagnetic fields. In quantum mechanics the dichotomy between particles and fields is upheld, although the wave-particle duality shows up. But in the relativistic quantum mechanics of particles one encounters contradictions. These are resolved in Quantum Field Theory (QFT). QFT deals with quantised fields, functions of space and time, where the values of the fields $f(\vec{r}, t)$ themselves become operators.

field $f \longrightarrow$ operator.

QFT is a quantum theory of many particles. In this lecture we consider the three most prominent relativistic field equations,

- Klein-Gordon equation for spin 0 particles,
- Dirac equation for spin 1/2 particles,
- Maxwell's equations for massless spin 1 particles.

There are other relativistic equations, too (Proca, etc). In QFT the spin of fundamental fields does not exceed 2.

1.2.1 Klein-Gordon equation

For a non-relativistic free particle the equation

$$E = \frac{\vec{p}^2}{2m} \quad (1.11)$$

together with de Broglie's plane wave ansatz

$$\psi = A e^{i(\vec{k} \cdot \vec{r} - \omega t)}, \quad E = \hbar\omega, \quad \vec{p} = \hbar\vec{k} \quad (1.12)$$

leads to the non-relativistic Schrödinger equation

$$i\hbar \frac{\partial}{\partial t} \psi = -\frac{\hbar^2}{2m} \nabla^2 \psi. \quad (1.13)$$

For a relativistic particle with $E = \hbar\omega = cp^0$ energy and momentum are components of a four-vector and our starting point is the equation for the square of the 4-momentum

$$E^2 = c^2 \vec{p}^2 + m^2 c^4, \quad (1.14)$$

which leads to

$$-\hbar^2 \frac{\partial^2}{\partial t^2} \phi = -c^2 \hbar^2 \nabla^2 \phi + m^2 c^4 \phi$$

or

$$\left(-\frac{\partial^2}{\partial (ct)^2} + \nabla^2 - \frac{m^2 c^2}{\hbar^2} \right) \phi = 0. \quad (1.15)$$

This is the Klein-Gordon equation, invented by Schrödinger, Fock and others, and rediscovered by Klein and Gordon.

Relativistic notations

$$x^0 = ct, \quad x^1 = x, \quad x^2 = y, \quad x^3 = z$$

$$x = (x^0, x^1, x^2, x^3) = (x^0, \vec{x}) = (x^\mu)$$

$$g_{\mu\nu} = \begin{pmatrix} 1 & & & \\ & -1 & & \\ & & -1 & \\ & & & -1 \end{pmatrix}$$

$$x \cdot y = x^0 y^0 - \vec{x} \cdot \vec{y} = x_\mu x^\mu, \quad x_\mu = g_{\mu\nu} x^\nu$$

$$\partial_\mu = \frac{\partial}{\partial x^\mu} = \left(\frac{1}{c} \frac{\partial}{\partial t}, \nabla \right)$$

$$\partial^\mu = \frac{\partial}{\partial x_\mu} = \left(\frac{1}{c} \frac{\partial}{\partial t}, -\nabla \right)$$

$$\square = -\partial_\mu \partial^\mu = -\frac{\partial^2}{\partial (ct)^2} + \Delta$$

$$p^\mu = \left(\frac{E}{c}, \vec{p} \right)$$

$$p^2 = p_\mu p^\mu = \frac{E^2}{c^2} - \vec{p}^2 = m^2 c^2$$

$$p^\mu \rightarrow i\hbar \partial^\mu$$

$$\left(\square - \frac{m^2 c^2}{\hbar^2} \right) \phi(x) = 0$$

Note³

de Broglie

Klein-Gordon

From now on we use natural units setting $\hbar = c = 1$.⁴

³The symbol p^2 is ambiguous. Its meaning must be determined from the context.

⁴To go back to SI-units in an equation one may analyse the dimension of the terms and insert $\hbar$ and/or c to get the right dimension.

Solution of the Klein-Gordon equation

Let $\phi(x)$ be a complex scalar field ($\phi \in \mathbb{C}$), that means, it is not quantised yet (ϕ is not operator-valued), Spin = 0 (ϕ is scalar). It will turn out that complex scalar fields describe particles with positive and negative charges. Examples are the mesons π^+ and π^- .

A general solution to the Klein-Gordon equation for free particles, being linear and of second order, is a superposition of two plane waves

$$\phi(x) = \int \frac{d^3k}{(2\pi)^3 2\omega_k} \left\{ a(k) e^{-ikx} + b^*(k) e^{ikx} \right\}. \quad (1.16)$$

Here we denoted

$$\omega_k = k^0 = \sqrt{\vec{k}^2 + m^2} \quad (1.17)$$

to be a **positive** frequency. The solution is verified by

$$\begin{aligned} \partial_\mu \partial^\mu e^{ik_\mu x^\mu} &= (\partial_0^2 - \partial_j \partial_j) e^{i(k^0 x^0 - k^j x^j)} \\ &= i^2 (k^0 k^0 - k^j k^j) e^{ik_\mu x^\mu} \\ &= -k_\mu k^\mu e^{ik_\mu x^\mu} \\ (\square - m^2) e^{ik_\mu x^\mu} &= (k_\mu k^\mu - m^2) e^{ik_\mu x^\mu} \\ &= \left\{ (k^0)^2 - (\vec{k}^2 + m^2) \right\} e^{ik_\mu x^\mu} = 0. \end{aligned}$$

Now let $\phi(x)$ be a *real* scalar field, $\phi \in \mathbb{R}$, being used for neutral spin 0 particles, like π^0 . Then the general solution is

$$\phi(x) = \int \frac{d^3k}{(2\pi)^3 2\omega_k} \left\{ a(k) e^{-ikx} + a^*(k) e^{ikx} \right\}. \quad (1.18)$$

The problem with negative frequencies

$$e^{i(\vec{k} \cdot \vec{x} - \omega t)} \Rightarrow E\phi = i\hbar \partial_t \phi = +\hbar\omega\phi$$

$$e^{-i(\vec{k} \cdot \vec{x} - \omega t)} \Rightarrow E\phi = i\hbar \partial_t \phi = -\hbar\omega\phi$$

So, free particles could have arbitrarily large negative energies, which is unphysical. In the presence of interactions, e.g. with the electromagnetic field, this would lead to instabilities, because a particle would jump to lower and lower states, emitting an unbounded amount of energy. This problem will be solved by field quantisation.

1.2.2 Dirac equation

The Dirac equation was found by P. A. M. Dirac in 1928. He was searching for a covariant version of the Schrödinger equation

$$i\partial_t\psi = H\psi. \quad (1.19)$$

To be manifestly covariant, it has to be of first order in the spatial derivatives, too.

$$H \text{ linear in } \frac{\partial}{\partial x^k} \quad (k = 1, 2, 3),$$

$$H = \sum_{k=1}^3 \alpha_k P^k + \beta m = - \sum_{k=1}^3 \alpha_k i\partial_k + \beta m. \quad (1.20)$$

We will now derive conditions for the constant terms α_k and β . Squaring both sides of the equation we get

$$\begin{aligned} -(\partial^0)^2\psi &= H^2\psi \\ &= \frac{1}{2} \sum_{j,k=1}^3 (\alpha_j\alpha_k + \alpha_k\alpha_j) P_j P_k \psi \\ &\quad - m \sum_{k=1}^3 (\alpha_k\beta + \beta\alpha_k) P_k \psi \\ &\quad + \beta^2 m^2 \psi. \end{aligned} \quad (1.21)$$

Consider a plane wave $\psi = e^{ipx}$, for which one should have

$$i\partial^0\psi = E\psi, \quad \vec{P}\psi = \vec{p}\psi, \quad E^2 = \vec{p}^2 + m^2. \quad (1.22)$$

This wave satisfies the Dirac equation only if

$$\alpha_j\alpha_k + \alpha_k\alpha_j = 2\delta_{jk} \mathbb{1} \quad (1.23)$$

$$\alpha_k\beta + \beta\alpha_k = 0 \quad (1.24)$$

$$\beta^2 = \mathbb{1}. \quad (1.25)$$

From this one concludes that α_k and β cannot be numbers. The relations can be satisfied by matrices, which must at least be of size 4 by 4. They can be composed by blocks of Pauli spin matrices

$$\alpha_k = \begin{pmatrix} 0 & \sigma_k \\ \sigma_k & 0 \end{pmatrix}, \quad \beta = \begin{pmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{pmatrix}. \quad (1.26)$$

By convention one uses the Dirac matrices γ^μ :

$$\gamma^0 := \beta, \quad \gamma^k := \beta \alpha_k, \quad (k = 1, 2, 3) \quad (1.27)$$

The Hamiltonian can be written

$$H = -\gamma^0 \sum_{k=1}^3 \gamma^k i \partial_k + \gamma^0 m,$$

and multiplying with γ^0 we get

$$\gamma^0 i \partial_0 \psi = \gamma^0 H \psi = - \sum_{k=1}^3 i \gamma^k \partial_k \psi + m \psi.$$

This is the Dirac equation, reading

$$\boxed{(i\gamma^\mu \partial_\mu - m)\psi(x) = 0.} \quad (1.28)$$

Equivalent notations of Dirac's equation are

$$(\gamma^\mu P_\mu - m)\psi(x) = (\not{p} - m)\psi(x) = 0.$$

The algebra of the γ -symbols is

$$\gamma^\mu \gamma^\nu + \gamma^\nu \gamma^\mu = 2g^{\mu\nu} \mathbf{1}. \quad (1.29)$$

The matrices given above are Dirac's representation of the γ 's. There are others representations, e.g. by Weyl or by Majorana. The Dirac matrices written in blocks of Pauli spin matrices are

$$\gamma^0 = \begin{pmatrix} \mathbf{1} & 0 \\ 0 & -\mathbf{1} \end{pmatrix}, \quad \gamma^k = \begin{pmatrix} 0 & \sigma_k \\ -\sigma_k & 0 \end{pmatrix}. \quad (1.30)$$

Solutions of Dirac's equation will be given by spinor wavefunctions or fields

$$\psi(x) = \begin{pmatrix} \psi_1(x) \\ \vdots \\ \psi_4(x) \end{pmatrix}. \quad (1.31)$$

These are made out of two kinds of plane waves, given by

$$\psi(x) = u(k) e^{-ikx}, \quad k^0 = \omega_k > 0 \quad (1.32)$$

with 2 independent spinors $u^{(r)}(k)$, $r = 1, 2$, and

$$\psi(x) = v(k) e^{ikx}, \quad k^0 = \omega_k > 0 \quad (1.33)$$

with another 2 independent spinors $v^{(r)}(k)$, $r = 1, 2$. The general solution is a superposition

$$\psi(x) = \int \frac{d^3k}{(2\pi)^3 2\omega_k} \sum_{r=1}^2 \left\{ b_r(\vec{k}) u^{(r)}(\vec{k}) e^{-ikx} + d_r^*(\vec{k}) v^{(r)}(\vec{k}) e^{ikx} \right\}. \quad (1.34)$$

Spin

The Dirac Hamiltonian H and the orbital angular momentum operator $\vec{L} = \vec{R} \times \vec{P}$ do not commute

$$[\vec{L}, H] \neq 0. \quad (1.35)$$

The angular momentum of free particles should be conserved! So there must be an additional hidden contribution to the angular momentum, which is called spin.

$$\vec{S} = \frac{\hbar}{2} \vec{\Sigma} = \frac{\hbar}{2} \begin{pmatrix} \vec{\sigma} & 0 \\ 0 & \vec{\sigma} \end{pmatrix}. \quad (1.36)$$

The algebra of $\vec{S}$ is given by

$$[S_k, S_l] = i S_m \quad (k, l, m) = (1, 2, 3) + \text{cycl.} \quad (1.37)$$

$$\vec{S}^2 = \frac{3}{4} \mathbb{1} = s(s+1) \mathbb{1} \quad \Rightarrow \quad s = \frac{1}{2}. \quad (1.38)$$

The total angular momentum $\vec{J} = \vec{L} + \vec{S}$ obeys

$$[\vec{J}, H] = 0. \quad (1.39)$$

Thus the total angular momentum is conserved.

Notice: The last commutator can be verified with the help of

$$\Sigma_1 = \frac{1}{2} [\gamma_2, \gamma_3]. \quad (1.40)$$

Covariant expressions

To write down Lorentz covariant expressions with ψ

$$\psi(x) = \begin{pmatrix} \psi_1(x) \\ \vdots \\ \psi_4(x) \end{pmatrix}, \quad \psi^\dagger = (\psi_1^*(x), \dots, \psi_4^*(x)), \quad (1.41)$$

we define the Dirac conjugate

$$\bar{\psi}(x) = \psi^\dagger(x) \gamma^0. \quad (1.42)$$

Covariant scalar and vector expressions are

$$\bar{\psi}(x)\psi(x), \quad \bar{\psi}(x)\gamma^\mu\psi(x); \quad (1.43)$$

Objects which transform under an antisymmetric tensor representation of the Lorentz group are

$$\bar{\psi}\sigma_{\mu\nu}\psi, \quad \sigma_{\mu\nu} = [\gamma_\mu, \gamma_\nu]. \quad (1.44)$$

With

$$\gamma_5 = \gamma^5 := i\gamma_1\gamma_2\gamma_3\gamma_4 = \begin{pmatrix} 0 & \mathbf{1} \\ \mathbf{1} & 0 \end{pmatrix} \quad (1.45)$$

pseudoscalars and pseudovectors are given by

$$\bar{\psi}\gamma_5\psi, \quad \bar{\psi}\gamma_5\gamma^\mu\psi. \quad (1.46)$$

1.2.3 Maxwell's equations

Maxwell's equations in the MKSA system read

$$\nabla \cdot \vec{E} = \frac{\rho}{\epsilon_0} \quad (1.47)$$

$$\nabla \cdot \vec{B} = 0 \quad (1.48)$$

$$\nabla \times \vec{E} = -\frac{\partial \vec{B}}{\partial t} \quad (1.49)$$

$$\nabla \times \vec{B} = \mu_0 \vec{j} + \mu_0 \epsilon_0 \frac{\partial \vec{E}}{\partial t} \quad (1.50)$$

In QFT often the Heaviside-Lorentz unit system is used. Conversion formulae are:

$$\vec{E}_H = \sqrt{\epsilon_0} \vec{E} \quad (1.51)$$

$$\vec{B}_H = \frac{1}{\mu_0} \vec{B} \quad (1.52)$$

$$\Phi_H = \sqrt{\epsilon_0} \Phi \quad (1.53)$$

$$\vec{A}_H = \frac{1}{\sqrt{\mu_0}} \vec{A} \quad (1.54)$$

$$\rho_H = \frac{1}{\sqrt{\epsilon_0}} \rho \quad (1.55)$$

$$\vec{j}_H = \frac{1}{\sqrt{\epsilon_0}} \vec{j}. \quad (1.56)$$

Now the Maxwell equations in Heaviside-Lorentz units read

$$\nabla \cdot \vec{E} = \rho \quad (1.57)$$

$$\nabla \cdot \vec{B} = 0 \quad (1.58)$$

$$\nabla \times \vec{E} + \frac{1}{c} \frac{\partial \vec{B}}{\partial t} = 0 \quad (1.59)$$

$$\nabla \times \vec{B} - \frac{1}{c} \frac{\partial \vec{E}}{\partial t} = \vec{j} \quad (1.60)$$

The fields can be derived from potentials

$$\vec{B} = \nabla \times \vec{A}, \quad \vec{E} = -\nabla \Phi - \frac{1}{c} \frac{\partial \vec{A}}{\partial t} \quad (1.61)$$

Equivalently the potentials are written in covariant form

$$A^\mu(x) := (\Phi(x), \vec{A}(x)). \quad (1.62)$$

From these the field strengths are derived by

$$F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu, \quad (\mu, \nu = 0, 1, 2, 3) \quad (1.63)$$

$$F^{\mu\nu} = \begin{pmatrix} 0 & -E_x & -E_y & -E_z \\ E_x & 0 & -B_z & B_y \\ E_y & B_z & 0 & -B_x \\ E_z & -B_y & B_x & 0 \end{pmatrix} \quad (1.64)$$

or

$$E_i = F^{i0}, \quad B_i = -\frac{1}{2} \epsilon_{ijk} F^{jk}. \quad (i, j, k \in \{1, 2, 3\}) \quad (1.65)$$

For the the 4-vector current density

$$j^\mu := (\rho, \vec{j}) \quad (1.66)$$

Maxwell's equations give the continuity equation

$$\partial_\mu j^\mu = 0 \quad (1.67)$$

or

$$\frac{\partial}{\partial t} \rho + \nabla \cdot \vec{j} = 0. \quad (1.68)$$

Maxwell's equations themselves may be written in covariant form also:
inhomogeneous equations

$$\partial_\mu F^{\mu\nu} = j^\nu \quad (1.69)$$

and homogeneous equations

$$\partial^\mu F^{\nu\rho} + \partial^\nu F^{\rho\mu} + \partial^\rho F^{\mu\nu} = 0. \quad (1.70)$$

Gauge freedom, Lorenz gauge

(Ludvig Lorenz, 1867; George F. FitzGerald, 1888)

$$\partial_\mu A^\mu = 0 \quad (1.71)$$

If one fixes the Lorenz gauge in one inertial frame, then it is fulfilled in all inertial frames. Let us consider again free fields,

$$j^\mu = 0, \quad (1.72)$$

$$\partial_\mu F^{\mu\nu} = 0. \quad (1.73)$$

Together with the Lorenz gauge we get

$$\begin{aligned} 0 = \partial_\mu F^{\mu\nu} &= \partial_\mu (\partial^\mu A^\nu - \partial^\nu A^\mu) = \partial_\mu \partial^\mu A^\nu - \partial^\nu \partial_\mu A^\mu = \partial_\mu \partial^\mu A^\nu, \\ \square A^\nu &= 0. \end{aligned} \quad (1.74)$$

To solve the wave equation we take the plane wave ansatz

$$A_\mu(x) = \epsilon_\mu^{(\lambda)} e^{ikx}. \quad (1.75)$$

From Lorenz gauge it follows

$$k \cdot k = 0, \quad k^0 = |\vec{k}| = \omega_k. \quad (1.76)$$

There are remaining superfluous degrees of freedom. The **Coulomb gauge** for a field free of sources fixes

$$\Phi = 0, \quad \nabla \cdot \vec{A} = 0. \quad (1.77)$$

For the plane wave solutions this implies

$$\epsilon_0 = 0, \quad \epsilon \cdot k = 0. \quad (1.78)$$

Thus there are 2 linearly independent solutions, representing the 2 transversal polarisations of radiation

$$\epsilon_\mu^{(1)}(k), \epsilon_\mu^{(2)}(k) \perp [(1, 0, 0, 0), k]. \quad (1.79)$$

The two transversal polarisations imply that the photon spin ($s = 1$) is in the direction of propagation.

For a massive particle moving in a certain direction and having its spin parallel to its velocity, a different inertial frame can be chosen such that in this frame the particle moves in the opposite direction and its spin is antiparallel to the velocity. Therefore the projection of its spin on the velocity is not invariant under Lorentz transformation. On the other hand, for massless particles travelling with the velocity of light, the projection of the spin on the velocity is Lorentz-invariant and is called “helicity”:

$$J_S = \pm 1. \quad (1.80)$$

The general solution of the electromagnetic wave equation in the Coulomb gauge is

$$A_\mu(x) = \int \frac{d^3k}{(2\pi)^3 2\omega_k} \sum_{\lambda=1}^2 \epsilon_\mu^{(\lambda)}(k) \left(a^{(\lambda)}(k) e^{-ikx} + a^{(\lambda)*}(k) e^{ikx} \right). \quad (1.81)$$

1.2.4 Lagrangian formalism for fields

Recapitulation: Classical mechanics

$$m\ddot{\vec{r}} = -\nabla V(\vec{r}), \quad (1.82)$$

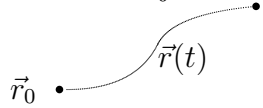
$$H = \frac{p^2}{2m} + V(\vec{r}) \quad \text{with} \quad \vec{p} = m\dot{\vec{r}}. \quad (1.83)$$

Hamilton’s equations give the equation of motion. Hamilton’s principle uses the action S , build from the Lagrangian L :

$$S = \int dt L(\vec{r}(t), \dot{\vec{r}}(t)) \quad (1.84)$$

$$L = \frac{m}{2} \dot{\vec{r}}^2 - V(\vec{r}). \quad (1.85)$$

The realised trajectories $\vec{r}(t)$



are such that the action S is stationary under infinitesimal variations $\delta\vec{r}(t)$ provided the endpoints $\vec{r}(t_0)$ and $\vec{r}(t_1)$ are fixed:

$$\vec{r}'(t) = \vec{r}(t) + \delta\vec{r}(t) \quad (1.86)$$

$$\delta\vec{r}(t_0) = \delta\vec{r}(t_1) = 0. \quad (1.87)$$

$$\boxed{\delta S = 0} \quad (1.88)$$

The calculus of variation leads to the equations of motion:

$$\delta S = \int_{t_0}^{t_1} dt \delta L = \int_{t_0}^{t_1} dt \sum_i \left(\frac{\partial L}{\partial x_i} \delta x_i + \frac{\partial L}{\partial \dot{x}_i} \delta \dot{x}_i \right) \quad (1.89)$$

$$= \int_{t_0}^{t_1} dt \sum_i \left(\frac{\partial L}{\partial x_i} - \frac{d}{dt} \frac{\partial L}{\partial \dot{x}_i} \right) \delta x_i = 0. \quad (1.90)$$

The fundamental lemma of the calculus of variation then yields the Euler-Lagrange equations

$$\frac{\partial L}{\partial x_i} - \frac{d}{dt} \frac{\partial L}{\partial \dot{x}_i} = 0. \quad (1.91)$$

This procedure can be taken over to field theory. An advantage is that symmetries in the action S directly show up as symmetries in the field equations. The Lagrangian density $\mathcal{L}$ in the field variables ϕ_a and their derivatives $\partial_\mu \phi_a$ shall be denoted by

$$\mathcal{L}(\phi_a(x), \partial_\mu \phi_a(x)); \quad (1.92)$$

here $\phi_a(x)$ and $\partial_\mu \phi_a(x)$ take over the rôle of infinitely many coordinates x_i and velocities $\dot{x}_i$, while the argument $x = (x^\mu)$ of $\phi_a(x)$ takes over the rôle of time t in mechanics. The action is

$$S = \int_G d^4x \mathcal{L}(\phi_a(x), \partial_\mu \phi_a(x)). \quad (1.93)$$

Consider now small variations of $\phi_a(x)$ fixed at the boundary ∂G of the domain of integration G . Hamilton's principle $\delta S = 0$ leads to

$$0 = \int_G d^4x \left(\frac{\partial \mathcal{L}}{\partial \phi_a(x)} \delta \phi_a(x) + \frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi_a(x))} \delta \partial_\mu \phi_a(x) \right) \quad (1.94)$$

$$= \int_G d^4x \left(\frac{\partial \mathcal{L}}{\partial \phi_a(x)} - \partial_\mu \frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi_a(x))} \right) \delta \phi_a(x). \quad (1.95)$$

Here we performed a partial integration and used the fact that the integrated part vanishes due to $\delta \phi_a = 0$ on the boundary ∂G . To see this in detail, let

$$B^\mu := \frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi_a(x))},$$

$$\partial_\mu (B^\mu \delta \phi_a) = (\partial_\mu B^\mu) \delta \phi_a + B^\mu \partial_\mu \delta \phi_a,$$

from Leibniz's product rule. From Stokes theorem we get

$$\int_G d^4x \partial_\mu (B^\mu \delta \phi_a) = \int_{\partial G} dx_\mu B^\mu \delta \phi_a = 0,$$

since $\delta\phi_a = 0$ on the boundary ∂G . Thus we can replace $B^\mu\partial_\mu\delta\phi_a$ by $-(\partial_\mu B^\mu)\delta\phi_a$ in the integral.

We end up with the Lagrange field equations

$$\frac{\partial\mathcal{L}}{\partial\phi_a(x)} - \partial_\mu \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a(x))} = 0. \quad (1.96)$$

Reasons for using Lagrangian densities:

- a) There is a single function $\mathcal{L}$ instead of many field equations.
- b) There are advances when non-Cartesian coordinates are used – similar as in mechanics.
- c) Symmetries can be expressed in a simple manner. Noether theorems lead to conservation laws.
- d) Gauge theories can be quantised in a simpler way.

Real scalar field

$$\mathcal{L} = \frac{1}{2}\partial_\mu\phi\partial^\mu\phi - \frac{m^2}{2}\phi^2 \quad (1.97)$$

The field equations are linear equations, therefore the Lagrangian has to be quadratic in the field and its derivatives. Here we have the simplest expression for $\mathcal{L}$ being quadratic and Lorentz invariant. We derive the Lagrange equations of motion:

$$\begin{aligned} \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)} &= \frac{1}{2} \left(\partial^\mu\phi + \partial_\lambda\phi \frac{\partial}{\partial(\partial_\mu\phi)} g^{\lambda\nu} \partial_\nu\phi \right) \\ &= \frac{1}{2} \left(\partial^\mu\phi + g^{\lambda\nu} \partial_\lambda\phi \frac{\partial(\partial_\nu\phi)}{\partial(\partial_\mu\phi)} \right) \\ &= \frac{1}{2} (\partial^\mu\phi + \partial^\nu\phi \delta_{\nu\mu}) \\ &= \partial^\mu\phi \end{aligned} \quad (1.98)$$

$$\partial_\mu \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)} = \partial_\mu\partial^\mu\phi = -\square\phi \quad (1.99)$$

$$\frac{\partial\mathcal{L}}{\partial\phi} = -m^2\phi \quad (1.100)$$

This gives the Klein-Gordon-equation

$$\boxed{(\partial_\mu\partial^\mu + m^2)\phi = 0} \quad (1.101)$$

Complex scalar field

$$\mathcal{L} = \partial_\mu \phi^* \partial^\mu \phi - m^2 \phi^* \phi \quad (1.102)$$

ϕ^* and ϕ are not totally independent complex-valued fields – there are not 4 independent real-valued fields. The derivatives $\partial\phi$ and $\partial\phi^*$ also do not give further freedom, as they are connected by Cauchy-Riemann differential equations.⁵

We separate $\mathcal{L}$ into independent parts by means of $\phi = \frac{1}{\sqrt{2}}(\phi_1 + i\phi_2)$, $\phi^* = \frac{1}{\sqrt{2}}(\phi_1 - i\phi_2)$:

$$\begin{aligned} \mathcal{L} &= \frac{1}{2} \partial_\mu \phi_1 \partial^\mu \phi_1 + \frac{1}{2} \partial_\mu \phi_2 \partial^\mu \phi_2 \\ &\quad + \frac{i}{2} \partial_\mu \phi_1 \partial^\mu \phi_2 - \frac{i}{2} \partial_\mu \phi_2 \partial^\mu \phi_1 \\ &\quad - \frac{1}{2} m^2 \phi_1^2 - \frac{1}{2} m^2 \phi_2^2 \\ \mathcal{L} &= \frac{1}{2} (\partial_\mu \phi_1 \partial^\mu \phi_1 - m^2 \phi_1^2) + \frac{1}{2} (\partial_\mu \phi_2 \partial^\mu \phi_2 - m^2 \phi_2^2). \end{aligned} \quad (1.103)$$

This is a Lagrangian density for two real scalar fields ϕ_1 and ϕ_2 , each giving a Klein-Gordon-equation for the real and imaginary parts of ϕ .

$$(\partial_\mu \partial^\mu + m^2) \phi_a = 0 \quad (a = 1, 2). \quad (1.104)$$

Sum and difference of both equations give two identical equations:

$$(\partial_\mu \partial^\mu + m^2) \phi = 0 \quad (1.105)$$

$$(\partial_\mu \partial^\mu + m^2) \phi^* = 0 \quad (1.106)$$

Dirac field

$$\psi(x) = (\psi_1(x), \dots, \psi_4(x))^\top \quad (1.107)$$

ψ is not a 4-vector like x or A^μ but a spinor with 4 complex-valued components. Thus ψ describes 8 real fields.

$$\mathcal{L}(\psi, \partial\psi) = \bar{\psi}(x)(i\gamma^\mu \partial_\mu - m)\psi(x), \quad (1.108)$$

⁵The Cauchy-Riemann differential equations with $z = x + iy$, $\phi(z) = \phi_1(z) + i\phi_2(z)$ read $\frac{\partial\phi_1}{\partial x} = \frac{\partial\phi_2}{\partial y} \wedge \frac{\partial\phi_1}{\partial y} = -\frac{\partial\phi_2}{\partial x}$. – Furthermore a term like $\frac{\partial\phi^*}{\partial\phi}$ must not be seen as a derivative in the complex plane, for that would not exist. Instead the *partial* derivative $\frac{\partial}{\partial\phi}$ should be considered as a vector $\frac{\partial}{\partial\phi} = (\frac{\partial}{\partial\phi_1}, \frac{\partial}{\partial\phi_2})$.

where

$$\bar{\psi} := \psi^\dagger \gamma^0 = (\psi_1^*, \psi_2^*, -\psi_3^*, -\psi_4^*). \quad (1.109)$$

In this form of $\mathcal{L}$ the derivatives act on ψ and not on $\bar{\psi}$. A more symmetric alternative would be

$$\mathcal{L} = \frac{1}{2} \left\{ \bar{\psi} \left(i\gamma^\mu \overrightarrow{\partial}_\mu - m \right) \psi + \bar{\psi} \left(-i\gamma^\mu \overleftarrow{\partial}_\mu - m \right) \psi \right\}, \quad (1.110)$$

where the arrows indicate whether the derivative acts to the right on ψ or to the left on $\bar{\psi}$. The two versions for $\mathcal{L}$ differ by a total derivative $\frac{1}{2}\partial_\mu(\bar{\psi}i\gamma^\mu\psi)$, which does not change the field equations. The resulting field equations are the Dirac equation and its Dirac-conjugate equation

$$(i\gamma^\mu \partial_\mu - m)\psi = 0, \quad (1.111)$$

$$i\partial_\mu \bar{\psi} \gamma^\mu + m\bar{\psi} = 0. \quad (1.112)$$

Taking the first Lagrangian, Eq. (1.108), we have

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial \psi} &= -m\bar{\psi}, & \partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu \psi)} &= \partial_\mu \bar{\psi} i\gamma^\mu, \\ \frac{\partial \mathcal{L}}{\partial \bar{\psi}} &= (i\gamma^\mu \partial_\mu - m)\psi(x), & \partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu \bar{\psi})} &= 0. \end{aligned}$$

This leads to the given Lagrangian equations (1.111) and (1.112).

Maxwell field

The expression of the field strengths in terms of the vector potential $A_\mu(x)$,

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu, \quad (1.113)$$

implies that the homogeneous Maxwell equations

$$\partial^\mu F^{\nu\rho} + \partial^\nu F^{\rho\mu} + \partial^\rho F^{\mu\nu} = 0 \quad (1.114)$$

are automatically fulfilled. The inhomogeneous equations

$$\partial_\mu F^{\mu\nu} = j^\nu \quad (1.115)$$

should be derived from the Lagrangian. Let us look for a Lorentz invariant and gauge invariant Lagrangian. The gauge transformations of electrodynamics, which leave the field strengths invariant,

$$\vec{A}' = \vec{A} + \nabla\chi \quad \text{and} \quad \Phi' = \Phi - \frac{1}{c}\dot{\chi},$$

read covariantly

$$A'^\mu = A^\mu - \partial^\mu \chi.$$

The Lagrangian cannot contain a mass term $m^2 A_\mu A^\mu$ since this is not gauge invariant, $A'_\mu A'^\mu \neq A_\mu A^\mu$. The field strengths are gauge invariant per definition, therefore the Lagrangian density

$$\mathcal{L}(A_\mu, \partial_\nu A_\mu) = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} \quad (1.116)$$

is both Lorentz invariant and gauge invariant. It is the only such choice being quadratic in the fields.

In the Lagrangian, $F^{\mu\nu}$ is to be considered as a function of $A_\mu(x)$ and $\partial_\nu A_\mu(x)$, thus

$$\mathcal{L} = -\frac{1}{4} (\partial_\mu A_\nu - \partial_\nu A_\mu) (\partial^\mu A^\nu - \partial^\nu A^\mu) \quad (1.117)$$

and ⁶

$$\frac{\partial \mathcal{L}}{\partial A_\mu} = 0, \quad \frac{\partial \mathcal{L}}{\partial (\partial_\mu A_\nu)} = F^{\nu\mu}. \quad (1.118)$$

From this the field equations are

$$\partial_\mu F^{\mu\nu} = 0. \quad (1.119)$$

The charges and currents, which enter the inhomogeneous equations, are themselves fields and should be dealt with by other field equations. On the

6

$$\begin{aligned} -\frac{\partial \mathcal{L}}{\partial (\partial_\alpha A_\beta)} &= \frac{1}{4} \frac{\partial}{\partial (\partial_\alpha A_\beta)} \{ (\partial_\mu A_\nu - \partial_\nu A_\mu) g^{\mu\lambda} g^{\nu\rho} (\partial_\lambda A_\rho - \partial_\rho A_\lambda) \} \\ &= \frac{1}{4} g^{\mu\lambda} g^{\nu\rho} \left[\frac{\partial}{\partial (\partial_\alpha A_\beta)} (\partial_\mu A_\nu - \partial_\nu A_\mu) \right] (\partial_\lambda A_\rho - \partial_\rho A_\lambda) \\ &\quad + \frac{1}{4} g^{\mu\lambda} g^{\nu\rho} (\partial_\mu A_\nu - \partial_\nu A_\mu) \left[\frac{\partial}{\partial (\partial_\alpha A_\beta)} (\partial_\lambda A_\rho - \partial_\rho A_\lambda) \right] \\ &= \frac{1}{4} g^{\mu\lambda} g^{\nu\rho} (\delta_{\alpha\mu} \delta_{\beta\nu} - \delta_{\alpha\nu} \delta_{\beta\mu}) (\partial_\lambda A_\rho - \partial_\rho A_\lambda) \\ &\quad + \frac{1}{4} g^{\mu\lambda} g^{\nu\rho} (\partial_\mu A_\nu - \partial_\nu A_\mu) (\delta_{\alpha\lambda} \delta_{\beta\rho} - \delta_{\alpha\rho} \delta_{\beta\lambda}) \\ &= \frac{1}{4} \{ (\delta_{\alpha\mu} \delta_{\beta\nu} - \delta_{\alpha\nu} \delta_{\beta\mu}) F^{\mu\nu} + F^{\lambda\rho} (\delta_{\alpha\lambda} \delta_{\beta\rho} - \delta_{\alpha\rho} \delta_{\beta\lambda}) \} \\ &= \frac{1}{4} (F^{\alpha\beta} - F^{\beta\alpha} + F^{\alpha\beta} - F^{\beta\alpha}) \\ &= F^{\alpha\beta} \end{aligned}$$

other hand, it is possible to introduce them as external sources j^μ in the Lagrangian. The sources obey a continuity equation

$$\partial_\mu j^\mu = 0. \quad (1.120)$$

With the Lagrangian

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} - j^\mu(x)A_\mu(x) \quad (1.121)$$

one gets the inhomogeneous Maxwell equations

$$\partial_\nu F^{\mu\nu} + j^\mu = 0. \quad (1.122)$$

The Lagrangian with external currents is not gauge invariant. The action, however, does not change under a gauge transformation, since

$$\int d^4x j^\mu \partial_\mu \chi = - \int d^4x \chi \partial_\mu j^\mu = 0. \quad (1.123)$$

Lagrangian for a massive vector field

The field, denoted by

$$B_\mu(x), \quad (1.124)$$

is a Lorentz vector. In analogy to the Maxwell field we define

$$G_{\mu\nu} := \partial_\mu B_\nu - \partial_\nu B_\mu, \quad (1.125)$$

and set

$$\mathcal{L} = -\frac{1}{4}G_{\mu\nu}G^{\mu\nu} + \frac{m^2}{2}B_\mu B^\mu. \quad (1.126)$$

With

$$\frac{\partial \mathcal{L}}{\partial B_\mu} = m^2 B^\mu, \quad \frac{\partial \mathcal{L}}{\partial(\partial_\mu B_\nu)} = G^{\nu\mu} \quad (1.127)$$

we get the field equations

$$\partial_\nu G^{\mu\nu} - m^2 B^\mu = 0, \quad (1.128)$$

and with $\partial_\nu G^{\mu\nu} = \partial_\nu \partial^\mu B^\nu - \partial_\nu \partial^\nu B^\mu$ this yields

$$(\partial_\nu \partial^\nu + m^2) B^\mu = \partial^\mu \partial_\nu B^\nu. \quad (1.129)$$

These are four field equations containing the Klein-Gordon operator. Taking the 4-divergence we get

$$(\partial_\nu \partial^\nu + m^2) \partial_\mu B^\mu = \partial_\mu \partial^\mu \partial_\nu B^\nu,$$

$$\Rightarrow \quad m^2 \partial_\mu B^\mu = 0. \quad (1.130)$$

Thus the field equations can be represented equivalently by four Klein-Gordon equations augmented by an equation which looks like a Lorenz gauge:

$$\left(\partial_\nu \partial^\nu + m^2 \right) B^\mu = 0, \quad (1.131)$$

$$\partial_\mu B^\mu = 0. \quad (1.132)$$

This field describes massive spin 1 particles, like the $\rho^{\pm,0}$ and ω mesons.

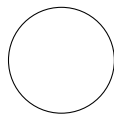
1.3 Symmetries

1.3.1 Symmetries and conservation laws

What does it mean to be symmetric? Hermann Weyl described it as follows. One needs three things:

- An object, which turns out to be symmetric.
- A procedure, doing something with this object.
- An observer, who states that after the procedure nothing has changed.

In physics we may say, a symmetry is “a mapping, which does not change the physics.” The meaning of the stated invariance depends on the structure that defines what “is the same”. For instance, in Euclidean geometry a circle is symmetric, any closed loop in general not.



circle



closed loop

On the other hand, in topology the closed loop is regarded as equivalent to the circle, the deformation of the circle to a closed loop is a symmetry transformation in topology, but not in Euclidean geometry.

Symmetry transformations can be concatenated to give a symmetry transformation again. As with any mapping an associative rule is valid. The inverse transformation is also a symmetry, so symmetries form groups. (With few exceptions, when a semigroup is also regarded as a symmetry.) The most prominent symmetry in everyday life, the mirror reflection symmetry, belongs to a very small group of only two elements: reflection and identity. In this lecture we consider continuous symmetry groups and their associated conservation laws, which Emmy Noether found in classical mechanics and field theory. Most of her work can be transferred to quantum mechanics. In quantum theory also discrete symmetry groups are being considered, for instance parity. (Here the word “parity” denotes both the symmetry and the conserved quantity.)

Lagrange formalism

In the Lagrange formalism symmetries can be dealt with by infinitesimal symmetry transformations, which is simpler than using finite transformations

although this would be possible, too. Let the fields undergo an infinitesimal transformation, which is very close to the identity.

$$\phi_a \longrightarrow \phi'_a = \phi_a + \delta\phi_a. \quad (1.133)$$

The change of the action is written

$$S \longrightarrow S' = S + \delta S. \quad (1.134)$$

If the transformation is a symmetry, the system should not change, which means that the action does not change. Thus symmetry is characterised by

$$\delta S = 0, \quad \text{if} \quad \phi_a \longrightarrow \phi_a + \delta\phi_a. \quad (1.135)$$

Here the field equations are not used. In this case one has:

if ϕ is a solution of the field equations,
then ϕ' is also a solution of the field equations.

The Noether theorem states that associated with such a symmetry there exists a conserved current $j^\mu(x)$,

$$\partial_\mu j^\mu(x) = 0. \quad (1.136)$$

The conserved quantity is given by

$$Q = \int d^3x j^0(x),$$

$$\frac{dQ}{dt} = \int_{\mathbb{R}^3} d^3x \partial_0 j^0(x) = - \int_{\mathbb{R}^3} d^3x \nabla \cdot \vec{j}(x) = - \int_{\partial\mathbb{R}^3} d^2\vec{x} \cdot \vec{j}(x) = 0.$$

Symmetries might involve space-time transformations, where $\delta x \neq 0$. For example, under translations the fields transform as

$$\phi'_a(x) = \phi_a(x - \delta x). \quad (1.137)$$

If the system is invariant under translations, the conserved quantity is the energy-momentum four-vector p^μ . Similarly, invariance under rotations gives the conservation of angular momentum. For the application of Noether's theorem to such space-time symmetries we refer to the textbooks.

Here we restrict our discussion to *internal symmetries*, for which $\delta x = 0$.

$$\phi_a(x) \longrightarrow \phi_a(x) + \delta\phi_a(x). \quad (1.138)$$

Symmetry means invariance of the action

$$S = \int d^4x \mathcal{L}(\phi_a(x), \partial\phi_a(x)). \quad (1.139)$$

We consider the stronger condition that the Lagrangian density is invariant: $\delta\mathcal{L} = 0$.

$$\begin{aligned} \delta\mathcal{L} &= \frac{\partial\mathcal{L}}{\partial\phi_a} \delta\phi_a + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)} \delta\partial_\mu\phi_a \\ &= \frac{\partial\mathcal{L}}{\partial\phi_a} \delta\phi_a - \left[\partial_\mu \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)} \right] \delta\phi_a + \partial_\mu \left[\frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)} \delta\phi_a \right] \\ &= 0. \end{aligned} \quad (1.140)$$

Here we have used $\delta\partial_\mu\phi_a = \partial_\mu\delta\phi_a$ and the Leibniz product rule for differentiation. The last term in the above equation already has the form of a 4-divergence $\partial_\mu(\)^\mu$. We write

$$\delta\phi_a = \epsilon f_a \quad (1.141)$$

with infinitesimal small ϵ and finite f_a and define

$$j^\mu(x) := \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)} f_a \quad (1.142)$$

to get

$$\partial_\mu j^\mu = \left\{ \partial_\mu \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)} - \frac{\partial\mathcal{L}}{\partial\phi_a} \right\} f_a. \quad (1.143)$$

Consequence: If the field equations hold, the current j^μ is conserved

$$\partial_\mu j^\mu = 0. \quad (1.144)$$

1.3.2 U(1) symmetry, electric charge

Let ϕ be a complex scalar ⁷ field.

$$\phi(x) \longrightarrow \phi'(x) = e^{-iq\alpha} \phi(x), \quad (q \in \mathbb{Z}, \alpha \in \mathbb{R}). \quad (1.145)$$

The transformation group has the representation

$$U(1) = \{c \in \mathbb{C} \mid c^*c = 1\} = \{c \in \mathbb{C} \mid c = e^{-i\gamma}, -\pi \leq \gamma < \pi\}. \quad (1.146)$$

⁷It could also be a Dirac field or a field for particles with higher spin.

$U(1)$ is an abelian group generated by the infinitesimal transformation

$$\begin{aligned}\phi'(x) &= (1 - iq\epsilon)\phi(x) \\ &= \phi(x) - iq\phi(x)\epsilon\end{aligned}\tag{1.147}$$

$$\delta\phi(x) = -iq\phi(x)\epsilon\tag{1.148}$$

$$\delta\phi^*(x) = iq\phi^*(x)\epsilon.\tag{1.149}$$

We assume a symmetry under $U(1)$,

$$\delta\mathcal{L} = 0,\tag{1.150}$$

as is the case with the Lagrangian for a complex scalar field,

$$\mathcal{L} = \frac{1}{2}\partial_\mu\phi^*\partial^\mu\phi - \frac{m^2}{2}\phi^*\phi.\tag{1.151}$$

The Lagrangian is in fact invariant under finite $U(1)$ -transformations. Now we have to find the Noether current.

$$\begin{aligned}j^\mu &= \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi_a)}f_a, \quad f_1 = -iq\phi, \quad f_2 = iq\phi^* \\ &= \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi)}(-iq\phi) + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\phi^*)}(iq\phi^*) \\ j^\mu &= iq(\phi^*\partial^\mu\phi - \phi\partial^\mu\phi^*).\end{aligned}\tag{1.152}$$

We note that the spatial part of this current has the same form as the probability current in Schrödinger theory, which is defined by $\vec{j} = \frac{\hbar}{2mi}(\psi\nabla\psi^* - \psi^*\nabla\psi)$. Let us check the conservation law

$$\frac{1}{iq}\partial_\mu j^\mu = \partial_\mu(\phi^*\partial^\mu\phi - \phi\partial^\mu\phi^*)\tag{1.153}$$

$$\begin{aligned}&= \partial_\mu\phi^*\partial^\mu\phi + \phi^*\partial_\mu\partial^\mu\phi - \partial_\mu\phi\partial^\mu\phi^* - \phi\partial_\mu\partial^\mu\phi^* \\ &= \phi^*\partial_\mu\partial^\mu\phi - \phi\partial_\mu\partial^\mu\phi^* \\ &= \phi^*(\partial_\mu\partial^\mu + m^2)\phi - \phi(\partial_\mu\partial^\mu + m^2)\phi^*.\end{aligned}\tag{1.154}$$

If the field equation holds, then

$$\partial_\mu j^\mu = 0.\tag{1.155}$$

The conserved charge is given by

$$Q = \int d^3x j^0 = iq \int d^3x (\phi^*\partial_0\phi - \phi\partial_0\phi^*),\tag{1.156}$$

q may be an integer number.

For the Dirac field we have a similar U(1) symmetry. The transformation is

$$\psi'(x) = e^{-iq\alpha}\psi(x), \quad (1.157)$$

$$\bar{\psi}'(x) = e^{iq\alpha}\bar{\psi}(x), \quad (1.158)$$

and the Lagrangian density

$$\mathcal{L} = \bar{\psi}(i\gamma^\mu\partial_\mu - m)\psi \quad (1.159)$$

again is invariant under this transformation. For the current and charge we get

$$j^\mu = \frac{\partial\mathcal{L}}{\partial(\partial_\mu\psi_\alpha)}(-iq\psi_\alpha) = q\bar{\psi}\gamma^\mu\psi, \quad (1.160)$$

$$Q = q \int d^3x \bar{\psi}\gamma^0\psi = q \int d^3x \psi^\dagger\psi, \quad (1.161)$$

$\psi^\dagger\psi$ is the charge density.

1.3.3 SU(2) symmetry, isospin

Consider neutron and proton, described by two Dirac fields

$$p(x) = (p_\alpha(x)), \alpha = 1, \dots, 4, \quad (1.162)$$

$$n(x) = (n_\alpha(x)), \alpha = 1, \dots, 4. \quad (1.163)$$

The nuclear forces are independent of the electric charge. They are the same for proton and neutron. The idea of isospin (Werner Heisenberg, Dmitri Ivanenko) is to describe this in terms of a symmetry. The situation is in analogy with the two spin states of the electron, which form a basis of a two-dimensional sub-Hilbert space

$$|\uparrow\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad |\downarrow\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \quad (1.164)$$

In the absence of a magnetic field both states carry the same energy, the Hamiltonian is invariant under a rotation in spin space, made up of the Pauli spinors

$$\psi(x) = \begin{pmatrix} \psi_+(x) \\ \psi_-(x) \end{pmatrix}. \quad (1.165)$$

The symmetry transformation shall be denoted by

$$\psi \longrightarrow U(\alpha)\psi \quad (1.166)$$

with angles α that parameterise the rotation.

In analogy to spin Heisenberg introduced *isotopic spin* for proton and neutron, sometimes called isobaric spin; today it is mostly called *isospin*. The nucleon is represented by

$$N = \begin{pmatrix} N_1(x) \\ N_2(x) \end{pmatrix}, \quad (1.167)$$

where N_1 and N_2 are Dirac spinors

$$N_{i,\alpha}(x), \quad i = 1, 2; \quad \alpha = 1, \dots, 4. \quad (1.168)$$

The pure proton or neutron states are (somewhat ambiguously) denoted as

$$p(x) = \begin{pmatrix} p(x) \\ 0 \end{pmatrix}, \quad n(x) = \begin{pmatrix} 0 \\ n(x) \end{pmatrix}, \quad (1.169)$$

and the nucleon field is

$$N(x) = \begin{pmatrix} p(x) \\ n(x) \end{pmatrix}. \quad (1.170)$$

The corresponding isospin observable is

$$\vec{I} = \frac{1}{2} \vec{\tau}. \quad (1.171)$$

$$\begin{aligned} I_1 &= \frac{1}{2} \tau_1, & \tau_1 &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \\ I_2 &= \frac{1}{2} \tau_2, & \tau_2 &= \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \\ I_3 &= \frac{1}{2} \tau_3, & \tau_3 &= \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \end{aligned}$$

Thus

$$I_3 p(x) = +\frac{1}{2} p(x), \quad I_3 n(x) = -\frac{1}{2} n(x). \quad (1.172)$$

Summary:

Isospin	Proton	Neutron
I	$\frac{1}{2}$	$\frac{1}{2}$
I_3	$+\frac{1}{2}$	$-\frac{1}{2}$

The symmetry group SU(2)

Symmetries may be approximate symmetries, but here we shall make the hypothesis that the Hamiltonian is invariant under a rotation in 2-dimensional

isospin space. Because of the form of the free Hamiltonian, the symmetry transformation must be unitary. We write

$$N \longrightarrow N' = UN. \quad (1.173)$$

The group of such transformations, $U(2)$, is made from 2×2 unitary matrices U .

$$U^\dagger U = \mathbb{1} \quad (1.174)$$

$$\det(U^\dagger U) = \det(U^\dagger) \det(U) = |\det(U)|^2 = 1 \quad (1.175)$$

$$|\det(U)| = 1 \quad (1.176)$$

The subgroup that has $\det U = +1$ is called $SU(2)$:

$$SU(2) = \{U \in GL_2(\mathbb{C}) \mid U^\dagger U = 1, \det U = +1\}. \quad (1.177)$$

$$U = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad (1.178)$$

$$U^\dagger U = \begin{pmatrix} a^* & c^* \\ b^* & d^* \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a^*a + c^*c & a^*b + c^*d \\ b^*a + d^*c & b^*b + d^*d \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

The unitarity of U implies that the vector norm of the columns has to be 1, the same holds for the rows of U since $U^\dagger$ is also unitary. Together we have

$$1 = \det U = ad - bc$$

$$1 = a^*a + c^*c = b^*b + d^*d = a^*a + b^*b = c^*c + d^*d$$

leading to $a^*a = d^*d$, $b^*b = c^*c$. From this we can restrict the general form of U to be ⁸

$$U = \begin{pmatrix} a & b \\ -b^* e^{i\beta} & a^* e^{i\alpha} \end{pmatrix}, \quad a^*a + b^*b = 1.$$

The determinant gives

$$1 = a^*a e^{i\alpha} + b^*b e^{i\beta} = e^{i\alpha}(a^*a + b^*b e^{i(\beta-\alpha)}).$$

Together with $a^*a + b^*b = 1$ this is only possible if $\alpha - \beta = 0$. Then $\alpha = 0$ follows, too, and we have

$$U = \begin{pmatrix} a & b \\ -b^* & a^* \end{pmatrix}, \quad a^*a + b^*b = 1. \quad (1.179)$$

⁸ $c = \frac{b^*b}{c^*}$, $|b| = |c|$ gives $c = -b^* e^{i\alpha}$ and $d = \frac{a^*a}{d^*}$, $|a| = |d|$ gives $d = a^* e^{i\beta}$.

Now we may choose four real parameters, together with one condition, to characterise U ⁹

$$U = \begin{pmatrix} a_0 - ia_3 & -a_2 - ia_1 \\ a_2 - ia_1 & a_0 + ia_3 \end{pmatrix}, \quad \sum_{k=0}^3 a_k^2 = 1. \quad (1.180)$$

This can be written as $U = a_0 \mathbb{1} - ia_1 \tau_1 - ia_2 \tau_2 - ia_3 \tau_3$. Since $\sum_{k=0}^3 a_k^2 = 1$, one may write

$$a_0 = \cos\left(\frac{\alpha}{2}\right), \quad 0 \leq \alpha < 2\pi, \quad (1.181)$$

$$(a_1, a_2, a_3) = \sin\left(\frac{\alpha}{2}\right) \vec{n}, \quad |\vec{n}| = 1. \quad (1.182)$$

Then

$$U = \cos\left(\frac{\alpha}{2}\right) \mathbb{1} - \sin\left(\frac{\alpha}{2}\right) (in_1 \tau_1 + in_2 \tau_2 + in_3 \tau_3). \quad (1.183)$$

With the definition

$$\vec{\alpha} := \alpha \vec{n} \quad (1.184)$$

we finally write it in a way standard for Lie group representations:

$$U = \exp\left(-i \frac{\alpha}{2} \sum_{k=1}^3 n_k \tau_k\right) = \exp(-i \vec{\alpha} \cdot \vec{I}). \quad (1.185)$$

That the two expressions for U , (1.183) and (1.185), are equal can easily be seen to first order in α :

$$\cos\left(\frac{\alpha}{2}\right) \mathbb{1} - i \sin\left(\frac{\alpha}{2}\right) (\vec{n} \cdot \vec{\tau}) \approx \mathbb{1} - i \frac{\alpha}{2} \vec{n} \cdot \vec{\tau} \approx \exp\left(-i \frac{\alpha}{2} \vec{n} \cdot \vec{\tau}\right).$$

It can be shown that they are equal in general.

$\vec{I} = \vec{\tau}/2$ is the Hermitian matrix for the isospin observable. The anti-Hermitian, traceless exponent is a common feature of Lie algebras.

Once again we state that the situation is similar to quantum mechanics, where $U(\vec{\alpha})$ in Pauli spinor space describes a rotation by an angle α with a rotation axis given by $\vec{n}$. The matrix elements in this representation of $SU(2)$ are determined by 3 real parameters $(\alpha_1, \alpha_2, \alpha_3) = \vec{\alpha}$.

⁹In general, the matrices representing the group

$$SU(N) = \{U \in GL_n \mid U^\dagger U = 1, \det U = 1\}$$

have $N^2 - 1$ real parameters.

Lie algebra

Consider an infinitesimal transformation, that means a transformation (1.185) with infinitesimal small α :

$$\delta\vec{\alpha} = \delta\alpha \vec{n}, \quad (1.186)$$

$$U(\delta\vec{\alpha}) = \mathbb{1} - i\delta\vec{\alpha} \cdot \vec{I}, \quad (1.187)$$

$$N' = N - i\delta\vec{\alpha} \cdot \vec{I}N, \quad (1.188)$$

$$\delta N = -i\delta\vec{\alpha} \cdot \vec{I}N. \quad (1.189)$$

The isospin operators (I_1, I_2, I_3) are called generators of the Lie group SU(2). From unitarity it follows that they are Hermitian and traceless,¹⁰

$$I_k^\dagger = I_k, \quad \text{tr}(I_k) = 0. \quad (1.190)$$

Their commutators are the same as those of the usual spin operators:

$$[I_k, I_l] = i\epsilon_{klm}I_m, \quad (1.191)$$

This is the Lie algebra of SU(2). ϵ_{klm} are the structure constants of the Lie algebra spanned by I_1, I_2, I_3 .

Isospin symmetry

Invariant expressions

Expressions which are invariant under isospin symmetry are useful as building blocks for an invariant Lagrangian. Here are some expressions containing $N = \begin{pmatrix} p \\ n \end{pmatrix}$.

$$N^\dagger N = p^\dagger p + n^\dagger n : \quad (UN)^\dagger UN = N^\dagger U^\dagger UN = N^\dagger N \quad (1.192)$$

$$N^\dagger \gamma^\mu N = p^\dagger \gamma^\mu p + n^\dagger \gamma^\mu n \quad (1.193)$$

$$\bar{N}N \quad (1.194)$$

$$\bar{N}\partial_\mu N^{11} \quad (1.195)$$

An invariant Lagrangian for free nucleons is

$$\begin{aligned} \mathcal{L} &= \bar{N}(x)(i\gamma^\mu \partial_\mu - m)N(x) \\ &= \bar{p}(x)(i\gamma^\mu \partial_\mu - m)p(x) + \bar{n}(x)(i\gamma^\mu \partial_\mu - m)n(x). \end{aligned} \quad (1.196)$$

¹⁰Every anti-Hermitian operator can be decomposed into a traceless part and an imaginary multiple of the identity $iH = iJ + i \text{tr}(H)\mathbb{1}$. Then $\exp(iH) = \exp(i \text{tr}(H)) \exp(iJ) \in \text{U}(1) \otimes \text{SU}(n)$ with $n = \text{dimension of the Hilbert(sub)space}$.

¹¹The isospin symmetry does not depend on space-time.

Isospin symmetry implies that $m = m_p = m_n$. Experimental values for proton and neutron masses are

$$m_p = 938.272 \text{ MeV} \quad (1.197)$$

$$m_n = 939.565 \text{ MeV}. \quad (1.198)$$

The isospin symmetry is nearly perfect, it is broken by the electromagnetic interaction and by differences between the masses of up- and down-quarks.

	p	n
I	$1/2$	$1/2$
I_3	$1/2$	$-1/2$

$$Q = \frac{1}{2} + I_3. \quad (1.199)$$

Other representations of SU(2)

Similar to higher spin quantum numbers, belonging to operators in higher dimensional spin subspaces, there are other representations of the isospin symmetry group. The group is represented by matrices in isospin space with an arbitrary dimension. The dimension of the isospin multiplets is given by $2I + 1$, where $I = 0, \frac{1}{2}, 1, \dots$ denotes the isospin quantum number.

Consider $I = 1$, $I_3 = -1, 0, 1$, giving an isospin triplet, e.g. the pion triplet

$$\pi = \begin{pmatrix} \pi^+ \\ \pi^0 \\ \pi^- \end{pmatrix}. \quad (1.200)$$

Here the quantum numbers for charge and isospin components are equal

$$Q = I_3. \quad (1.201)$$

In this $I = 1$ representation the generators of SU(2) have the following form¹²

$$I_1^{(1)} = \frac{1}{2} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad (1.202)$$

$$I_2^{(1)} = \frac{1}{2} \begin{pmatrix} 0 & -i & 0 \\ i & 0 & -i \\ 0 & i & 0 \end{pmatrix}, \quad (1.203)$$

¹²This follows in the same way as one gets the matrices for angular momentum operators: $L_{\pm} = L_x \pm iL_y$, $L_{\pm}|l, m\rangle = \sqrt{l(l+1) \mp m(m \pm 1)}|l, m \pm 1\rangle$. From this the images $L_{x,y}|l, m\rangle$ as matrix-columns are found using $L_x = \frac{1}{2}(L_+ + L_-)$, $L_y = \frac{1}{2i}(L_+ - L_-)$.

$$I_3^{(1)} = \frac{1}{2} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}. \quad (1.204)$$

Quarks

From the composition of the nucleons or the pions one can deduce the isospin quantum numbers for the quarks. This is true under the assumption that the quantum numbers of quark content add up to the total quantum number for the composed particles. From

$$\begin{aligned} \pi^+ &= u\bar{d} \\ \pi^0 &= u\bar{u} - d\bar{d} \\ \pi^- &= d\bar{u} \\ p &= uud \\ n &= dd\bar{u} \end{aligned}$$

we infer that for quarks

$$I = \frac{1}{2}, \quad I_3 u = \frac{1}{2}u, \quad I_3 d = -\frac{1}{2}d. \quad (1.205)$$

Thus the u , d quarks belong to an isospin doublet and their electric charges Q obey

$$Q = \frac{1}{6} + I_3. \quad (1.206)$$

In general, one defines a *hypercharge* Y by ¹³

$$Q = I_3 + \frac{1}{2}Y, \quad Y = 2(Q - I_3). \quad (1.207)$$

	Q	I	I_3	Y
u	2/3	1/2	1/2	1/3
d	-1/3	1/2	-1/2	1/3
p	1	1/2	1/2	1
n	0	1/2	-1/2	1
π^+	1	1	1	0
π^0	0	1	0	0
π^-	-1	1	-1	0

¹³Here Y is equal to the baryon number B . In general $Y = B + S - \frac{1}{3}C$, with S and C expressing strangeness and charm.

Noether currents

The Noether procedure leads to conservation of currents in the form $\partial_\mu j^\mu = 0$. For the SU(2) isospin symmetry the field $\phi(x)$ has components $\phi_a(x)$, one for each particle of the multiplet

$$(\phi_a) = \begin{pmatrix} u \\ d \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} p \\ n \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} \pi^+ \\ \pi^0 \\ \pi^- \end{pmatrix}. \quad (1.208)$$

The infinitesimal transformation is ¹⁴

$$\delta\phi = -i\delta\vec{\alpha} \cdot \vec{I}^{(I)} \phi. \quad (1.209)$$

The superscript (I) stands for the isospin quantum number. In components of the multiplet this reads

$$\delta\phi_a = -i\delta\alpha_k (I_k^{(I)})_{ab} \phi_b. \quad (1.210)$$

From this we find the current for each of the three isospin components

$$j_k^\mu = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_a)} \frac{\delta\phi_a}{\delta\alpha_k} = -i \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_a)} (I_k^{(I)})_{ab} \phi_b \quad (k = 1, 2, 3). \quad (1.211)$$

As an example consider a quark isospin doublet $q = \begin{pmatrix} u \\ d \end{pmatrix}$ with Lagrangian ¹⁵

$$\mathcal{L} = \bar{q}(x)(i\gamma^\mu \partial_\mu - m)q(x). \quad (1.212)$$

$$j_k^\mu(x) = \bar{q}(x)\gamma^\mu I_k^{(1/2)} q(x), \quad (1.213)$$

$$j_1^\mu = \frac{1}{2} (\bar{u}\gamma^\mu d + \bar{d}\gamma^\mu u), \quad (1.214)$$

$$j_2^\mu = -\frac{i}{2} (\bar{u}\gamma^\mu d - \bar{d}\gamma^\mu u), \quad (1.215)$$

$$j_3^\mu = \frac{1}{2} (\bar{u}\gamma^\mu u - \bar{d}\gamma^\mu d). \quad (1.216)$$

The conserved charge $\hat{I}_3$ is

$$\hat{I}_3 = \int d^3x j_3^0(x) = \frac{1}{2} \int d^3x (\bar{u}\gamma^0 u - \bar{d}\gamma^0 d). \quad (1.217)$$

¹⁴See the procedure for the charge conservation by U(1) symmetry, Eqs. (1.145) – (1.156).

¹⁵See the Lagrangian for the Dirac field in subsection 1.2.4, Eq. (1.108), page 25.

1.3.4 SU(3) flavour symmetry

The experimentally found hadrons motivated to extend the approximate isospin SU(2) symmetry to an approximate SU(3) flavour symmetry. It should be noted that the SU(3) flavour symmetry is a global symmetry, in contrast to the SU(3)_{colour} symmetry, which is a local gauge symmetry. In nature the SU(3)_{flavour} is broken by the electromagnetic interaction and through the differences between the up-, down- and strange-quark masses. This leads to considerable mass differences within the SU(3) multiplets. Because the mass differences to the heavier charm-, top- and bottom-quarks are even much bigger, one cannot speak of an approximate higher SU(N), $N \geq 4$, flavour symmetry.¹⁶ Arranging the quarks by their masses, we depict the symmetries by

$$\begin{array}{c} \text{SU(2)} \\ \underbrace{u, d, s}_{\text{SU(3)}}, c, b, t. \end{array} \quad (1.218)$$

Let the flavour symmetry transformation be

$$q \longrightarrow q' = Uq, \quad (1.219)$$

$$U \in \text{SU}(3) = \{U \in \text{GL}_3(\mathbb{C}) \mid U^\dagger U = 1, \det U = 1\}. \quad (1.220)$$

The group elements are characterised by $n^2 - 1 = 8$ independent real parameters α_k , which we write as $\vec{\alpha} = (\alpha_1, \dots, \alpha_8)$. For the representation of the group by its corresponding algebra we note:

$$\begin{array}{ll} \text{Algebra} \longrightarrow \text{Group} & \text{by exponentiation} \\ \text{Group} \longrightarrow \text{Algebra} & \text{by infinitesimal transformations} \end{array}$$

$$U(\vec{\alpha}) = \exp \left(-i \sum_{k=1}^8 \alpha_k T_k \right), \quad (1.221)$$

$$U(\delta\vec{\alpha}) = \mathbb{1} - i \sum_{k=1}^8 \delta\alpha_k T_k. \quad (1.222)$$

$$T_k^\dagger = T_k, \quad \text{tr}(T_k) = 0. \quad (1.223)$$

¹⁶Figs. 1 and 2 on pages 6 – 7 show SU(4) multiplets. There the SU(3) multiplets have been extended in a third, vertical dimension to incorporate charm quantum numbers $C = 1, 2, 3$. That gives 20-plets for baryons and 16-plets for mesons.

The structure constants of the algebra shall be denoted by f

$$[T_k, T_l] = f_{klm} T_m. \quad (1.224)$$

Analogous to the Pauli matrices σ_k , the Gell-Mann matrices are defined by $\lambda_k = \frac{1}{2} T_k$ for $k = 1, \dots, 8$. An additional property of the SU(3) generators is

$$\text{tr}(T_k T_l) = \frac{1}{2} \delta_{kl}. \quad (1.225)$$

The group SU(2) is a subgroup of SU(3), the elements of SU(2) have the representation

$$U = \left(\begin{array}{cc|c} & & 0 \\ U' & & \\ \hline 0 & 0 & 1 \end{array} \right). \quad (1.226)$$

The generators of $\{U'\}$ form a subalgebra of the algebra of SU(3). Using the isospin operators of SU(2) one can express the corresponding generators of SU(3) as

$$T_1 = \left(\begin{array}{cc|c} & & 0 \\ I_1 & & \\ \hline 0 & 0 & 0 \end{array} \right), \quad T_2 = \left(\begin{array}{cc|c} & & 0 \\ I_2 & & \\ \hline 0 & 0 & 0 \end{array} \right), \quad T_3 = \left(\begin{array}{cc|c} & & 0 \\ I_3 & & \\ \hline 0 & 0 & 0 \end{array} \right). \quad (1.227)$$

A maximum of 2 generators of SU(3) can be chosen such that they commute with each other and can thus be diagonalised simultaneously. Such a maximal abelian subalgebra is called Cartan subalgebra. Usually one chooses T_3 and T_8 . Explicitly one writes

$$T_3 = \frac{1}{2} \left(\begin{array}{cc|c} 1 & 0 & 0 \\ 0 & -1 & 0 \\ \hline 0 & 0 & 0 \end{array} \right), \quad T_8 = \frac{1}{2\sqrt{3}} \left(\begin{array}{cc|c} 1 & 0 & 0 \\ 0 & 1 & 0 \\ \hline 0 & 0 & -2 \end{array} \right). \quad (1.228)$$

We see that $[T_3, T_8] = 0$. The quantum numbers belonging to T_3 and T_8 are isospin and the additional quantum number Y , called hypercharge. It is defined by

$$T_8 = \frac{\sqrt{3}}{2} Y, \quad Y = \begin{pmatrix} 1/3 & 0 & 0 \\ 0 & 1/3 & 0 \\ 0 & 0 & -2/3 \end{pmatrix}. \quad (1.229)$$

In this 3-dimensional representation hypercharge Y can have the values $1/3$ and $-2/3$. In addition to the two up- and down-quark states, having isospin $I_3 = \pm 1/2$ and hypercharge $Y = 1/3$, one introduces a third state with isospin $I_3 = 0$ and hypercharge $Y = -2/3$, called strange quark. The quantum number strangeness S is defined to be zero for up- and down-quarks and

$S = -1$ for strange quarks. Here we list the quantum numbers which are assigned to the three light quarks.

q	Q	I	I_3	Y	S	B
u	$2/3$	$1/2$	$1/2$	$1/3$	0	$1/3$
d	$-1/3$	$1/2$	$-1/2$	$1/3$	0	$1/3$
s	$-1/3$	0	0	$-2/3$	-1	$1/3$

We see that

$$Q = I_3 + \frac{1}{2}Y, \quad Y = B + S.^{17} \quad (1.231)$$

The (partially) conserved quantities now are

Quantity	conserved by interactions	broken by interactions
B, Q	all	none
I	strong	em, weak
S	strong, em	weak

The remaining Gell-Mann matrices for SU(3) are

$$\lambda_4 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \quad \lambda_5 = \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix} \quad (1.232)$$

$$\lambda_6 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \quad \lambda_7 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix} \quad (1.233)$$

1.3.5 Some comments about symmetry

As we have seen, symmetries play an important rôle in the physics of elementary particles. The U(1) symmetry leads to charge conservation. The SU(2) isospin symmetry is a property of the strong nuclear interactions and is associated with the equality of the masses of proton and neutron or of the pion masses. This symmetry is only approximately true, but approximate symmetries give a starting point for perturbative calculations. Symmetries help to find more or less good quantum numbers and to classify particles.

¹⁷Including the charm quark the hypercharge is

$$Y = B + S - \frac{1}{3}C \quad (1.231)$$

For extensions of the Standard Model, the established symmetries must be taken into account.

There are two kinds of symmetries: firstly, there are the space-time symmetries like translations, rotations, Lorentz transformations, assembled in the inhomogeneous Lorentz group, also called Poincaré group. Secondly, we have internal symmetries: $U(1)$, $SU(2)_{\text{flavour}}$, $SU(3)_{\text{flavour}}$, the local gauge symmetry $SU(3)_{\text{colour}}$ and the chiral symmetry.

The integration of these two kinds of symmetries into a unified framework requires an extension of the concept of symmetry, namely supersymmetry (SUSY).

1.4 Field Quantisation

So far we have considered *relativistic classical fields*, e.g. the Klein-Gordon field or Maxwell field, which are consistent with special relativity. In this chapter we develop field theory in the framework of quantum theory.

In classical physics, fields are continuous systems, described by functions of space and time. Examples are elastic media, electric and magnetic force fields.

In the context of quantum theory, Schrödinger's wave function can be considered as a field. Its interpretation as a probability amplitude is, however, different from classical fields. The probability density is $\rho(x) = \psi^*(x)\psi(x)$.

The attempts to reconcile quantum mechanics and special relativity by introducing relativistic wave equations, i.e. the Klein-Gordon and Dirac-equation, led to problems:

- An infinite number of negative energy states arouse.
- When analysing scattering on quantum wells, it turned out that negative probabilities, as well as probabilities larger than 1, occurred (Klein's paradox). This is related to particle creation in strong fields.

The solution of these problems lies in the quantisation of relativistic field equations. The fields are then not considered as wave functions, but as physical systems with an infinite number of degrees of freedom, and are subject to quantisation. It turns out that this leads to quantum theory of many particles.

Quantisation in quantum mechanics and in quantum field theory

Quantisation of a classical theory implies the replacement of observables by operators. The fundamental Poisson brackets are replaced by commutators. In classical mechanics the Poisson brackets are defined by

$$\{f, g\} = \sum_i \left(\frac{\partial f}{\partial q_i} \frac{\partial g}{\partial p_i} - \frac{\partial g}{\partial q_i} \frac{\partial f}{\partial p_i} \right). \quad (1.234)$$

For fields the analogous definition is

$$\{F, G\} = \int d^3x \sum_a \left\{ \frac{\delta F}{\delta \phi_a(x^0, \vec{x})} \frac{\delta G}{\delta \pi_a(x^0, \vec{x})} - \frac{\delta G}{\delta \phi_a(x^0, \vec{x})} \frac{\delta F}{\delta \pi_a(x^0, \vec{x})} \right\}, \quad (1.235)$$

where F and G are functionals of the fields $\phi_a(x^0, \vec{x})$ and their conjugate momenta $\pi_a(x^0, \vec{x})$ at fixed time x^0 .

Using the functional derivatives

$$\frac{\delta\phi_a(x^0, \vec{x})}{\delta\phi_b(x^0, \vec{x}')} = \delta_{ab} \delta^3(\vec{x} - \vec{x}') \quad (1.236)$$

one gets the fundamental Poisson brackets for fields

$$\{\pi_a(x^0, \vec{x}), \phi_b(x^0, \vec{x}')\} = -\delta_{ab} \delta^3(\vec{x} - \vec{x}') \quad (1.237)$$

Below we summarise variables in the Hamiltonian formalism of classical physics and their quantum counterparts.

Mechanical description	→	quantum mechanical description
variables		operators
q_i, p_i		Q_i, P_i
$p_i = \frac{\partial \mathcal{L}}{\partial \dot{q}_i}$		
Poisson bracket		commutator/ $i\hbar$
$\{q_i, p_k\} = \delta_{ik}$		$[Q_i, P_k] = i\hbar \delta_{ik}$
classical field	→	quantum field
field $\phi_a(x)$		field operator $\phi_a(x)$
$\pi(x) = \frac{\partial \mathcal{L}}{\partial \dot{\phi}(x)}$		
$\{ , \}$		$\frac{1}{i\hbar} [,]$
		states $ \psi\rangle \in \mathcal{H}$

1.4.1 Quantisation of the real scalar field

With the expression of the Lagrangian density for a real scalar field

$$\mathcal{L} = \frac{1}{2} \partial_\mu \phi \partial^\mu \phi - \frac{m^2}{2} \phi^2 \quad (1.238)$$

the conjugate momentum density is

$$\pi(x) = \frac{\partial \mathcal{L}}{\partial(\partial_0 \phi(x))} = \partial^0 \phi(x) = \dot{\phi}(x). \quad (1.239)$$

The fundamental Poisson brackets for classical fields are

$$\{\pi(x^0, \vec{x}), \phi(x^0, \vec{x}')\} = -\delta^3(\vec{x} - \vec{x}') \quad (1.240)$$

$$\{\phi(x^0, \vec{x}), \phi(x^0, \vec{x}')\} = \{\pi(x^0, \vec{x}), \pi(x^0, \vec{x}')\} = 0. \quad (1.241)$$

Their counterparts in quantum field theory will be

$$\left[\pi(x^0, \vec{x}), \phi(x^0, \vec{x}')\right] = \frac{\hbar}{i} \delta^3(\vec{x} - \vec{x}'), \quad (1.242)$$

$$\left[\pi(x^0, \vec{x}), \pi(x^0, \vec{x}')\right] = \left[\phi(x^0, \vec{x}), \phi(x^0, \vec{x}')\right] = 0. \quad (1.243)$$

A Legendre transformation gives the Hamiltonian density

$$\mathcal{H} = \pi \dot{\phi} - \mathcal{L} \quad (1.244)$$

$$= \frac{1}{2} \pi^2 + \frac{1}{2} (\nabla \phi)^2 + \frac{m^2}{2} \phi^2, \quad (1.245)$$

and the Hamiltonian

$$H = \frac{1}{2} \int d^3x \left(\pi^2 + (\nabla \phi)^2 + m^2 \phi^2 \right). \quad (1.246)$$

The classical canonical equations of motion then read

$$\dot{\phi}(x) = \{\phi(x), H\} = \pi(x), \quad (1.247)$$

$$\dot{\pi}(x) = \{\pi(x), H\} = \Delta \phi(x) - m^2 \phi(x) \quad (1.248)$$

Combining them yields

$$\ddot{\phi}(x) = (\Delta - m^2) \phi(x), \quad (1.249)$$

which is the Klein-Gordon equation, $(\partial_\mu \partial^\mu + m^2) \phi = 0$. As we see, the resulting field equation is Lorentz invariant, although we introduced a distinction of space and time with the definition of the canonical momentum π , and by using only equal time Poisson brackets or – after quantisation – equal time commutators.¹⁸

The corresponding quantum canonical equations read

$$\dot{\phi}(x) = \frac{1}{i\hbar} [\phi(x), H] = \pi(x), \quad (1.250)$$

$$\dot{\pi}(x) = \frac{1}{i\hbar} [\pi(x), H] = \Delta \phi(x) - m^2 \phi(x). \quad (1.251)$$

¹⁸This approach to field quantisation was introduced by Heisenberg and Pauli. They proceeded from classical mechanics to field theory by dividing space into small cells, to each of which was associated a pair of conjugate generalised coordinates (q_i, p_i) . These undergo time evolution like the many conjugate variables of a many particle mechanical systems. Going to the limit of a continuum of infinitely small cells one arrives at the time evolution of fields.

The quantised scalar field obeys the Klein-Gordon equation. Therefore, like in the classical case it can be decomposed into plane wave solutions

$$\phi(x) = \int \frac{d^3k}{(2\pi)^3 2\omega_k} \left(a(\vec{k}) e^{-ik \cdot x} + a^\dagger(\vec{k}) e^{ik \cdot x} \right), \quad (1.252)$$

where

$$k_0 = \omega_k = \sqrt{\vec{k}^2 + m^2}. \quad (1.253)$$

In contrast to the classical case, here the coefficients $a(\vec{k}), a^\dagger(\vec{k})$ are operators. In the following we just denote them $a(k), a^\dagger(k)$, where it is to be understood that $k_0 = \omega_k$. From the definition of the canonical conjugate momentum, Eq. (1.239), we have

$$\pi(x) = \int \frac{d^3k}{(2\pi)^3 2\omega_k} (-i\omega_k) \left(a(k) e^{-ik \cdot x} - a^\dagger(k) e^{ik \cdot x} \right). \quad (1.254)$$

We can invert the mode expansions of the field to get an expression for $a(k)$:

$$\begin{aligned} \pi(x) - i\omega_k \phi(x) &= \int \frac{d^3k}{(2\pi)^3 2\omega_k} (-2i\omega_k) a(k) e^{-ik \cdot x}, \\ a(k) &= i \int d^3x e^{ikx} \left(\pi(x) - i\omega_k \phi(x) \right) \Big|_{x_0=0}. \end{aligned} \quad (1.255)$$

From these mode decompositions and the commutator rules for the fields, Eq. (1.242) and (1.243), we find that $a(k), a(k^\dagger)$ satisfy the commutation rules

$$[a(k), a^\dagger(k')] = (2\pi)^3 2\omega_k \delta^3(\vec{k} - \vec{k}') \quad (1.256)$$

$$[a(k), a(k')] = [a^\dagger(k), a^\dagger(k')] = 0. \quad (1.257)$$

We can express the Hamiltonian (1.246) in these operators, too. Let us do

this in detail for the three terms. Using (1.252) and (1.254) we obtain

$$\begin{aligned}
\nabla\phi(x) &= \int \frac{d^3k}{(2\pi)^3 2\omega_k} (i\vec{k}) \left(a(k) e^{-ik \cdot x} - a^\dagger(k) e^{ik \cdot x} \right), \\
\frac{1}{2} \int d^3x \pi^2 &= \frac{1}{2} \int \frac{d^3k}{(2\pi)^3 2\omega_k} \int \frac{d^3k'}{(2\pi)^3 2\omega_{k'}} \int d^3x (-i\omega_k)(-i\omega_{k'}) \\
&\quad \left(a(k) a(k') e^{-i(k+k') \cdot x} + a^\dagger(k) a^\dagger(k') e^{i(k+k') \cdot x} \right. \\
&\quad \left. - a(k) a^\dagger(k') e^{-i(k-k') \cdot x} - a^\dagger(k) a(k') e^{i(k-k') \cdot x} \right) \\
&= \frac{1}{8} \int \frac{d^3k}{(2\pi)^3} \left(-a(k) a(-k) - a^\dagger(k) a^\dagger(-k) + a(k) a^\dagger(k) + a^\dagger(k) a(k) \right) \\
\frac{1}{2} \int d^3x (\nabla\phi)^2 &= \frac{1}{2} \int \frac{d^3k}{(2\pi)^3 2\omega_k} \int \frac{d^3k'}{(2\pi)^3 2\omega_{k'}} \int d^3x (i\vec{k}) \cdot (i\vec{k}') \\
&\quad \left(a(k) a(k') e^{-i(k+k') \cdot x} + a^\dagger(k) a^\dagger(k') e^{i(k+k') \cdot x} \right. \\
&\quad \left. - a(k) a^\dagger(k') e^{-i(k-k') \cdot x} - a^\dagger(k) a(k') e^{i(k-k') \cdot x} \right) \\
&= \frac{1}{8} \int \frac{d^3k}{(2\pi)^3} \frac{\vec{k}^2}{\omega_k^2} \left(a(k) a(-k) + a^\dagger(k) a^\dagger(-k) + a(k) a^\dagger(k) + a^\dagger(k) a(k) \right) \\
\frac{m^2}{2} \int d^3x \phi^2 &= \frac{m^2}{2} \int \frac{d^3k}{(2\pi)^3 2\omega_k} \int \frac{d^3k'}{(2\pi)^3 2\omega_{k'}} \int d^3x \\
&\quad \left(a(k) a(k') e^{-i(k+k') \cdot x} + a^\dagger(k) a^\dagger(k') e^{i(k+k') \cdot x} \right. \\
&\quad \left. + a(k) a^\dagger(k') e^{-i(k-k') \cdot x} + a^\dagger(k) a(k') e^{i(k-k') \cdot x} \right) \\
&= \frac{1}{8} \int \frac{d^3k}{(2\pi)^3} \frac{m^2}{\omega_k^2} \left(a(k) a(-k) + a^\dagger(k) a^\dagger(-k) + a(k) a^\dagger(k) + a^\dagger(k) a(k) \right)
\end{aligned} \tag{1.258}$$

Using $\omega_k^2 = \vec{k}^2 + m^2$ and adding the terms up, we finally find

$$H = \frac{1}{2} \int \frac{d^3k}{(2\pi)^3 2\omega_k} \omega_k \left(a^\dagger(k) a(k) + a(k) a^\dagger(k) \right) \tag{1.259}$$

$$= \int \frac{d^3k}{(2\pi)^3 2\omega_k} \omega_k \left(a^\dagger(k) a(k) + \frac{1}{2} [a(k), a^\dagger(k)] \right) \tag{1.260}$$

$$= \int \frac{d^3k}{(2\pi)^3 2\omega_k} \omega_k a^\dagger(k) a(k) + \int d^3k \frac{1}{2} \omega_k \delta^3(0). \tag{1.261}$$

The second integral is an infinite constant. We may interpret it as a divergent zero point energy. The Hamiltonian can be compared with the Hamiltonian of an infinite number of harmonic oscillators with creation operators $a_i^\dagger$ and

annihilation operators a_i

$$H = \sum_k \hbar \omega_k (a_k^\dagger a_k + \frac{1}{2}). \quad (1.262)$$

The ground state $|0\rangle$ obeys $a_k|0\rangle = 0$ for all k , and the excited states are given by expressions like

$$a_j^\dagger|0\rangle, \quad a_j^\dagger a_k^\dagger|0\rangle, \quad a_j^\dagger a_k^\dagger a_l^\dagger|0\rangle, \quad \dots$$

The “quantum number” operator is

$$N = \sum_k a_k^\dagger a_k. \quad (1.263)$$

In our field theoretical context we interpret

$$a^\dagger(k) \quad \text{as particle creation operator,} \quad (1.264)$$

$$a(k) \quad \text{as particle annihilation operator.} \quad (1.265)$$

We deal with the zero point energy by requiring that $H|0\rangle = 0$. This is achieved by removing the infinite constant from H :

$$H \rightarrow \int \frac{d^3k}{(2\pi)^3 2\omega_k} \omega_k a^\dagger(k) a(k). \quad (1.266)$$

As this is a fixing of the absolute scale of the energy, physical meaningful energy differences are not affected by it.

The subtraction of the zero point energy can be expressed in terms of normal ordering. Normal ordering means, that one has to put each creation operator to the left of every annihilation operator. Normal ordering is symbolised by colons, e.g.

$$:a^\dagger a: = a^\dagger a, \quad :aa^\dagger: = a^\dagger a. \quad (1.267)$$

Then we define the Hamiltonian to be $:H:$, such that

$$:H:|0\rangle = 0. \quad (1.268)$$

In the following we leave out the colons and understand H to be normal ordered. Applying creation operators onto the ground state yields excited states

$$H a^\dagger(k)|0\rangle = \omega_k a^\dagger(k)|0\rangle, \quad (1.269)$$

$$H a^\dagger(k_1) a^\dagger(k_2)|0\rangle = (\omega_{k_1} + \omega_{k_2}) a^\dagger(k_1) a^\dagger(k_2)|0\rangle. \quad (1.270)$$

From these equations we see that the energies of these states are the energies of non-interacting relativistic multi-particle states. Each creation operator creates one particle. The ground state is empty and is the vacuum state.

The representation of the field operators by particle creation and annihilation operators is called the *Fock representation*; the states created in this way out of the vacuum state are Fock states. The Hilbert space spanned by all these multi-particle states is the direct sum of n -particle Hilbert spaces $\mathcal{H}_n$,

$$\mathcal{H} = \bigoplus_{n=0}^{\infty} \mathcal{H}_n \quad (1.271)$$

and is called Fock space.

A remark about the zero-point energy

The zero-point energy changes, if there are boundary conditions. An example is the *Casimir effect*, where two parallel conducting plates feel a force pulling them together, although they are not electrically charged. Between the conductors the possible “cavity modes” are restricted by boundary conditions. Their number increases, when the distance a of the plates is increased. Thus the zero-point energy in between the plates grows with distance, while outside there is no restriction on the possible modes. The increasing zero-point energy leads to an attractive force on the plates,

$$\frac{\partial}{\partial a} \text{Energy} = \text{attractive force} \sim \frac{1}{a^4}. \quad (1.272)$$

1.4.2 Quantisation of the complex scalar field

The complex scalar field

$$\phi(x) = \int \frac{d^3k}{(2\pi)^3 2\omega_k} \left(a(k) e^{-ikx} + b^\dagger(k) e^{ikx} \right) \quad (1.273)$$

can be written as a combination of two real scalar fields

$$\phi = \frac{1}{\sqrt{2}} (\phi_1 + i\phi_2), \quad (1.274)$$

so that

$$\begin{aligned} \phi_j &= \int \frac{d^3k}{(2\pi)^3 2\omega_k} \left(a_j(k) e^{-ikx} + a_j^\dagger(k) e^{ikx} \right), \quad (j = 1, 2.) \\ \phi &= \int \frac{d^3k}{(2\pi)^3 2\omega_k} \left(\underbrace{\frac{1}{\sqrt{2}} (a_1(k) + i a_2(k))}_{=a(k)} e^{-ikx} + \underbrace{\frac{1}{\sqrt{2}} (a_1^\dagger(k) + i a_2^\dagger(k))}_{=b^\dagger(k)} e^{ikx} \right). \end{aligned}$$

This leads to creation and annihilation operators

$$a(k) = \frac{1}{\sqrt{2}}(a_1(k) + ia_2(k)), \quad (1.275)$$

$$b^\dagger(k) = \frac{1}{\sqrt{2}}(a_1^\dagger(k) + ia_2^\dagger(k)), \quad (1.276)$$

$$b^\dagger(k) \neq a^\dagger(k). \quad (1.277)$$

The complex scalar field is equivalent to a pair of two independent real scalar fields. The commutation relations are

$$[a(k), a^\dagger(k')] = [b(k), b^\dagger(k')] = (2\pi)^3 2\omega_k \delta^3(\vec{k} - \vec{k}'), \quad (1.278)$$

$$[a, b] = [a^\dagger, b^\dagger] = [a, b^\dagger] = [a^\dagger, b] = 0. \quad (1.279)$$

The field describes two sorts of particles, which can be created from the vacuum state $|0\rangle$, obeying

$$a(k)|0\rangle = b(k)|0\rangle = 0. \quad (1.280)$$

For example, one gets Fock states like

$$\text{one-particle states} \quad a^\dagger(k)|0\rangle, \quad b^\dagger(k)|0\rangle,$$

$$\text{two-particle states} \quad a^\dagger(k)a^\dagger(k')|0\rangle, \quad a^\dagger(k)b^\dagger(k')|0\rangle.$$

The conjugate momentum and the Hamiltonian density are

$$\pi = \frac{\partial}{\partial(\partial_0\phi)} \left((\partial_\mu\phi^\dagger)(\partial^\mu\phi) - m^2\phi^\dagger\phi \right) = \partial^0\phi^\dagger = \dot{\phi}^\dagger, \quad \pi^\dagger = \dot{\phi}, \quad (1.281)$$

$$\mathcal{H} = \pi\pi^\dagger - \mathcal{L}. \quad (1.282)$$

The Hamiltonian looks like the classic Hamiltonian with the fields replaced by field operators,

$$H = \int d^3x \left(\pi\pi^\dagger + (\nabla\phi^\dagger)(\nabla\phi) + m^2\phi^\dagger\phi \right). \quad (1.283)$$

In the Fock representation after normal ordering this is

$$:H: = \int \frac{d^3k}{(2\pi)^3 2\omega_k} \omega_k \left(a^\dagger(k)a(k) + b^\dagger(k)b(k) \right). \quad (1.284)$$

In a similar way one finds an expression for the charge. From the U(1) symmetry and the Noether theorem a conserved charge Q follows, which can be expressed by field operators

$$Q = iq \int d^3x \left(\phi^\dagger \dot{\phi} - \dot{\phi}^\dagger \phi \right). \quad (1.285)$$

In the Fock representation this is

$$Q = q \int \frac{d^3k}{(2\pi)^3 2\omega_k} (a^\dagger(k)a(k) - b^\dagger(k)b(k)). \quad (1.286)$$

This immediately gives

$$Q a^\dagger(k)|0\rangle = q a^\dagger(k)|0\rangle \quad \text{for a particle state of sort } a, \quad (1.287)$$

$$Q b^\dagger(k)|0\rangle = -q b^\dagger(k)|0\rangle \quad \text{for a particle state of sort } b. \quad (1.288)$$

We see that the state $b^\dagger(k)|0\rangle$ represents a particle with opposite charge than the particle represented by $a^\dagger(k)|0\rangle$, and we call it the corresponding antiparticle. We define a charge conjugation operator C , which exchanges a and b :

$$C a^\dagger C = b^\dagger, \quad C b^\dagger C = a^\dagger. \quad (1.289)$$

Then, since $CC = 1$,

$$C a^\dagger(k)|0\rangle = C a^\dagger(k)CC|0\rangle = b^\dagger(k)C|0\rangle = b^\dagger(k)|0\rangle. \quad (1.290)$$

1.4.3 Quantisation of the Dirac field

Similar as for the complex scalar field, the expansion of the Dirac field in terms of plane waves, Eq. (1.34), contains two sorts of coefficients, labelled b and d . Quantisation turns them into creation and annihilation operators, such that

$$\begin{aligned} b_r^\dagger(k) & \text{ creates particles, e.g. electrons,} \\ b_r(k) & \text{ annihilates particles,} \\ d_r^\dagger(k) & \text{ creates antiparticles, e.g. positrons,} \\ d_r(k) & \text{ annihilates antiparticles.} \end{aligned}$$

Starting from the Lagrangian (1.110)

$$\mathcal{L} = \bar{\psi}(i\gamma^\mu \partial_\mu - m)\psi$$

one defines the conjugate momentum field

$$\Pi(x) = \frac{\partial \mathcal{L}}{\partial(\partial_0 \psi)} = i\psi^\dagger. \quad (1.291)$$

Then the Hamiltonian is

$$H = \int d^3x (\Pi \dot{\psi} - \mathcal{L}) = \int d^3x (\psi^\dagger i\partial_0 \psi - \mathcal{L}) \quad (1.292)$$

$$= \int \frac{d^3k}{(2\pi)^3 2\omega_k} \omega_k \sum_{r=1}^4 (b_r^\dagger(k)b_r(k) - d_r(k)d_r^\dagger(k)). \quad (1.293)$$

In contrast to the scalar field, we have to impose anticommutation rules for the Dirac field, because it describes Fermions:

$$\left[\psi_\alpha(x^0, \vec{x}), \psi_\beta^\dagger(x^0, \vec{x}') \right]_+ = \delta_{\alpha\beta} \delta^3(\vec{x} - \vec{x}'). \quad (1.294)$$

This leads to

$$\left[b_r(k), b_{r'}^\dagger(k') \right]_+ = \left[d_r(k), d_{r'}^\dagger(k') \right]_+ = (2\pi)^3 2\omega_k \delta_{r,r'} \delta^3(\vec{k} - \vec{k}'), \quad (1.295)$$

$$\left[b_r(k), b_{r'}(k') \right]_+ = \left[b_r^\dagger(k), b_{r'}^\dagger(k') \right]_+ = 0, \quad (1.296)$$

$$\left[d_r(k), d_{r'}(k') \right]_+ = \left[d_r^\dagger(k), d_{r'}^\dagger(k') \right]_+ = 0, \quad (1.297)$$

$$\left[b_r(k), d_{r'}(k') \right]_+ = \text{mixed anticommutators} = 0, \quad (1.298)$$

The definition of normal ordering contains a minus sign, when two operators are interchanged. Therefore the normal ordered Hamiltonian

$$H = \int \frac{d^3k}{(2\pi)^3 2\omega_k} \omega_k \sum_{r=1}^4 \left(b_r^\dagger(k) b_r(k) + d_r^\dagger(k) d_r(k) \right) \quad (1.299)$$

is positive for all states in the Fock space.

1.4.4 Quantisation of the Maxwell field

The vector potential $A_\mu(x)$ is only determined up to gauge transformations. Therefore it contains redundant, unphysical degrees of freedom, and the quantisation procedure is non-trivial. There are two possibilities, which are considered in this context. The first possibility is to impose the Coulomb (or radiation) gauge on the fields. In this way, however, manifest Lorentz invariance is lost. To show the relativistic invariance of the physical results of the theory is a non-trivial task. The second possibility is to keep manifest Lorentz invariance and to quantise the theory covariantly. But then the gauge freedom in the field variables leads to unphysical states, which must be removed afterwards. This second possibility is called the Gupta-Bleuler quantisation.

Coulomb or radiation gauge

When external currents are absent, in the radiation gauge the field only contains transversal modes with amplitudes $a^{(\lambda)}(k)$ and $a^{(\lambda)*}(k)$, where $\lambda = 1, 2$ labels two polarisation states or two helicity modes of the photon, $k^0 = \omega_k$ and $\vec{k} \perp \vec{A}$, see Eq. (1.81). Quantisation promotes the amplitudes to operators $a^{(\lambda)}(k)$, $a^{(\lambda)\dagger}(k)$. They are photon annihilation and creation operators, and we have

$$H a^{(\lambda)\dagger}(k) |0\rangle = k_0 a^{(\lambda)\dagger}(k) |0\rangle. \quad (1.300)$$

Let us study the conjugate momenta of the field. We have

$$\frac{\partial \mathcal{L}}{\partial(\partial_\mu A_\nu)} = -F^{\mu\nu}, \quad (1.301)$$

so that

$$\Pi^0(x) = \frac{\partial \mathcal{L}}{\partial \dot{A}_0(x)} = 0, \quad (1.302)$$

$$\Pi^i(x) = \frac{\partial \mathcal{L}}{\partial \dot{A}_i(x)} = -F^{0i} = E^i. \quad (1.303)$$

Canonical commutation relations

$$[A^i(x^0, \vec{x}), \Pi^j(x^0, \vec{x}')] = -i \delta_{ij} \delta^3(\vec{x} - \vec{x}') \quad (1.304)$$

would be in contradiction with the transversality $\partial_i A^i(x) = 0$, since

$$0 = [\partial_i A^i(x^0, \vec{x}), \Pi^j(x^0, \vec{x}')] = -i \partial_j \delta^3(\vec{x} - \vec{x}') \neq 0. \quad (1.305)$$

Instead, the standard commutators for $a^{(\lambda)}(k)$ and $a^{(\lambda)\dagger}(k)$ lead to

$$[A^i(x^0, \vec{x}), \Pi^j(x^0, \vec{x}')] = -i \tilde{\delta}_{ij}(\vec{x} - \vec{x}'), \quad (1.306)$$

where the transverse δ -function is defined by

$$\tilde{\delta}_{ij}(\vec{x} - \vec{x}') := \int \frac{d^3 k}{(2\pi)^3} e^{i\vec{k} \cdot (\vec{x} - \vec{x}')} \left(\delta_{ij} - \frac{k_i k_j}{\vec{k}^2} \right). \quad (1.307)$$

Covariant quantisation

Covariant quantisation starts with a Lagrangian

$$\mathcal{L} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} - \frac{1}{2} (\partial_\mu A^\mu)^2, \quad (1.308)$$

which is manifestly Lorentz invariant, but not gauge invariant. Covariant commutators are imposed on the fields and momenta. There are four types of annihilation and creation operators $a^{(\lambda)}(k)$, $a^{(\lambda)\dagger}(k)$ for $\lambda = 0, \dots, 3$. The Fock space contains unphysical states, e.g. with negative norm. The physical states are restricted by

$$(\partial_\mu A^\mu)^+ |\text{physical state}\rangle = 0, \quad (1.309)$$

where $(\partial_\mu A^\mu)^+$ is the positive frequency part of $\partial_\mu A^\mu$. With this formalism all physical, gauge invariant results are the same as with the radiation gauge. To summarise

- explicit Lorentz invariance is kept,
- unphysical states in the Fock space have to be removed by constraints.

1.4.5 Symmetries and Noether charges

According to the Noether theorem, to each continuous symmetry belongs a conserved charge. Quantisation turns it into a charge operator. Consider, for example, the isospin for up and down quarks:

$$\hat{I}_3 := \int d^3x \bar{q}(x) \gamma^0 I_3 q(x) = \frac{1}{2} \int d^3x \left(\bar{u}(x) \gamma^0 u(x) - \bar{d}(x) \gamma^0 d(x) \right). \quad (1.310)$$

In the quantised field theory the quark fields $\bar{u}$, u , $\bar{d}$ and d are field operators. In Fock space with creation and annihilation operators $b^{(u)\dagger}$, $b^{(u)}$, $b^{(d)\dagger}$, $b^{(d)}$ and states

$$|u\rangle = b^{(u)\dagger}|0\rangle, \quad |d\rangle = b^{(d)\dagger}|0\rangle, \quad (1.311)$$

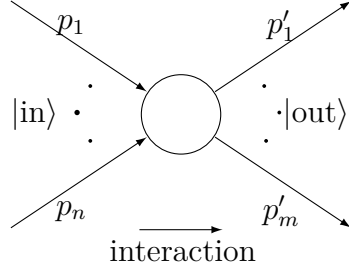
the charge operator acts on these states as

$$\hat{I}_3|u\rangle = \frac{1}{2}|u\rangle, \quad \hat{I}_3|d\rangle = -\frac{1}{2}|d\rangle. \quad (1.312)$$

compare Eqs. (1.286) ff.

1.5 Interacting Fields

In the previous sections free field theories have been considered. They describe particles that don't interact with each other. Now we turn to the consideration of interactions. Consider a scattering process between particles, as indicated in the picture.



A number of n particles approach each other and interact with each other. They form the ingoing state. After the interaction there are particles in an outgoing state. During the scattering process the particles are in a highly complicated state. But in the far past and in the far future they are far away from each other and can be considered as non-interacting (we neglect self-interactions here). The corresponding asymptotic states describe free particles.

The transition probability from the ingoing state $|\text{in}\rangle$ to an outgoing state $|\text{out}\rangle$ will be described by the matrix element of a unitary time evolution operator $U(t_1, t_0)$,

$$\langle \text{out} | U(+\infty, -\infty) | \text{in} \rangle. \quad (1.313)$$

1.5.1 Interaction picture

Here is a short reminder about the interaction picture. In the Schrödinger picture the states are time dependent, their time evolution is given by a unitary operator:

$$|\psi_S(t)\rangle = U(t, t_0) |\psi_S(t_0)\rangle = e^{-iH(t-t_0)} |\psi_S(t_0)\rangle. \quad (1.314)$$

This holds, provided the Hamiltonian has no explicit time dependence. In the Heisenberg picture the time dependence is shifted from the states to the operators:

$$O_H(t) = U^\dagger(t, t_0) O_S U(t, t_0). \quad (1.315)$$

In the free field theories considered so far, the field operators and the operators formed out of them are understood to be in the Heisenberg picture, e.g.

the Hamiltonian

$$:H_0: = \frac{1}{2} \int d^3x : [\pi^2 + (\nabla\phi)^2 + m^2\phi^2] := \int \frac{d^3k}{(2\pi)^3 2\omega_k} \omega_k a_k^\dagger a_k. \quad (1.316)$$

For free fields the Hamiltonians in the different pictures are identical:

$$H_0^{(H)} = H_0^{(S)} = H_0. \quad (1.317)$$

In free theories there are no non-trivial transition probabilities

$${}_S\langle \text{out}, t | \text{in}, t_0 \rangle_S = {}_H\langle \text{out} | U(t, t_0) | \text{in} \rangle_H. \quad (1.318)$$

As for the harmonic oscillator, where the time dependence of the ladder operators in the Heisenberg picture is given by

$$\dot{a}_H(t) = i[H, a_H] = i\omega[a^\dagger a, a_H] = -i\omega a_H(t), \quad (1.319)$$

resulting in

$$\begin{aligned} a_H(t) &= \exp(i\omega a^\dagger a t) a \exp(-i\omega a^\dagger a t) = a \exp(-i\omega t) \\ a_H^\dagger(t) &= \exp(i\omega a^\dagger a t) a^\dagger \exp(-i\omega a^\dagger a t) = a_H^\dagger \exp(i\omega t), \end{aligned}$$

the time dependence of annihilation and creation operators in field theory is given by

$$\begin{aligned} a_{k,H}(t) &= \exp(iH_0 t) a_{k,S} \exp(-iH_0 t) = a_{k,S} \exp(-i\omega_k t) \\ a_{k,H}^\dagger(t) &= \exp(iH_0 t) a_{k,S}^\dagger \exp(-iH_0 t) = a_{k,S}^\dagger \exp(i\omega_k t), \end{aligned}$$

and we find that in the transition probabilities only terms of the type

$$e^{-i\omega_k(t-t_0)} \delta_{\text{in}, \text{out}} \quad (1.320)$$

survive. So there are no transitions in the free theory.

Now we include interactions. In the Schrödinger picture we have an interaction Hamiltonian $H_S^{(I)}$,

$$i\partial_t |\psi(t)\rangle = (H_0 + H_S^{(I)}) |\psi(t)\rangle. \quad (1.321)$$

If one thinks of small interactions one may define slowly varying states $|\phi(t)\rangle$

$$|\phi(t)\rangle = e^{iH_0 t} |\psi(t)\rangle = U_0^{-1} |\psi(t)\rangle, \quad (1.322)$$

and with $i\partial_t U_0^{-1} = -H_0 U_0^{-1}$ one finds

$$i\partial_t U_0^{-1}|\psi(t)\rangle = -H_0 U_0^{-1}|\psi(t)\rangle + U_0^{-1}(H_0 + H_S^{(I)})\psi(t)\rangle \quad (1.323)$$

$$= U_0^{-1}H_S^{(I)}U_0 U_0^{-1}|\psi(t)\rangle. \quad (1.324)$$

With $H_I(t) := U_0^{-1}(t)H_S^{(I)}U_0(t)$ we get the time dependence of states in the interaction picture

$$i\partial_t |\phi(t)\rangle = H_I(t)|\phi(t)\rangle. \quad (1.325)$$

Operators in the interaction picture are consistently defined to be

$$O_I(t) = U_0^\dagger(t)O_S U_0(t). \quad (1.326)$$

They evolve like Heisenberg operators in the free theory,

$$\dot{O}_I(t) = i[H_0, O_I(t)]. \quad (1.327)$$

The advantages of the interaction picture are

- one keeps the formulations of the free theory for the operators,
- the $|\text{in}\rangle, |\text{out}\rangle$ states can be prepared as states of the free theory.

While at times $t = \pm\infty$ the states are simple eigenstates of H_0 , the time evolution during the interacting becomes complicated, since it involves $H_I(t)$. Therefore, one has to find approximations. Integrating the Schrödinger equation in the interaction picture (1.325) and iterating the result leads to

$$\begin{aligned} |\phi(t)\rangle &= \phi(t_0)\rangle + (-i) \int_{t_0}^t dt_1 H_I(t_1)|\phi(t_1)\rangle \\ &= \phi(t_0)\rangle + (-i) \int_{t_0}^t dt_1 H_I(t_1)|\phi(t_0)\rangle \\ &\quad + (-i)^2 \int_{t_0}^t dt_1 \int_{t_0}^{t_1} dt_2 H_I(t_1)H_I(t_2)|\phi(t_0)\rangle \\ &\quad + (-i)^3 \int_{t_0}^t dt_1 \int_{t_0}^{t_1} dt_2 \int_{t_0}^{t_2} dt_3 H_I(t_1)H_I(t_2)H_I(t_3)|\phi(t_0)\rangle \\ &\quad \dots\dots\dots \end{aligned}$$

The limits of integration obviously underlie the restriction $t > t_1 > t_2 > t_3 > \dots > t_0$. The restriction can be implemented by defining an operator $\mathcal{T}$ that generates time ordered products from arbitrary products of operators, i.e.

$$\mathcal{T}[O(t_1)O(t_2)] = \begin{cases} O(t_1)O(t_2), & \text{if } t_1 > t_2, \\ O(t_2)O(t_1) & \text{else.} \end{cases} \quad (1.328)$$

In the same manner $\mathcal{T}$ orders higher operator products with respect to their time arguments. With the aid of $\mathcal{T}$ the integral over the simplex $t > t_1 > t_2 > t_3 > \dots > t_0$ can be transformed into an integral over a hypercube

$$|\phi(t)\rangle = \sum_{i=0}^{\infty} \frac{(-i)^n}{n!} \int_{t_0}^t \int_{t_0}^t \dots \int_{t_0}^t dt_1 dt_2 \dots dt_n \mathcal{T}[H_I(t_1)H_I(t_2)\dots H_I(t_n)] |\phi(t_0)\rangle \quad (1.329)$$

$$= \mathcal{T} \exp \left(-i \int_{t_0}^t dt' H_I(t') \right) |\phi(t_0)\rangle. \quad (1.330)$$

This series is called Dyson series. It represents the unitary time evolution operator in the interaction picture,

$$|\phi(t)\rangle = U_I(t, t_0) |\phi(t_0)\rangle. \quad (1.331)$$

1.5.2 The S-matrix

In the limit $t_0 \rightarrow -\infty$, $t \rightarrow \infty$, the time evolution operator contains the description of scattering processes. The S-matrix is defined by

$$S := U_I(+\infty, -\infty) = \mathcal{T} \exp \left(-i \int_{-\infty}^{\infty} dt H_I(t) \right). \quad (1.332)$$

It gives the transition probability $\langle \text{out} | S | \text{in} \rangle$. In a free theory the S-matrix is the identity.

In general the interaction Hamiltonian contains field operators

$$\phi_I(x) = \underbrace{\int \frac{d^3k}{(2\pi)^3 2\omega_k} a_k e^{-ikx}}_{\phi^+} + \underbrace{\int \frac{d^3k}{(2\pi)^3 2\omega_k} a_k^\dagger e^{-ikx}}_{\phi^-}, \quad (1.333)$$

where ϕ^+ and ϕ^- are the positive and negative frequency parts, respectively.

Let us assume that $H_I = -\int d^3x \mathcal{L}_I$ arises from a Lagrangian density of the form $\mathcal{L}_I = \lambda F(\phi(x))$ with small λ . Then one can try to expand S in powers of λ ,

$$S = \mathbb{1} + S^{(1)} + S^{(2)} + \dots \quad (1.334)$$

$$S^{(0)} = \mathbb{1} \quad (1.335)$$

$$S^{(1)} = i\lambda \int d^4x [F(\phi(x))] \quad (1.336)$$

$$S^{(2)} = \frac{(i\lambda)^2}{2} \iint d^4x_1 d^4x_2 \mathcal{T}[F(\phi(x_1)) F(\phi(x_2))]. \quad (1.337)$$

For the calculation of matrix elements $\langle \text{out} | S^{(n)} | \text{in} \rangle$, it is convenient to bring this into normal ordered form by use of the commutation relations. The terms in the matrix element are of the form

$$\begin{aligned} & (\text{constants, integrals}) \\ & \langle 0 | (\text{creation operators for out states})^\dagger \\ & \mathcal{T}[\text{creation and annihilation operators from } H_I] \\ & (\text{creation operators for in states}) | 0 \rangle. \end{aligned}$$

To evaluate this one uses commutation or anticommutation relations to arrive at

$$\begin{aligned} & (\text{constants, integrals}) \text{functions}(x_1, \dots, x_n) \langle 0 | 0 \rangle \\ & + (\text{constants, integrals}) \text{functions}(x_1, \dots, x_n) \text{from } [,]_\pm \\ & \times \langle 0 | N[\text{creation and annihilation operators}] | 0 \rangle. \end{aligned}$$

Since the vacuum expectation values of all terms, which contain normal ordered operators vanish, this transformation of time ordered operator products to normal ordered products is especially useful. It can be accomplished with the help of Wick's theorem.

1.5.3 Wick's theorem

Define contractions $\overline{\phi(x_1)\phi(x_2)}$ as the difference between time ordered products and normal ordered products:

$$\overline{\phi(x_1)\phi(x_2)} = \mathcal{T}[\phi(x_1)\phi(x_2)] - N[\phi(x_1)\phi(x_2)] \quad (1.338)$$

$$= \langle 0 | \mathcal{T}[\phi(x_1)\phi(x_2)] | 0 \rangle \quad (1.339)$$

$$=: i\Delta_F(x_1 - x_2). \quad (1.340)$$

$\Delta_F(x)$ is called *Feynman propagator*

$$\Delta_F(x) = \lim_{\epsilon \rightarrow 0^+} \int \frac{d^4 k}{(2\pi)^4} \frac{1}{k^2 - m^2 + i\epsilon} e^{-ikx}. \quad (1.341)$$

Wick's theorem expresses time ordered products in terms of normal ordered products and contractions. Some examples are

$$\begin{aligned}\mathcal{T}[\phi_1\phi_2\phi_3] &= N[\phi_1\phi_2\phi_3] + \overline{\phi_1}\phi_2\phi_3 + \phi_1\overline{\phi_2}\phi_3 + \overline{\phi_1}\phi_2\overline{\phi_3} \\ \mathcal{T}[\phi_1\phi_2\phi_3\phi_4] &= N[\phi_1\phi_2\phi_3\phi_4] \\ &\quad + \overline{\phi_1}\phi_2N[\phi_3\phi_4] + \overline{\phi_1}\phi_3N[\phi_2\phi_4] \\ &\quad + \overline{\phi_1}\phi_4N[\phi_2\phi_3] + \overline{\phi_2}\phi_3N[\phi_1\phi_4] \\ &\quad + \overline{\phi_2}\phi_4N[\phi_1\phi_3] + \overline{\phi_3}\phi_4N[\phi_1\phi_2] \\ &\quad + \overline{\phi_1}\phi_2\overline{\phi_3}\phi_4 + \overline{\phi_1}\phi_2\phi_3\overline{\phi_4} + \overline{\phi_1}\phi_2\phi_3\phi_4,\end{aligned}$$

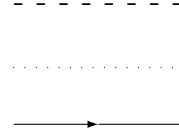
where we write ϕ_i for $\phi(x_i)$. The general rule is

$$\mathcal{T}[\text{fields}] = \sum_{\substack{\text{subsets} \\ \text{of fields}}} N[\text{subset of fields}] \times (\text{all possible contractions of other fields}). \quad (1.342)$$

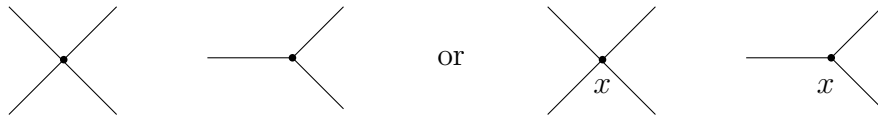
From Wick's theorem it follows that in $\langle 0 | \dots | 0 \rangle$ only the complete contractions survive.

1.5.4 Feynman diagrams

The contractions resulting from Wick's theorem are represented by lines. The lines for bosonic fields are drawn dotted or dashed.



The integral $-i\lambda \int d^4x$ in the interaction Hamiltonian gives a vertex, with lines connected to it like the following



Sometimes the integration variable (x) is denoted near the vertex. It remains to take into account the operators that generate the *in* and *out* states out of the vacuum. The contraction of these operators leads to diagram lines which come from outside. These lines are called *external lines*, while *internal lines* arise from contractions of field operators in the interaction Hamiltonian part. By means of Fourier transform the x -integrations can be evaluated in momentum space. External lines are associated with solutions of the free field equations and are subject to the restriction

$$k_0 = \sqrt{\vec{k}^2 + m^2}; \quad (1.343)$$

the corresponding propagators are denoted to be *on the mass shell*. At each vertex the sum of incoming momenta minus the sum of outgoing momenta has to be zero.

We illustrate the Feynman diagrams for the case of a real scalar field with quartic self-interaction. Let the Lagrangian be

$$\mathcal{L}_0 = \frac{1}{2}(\partial_\mu \phi \partial^\mu \phi - m^2 \phi^2), \quad (1.344)$$

$$\mathcal{L}_I = -g\phi^4. \quad (1.345)$$

Here g is for a small, real coupling constant. Other powers of ϕ in the interaction terms are not useful, for

ϕ^1 would give only a constant shift,
 ϕ^2 gives an extra mass,
 ϕ^3 would lead to instabilities.

The Hamiltonian density for the interaction is

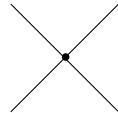
$$\mathcal{H}_I = g\phi^4(x) \quad (1.346)$$

and for two-particle scattering to first order in g one has to evaluate a term of the type

$$\langle \text{out} | \int d^4x g\phi^4(x) | \text{in} \rangle. \quad (1.347)$$

This is represented by

1 vertex for $\int d^4x g$,
 4 external lines.



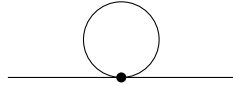
The graphs for the lowest contributions to two-particle scattering are

$$\begin{aligned}
 S^{(0)} &= \text{two parallel horizontal lines} + \text{cross with a small loop at the intersection} \\
 S^{(1)} &= \text{cross with a dot at the intersection} \\
 S^{(2)} &= \text{fish diagram} + \text{cross with a self-energy loop on one arm} + \dots
 \end{aligned}$$

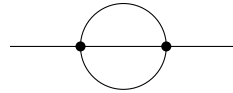
Classification of graphs

Let us comment on some particular types of graphs.

- Graphs with no external lines, the so called “vacuum bubbles”. It can be shown that they can be neglected in S-matrix elements.
- Graphs with 2 external lines are called self-energy graphs.



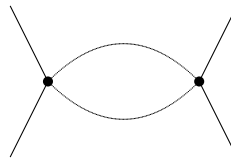
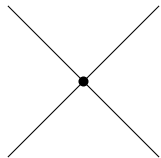
Tadpole (Kaulquappe)



Sunset

the full particle propagator and lead to corrections to the mass of the particle. They contribute to

- Scattering graphs with more than 2 external lines.



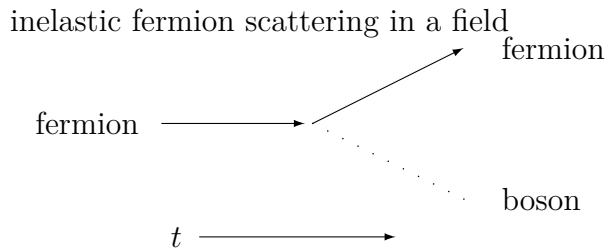
- Connected graphs, on which it is possible to go from each vertex to each other vertex by moving on lines, like the two previous ones.

- One-particle-irreducible graphs, which cannot be decomposed into 2 disconnected graphs by removing a single line. All graphs can be composed of these.

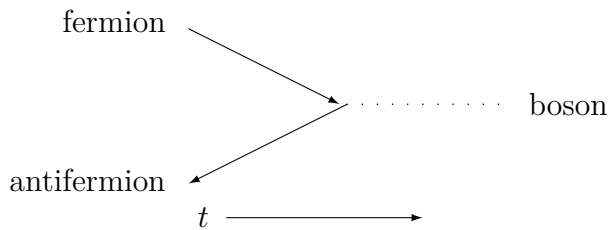
1.5.5 Fermions

For fermions the calculations explained above can be performed, too. It has to be observed that commutators between the basic fields have to be replaced by anticommutators. The resulting Feynman rules are analogous to the bosonic case. Due to the anticommuting nature of fermions, closed loops of fermion lines get an extra factor of (-1) .

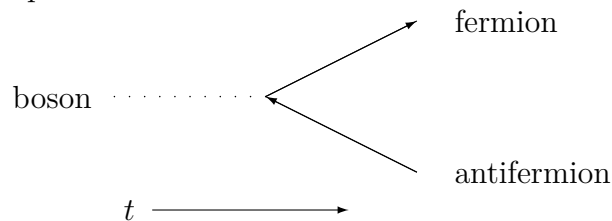
Examples for graphs including fermions are



e^+e^- annihilation



pair production



Yukawa coupling

The Lagrangian for the simplest Yukawa model contains a real scalar field $\phi(x)$ for a π^0 -meson, a Dirac field for a proton, and an interaction term:

$$\mathcal{L} = \mathcal{L}_0^\phi + \mathcal{L}_0^\psi + \mathcal{L}_I^{\phi\psi}, \quad (1.348)$$

$$\mathcal{L}_0^\phi = \frac{1}{2} \left(\partial_\mu \phi \partial^\mu \phi - M^2 \phi^2 \right), \quad (1.349)$$

$$\mathcal{L}_0^\psi = \bar{\psi} (i\gamma^\mu \partial_\mu - m) \psi, \quad (1.350)$$

$$\mathcal{L}_I^{\phi\psi} = -G : \bar{\psi}(x) \psi(x) \phi(x) = -\mathcal{H}_I. \quad (1.351)$$

A self-interaction for the pions, like $\frac{\lambda}{4!} \phi^4$ can also be included.

The propagator for fermion fields is given by

$$S_F(x) := \lim_{\epsilon \rightarrow 0^+} i \int \frac{d^4 k}{(2\pi)^4} \frac{\gamma^\mu k_\mu + m}{k^2 - m^2 + i\epsilon} e^{-ikx}. \quad (1.352)$$

In S_F , the time ordering for fermions is defined by

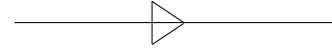
$$\mathcal{T}[\psi(x), \bar{\psi}(y)] = \begin{cases} \psi(x) \bar{\psi}(y), & \text{if } x^0 > y^0, \\ -\bar{\psi}(y) \psi(x), & \text{if } x^0 < y^0. \end{cases} \quad (1.353)$$

The contractions are defined by

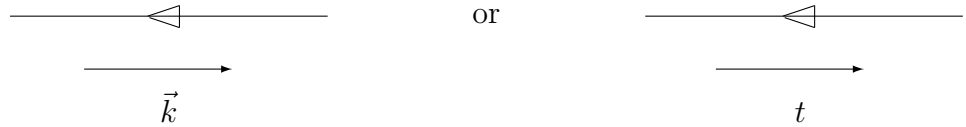
$$\overline{\psi(x) \bar{\psi}(y)} = \mathcal{T}[\psi(x) \bar{\psi}(y)] - N[\psi(x) \bar{\psi}(y)] = S_F(x - y). \quad (1.354)$$

Based on this we can draw Feynman graphs for the perturbative contributions.

The fermion contraction is represented by a line with an arrow.



The arrow gives the direction of charge flow. If charge flows in the same direction as momentum, the particle is a fermion, else it is the antiparticle. Sometimes the flow of momentum or the direction of time is indicated additionally as in these graphs for an incoming antifermion.



To distinguish fermions from bosons one draws dotted or dashed lines for bosons.

1.5.6 Limitations of the perturbative approach

- The coupling constant, e.g. λ , should be small, higher orders in the Dyson series should be less relevant. In various interesting cases, e.g. in strong interactions, this is not the case.
- At higher orders the number of graphs increases rapidly. They can only be managed by computer programs.
- Complicated integrals need special technical tricks.
- Some integrals are divergent. Renormalisation is needed.
- In general the perturbative series are not convergent, but only asymptotic. For practical purposes they have to be truncated.

2 Quantum Electrodynamics (QED)

2.1 Local U(1) Gauge Symmetry

QED deals with the interactions of electrons and positrons with the Maxwell field. Let us begin with the Dirac Lagrangian

$$\mathcal{L} = \bar{\psi}(x)(i\gamma^\mu \partial_\mu - m)\psi(x), \quad (2.1)$$

which has a *global* U(1) symmetry:

$$\psi(x) \longrightarrow \psi'(x) = e^{-iq\alpha}\psi(x). \quad (2.2)$$

If instead of a constant α , one considers $\alpha(x)$ to depend on x , one has *local* transformations

$$\psi(x) \longrightarrow \psi'(x) = e^{-iq\alpha(x)}\psi(x). \quad (2.3)$$

But since $\partial_\mu\psi$ does not transform like ψ ,

$$\partial_\mu\psi'(x) = e^{-iq\alpha(x)}\partial_\mu\psi(x) - iq(\partial_\mu\alpha(x))e^{-iq\alpha(x)}\psi(x),$$

the Lagrangian is not any more invariant under this transformation:

$$\mathcal{L}' = \mathcal{L} + \bar{\psi}i\gamma^\mu(-iq\partial_\mu\alpha)\psi = \mathcal{L} + q\bar{\psi}\gamma^\mu\psi\partial_\mu\alpha = \mathcal{L} + j^\mu\partial_\mu\alpha. \quad (2.4)$$

The invariance of the Lagrangian can be restored by coupling the Dirac field to the Maxwell field:

$$\begin{aligned} \mathcal{L} &= \bar{\psi}(i\gamma^\mu\partial_\mu - m)\psi - j^\mu A_\mu \\ &= \bar{\psi}(i\gamma^\mu\partial_\mu - m)\psi - q\bar{\psi}\gamma^\mu\psi A_\mu \\ &= \bar{\psi}(i\gamma^\mu(\partial_\mu + iqA_\mu) - m)\psi. \end{aligned}$$

Under gauge transformations the Maxwell field transforms as

$$A_\mu(x) \longrightarrow A'_\mu(x) = A_\mu(x) + \partial_\mu\alpha(x). \quad (2.5)$$

Thus we find that $\mathcal{L}$ is invariant:

$$\mathcal{L}' = \mathcal{L} + q\bar{\psi}\gamma^\mu\psi\partial_\mu\alpha - q\bar{\psi}\gamma^\mu\psi\partial_\mu\alpha = \mathcal{L}. \quad (2.6)$$

The Lagrangian contains the *covariant derivative*

$$D_\mu := \partial_\mu + iqA_\mu. \quad (2.7)$$

The covariant derivative of the fermion field transforms as

$$\begin{aligned}
D_\mu \psi &\longrightarrow D'_\mu \psi'(x) \\
&= (\partial_\mu + iqA'_\mu) \psi'(x) \\
&= (\partial_\mu + iq(A_\mu(x) + \partial_\mu \alpha(x))) e^{-iq\alpha(x)} \psi(x) \\
&= e^{-iq\alpha(x)} (\partial_\mu - iq(\partial_\mu \alpha(x)) + iq(A_\mu(x) + \partial_\mu \alpha(x))) \psi(x) \\
&= e^{-iq\alpha(x)} (\partial_\mu + iqA_\mu(x)) \psi(x) \\
&= e^{-iq\alpha(x)} (\partial_\mu + iqA_\mu(x)) e^{iq\alpha(x)} \psi'(x) \\
&= e^{-iq\alpha(x)} D_\mu e^{iq\alpha(x)} \psi'(x).
\end{aligned}$$

So we get the following transformation of the covariant derivative operator

$$D'_\mu = e^{-iq\alpha(x)} D_\mu e^{iq\alpha(x)}. \quad (2.8)$$

The case considered here is the simplest example of a more general situation, where a gauge field is coupled to matter fields.

The general picture

We start with a global symmetry. ($U(1)$)
This gives a Noether current. ($j^\mu = q\bar{\psi}\gamma^\mu\psi$)
We generalise to a local symmetry. ($\alpha \rightarrow \alpha(x)$)
We restore the invariance of the action by adding
a gauge field coupled to this Noether current.
Thus a symmetry determines the interaction. ($-q\bar{\psi}\gamma^\mu\psi A_\mu$)

To complete the Lagrangian, we have to add the dynamics of the free Maxwell field

$$\mathcal{L}_{(A_\mu)} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu}, \quad (2.9)$$

with the field strength

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu = -\frac{i}{q} [D_\mu, D_\nu]. \quad (2.10)$$

The complete Lagrangian density of QED thus reads

$$\mathcal{L} = \bar{\psi}(i\gamma^\mu D_\mu - m)\psi - \frac{1}{4} F_{\mu\nu} F^{\mu\nu}. \quad (2.11)$$

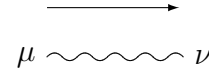
2.2 Quantum Electrodynamics

Now we consider quantisation of the theory. The Hamiltonian contains the free parts for fermions and photons and an interaction term

$$\mathcal{H}_I = -e_0 \bar{\psi} \gamma^\mu \psi A_\mu. \quad (2.12)$$

The free parts give

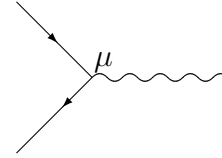
a fermion (electron or positron) propagator
and a photon propagator



The photon propagator in momentum space is given by

$$\frac{-ig^{\mu\nu}}{k^2 + i\epsilon}. \quad (2.13)$$

The interaction vertex is $ie_0 \gamma^\mu (\int d^4x)$.



The external photon lines get polarisation vectors

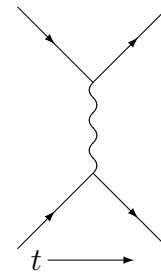
$$\epsilon_\mu(p), \epsilon_\mu^*(p) \perp \vec{p}. \quad (2.14)$$

Some QED processes

$e^- e^-$ scattering

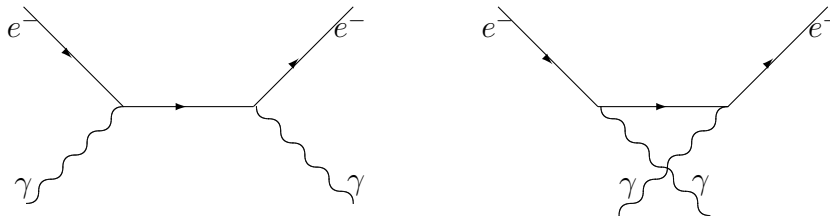
This graph represents the leading order contribution. In the non-relativistic limit this gives the Coulomb interaction with potential¹⁹

$$V = -\frac{\alpha}{r}, \quad \alpha = \frac{e^2}{4\pi\hbar c}$$



Compton scattering

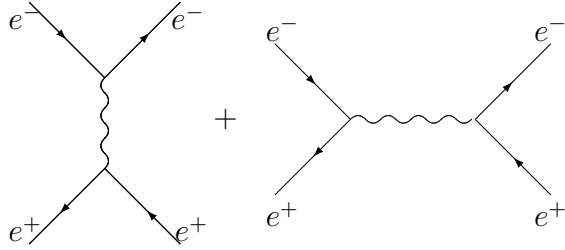
There are two graphs in leading order ($O(\alpha)$), resulting in two interaction processes.



¹⁹In SI units $\alpha = e^2/4\pi\epsilon_0\hbar c = 1/137.036$

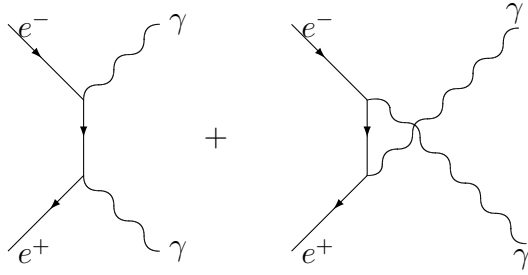
$e^+ e^-$ scattering

The leading order graphs are

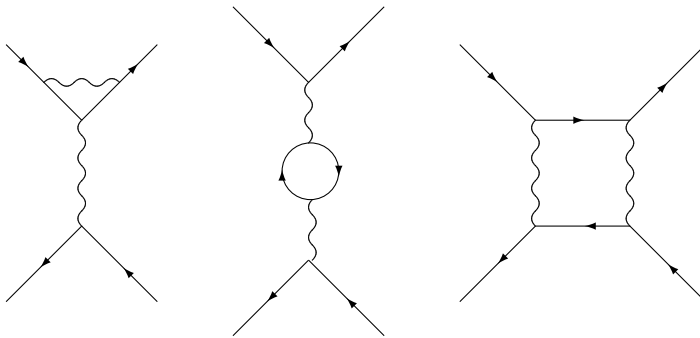


$e^+ e^-$ annihilation

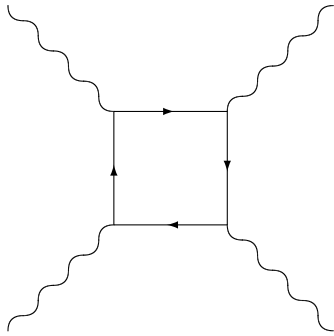
The leading order graphs describe $e^+ + e^- \rightarrow 2\gamma$.



The graphs considered so far do not contain closed loops of lines – these are tree level graphs. In higher orders graphs with loops give corrections to the leading order, or they may correspond to new processes. For example, corrections to $e^+ e^-$ scattering are



The following loop graph contributes to photon–photon scattering by exchange of virtual electrons. This nonlinear breaking of the superposition principle is also called *vacuum birefringence*.



3 Non-abelian Gauge Theory

3.1 Local Gauge Invariance

As in the case of QED, we start by considering a global symmetry. This time, it is given by the non-abelian group $SU(N)$. Consider a field having N complex components:

$$\phi(x) = \begin{pmatrix} \phi_1(x) \\ \vdots \\ \phi_N(x) \end{pmatrix}, \quad \phi^+(x) = (\phi_1^*(x), \dots, \phi_N^*(x)). \quad (3.1)$$

The scalar product between two complex n -component vectors ϕ and ϕ' is denoted

$$\phi^+ \cdot \phi' := \sum_{a=1}^N \phi_a^* \phi'_a. \quad (3.2)$$

Let the Lagrangian for the free field be

$$\mathcal{L} = \partial_\mu \phi^+(x) \cdot \partial^\mu \phi(x) - m^2 \phi^+(x) \cdot \phi(x). \quad (3.3)$$

It is invariant under the transformation

$$\phi \longrightarrow \phi' = U \cdot \phi, \quad (3.4)$$

where U is a $N \times N$ matrix obeying $U^+ \cdot U = 1$, namely

$$\partial_\mu \phi'^+ \cdot \partial^\mu \phi' = \partial_\mu (U \cdot \phi)^+ \cdot \partial^\mu (U \cdot \phi) = \partial_\mu \phi^+ \cdot U^+ \cdot U \partial^\mu \phi = \partial_\mu \phi^+ \cdot \partial^\mu \phi \quad (3.5)$$

and analogously

$$\phi'^+ \cdot \phi' = \phi^+ \cdot \phi. \quad (3.6)$$

In particular, matrices U which are in the group $SU(N)$ obey the condition, so that $SU(N)$ is a symmetry group of the Lagrangian.

Let us now consider *local* transformations $U(x) \in SU(N)$

$$U(x) = \exp \left(-i \sum_{a=1}^{N^2-1} \alpha^a(x) T_a \right), \quad (3.7)$$

where T_a are the $N^2 - 1$ generators of $SU(N)$. The spacetime dependency of the local transformations $U(x)$ gives an additional inhomogeneous term in the derivative of $U(x) \cdot \phi(x)$,

$$\partial_\mu \phi'(x) = U(x) \cdot \partial_\mu \phi(x) + (\partial_\mu U(x)) \phi(x), \quad (3.8)$$

such that the Lagrangian is no longer invariant. Again, the invariance can be restored by replacing the derivative by the covariant derivative

$$D_\mu \phi(x) := \left(\partial_\mu - ig \sum_{a=1}^{N^2-1} A_\mu^a(x) T_a \right) \phi(x). \quad (3.9)$$

This introduces N^2-1 gauge fields $A^a(x)$, one for each generator T_a of $SU(N)$. The coupling constant, denoted by $-g$, replaces the coupling constant $+q$ of QED.

A more compact notation is

$$A_\mu := A_\mu^a T_a := \sum_{a=1}^{N^2-1} A_\mu^a T_a. \quad (3.10)$$

The fields $A_\mu(x)$ are elements of the Lie algebra of $SU(N)$. Then the covariant derivation reads

$$D_\mu = \partial_\mu - ig A_\mu. \quad (3.11)$$

In terms of the covariant derivatives the Lagrangian is written

$$\mathcal{L} = D_\mu \phi^\dagger(x) \cdot D^\mu \phi(x) - m^2 \phi^\dagger(x) \cdot \phi(x). \quad (3.12)$$

Requiring the invariance of the Lagrangian, we can determine the transformation law for the gauge fields. The covariant derivative should obey

$$D'_\mu \phi'(x) \stackrel{!}{=} U(x) \cdot D_\mu \phi(x), \quad (3.13)$$

which is

$$D'_\mu \phi'(x) \stackrel{!}{=} U(x) \cdot D_\mu U^{-1}(x) \phi(x). \quad (3.14)$$

As this should hold for all $\phi'(x)$, we get

$$D'_\mu \stackrel{!}{=} U(x) \cdot D_\mu U^{-1}(x). \quad (3.15)$$

With the explicit expression for D_μ , Eq. (3.14) reads

$$(\partial_\mu - ig A'_\mu(x)) \cdot \phi'(x) \stackrel{!}{=} U(x) \cdot (\partial_\mu - ig A_\mu(x)) \cdot U^{-1}(x) \phi(x). \quad (3.16)$$

which implies

$$-ig A'_\mu(x) = U(x) \cdot \partial_\mu U^{-1}(x) - ig U(x) A_\mu(x) U^{-1}(x).$$

or

$$A'_\mu(x) = U(x) A_\mu(x) U^{-1}(x) + \frac{i}{g} U(x) \partial_\mu U^{-1}(x). \quad (3.17)$$

This is the transformation law for the gauge fields. It generalises the transformation law for the potentials $A_\mu(x)$ in QED.

Let us compare the formulae for gauge groups $U(1)$ and $SU(N)$:

$$\begin{aligned}
 U(1): \quad & U(x) = \exp(-iq\alpha(x)\mathbb{1}) \\
 & A_\mu(x) = \text{Maxwell field} \\
 & D_\mu = \partial_\mu + iqA_\mu(x) \\
 & A'_\mu(x) = A_\mu(x) + \partial_\mu\alpha(x) \\
 SU(N): \quad & U(x) = \exp(-i\alpha^a(x)T_a), \text{ and } [T_a, T_b] \neq 0 \text{ in general} \\
 & A_\mu(x) = A_\mu^a(x)T_a \\
 & D_\mu = \partial_\mu - igA_\mu(x) \\
 & A'_\mu(x) = U(x)A_\mu(x)U^{-1}(x) + \frac{i}{g}U(x)\partial_\mu U^{-1}(x).
 \end{aligned}$$

Infinitesimal transformations

It is also instructive to consider the formulae for the case of infinitesimal transformations

$$U(x) = \mathbb{1} - i\delta\alpha^a(x)T_a, \quad (3.18)$$

$$\phi'(x) = \phi(x) - i\delta\alpha^a(x)T_a\phi(x), \quad (3.19)$$

$$A'_\mu(x) = A_\mu(x) - i\delta\alpha^a(x)[T_a, A_\mu(x)] - \frac{1}{g}\partial_\mu\delta\alpha^a(x)T_a + O((\delta\alpha)^2). \quad (3.20)$$

Remembering that $A_\mu = A_\mu^a T_a$, we see that the commutators of the generators T_a are involved. These commutators are given by the structure constants of the Lie algebra associated with the group $SU(N)$

$$[T_a, T_b] = if_{ab}^c T_c. \quad (3.21)$$

With this, the transformation rule for the gauge fields are

$$\begin{aligned}
 A'^a_\mu T_a &= A_\mu^a T_a - i\delta\alpha^a A_\mu^b [T_a, T_b] - \frac{1}{g}\partial_\mu\delta\alpha^a(x)T_a \\
 &= A_\mu^a T_a + \delta\alpha^a A_\mu^b f_{ab}^c T_c - \frac{1}{g}\partial_\mu\delta\alpha^a(x)T_a \\
 &= (A_\mu^a + \delta\alpha^b A_\mu^c f_{bc}^a - \frac{1}{g}\partial_\mu\delta\alpha^a(x))T_a, \\
 A'^a_\mu(x) &= A_\mu^a(x) + f_{bc}^a \delta\alpha^b(x)A_\mu^c(x) - \frac{1}{g}\partial_\mu\delta\alpha^a(x). \quad (3.22)
 \end{aligned}$$

Field strengths and Yang-Mills Lagrangian

The missing piece is now the Lagrangian for the gauge fields. Motivated by QED we attempt to express it in terms of field strengths. We have to find appropriate expressions for the field strengths. Just defining them by $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ does not work, because this expression does not transform in such a way that a gauge invariant expression can be built from it. Considering the commutators of the covariant derivatives leads to the proper expression:

$$\begin{aligned}
 [D_\mu, D_\nu]\phi(x) &= [(\partial_\mu - igA_\mu), (\partial_\nu - igA_\nu)]\phi(x) \\
 &= (-ig[A_\mu, \partial_\nu] - ig[\partial_\mu, A_\nu] + (-ig)^2[A_\mu, A_\nu])\phi(x) \\
 &= -ig\{A_\mu\partial_\nu - \partial_\nu A_\mu + \partial_\mu A_\nu - A_\nu\partial_\mu - ig[A_\mu, A_\nu]\}\phi(x) \\
 &= -ig\{(\partial_\mu A_\nu - \partial_\nu A_\mu) - ig[A_\mu, A_\nu]\}\phi(x)
 \end{aligned}$$

In the last line, the derivatives act on the A_μ only, and dividing by $\phi(x)$ we define the field strengths through

$$F_{\mu\nu} = \frac{i}{g}[D_\mu, D_\nu] = (\partial_\mu A_\nu - \partial_\nu A_\mu) - ig[A_\mu, A_\nu]. \quad (3.23)$$

Expanding the A_μ in the generators yields

$$\begin{aligned}
 F_{\mu\nu} &= \partial_\mu A_\nu^a T_a - \partial_\nu A_\mu^a T_a + gA_\mu^a A_\nu^b f_{ab}^c T_c \\
 &= (\partial_\mu A_\nu^a - \partial_\nu A_\mu^a + gA_\mu^b A_\nu^c f_{bc}^a) T_a \\
 &=: F_{\mu\nu}^a T_a.
 \end{aligned} \quad (3.24)$$

So the components of $F_{\mu\nu}$ are

$$F_{\mu\nu}^a = \partial_\mu A_\nu^a - \partial_\nu A_\mu^a + gf_{bc}^a A_\mu^b A_\nu^c. \quad (3.25)$$

How does the field strength transform under gauge transformations? From Eq. (3.15) we obtain

$$F'_{\mu\nu} = \frac{i}{g}[D'_\mu, D'_\nu] = \frac{i}{g}[UD_\mu U^{-1}, UD_\nu U^{-1}] = \frac{i}{g}U[D_\mu, D_\nu]U^{-1} = UF_{\mu\nu}U^{-1}. \quad (3.26)$$

This homogeneous transformation law allows us to form gauge invariant expressions from $F_{\mu\nu}$. In particular, we can write down a gauge invariant Lagrangian for the gauge fields, which parallels the Lagrangian of the Maxwell

theory. This is the Yang-Mills Lagrangian: $\mathcal{L}_{YM}$

$$\mathcal{L}_{YM} = -\frac{1}{2} \text{tr}(F_{\mu\nu} F^{\mu\nu}) \quad (3.27)$$

$$\begin{aligned} &= -\frac{1}{2} \text{tr}(F_{\mu\nu}^a F^{a,\mu\nu} T_a T_b) \\ &= -\frac{1}{2} \text{tr}(F_{\mu\nu}^a F^{a,\mu\nu} \frac{1}{2} \delta_{ab}) \\ &= -\frac{1}{4} F_{\mu\nu}^a F^{a,\mu\nu}. \end{aligned} \quad (3.28)$$

In fact, it is invariant, as can be checked by using the cyclicity of the trace.

Remarks:

- A mass term like $m^2 A_\mu^a A^{a,\mu}$ is forbidden by gauge invariance.
- $\mathcal{L}_{YM}$ contains cubic ($A_\lambda^a A_\mu^b A_\nu^c$) and quartic ($A_\kappa^a A_\lambda^b A_\mu^c A_\nu^d$) terms. They represent self-interactions of the gauge field.

Field equations

We can now formulate the complete Lagrangian for an N -component complex scalar field interacting with non-abelian gauge fields, including an optional ϕ^4 -term:

$$\mathcal{L} = (D_\mu \phi)^\dagger \cdot D_\mu \phi - m^2 \phi^\dagger \cdot \phi - \lambda (\phi^\dagger \cdot \phi)^2 - \frac{1}{4} F_{\mu\nu}^a F^{a,\mu\nu}. \quad (3.29)$$

From this we may derive field equations for the field $\phi(x)$ and for the gauge field strength $F^{a,\mu\nu}(x)$:

$$(D_\mu D^\mu + m^2) \phi(x) = 0, \quad (3.30)$$

$$\partial_\mu F^{\mu\nu} - ig[A_\mu, F^{\mu\nu}] = j^\nu, \quad (3.31)$$

where

$$j^\nu := j^{a,\nu} T_a \quad (3.32)$$

is the current formed out of the scalar fields. The first equation is a gauge covariant generalisation of the Klein-Gordon equation. The second equation is the non-abelian generalisation of the inhomogeneous Maxwell equations. The analogue of the homogeneous Maxwell equations is

$$[D_\rho, F_{\mu\nu}] + [D_\mu, F_{\nu\rho}] + [D_\nu, F_{\rho\mu}] = 0. \quad (3.33)$$

It holds identically due to the definition of the field strengths. This equation is called *Bianchi identity*, because it has the same structure as the first

Bianchi identity, which expresses a symmetry of the curvature tensor in differential geometry. (Since $F_{\mu\nu}$ is a Lie bracket, the Bianchi identity can also be considered as a Jacobi identity.)

The second field equation can also be written as

$$[D_\mu, F^{\mu\nu}] = j^\nu, \quad (3.34)$$

because

$$\begin{aligned} [D_\mu, F^{\mu\nu}] &= [\partial_\mu - igA_\mu, F^{\mu\nu}] \\ &= \partial_\mu F^{\mu\nu} - F^{\mu\nu} \partial_\mu - ig[A_\mu, F^{\mu\nu}] \\ &= (\partial_\mu F^{\mu\nu}) - ig[A_\mu, F^{\mu\nu}]. \end{aligned} \quad (3.35)$$

The current j^μ is not conserved, $\partial_\mu j^\mu \neq 0$, since the gauge field itself is charged and contributes to the total current.

Historical remarks on gauge theory

The concept of local gauge invariance has its origin in the work of Hermann Weyl in 1918. He tried to extend general relativity by allowing spacetime dependent scale changes, “Umeichungen”, of the meter stick and the time-normal. Thus the metric tensor undergoes a local gauge transformation

$$g_{\mu\nu}(x) \longrightarrow \Omega(x)g_{\mu\nu}(x). \quad (3.36)$$

In Weyl’s approach $0 < \Omega(x) = e^{\alpha(x)}$. Einstein pointed out, however, that the length of a meter stick or the frequency of a clock could change if it is moved around a closed path and returned to its origin, which is physically unacceptable. After the introduction of Schrödinger’s wave function, London and others found that by considering purely imaginary $\alpha(x)$ in Weyl’s gauge transformations, and transforming the wave function correspondingly instead of the metric, the coupling to the electromagnetic field is recovered.

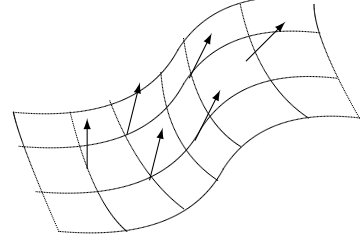
Here are some cornerstones in the development of gauge theory.

H. Weyl	1918	
V. Fock	1926	Gauge covariance of the Schrödinger equation.
F. London	1927	$\Omega(x) = e^{i\alpha(x)}$.
H. Weyl	1929	Local gauge theory for the Maxwell and matter fields.
O. Klein	1938	Conference in Warsaw. Non-abelian gauge field with covariant derivative, no Lagrangian.
W. Pauli	1953	Non-abelian gauge theory, based on non-abelian Kaluza-Klein theory. Unpublished, because there is no mass term in the theory.
C. N. Yang, and R. Mills	1954	Phys. Rev. 96 (1954) 191. Gauge theory for SU(2).
R. Shaw	1955	PhD thesis as student of A. Salam, Cambridge (UK).
R. Utiyama	1957	Phys. Rev. 101 (1957) 1597. Gauge model of gravity.

3.2 Geometry of Gauge Fields

3.2.1 Differential geometry

A manifold M is, roughly spoken, a space with a curvilinear coordinate system. The coordinates are sufficiently smooth. To each point x of the manifold is associated a tangent space. Its elements are tangent vectors. A vector field associates a particular vector $v(x)$ to each point x in a smooth way.



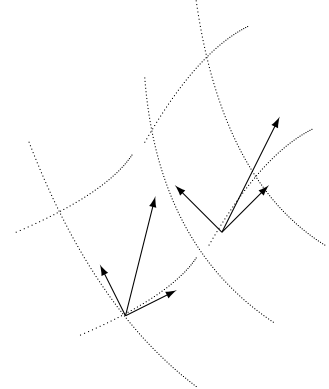
Vectors can be represented by their components $v_\mu(x)$ relative to a basis in the tangent space at x .

Now we would like to define how to take the derivative of a vector field along a particular curve. One is tempted to define the derivative via

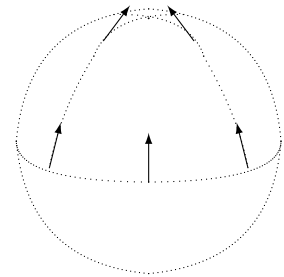
$$dv_\mu(x) = v_\mu(x + dx) - v_\mu(x), \quad (3.37)$$

where dx is an infinitesimal line element along the curve.

The problem with this is, however, that the direction of the basis vectors, which determine the components v_μ , in general can change from point to point. So, we need an appropriate way to compare vectors at different points. This can be done by taking the vector at a point x , $v(x)$, and *first* shift it by parallel transport to the point $x + dx$ and *then* compare it with the vector $v(x + dx)$.



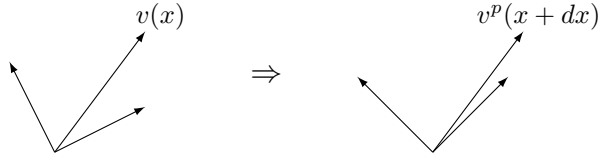
In order to put this into practice, a notion of parallel transport for vectors in the surface is needed. Parallel transport does not mean that the transported vectors are parallel in the embedding space. This is illustrated in the picture, where vectors are parallel transported along great circles on a sphere. The embedding space is not constitutive for the vector fields. We have introduced it here to illustrate the concepts. Parallel transport must be defined without referring to an embedding space.



Thus we must specify a rule for parallel transport. A vector $v(x)$ which is parallel transported from x to the point $x' = x + dx$ is a vector $v^p(x + dx)$,

$$v(x) \longrightarrow v^p(x + dx). \quad (3.38)$$

Although $v^p(x + dx)$ is parallelly transported from $v(x)$, its components change due to, say a rotation of the vector space basis, when going from x to $x + dx$.



Therefore, associated with an infinitesimal parallel transportation is an infinitesimal rotation.

$$v_\mu^p(x + dx) = \underbrace{\left(\mathbb{1} - \Gamma_\lambda dx^\lambda \right)_\mu^\nu}_{\text{infinitesimal rotation matrix}} v_\nu(x). \quad (3.39)$$

The elements of the infinitesimal rotation matrix are known as Christoffel symbols. With

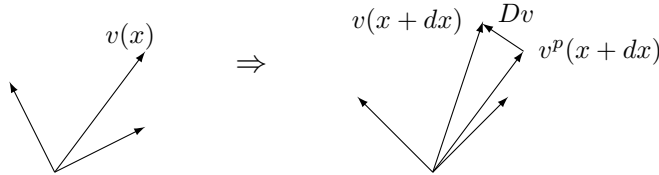
$$v_\mu^p(x + dx) = v_\mu(x) + \delta v_\mu(x) \quad (3.40)$$

one writes

$$\delta v_\mu(x) = -\Gamma_{\mu\lambda}^\nu(x) v_\nu(x) dx^\lambda. \quad (3.41)$$

Differentiation

Because $v(x)$ and $v(x + dx)$ belong to different vector spaces, there is a priori no way to add or subtract these vectors. To define differentiation, one may instead use the difference between $v_\mu(x + dx)$ and the components of the parallel transported v whose coordinates are denoted by $v_\mu^p(x + dx)$.



In order to distinguish it from the naive difference $dv_\mu(x) = v_\mu(x + dx) - v_\mu(x)$, it is denoted by a capital D :

$$Dv_\mu(x) := v_\mu(x + dx) - v_\mu^p(x + dx). \quad (3.42)$$

Using $\delta v_\mu(x)$, introduced above, we find

$$\begin{aligned} Dv_\mu(x) &= dv_\mu(x) - \delta v_\mu(x) \\ &= \partial_\lambda v_\mu(x) dx^\lambda + \Gamma_{\mu\lambda}^\nu(x) v_\nu(x) dx^\lambda \\ &= [\partial_\lambda v_\mu(x) + \Gamma_{\mu\lambda}^\nu(x) v_\nu(x)] dx^\lambda \\ &=: D_\lambda v_\mu(x) dx^\lambda, \end{aligned} \tag{3.43}$$

with the *covariant derivative*

$$D_\lambda v_\mu = \partial_\lambda v_\mu + \Gamma_{\mu\lambda}^\nu v_\nu. \tag{3.44}$$

$D_\lambda v_\mu$ transforms covariantly under coordinate transformations. It is a tensor, while $\partial_\lambda v_\mu$, $\Gamma_{\mu\lambda}^\nu v_\nu$, and $\Gamma_{\lambda\nu}^\nu$ are not tensors.

3.2.2 Gauge Theory

In contrast to differential geometry, the vector fields $\phi(x)$ of gauge theory are not vectors in the tangent space of a manifold. They belong to an “internal” vector space V_x like isospin, flavour, and colour vector space ²⁰.

$$\phi(x) = \begin{pmatrix} \phi_1(x) \\ \vdots \\ \phi_N(x) \end{pmatrix}. \tag{3.45}$$

The local gauge transformation with gauge group elements $U(x)$, denoted by

$$\phi(x) \longrightarrow \phi'(x) = U(x)\phi(x), \tag{3.46}$$

can be considered as a spacetime dependent change of the basis in V_x . It is a passive transformation, the components ϕ_a change because the basis vectors change.

Physics should not depend on the local choice of the basis. Therefore, differentiation has to be defined based on the change of ϕ relative to the parallel transported ϕ^p . Similar to the geometric case of the last section we write

$$\delta\phi_a(x) = \phi_a^p(x + dx) - \phi_a(x). \tag{3.47}$$

In this case we do not have Christoffel symbols, however, we can write the equation, which fixes the meaning of parallel transport, in an analogous way:

$$\delta\phi(x) = igA_\mu(x) \cdot \phi(x) dx^\mu. \tag{3.48}$$

²⁰The mathematics of gauge theory and fiber bundles was developed by Élie Cartan, and Charles Ehresmann.

Here $A_\mu(x)$ is a Hermitian matrix, and $igA_\mu(x)dx^\mu$ describes an infinitesimal rotation in the internal vector space. $A_\mu(x)$ is an element of the Lie algebra belonging to the gauge group element $U(x)$. We define

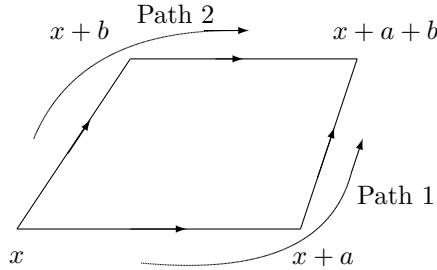
$$\begin{aligned}
 D\phi(x) &:= \phi(x+dx) - (\phi(x) + \delta\phi(x)) \\
 &= d\phi(x) - \delta\phi(x) \\
 &= \partial_\mu \phi(x) dx^\mu - igA_\mu(x)\phi(x)dx^\mu \\
 &= \underbrace{(\partial_\mu - igA_\mu(x))}_{=: D_\mu} \phi(x) dx^\mu.
 \end{aligned} \tag{3.49}$$

The gauge covariant derivative

$$D_\mu = \partial_\mu - igA_\mu(x) \tag{3.50}$$

corresponds to the covariant derivative of geometry.

There is another object of geometry that can be taken over to gauge theory, the curvature. As the picture about parallel displacement of vectors on a sphere showed, curvature is related to the differing results of parallel transportation, when the vectors are transported along different paths. Consider two different infinitesimal paths from a point to a neighbouring point, along which vectors are parallelly transported.



A vector $v(x)$, which is transported from x to $x+a+b$ along one path P_1 differs from the vector, which is transported along path P_2 .

Along path 1 the coordinates change according to

$$\begin{aligned}
 v_\mu(x) &\rightarrow v_\mu(x) - \Gamma_{\mu\alpha}^\nu(x)v_\nu(x)a^\alpha \\
 &\rightarrow v_\mu(x) - \Gamma_{\mu\alpha}^\nu(x)v_\nu(x)a^\alpha - \Gamma_{\mu\beta}^\lambda(x+a)\{v_\lambda(x) - \Gamma_{\lambda\alpha}^\nu(x)v_\nu(x)a^\alpha\}b^\beta \\
 &= v_\mu(x) - \Gamma_{\mu\alpha}^\nu(x)v_\nu(x)a^\alpha - \Gamma_{\mu\beta}^\lambda(x)v_\lambda(x)b^\beta - \partial_\alpha \Gamma_{\mu\beta}^\lambda(x)a^\alpha v_\lambda(x)b^\beta \\
 &\quad + \Gamma_{\mu\beta}^\lambda(x)\Gamma_{\lambda\alpha}^\nu(x)v_\nu(x)a^\alpha b^\beta + O(a^2b).
 \end{aligned}$$

Along path 2, after having exchanged a and b , the parallel transported vector is

$$\begin{aligned} v_\mu(x) \rightarrow & v_\mu(x) - \Gamma_{\mu\alpha}^\nu(x)v_\nu(x)b^\alpha - \Gamma_{\mu\beta}^\lambda(x)v_\lambda(x)a^\beta - \partial_\alpha\Gamma_{\mu\beta}^\lambda(x)b^\alpha v_\lambda(x)a^\beta \\ & + \Gamma_{\mu\beta}^\lambda(x)\Gamma_{\lambda\alpha}^\nu(x)v_\nu(x)b^\alpha a^\beta + O(a^2b). \end{aligned}$$

Then both displaced vectors differ by

$$\begin{aligned} \Delta v_\mu &= \delta(\text{path 2}) - \delta(\text{path 1}) \\ &= \partial_\alpha\Gamma_{\mu\beta}^\lambda(x)a^\alpha v_\lambda(x)b^\beta - \partial_\alpha\Gamma_{\mu\beta}^\lambda(x)b^\alpha v_\lambda(x)a^\beta \\ &+ \Gamma_{\mu\beta}^\lambda(x)\Gamma_{\lambda\alpha}^\nu(x)v_\nu(x)b^\alpha a^\beta - \Gamma_{\mu\beta}^\lambda(x)\Gamma_{\lambda\alpha}^\nu(x)v_\nu(x)a^\alpha b^\beta. \end{aligned}$$

The result is

$$\Delta v_\mu = R_{\mu\alpha\beta}^\nu v_\nu a^\alpha b^\beta, \quad (3.51)$$

where we introduced the Riemann-Christoffel curvature tensor $R_{\mu\alpha\beta}^\nu$

$$R_{\mu\alpha\beta}^\nu = \partial_\alpha\Gamma_{\mu\beta}^\nu - \partial_\beta\Gamma_{\mu\alpha}^\nu + \Gamma_{\mu\alpha}^\lambda\Gamma_{\lambda\beta}^\nu - \Gamma_{\mu\beta}^\lambda\Gamma_{\lambda\alpha}^\nu. \quad (3.52)$$

In a flat manifold both parallel transports must give the same result $\Delta v = 0$, and thus $R_{\mu\alpha\beta}^\nu = 0$. Manifolds can be flat locally, i.e. in an infinitesimal neighbourhood of a point, as is the case with a saddle point on a two dimensional surface.

Curvature of gauge fields

In gauge theory we have the same algebraic structure for parallel transporting the vector field ϕ

$$\delta\phi(x) = igA_\mu(x) \cdot \phi(x)dx^\mu. \quad (3.53)$$

Writing the matrix elements of A_μ as $A_{c\mu}^d$, we have

$$\delta\phi_c(x) = igA_{c\mu}^d(x)\phi_d(x)dx^\mu. \quad (3.54)$$

Comparing this with

$$\delta v_\nu(x) = -\Gamma_{\nu\mu}^\lambda(x)v_\lambda(x)dx^\mu, \quad (3.55)$$

we see that the analogue of the Christoffel symbol Γ is $-igA$. The difference of the parallel transport along two infinitesimally paths is then

$$\Delta\phi_c = -ig \left[\partial_\alpha A_{c\beta}^d - \partial_\beta A_{c\alpha}^d - igA_{c\alpha}^\lambda A_{\lambda\beta}^d + igA_{c\beta}^\lambda A_{\lambda\alpha}^d \right] \phi_d a^\alpha b^\beta, \quad (3.56)$$

or using matrix notation

$$\begin{aligned}
\Delta\phi &= -ig \{ \partial_\alpha A_\beta - \partial_\beta A_\alpha - ig A_\alpha \cdot A_\beta + ig A_\beta \cdot A_\alpha \} \cdot \phi a^\alpha b^\beta \\
&= \{ \underbrace{[\partial_\alpha, \partial_\beta]}_{=0} - [\partial_\alpha, ig A_\beta] + [\partial_\beta, ig A_\alpha] + [ig A_\alpha, ig A_\beta] \} \cdot \phi a^\alpha b^\beta \\
&= [\partial_\alpha - ig A_\alpha, \partial_\beta - ig A_\beta] \cdot \phi a^\alpha b^\beta \\
&= [D_\alpha, D_\beta] \cdot \phi a^\alpha b^\beta.
\end{aligned}$$

As we have already seen, the commutator of the D 's gives the field strengths $[D_\mu, D_\nu] = -ig F_{\mu\nu}$, therefore,

$$\Delta\phi(x) = -ig F_{\mu\nu}(x) \phi(x) \underbrace{a^\mu b^\nu}_{:=f^{\mu\nu}}, \quad (3.57)$$

where $f^{\mu\nu}$ is a surface element. Based on the analogy with differential geometry, one calls

$$F_{\mu\nu} = \text{“Curvature of the gauge field”}, \quad (3.58)$$

and if $F_{\mu\nu} \equiv 0$, one says “the gauge field is flat.” In this case A_μ can be transformed to 0 by a special choice of gauge transformation.

Let us consider parallel transport at finite distances. To see, how this works, we will look at the abelian case with symmetry group $U(1)$.



Then an infinitesimal parallel transport along a curve is given by

$$\delta\phi(x) = -iq A_\mu(x) dx^\mu, \quad (3.59)$$

which formally may be integrated

$$\phi^p(x') = \exp \left\{ -iq \int A_\mu(z) dz^\mu \right\} \phi(x). \quad (3.60)$$

We check this by returning to infinitesimal steps

$$\begin{aligned}
\phi^p(x + dx) &= \{1 - iq A_\mu(x) dx^\mu\} \phi(x) \\
&= \phi(x) - iq A_\mu(x) dx^\mu \phi(x) \\
&= \phi(x) + \delta\phi(x).
\end{aligned}$$

In the non-abelian case the exponentiated line integral $\int_x^{x'} A_\mu(z) dz^\mu$ has to be path ordered in much the same way as the time ordered expressions for the S-matrix.

For a *closed* curve $\mathcal{C}$ the change $\Delta\phi$ in the abelian case is given by

$$\phi(x) + \Delta\phi(x) = \exp\{-iq \oint_{\mathcal{C}} A_{\mu}(z) dz^{\mu}\} \phi(x). \quad (3.61)$$

From Stokes theorem

$$\oint_{\mathcal{C}} A_{\mu}(z) dz^{\mu} = \int_A F_{\mu\nu} df^{\mu\nu} = \text{flux through area } A \quad ^{21} \quad (3.62)$$

we have

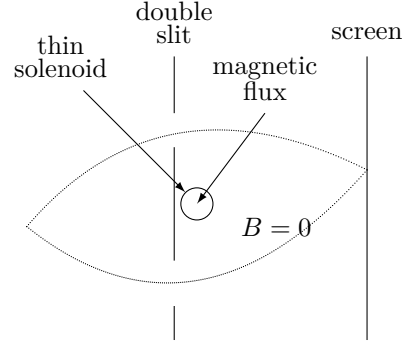
$$\phi(x) + \Delta\phi(x) = \exp\left\{-iq \int F_{\mu\nu} df^{\mu\nu}\right\} \phi(x). \quad (3.63)$$

This is a gauge invariant expression and consequently the line integral of A_{μ} along a closed loop is gauge invariant.

In electrodynamics the phase of the electron wavefunction changes along a curve due to the gauge dependent vector potential. But the phase is only measurable through interference, which means that one needs two different paths to the point of measurement. The resulting phase difference equals the phase of $\Delta\phi$ for a closed curve. The Aharonov-Bohm interference relies on this gauge invariant phase difference.

The infinitesimal version of Eq. (3.63) is

$$\Delta\phi(x) = -iq F_{\mu\nu}(x) f^{\mu\nu} \phi(x). \quad (3.64)$$



²¹Remind e.g. the magnetic flux

$$\int \vec{B} \cdot d\vec{f} = \int \nabla \times \vec{A} \cdot d\vec{f} = \oint \vec{A} \cdot d\vec{r}.$$

4 Quantum Chromodynamics (QCD)

4.1 Lagrangian Density and Symmetries

Quarks are represented by Dirac fields, which we now denote q instead of ψ .

$$q = (q_\alpha) = \begin{pmatrix} q_1 \\ \cdots \\ q_4 \end{pmatrix}. \quad (4.1)$$

Today, six “flavours” of quarks are known:

$$\begin{pmatrix} u \\ d \end{pmatrix}, \quad \begin{pmatrix} c \\ s \end{pmatrix}, \quad \begin{pmatrix} t \\ b \end{pmatrix},$$

and we write

$$q_f = (q_{\alpha,f}), \quad f = 1, 2, \dots, N_f, \quad N_f = 6. \quad (4.2)$$

Finally, there are three “colours” for each quark, which have been introduced originally to solve the problem of quark statistics:

$$q_i = (q_{\alpha,i,f}), \quad i = 1, 2, 3, \quad \text{or} \quad i = \text{red, green, blue}. \quad (4.3)$$

Thus in total we characterise a quark state by the three quantum numbers α, i, f , which gives $4 \times 3 \times N_f$ possibilities.

Let us begin with the Lagrangian density of free quarks

$$\mathcal{L} = \sum_{i,f} \bar{q}_{if} (i\gamma^\mu \partial_\mu - m_f) q_{if}. \quad (4.4)$$

For clarity we have explicitly written out the summation over the colour and flavour indices. The spinor indices are hidden. From this Lagrangian one gets $3 \times N_f$ Dirac equations, one for each colour and flavour.

Symmetries

Symmetries put very strong restrictions on the structure of physical objects or relations. For instance, symmetry around one axis of rotation restricts objects to look as being made on a turning lathe like some chess figures, and symmetry under the full rotation group allows only spherical objects. Symmetries also restrict interactions, which to a large extent are fixed by symmetry requirements. An example is the term $F_{\mu\nu} F^{\mu\nu}$ in the Lagrangian. Higher powers of $F_{\mu\nu} F^{\mu\nu}$ would show the same symmetry, but they must be excluded, because they lead to non-renormalisable theories.

Global SU(3) colour symmetry:

$$q_{if}(x) \longrightarrow q'_{if}(x) = (Uq)_{if} = \sum_{i,f} U_{ij} q_{jf}(x) \quad (4.5)$$

U is an unitary 3×3 matrix. Since there is another U(1) symmetry, U can be restricted to $\det U = +1$, so we have $U \in \text{SU}(3)$. U commutes with operators on flavour and spinor subspace, therefore, the symmetry of the Lagrangian is proven by

$$\begin{aligned} \bar{q}'_{if}(\mathrm{i}\gamma^\mu \partial_\mu - m_f)q'_{if} &= \bar{q}_{if}(U^\dagger)_{ij}(\mathrm{i}\gamma^\mu \partial_\mu - m_f)U_{jk}q_{kf} \\ &= \bar{q}_{if}(U^\dagger)_{ij}U_{jk}(\mathrm{i}\gamma^\mu \partial_\mu - m_f)q_{kf} \\ &= \bar{q}_{if}(\mathrm{i}\gamma^\mu \partial_\mu - m_f)q_{if}. \end{aligned}$$

U(1) baryon symmetry

$$q_{if}(x) \longrightarrow q'_{if}(x) = e^{-\mathrm{i}\alpha} q_{if}(x) \quad (4.6)$$

holds for the same reason. The corresponding Noether current is

$$j^\mu(x) = \bar{q}(x)\gamma^\mu q(x) = \sum_{i,f} \bar{q}_{if}(x)\gamma^\mu q_{if}(x). \quad (4.7)$$

The conserved charge is the integral of the zero component of the Noether current, which counts the number of quarks. By convention the baryon number is

$$B = \frac{1}{3} \int d^3x j^0(x). \quad (4.8)$$

Approximate flavour symmetry SU(N_f)

This symmetry would be exact, if all quark masses m_f were equal. We shall give the transformation laws later, when discussing interactions.

4.1.1 Local SU(3) colour symmetry

This symmetry was introduced in the early 1970's. The idea that colour is the source of a SU(3) gauge field, was formulated by Fritzsche, Gell-Mann and Leutwyler. Local SU(3) gauge symmetry is introduced in the following way.

1. Replace the derivative

$$\partial_\mu \longrightarrow D_\mu = \partial_\mu - \mathrm{i}g A_\mu^a(x) T_a, \quad (4.9)$$

$$T_a = \frac{1}{2} \lambda_a, \quad (a = 1, \dots, 8). \quad (4.10)$$

2. This introduces 8 gauge fields, called gluon fields. In the quantised theory they are associated with 8 new particles, the gluons. The field strengths are denoted $G_{\mu\nu}^a$,

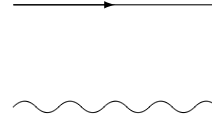
$$G_{\mu\nu}^a = \partial_\mu A_\nu^a - \partial_\nu A_\mu^a + gf_{abc}A_\mu^b A_\nu^c. \quad (4.11)$$

3. The Lagrangian of QCD now gets an extra Yang-Mills-action part

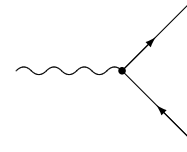
$$\mathcal{L}_{YM} = -\frac{1}{4} G_{\mu\nu}^a G^{\mu\nu,a}, \quad (4.12)$$

$$\mathcal{L}_{QCD} = \sum \bar{q}(i\gamma_\mu D_\mu - m_f)q - \frac{1}{4} G_{\mu\nu}^a G^{\mu\nu,a}. \quad (4.13)$$

The gluon fields lead to new Feynman graphs. The quark propagators are fermionic lines, and the gluon propagators are similar to photon lines.



The term $\bar{q}i\gamma^\mu D_\mu q$ leads to quark-gluon vertices, similar to QED.



The quadratic term in the gluon field strengths contains self-interaction of gluons.

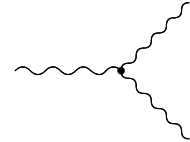
$$G_{\mu\nu}^a G^{\mu\nu,a}$$

$$= -\frac{1}{4}(\partial_\mu A_\nu^a - \partial_\nu A_\mu^a + gf_{abc}A_\mu^b A_\nu^c)(\partial^\mu A^{\nu,a} - \partial^\nu A^{\mu,a} + gf_{ade}A^{\mu,d} A^{\nu,e})$$

$$= -\frac{1}{4}(\partial_\mu A_\nu^a - \partial_\nu A_\mu^a)(\partial^\mu A^{\nu,a} - \partial^\nu A^{\mu,a})$$

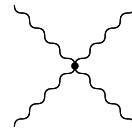
$$- \frac{1}{2}gf_{abc}A_\mu^b A_\nu^c(\partial^\mu A^{\nu,a} - \partial^\nu A^{\mu,a}),$$

3-gluon vertex

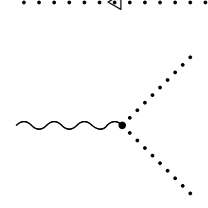


$$- \frac{1}{4}g^2 f_{abc}f_{ade}A_\mu^b A_\nu^c A^{\mu,d} A^{\nu,e}.$$

4-gluon vertex



Quantisation of gluon fields in perturbation theory requires gauge fixing. This introduces new fields, the *ghost fields*. Ghost propagators are represented by dotted lines. There are also ghost-gluon vertices.²²



4.1.2 Global flavour symmetry

While the masses of up and down quarks are nearly the same, the other quarks are much more massive. As a consequence the *flavour symmetry* $SU(N_f)$ is strongly broken. In spite of this, we can look not only for approximate symmetries but also for symmetries in subspaces of the quark states. For $U \in SU(N_f)$ we write

$$q_f \longrightarrow q'_{f'}(x) = U_{f'f} q_f(x). \quad (4.14)$$

Here, as elsewhere, we hide indices, which are unimportant in the context. Some examples are:

for quarks	q	means	$(q_{\alpha if}(x)),$
	$(\gamma^\mu q)_{\alpha if}$	means	$\gamma^\mu_{\alpha\beta} q_{\beta if},$
if $U \in SU(N_f),$	$U \cdot q$	means	$(U \cdot q)_{\alpha if} = U_{ff'} q_{\alpha if'}(x),$
if $U \in SU(3)_{\text{colour}},$	$U \cdot q$	means	$(U \cdot q)_{\alpha if} = U_{ij} q_{\alpha jf}(x).$

The term $\bar{q}(i\gamma^\mu D_\mu - m)q$ in the Lagrangian $\mathcal{L}_{QCD}$ transforms under $U \in SU(N_f)$ as

$$\begin{aligned} \bar{q}(i\gamma^\mu D_\mu - m)q &\longrightarrow \bar{q}'(i\gamma^\mu D_\mu - m)q' \\ &= \bar{q}(U^\dagger i\gamma^\mu D_\mu U - U^\dagger m U)q \\ &= \bar{q}(U^\dagger U i\gamma^\mu D_\mu - m U^\dagger U)q \\ &= \bar{q}(i\gamma^\mu D_\mu - m)q. \end{aligned}$$

Since the quark masses of different flavours are not the same, the mass term in the Lagrangian is

$$\sum_f m_f \bar{q}_f q_f = \bar{q} M q, \quad (4.15)$$

²²See Faddeev-Popov ghosts.

with the quark mass matrix

$$M = \begin{pmatrix} m_u & 0 & 0 & \dots & 0 \\ 0 & m_d & 0 & \dots & 0 \\ 0 & 0 & m_s & \dots & 0 \\ 0 & 0 & 0 & \ddots & \vdots \\ 0 & 0 & 0 & \dots & m_b \end{pmatrix}. \quad (4.16)$$

A total $SU(N_f)$ flavour symmetry would arise, if

$$\bar{q}' M q' = \bar{q} U^\dagger M U q = \bar{q} U^\dagger U M q = \bar{q} M q. \quad (4.17)$$

This would be the case, if all U commute with M , which means $M = m\mathbb{1}$. Then all quark masses are equal. If some quark masses are approximately equal, like $m_u \approx m_d$, one has an approximate symmetry in their subspace, e.g. the $SU(2)$ -isospin symmetry or the $SU(3)$ -flavour symmetry for $m_u \approx m_d \approx m_s$. In these cases the unitary transformation has the form

$$U = \left(\begin{array}{c|c} V & 0 \\ \hline 0 & \mathbb{1} \end{array} \right), \quad V \in SU(n). \quad (4.18)$$

4.1.3 Chiral symmetry

For processes at very high energies the masses of the quarks are negligible. This motivates another symmetry, which holds, if all quark masses vanish, $m_f = 0$. The transformations, which act in flavour and Dirac space, are called axial transformations and read

$$q_f \longrightarrow q'_f = [\exp(-i\omega^a T_a \gamma_5)]_{ff'} q_{f'}. \quad (4.19)$$

The invariance of the kinetic term relies on the relation

$$[\gamma^\mu, \gamma_5]_+ = 0. \quad (4.20)$$

Let us check this:

$$\begin{aligned} \gamma^0 \gamma_5 &= \begin{pmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{pmatrix} \begin{pmatrix} 0 & \mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix} = \begin{pmatrix} 0 & \mathbb{1} \\ -\mathbb{1} & 0 \end{pmatrix} = - \begin{pmatrix} 0 & \mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix} \begin{pmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{pmatrix} = -\gamma_5 \gamma^0 \\ \gamma^k \gamma_5 &= \begin{pmatrix} 0 & \sigma_k \\ -\sigma_k & 0 \end{pmatrix} \begin{pmatrix} 0 & \mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix} = \begin{pmatrix} \sigma_k & 0 \\ 0 & -\sigma_k \end{pmatrix} = - \begin{pmatrix} 0 & \mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix} \begin{pmatrix} 0 & \sigma_k \\ -\sigma_k & 0 \end{pmatrix} = -\gamma_5 \gamma^k \end{aligned}$$

The axial transformations do not form a group:

$$e^{-i\omega^a T_a \gamma_5} e^{-i\omega'^b T_b \gamma_5} \neq e^{-i\omega''^a T_a \gamma_5}. \quad (4.21)$$

They are elements of a larger group, the chiral symmetry group, which we shall discuss now.

In order to find the chiral symmetry group, we introduce the notion of chirality. The *chiral projections* are defined by

$$P_L : \quad q(x) \longrightarrow q_L(x) := \frac{1}{2}(1 - \gamma_5)q, \quad (4.22)$$

$$P_R : \quad q(x) \longrightarrow q_R(x) := \frac{1}{2}(1 + \gamma_5)q. \quad (4.23)$$

We speak of *left-handed* fields q_L or *right-handed* fields q_R .

We see immediately that

$$q(x) = q_L(x) + q_R(x) \quad \text{or} \quad P_L + P_R = \mathbf{1}. \quad (4.24)$$

P_L and P_R are projection operators

$$P_{R,L}^2 = \frac{1}{4}(1 \pm \gamma_5 \pm \gamma_5 + \underbrace{\gamma_5 \gamma_5}_{=1}) = \frac{1}{4}(2 \pm 2\gamma_5) = P_{R,L}. \quad (4.25)$$

One also has

$$P_L P_R = P_R P_L = 0, \quad (4.26)$$

from

$$(1 - \gamma_5)(1 + \gamma_5) = 1 - \gamma_5^2 = 0. \quad (4.27)$$

For high energetic particles described by a Dirac wavefunction, $P_L \psi = \psi$ or $P_R \psi = \psi$ means that the momentum of the particle is parallel (right-handed) or antiparallel (left-handed) to the direction of motion.

We shall make use of the following relations.

$$\bar{q}_L = \bar{q} \cdot \frac{1}{2}(1 + \gamma_5), \quad (4.28)$$

$$\bar{q}_R = \bar{q} \cdot \frac{1}{2}(1 - \gamma_5), \quad (4.29)$$

Check: since $1 - \gamma_5$ is Hermitian and $\gamma_5 \gamma^0 = -\gamma^0 \gamma_5$, one has

$$\bar{q}_L = (\frac{1}{2}(1 - \gamma_5)q)^\dagger \gamma^0 = q^\dagger \frac{1}{2}(1 - \gamma_5) \gamma^0 = q^\dagger \gamma^0 \frac{1}{2}(1 + \gamma_5) = \bar{q} \frac{1}{2}(1 + \gamma_5),$$

and similarly for $\bar{q}_R$. Also

$$\bar{q}_L q_L = \bar{q}_R q_R = 0 \quad (4.30)$$

holds due to $(1 + \gamma_5)(1 - \gamma_5) = 0$. The mass term can thus be decomposed as

$$\bar{q}q = (\bar{q}_L + \bar{q}_R)(q_L + q_R) = \bar{q}_L q_R + \bar{q}_R q_L. \quad (4.31)$$

On the other hand, for the terms in the kinetic part of the Lagrangian we find

$$\begin{aligned} \bar{q}_L \gamma^\mu q_R &= \frac{1}{4} \bar{q} (1 + \gamma_5) \gamma^\mu (1 + \gamma_5) q = \frac{1}{4} \bar{q} (1 + \gamma_5) (1 - \gamma_5) \gamma^\mu q = 0, \\ \bar{q}_R \gamma^\mu q_L &= \frac{1}{4} \bar{q} (1 - \gamma_5) \gamma^\mu (1 - \gamma_5) q = \frac{1}{4} \bar{q} (1 - \gamma_5) (1 + \gamma_5) \gamma^\mu q = 0. \end{aligned}$$

Using these relations the Lagrangian density for QCD can be written in terms of the left-handed and right-handed fields as

$$\mathcal{L} = \sum_f \{ \bar{q}_{fL} i \gamma^\mu D_\mu q_{fL} + \bar{q}_{fR} i \gamma^\mu D_\mu q_{fR} - m_f (\bar{q}_{fL} q_{fR} + \bar{q}_{fR} q_{fL}) \} - \frac{1}{4} G_{\mu\nu}^a G^{\mu\nu,a}. \quad (4.32)$$

The mass terms couple the left-handed to the right-handed fields. If all quark masses vanish, the left-handed and right-handed fields decouple. In this case there is a larger symmetry, which acts on the left-handed and the right-handed fields independently. The symmetry transformations are

$$q_L \longrightarrow q'_L = e^{-i\omega_L^a T_a} q_L, \quad e^{-i\omega_L^a T_a} \in \text{SU}(N_f)_L, \quad (4.33)$$

$$q_R \longrightarrow q'_R = e^{-i\omega_R^a T_a} q_R, \quad e^{-i\omega_R^a T_a} \in \text{SU}(N_f)_R. \quad (4.34)$$

Both groups $\text{SU}(N_f)_L$ and $\text{SU}(N_f)_R$ are isomorphic to $\text{SU}(N_f)$. The indices L or R just indicate that the symmetry transformations act on the left-handed and the right-handed fields, respectively. The invariance of the kinetic part of the Lagrangian holds because of

$$\bar{q}'_L i \gamma^\mu D_\mu q'_L = \bar{q}_L i \gamma^\mu D_\mu q_L, \quad (4.35)$$

$$\bar{q}'_R i \gamma^\mu D_\mu q'_R = \bar{q}_R i \gamma^\mu D_\mu q_R, \quad (4.36)$$

whereas in the mass terms the exponentials do not cancel, i.e.

$$\bar{q}'_L q'_R = \bar{q}_L e^{i\omega_L^a T_a} e^{-i\omega_R^a T_a} q_R \neq \bar{q}_L q_R.$$

So, for $m_f = 0$ there is a symmetry of the Lagrangian under the group

$$\text{SU}(N_f)_L \otimes \text{SU}(N_f)_R, \quad (4.37)$$

which has $2(N_f^2 - 1)$ parameters. This symmetry is called chiral symmetry.

The currents, belonging to this symmetry according to Noether's theorem are

$$j_{La}^\mu = \bar{q}_L \gamma^\mu T_a q_L, \quad (4.38)$$

$$j_{Ra}^\mu = \bar{q}_R \gamma^\mu T_a q_R. \quad (4.39)$$

For example, for the $SU(N_f)_L$ transformations we find

$$j_{La}^\mu = \frac{\partial \mathcal{L}}{\partial (\partial_\mu q_{Lf})} \frac{\delta q_{Lf}}{\delta \omega_L^a} = \bar{q}_{Lf} i \gamma^\mu (-i T_a) q_{Lf} = \bar{q}_L \gamma^\mu T_a q_L,$$

and similarly for the right handed transformations. Using

$$P_R \gamma^\mu P_L = \frac{1}{2} (1 + \gamma_5) \gamma^\mu \frac{1}{2} (1 - \gamma_5) = \gamma^\mu \frac{1}{2} (1 - \gamma_5) \frac{1}{2} (1 - \gamma_5) = \gamma^\mu P_L^2 = \gamma^\mu P_L,$$

the currents can be written

$$j_{La}^\mu = \bar{q} \gamma^\mu T_a \frac{1}{2} (1 - \gamma_5) q, \quad (4.40)$$

$$j_{Ra}^\mu = \bar{q} \gamma^\mu T_a \frac{1}{2} (1 + \gamma_5) q. \quad (4.41)$$

Relation to flavour symmetry

The flavour and chiral symmetries are related,

$$\text{flavour } SU(N_f) \longleftrightarrow \text{chiral } SU(N_f)_L \otimes SU(N_f)_R. \quad (4.42)$$

To see this, consider the infinitesimal transformations

$$\begin{aligned} q'_L &= (\mathbb{1}_L - i \omega_L^a T_a) q_L, \\ q'_R &= (\mathbb{1}_R - i \omega_R^a T_a) q_R, \\ q' &= (q'_L + q'_R) = (\mathbb{1} - i T_a [\omega_L^a P_L + \omega_R^a P_R]) q. \end{aligned} \quad (4.43)$$

If we restrict the chiral transformations to

$$\omega_L^a = \omega_R^a \equiv \omega^a, \quad (4.44)$$

we obtain the infinitesimal flavour transformations

$$q' = (\mathbb{1} - i \omega^a T_a) q. \quad (4.45)$$

Therefore, the flavour symmetry group is a subgroup of the chiral symmetry group $SU(N_f)_L \otimes SU(N_f)_R$. It is called the *diagonal subgroup*, because it consists of those group elements $U_L \otimes U_R$ which obey $U_L = U_R$.

$$\begin{aligned} \text{flavour group} &= \{U_L \otimes U_R \mid U_L, U_R \in SU(N_f), U_L = U_R\} \\ &\simeq SU(N_f) \subset SU(N_f) \otimes SU(N_f) \end{aligned} \quad (4.46)$$

The *flavour current* belonging to the flavour symmetry is

$$j_a^\mu = \bar{q}\gamma^\mu T_a q = j_{La}^\mu + j_{Ra}^\mu. \quad (4.47)$$

The flavour current is also called vector current, since it transforms as a Lorentz vector.

Infinitesimal *axial transformations* are defined by

$$\omega_L^a = -\omega_R^a \equiv -\Omega^a \quad (4.48)$$

and read ($P_L - P_R = -\gamma_5$)

$$\begin{aligned} q' &= (1 - iT_a[\omega_L P_L + \omega_R P_R])q \\ &= (1 - iT_a(-\Omega^a)(P_L - P_R))q \\ &= (1 - i\Omega^a \gamma_5 T_a)q. \end{aligned} \quad (4.49)$$

The corresponding *axial current* or axial vector current is given by

$$j_{5a}^\mu := \bar{q}\gamma^\mu \gamma_5 T_a q = j_{Ra}^\mu - j_{La}^\mu. \quad (4.50)$$

It transforms as an axial vector under Lorentz transformations, i.e. if the axial vector current is looked at through a mirror, symbolised by the space reflection (parity) operator P , then left and right would be interchanged, and the current changes sign: $P(j_{Ra}^\mu - j_{La}^\mu) = -j_{Ra}^\mu + j_{La}^\mu$.

4.1.4 Broken chiral symmetry

Symmetries of a physical theory are not just mathematical curiosities. In general they give strong restrictions on the structure and behaviour of the corresponding models. In case of a field theory, if the Lagrangian possesses a certain symmetry and if this symmetry is also realised by the physical states, then as a consequence various relations between observables emerge.

Now we shall consider the question, in which way chiral symmetry of the QCD Lagrangian is realised in nature. If chiral symmetry were realised on the physical states, then e. g. the mesons and baryons should form multiplets of the symmetry group, such that the members of a multiplet have the same mass.

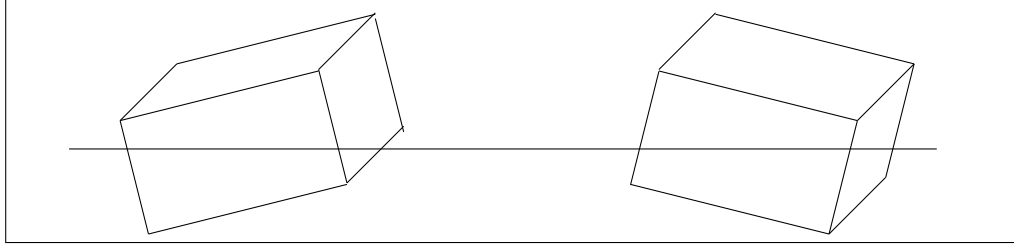
Consider the case $N_f = 2$ with the chiral $SU(2)_L \otimes SU(2)_R$ symmetry. The flavour symmetry subgroup, the $SU(2)$ isospin symmetry, is an approximate symmetry in nature, as is shown by the up and down quarks, by the nucleons p , n , and by the pions $\pi^{\pm,0}$.

On the other hand the chiral $SU(2) \otimes SU(2)$ group is not found as a symmetry group in nature. One can show that this symmetry would imply that to each isospin multiplet of particles with definite parity there should belong another isospin multiplet with the same mass but opposite parity. This is, however, not observed in nature. For instance, the pions form an isospin triplet of pseudoscalar particles, but there is no isospin triplet of scalar particles with masses near the pion masses. The same holds for other $SU(N)$ -multiplets.

If the Lagrangian $\mathcal{L}$ possesses a symmetry that is not realised in nature, one speaks of a “spontaneously broken symmetry”, in contrast to an explicitly broken symmetry, where the breaking results from symmetry breaking terms in the Lagrangian.

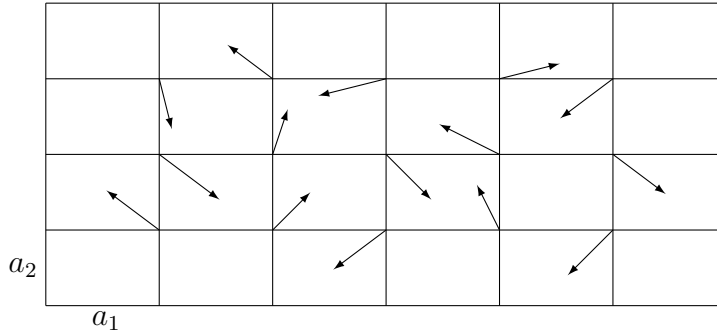
Spontaneously broken symmetries are already familiar from everyday life and classical physics, as the following examples show.

a) Swimming



If one puts a rectangular piece of wood in water, it can take two positions, in which one end of the wood is immersed deeper than the other (supposed the density of the wood is not too different from the density of water). There is obviously a reflection symmetry in the Lagrangian, which is broken by the outcome of the experiment.

b) Ferromagnet



A ferromagnet shows spontaneous magnetisation below the Curie point, $T < T_C$. The directions of the molecular magnetic moments $\vec{m}(\vec{r})$ give up their random distribution in favour of one macroscopic direction, although the

Hamiltonian

$$H = -J \sum_{\vec{r}, \mu=1,2,3} \vec{m}(\vec{r}) \cdot \vec{m}(\vec{r} + \vec{a}_\mu) \quad (4.51)$$

obeys rotation symmetry according to

$$\vec{m}(\vec{r}) \longrightarrow R \vec{m}(\vec{r}), \quad R \in SO(3). \quad (4.52)$$

The average values of the molecular magnetic moments turn out to be non-zero:

$$\langle \vec{m}(\vec{r}) \rangle = \vec{m}_0 \neq 0. \quad (4.53)$$

In addition to explicitly and spontaneously broken symmetries, there is another kind of symmetry breaking in quantum theory, which originates from a pure quantum effect. An example of such a broken symmetry is given by the axial $U(1)_A$ symmetry

$$q(x) \longrightarrow q'(x) = e^{-i\alpha\gamma_5} q(x). \quad (4.54)$$

The corresponding current density

$$j_5^\mu(x) = \bar{q}(x) \gamma^\mu \gamma_5 q(x) \quad (4.55)$$

is an axial vector current. If $m_f = 0$ then $U(1)_A$ is a symmetry of the Lagrangian. Classically the axial vector current is conserved. But this symmetry is not realised in nature. (We shall not discuss the arguments for this fact here.) It turns out that in the quantum theory the conservation of the axial vector current is violated,

$$\partial_\mu j_5^\mu \neq 0. \quad (4.56)$$

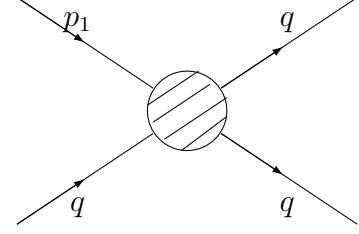
A situation, where a classically conserved current is not conserved in the quantum theory, is called an “anomaly”. So, anomalies are different from spontaneous symmetry breaking, they are quantum effects.

4.2 Running Coupling

The running of the QCD coupling is very important for phenomenology and experiments. A thorough discussion would require to treat renormalisation and the Callan-Symanzik equations. Here we just give a short outline of this topic, presenting some results and experimental evidence for the varying coupling “constants” of QCD.

4.2.1 Quark-quark scattering

We start with a perturbative theory along the lines of QED. There is no 4-quark diagram, so we expand the graph for quark-quark scattering in increasing order of the coupling. The simplest term corresponds to the Coulomb scattering of electrically charged particles, i. e. Rutherford scattering.

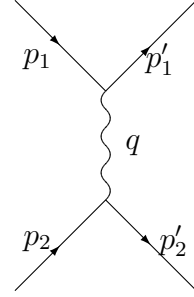


The simplest, tree level diagram describes a one gluon exchange. The mediating gluon carries 4-momentum

$$q = p'_1 - p_1 = -(p'_2 - p_2). \quad (4.57)$$

q is a spacelike vector,

$$q^2 < 0. \quad (4.58)$$



The scattering amplitude is determined by the gluon propagator and two vertices

$$\mathcal{M} = -g^2 \bar{u}(p'_1) \gamma_\mu u(p_1) \frac{1}{q^2} \bar{u}(p'_2) \gamma_\mu u(p_2) \frac{1}{3} \begin{cases} 1 & \text{rep. 6} \\ -2 & \text{rep. 3}^* \end{cases}, \quad (4.59)$$

where $u(p_i)$ are spinors, denoting solutions of the Dirac equation, and the simple form of the gluon propagator $1/q^2$ is due to Feynman gauge. This factor is dominant for the behaviour of the scattering.

Let us compare this with the $e^- \mu^-$ scattering in QED, where the muon mass is about 200 times the electron mass, so we are near the static limit of Rutherford scattering. Therefore, $|p_1| = |p'_1| =: k$ and

$$\mathcal{M} \sim \frac{1}{q^2} \approx -\frac{1}{4k^2 \sin^2(\frac{\theta}{2})}. \quad (4.60)$$

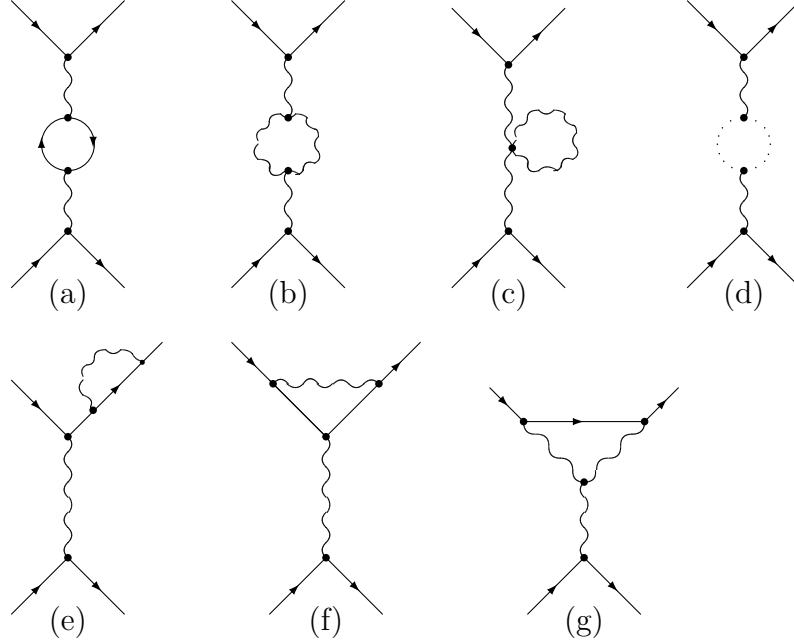
The differential cross-section then obeys

$$\frac{d\sigma}{d\Omega} \sim |\mathcal{M}|^2 \sim \frac{1}{\sin^4(\frac{\theta}{2})}. \quad (4.61)$$

Quantum mechanics tells us that in the Born-approximation $\mathcal{M}$ is proportional to the Fourier transformation (FT) of the potential, thus

$$\frac{1}{q^2} \sim FT(V(r)) \Rightarrow V(r) \sim \frac{1}{r}. \quad (4.62)$$

Thus, in lowest order of perturbation theory the quark-quark potential is a Coulomb potential. Since we only considered one-gluon exchange, we should not be surprised to get the same result as in first order of perturbative QED. In QCD, higher orders of perturbation theory add corrections to $1/q^2$, which will modify this result. The 1-loop corrections are the following:



The diagrams are denoted as: quark-loop (a), gluon-loop (b), gluon-tadpole (c), ghost-loop (d), quark-propagator correction (e), and vertex corrections (f), (g). Additional graphs, which are similar to (e) – (f) are not shown.

The result of these calculations in the limit of large momentum transfer,

$$Q^2 := -q^2 \quad \text{large}, \quad (4.63)$$

is essentially a modification of the one-gluon exchange case, with a corrected coupling constant

$$g^2 \longrightarrow g^2 \left\{ 1 - b_0 \frac{g^2}{16\pi^2} \ln \left(\frac{Q^2}{M^2} \right) \right\}, \quad (4.64)$$

with

$$b_0 = \frac{1}{3}(11N_c - 2N_f) = 11 - \frac{2}{3}N_f > 0 \quad (\text{for } N_c = 3 \text{ colours}). \quad (4.65)$$

b_0 is positive as long as the number of flavours is not larger than 16, and there is no experimental evidence for that. The number 11 comes from the graphs (b) to (g), while $\frac{2}{3}N_f$ stems from graph (a).

The expression contains a mass M . It enters the result due to renormalisation. The loops in the Feynman diagrams involve divergent integrals over the momenta of the type

$$\int \frac{d^4 p}{p^4}. \quad (4.66)$$

The renormalisation procedure removes the divergencies in two steps,

1. Regularisation: introduces a cutoff for the momentum integration, or analogous divergent integrals. Typical regularisations introduce a cutoff mass M , which characterises the largest allowed momenta. Loop integrals are then finite, but depend on M .
2. Renormalisation.

4.2.2 Renormalisation

The coupling g is a parameter in the Lagrangian and is not directly measurable. We call it “bare” coupling and from now on denote it by g_0 . A measurable, physical relevant coupling should be defined or fixed through some process. Here, for example, we fix it by choosing a fixed mass scale μ and considering quark-quark scattering at a momentum transfer $Q^2 = \mu^2$. The modified coupling at this momentum transfer is defined to be the renormalised coupling

$$g_R^2(\mu) = g_0^2 \left\{ 1 - b_0 \frac{g_0^2}{16\pi^2} \ln \left(\frac{\mu^2}{M^2} \right) + O(g_0^4) \right\}. \quad (4.67)$$

Renormalisation now means

- express everything in terms of g_R in place of g_0 ,
- then remove the cutoff by taking the limit $M \rightarrow \infty$.

Renormalisable theories are those which give finite physical results by this procedure. (Eventually other parameters like quark masses, etc., have to be renormalised, too.)

In our case, expressing g_0 in terms of $g_R(\mu)$ and expanding in powers of $g_R(\mu)$, one gets the scattering amplitude

$$\mathcal{M} \sim \frac{1}{Q^2} g_R^2(\mu) \left\{ 1 - b_0 \frac{g_R^2(\mu)}{16\pi^2} \ln \left(\frac{Q^2}{\mu^2} \right) + O(g_R^4) \right\}. \quad (4.68)$$

Indeed, $\mathcal{M}$ does not longer contain the cutoff M , and is finite.

4.2.3 Running coupling

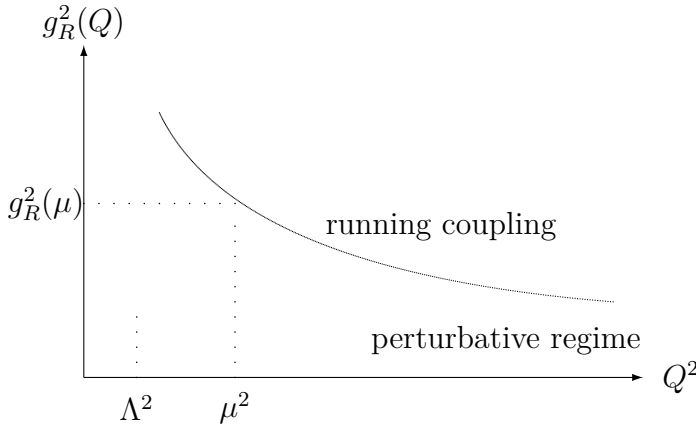
The result for $\mathcal{M}$ suggests to define an *effective coupling* for this process,

$$g_R^2(Q) := g_R^2(\mu) \left\{ 1 - b_0 \frac{g_R^2(\mu)}{16\pi^2} \ln \left(\frac{Q^2}{\mu^2} \right) + O(g_R^4) \right\}, \quad (4.69)$$

which indicates that the scattering at momentum transfer Q^2 is like Rutherford scattering with an effective coupling $g_R^2(Q)$. We see that the effective coupling is not constant. Therefore it is called “running coupling”. Moreover, in the one-loop expression above the corrections get large for large Q^2 . (Throughout our calculation we have made the assumption of large Q^2 .) Therefore, we have to extend the perturbation series. Calculating higher orders of perturbation, the dominant contributions can be found to lead to a geometric series, so

$$g_R^2(Q) \sim g_R^2(\mu) \left[1 + b_0 \frac{g_R^2(\mu)}{16\pi^2} \ln \left(\frac{Q^2}{\mu^2} \right) \right]^{-1}. \quad (4.70)$$

The figure shows the dependence of the running coupling on Q^2 .



The coupling gets weaker for higher momentum transfer, supporting the suggestion that perturbation theory is reliable at large Q .

A more solid consideration is based on the renormalisation group, in particular the Callan-Symanzik equations, which are beyond the scope of this lecture. The renormalisation group equation for the running coupling is

$$Q \frac{d}{dQ} g_R = \beta(g_R) \quad (4.71)$$

$$= -b_0 \frac{g_R^3}{16\pi^2} + \dots \text{ (odd powers),} \quad (4.72)$$

where $\beta(g_R)$ is the famous beta-function of QCD. Thus we have

$$Q \text{ increases} \longrightarrow g_R \text{ decreases.} \quad (4.73)$$

So, perturbation theory is reliable at large Q .

Here, a remark is in order. The perturbation series of QCD or QED are not convergent in the mathematical sense. However, the error made by truncating the series at some finite order can be controlled due to Borel summability. This allows to get useful results from finite orders of perturbation theory in a suitable regime of applicability, which in QCD is the high energy regime.

Let us now consider the region of small Q , or the low energy range. In QCD, $b_0 > 0$, therefore, for small enough Q , $\ln(Q^2/\mu^2)$ becomes negative, and in the considered approximation, $g_R(Q)$ gets a pole at a certain point $Q = \Lambda$.

The QCD Λ -parameter is (roughly) in one-loop order defined by

$$b_0 \frac{g_R^2(\mu)}{16\pi^2} \ln \left(\frac{\Lambda^2}{\mu^2} \right) = -1. \quad (4.74)$$

Its value (in a particular scheme, the $\overline{MS}$ scheme) is about $\Lambda_{\overline{MS}} \approx 200$ MeV.

The apparent pole at $Q = \Lambda$ is actually an artefact of the approximation and is not present for the full $g_R(Q)$. A more precise definition of Λ , which requires two-loop terms, can nevertheless be given.

Using the Λ -parameter, for $\mu \approx \Lambda$ and $Q^2 \gg \Lambda^2$, the high energy behaviour of the coupling constant can be written as

$$g_R^2(\mu) = \frac{16\pi^2}{b_0 \ln \left(\frac{Q^2}{\Lambda^2} \right)}. \quad (4.75)$$

This decrease of the running coupling towards zero is called “*asymptotic freedom*”,

$$\lim_{Q \rightarrow \infty} g_R(Q) = 0. \quad (4.76)$$

Asymptotic freedom in QCD was first realised by Gerard 't Hooft in the work on his thesis as a student of Martinus Veltman, but was not published. In 1973 asymptotic freedom in field theories has been discovered by D. J. Gross, F. Wilczek and H. D. Politzer, who were awarded the Nobel prize in 2004.

4.2.4 Discussion

a) Asymptotic freedom guaranties a good applicability of perturbation theory on processes at high energies or momenta. At low energies perturbation calculation fails.

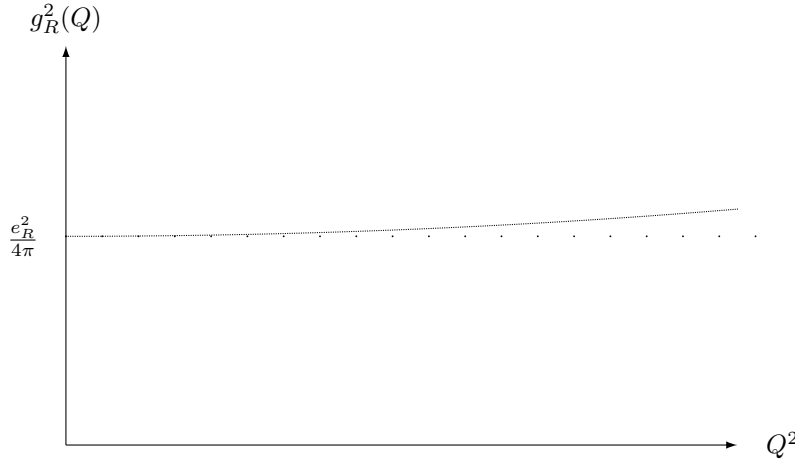
b) In QED the renormalisation group parameter is $b_0 < 0$. This comes from the 1-loop calculation of $e^- e^+$ scattering. Taking the electric charge for the coupling constant, one has

$$e_R^2(Q) = e_R^2 \left[1 - |b_0| \frac{e_R^2}{16\pi^2} \ln \left(\frac{Q^2}{m_e^2} \right) \right]^{-1}. \quad (4.77)$$

In QED the running fine-structure constant

$$\alpha = \frac{e^2(Q)}{4\pi\epsilon_0\hbar c} = \frac{e_R^2(Q)}{4\pi} \approx \frac{1}{137.036}$$

shows a tiny increase with momentum.



Measurements at 100 GeV have given a value of $\alpha^{-1} \approx 120$.

The Λ -point (“Landau pole”) belongs to a momentum already beyond the Planck-scale.

c) Interpretation

Since $Q \sim \frac{1}{r}$ and $b_0 < 0$ in QED, the electric charge $e_R(1/r)$ increases as $r \rightarrow 0$. The decrease of the measured charge, when going away from $r = 0$, has been interpreted as a screening of charges due to vacuum polarisation.

In QCD on the other hand, where $b_0 > 0$, there are also gluon loops and the gluons have self-interactions. This leads to an anti screening effect.

d) Experiments

The strong “fine-structure constant” $\alpha_S = \frac{g_R^2(Q)}{4\pi}$ of QCD has been measured in several experiments:

- deep inelastic $e^- p$ scattering,

- τ -lepton decay,
- e^+e^- scattering, which has contributions of virtual $q\bar{q}$ processes,
- Ψ , Υ (heavy quarkonia) spectroscopy,
- Z^0 -decay.

The results are in very good agreement with the prediction from asymptotic freedom.

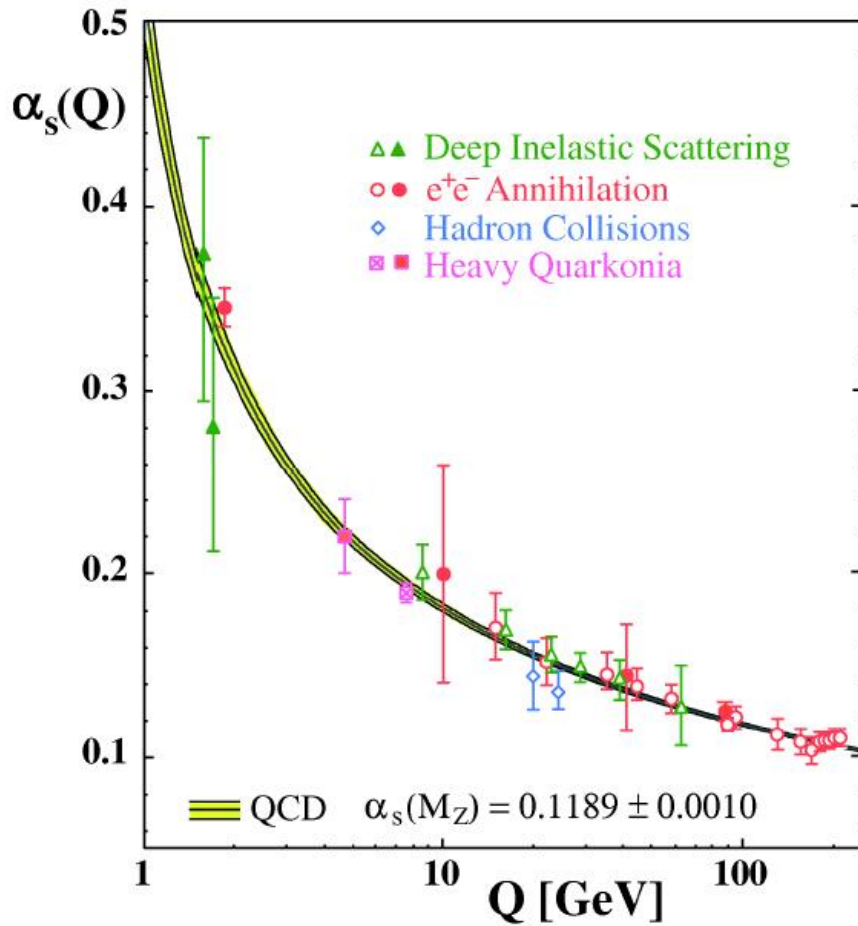


Figure 3: Summary of measurements of $\alpha_s(Q)$, from S. Bethke, Prog. Part. Nucl. Phys. 58 (2007) 351–386, hep-ex/0606035.

4.3 Confinement of Quarks and Gluons

In the QCD Lagrangian the quarks and gluons carry colour charge, but these particles do not exist as free particles. Accessible particles are the hadrons: mesons ($q\bar{q}$) as quark-antiquark states and baryons (qqq), made of three quarks. All these particles are colour neutral. Therefore, one states the

Confinement Hypothesis:

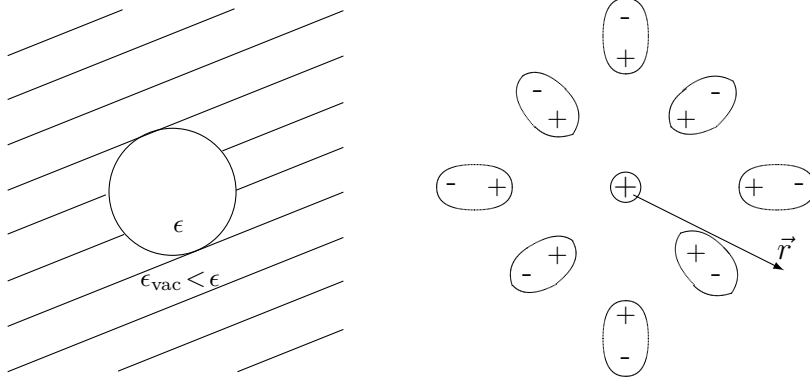
Physical states are colour neutral;

in particular, quarks and gluons do not exist as free particles.

Since perturbation theory works with quark and gluon fields, and assumes these constituents as ingoing and outgoing particles, confinement cannot be proven in the framework of perturbation theory. To demonstrate confinement is a non-perturbative problem. A rigorous proof of confinement is still missing.

Instead, there are some phenomenological descriptions.

- The bag model
is based on antiscreening effects. The particles are assumed to be confined in a region with boundary conditions similar to boundary conditions of electrodynamics. In electrodynamics or optics one has the polarisability of the matter, which in the bag model is replaced by the polarisability of vacuum.



- In the string picture
the quark-antiquark potential at large distances grows proportional to the distance ²³,

$$V(r) = kr. \quad (4.78)$$

²³For the static QCD potential see G.S. Ball, Phys. Rept. **343** (2001) 1.

The force, therefore, remains constant, and when one tries to separate the quarks, energy grows until it suffices to generate a new quark-antiquark pair. The situation may be depicted by lines of force, similar to electrodynamics, with the exception that the vacuum compresses the lines to flux tubes, which eventually split. Thus the *chromoelectric flux tubes* are squeezed by being “repelled” from the vacuum.

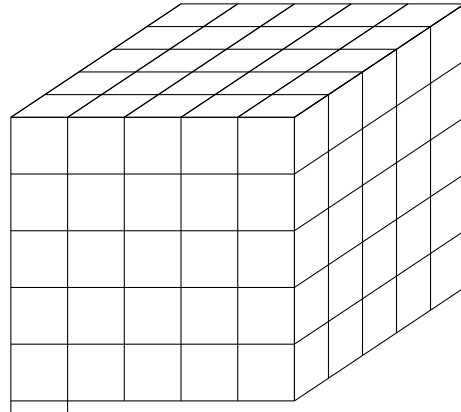


The bag model and the string model are based on postulated dielectric properties of the vacuum with a (relative) dielectric constant

$$\epsilon_r < 0 \longleftrightarrow \text{antiscreening.} \quad (4.79)$$

Lattice QCD

is able to provide the spectrum of particles and many other physical quantities from first principles. The calculated hadron masses are in very good agreement with the experimental values. Practical limitations are the number of lattice points and/or the lower limit to the lattice spacing, which typically is of the order of 0.1 fm. This also puts an upper limit to momenta.



$64^3 \times 128$
lattice points

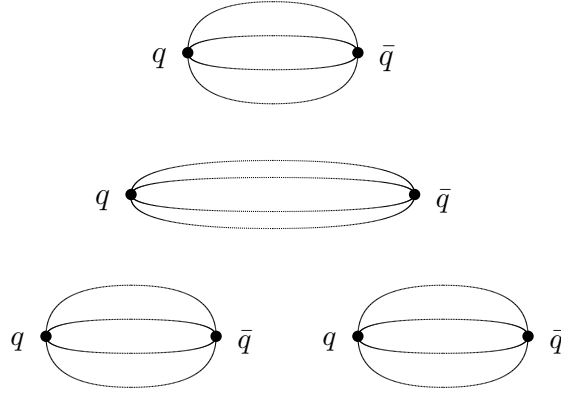
is

In *pure gauge theory*, for a gluon field with fixed quarks, the static quark-antiquark potential can rather well be described by

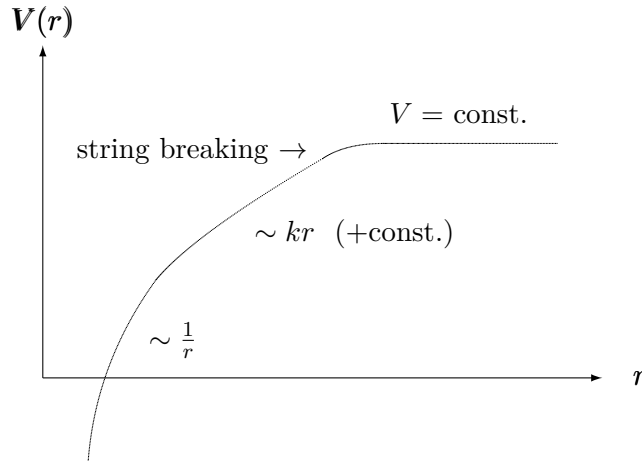
$$V(r) = \frac{c}{r} + k \cdot r + \text{const.} \quad (4.80)$$

Here the c/r term comes from perturbation theory, which is valid at small distances. Lattice calculations at large r lead to the static confinement $V(r) \sim r$.

Lattice QCD with quarks supports the string picture as it results in string breaking at larger quark distance.



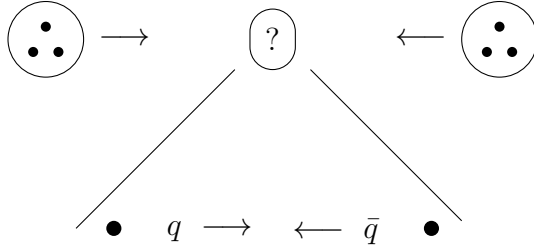
In this case the quark-quark potential is of the form



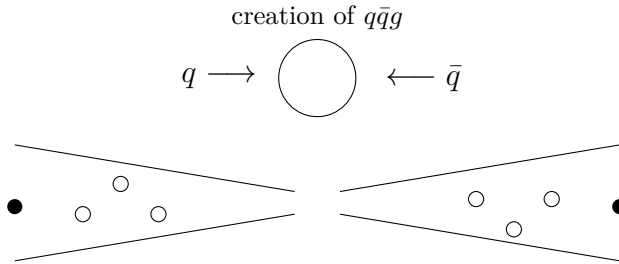
Lattice simulations in pure gauge theory reveal pure gluonic, colour neutral particles, the so-called glue-balls. These also exist in full QCD with quarks.

Jets

In perturbation theory, scattering processes in QCD are based on the scattering of quarks and gluons. On the other hand, all experimental information stems from the interaction of observable particles, namely hadrons. Therefore, the perturbative results for scattering processes in QCD have to be “hadronised”. Consider for example proton-antiproton scattering.



Perturbation theory gives cross-sections for quark-quark scattering. Hadronisation turns the outgoing quarks and antiquarks into hadrons. This leads to the formation of jets of hadrons (S. Weinberg, G. Sterman). At high momenta the primary quarks and antiquarks combine with newly created quarks and antiquarks and form jets.



The formation of jets is a non-perturbative process.

4.4 Experimental Evidence for QCD

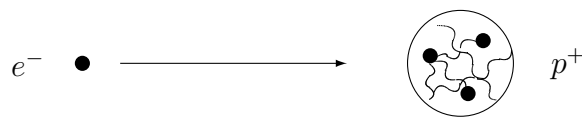
In general, there is a good consistency between the theoretical and experimental results for QCD. In the following we list some examples, where experiments agree well with theoretical modelling and/or predictions.

- Running coupling $\alpha_S(Q)$.
- Jet distribution.

The statistical angular distribution and the momentum distributions of the particles in jets were quantitatively predicted by QCD.

- Evidence for gluons.

– Deep inelastic e^-p^+ scattering at HERA.



It is found that quarks carry only about 50 % of the momentum of the nucleon. The other half has to come from particles, which are neutral with respect to electroweak interactions. This is an indication for gluons.

- Three-jet events.

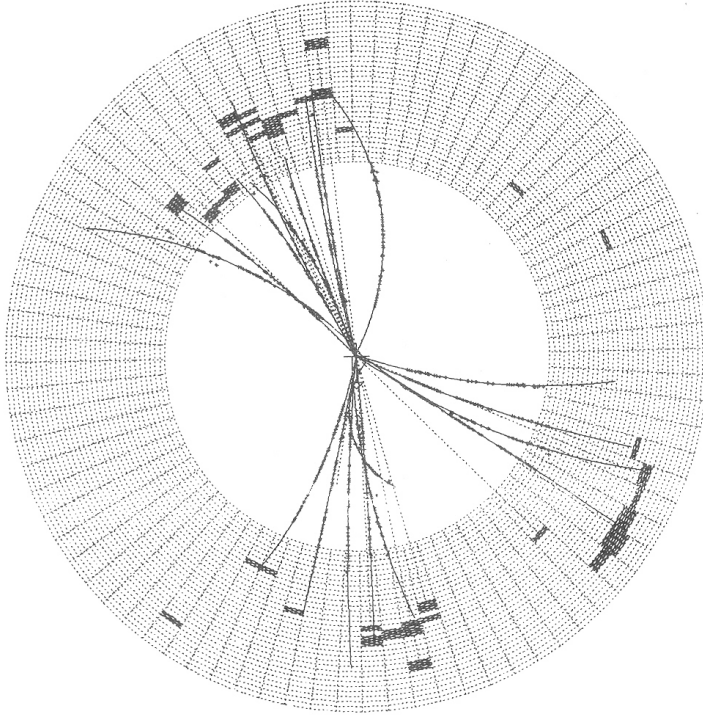


Figure 4: A three-jet event observed at PETRA. The figure is from: F. Halzen, A. D. Martin, *Quarks & Leptons: An Introductory Course in Modern Particle Physics*, J. Wiley & Sons (1984).

They are explained by additional particles. The angular distribution of the jets leads to spin 1 particles, which are identified with the gluons, in agreement with the QCD prediction.

- Lattice QCD.

The masses of stable hadrons have been calculated in lattice QCD. The masses $m_u = m_d$, m_s of up, down, and strange quarks, which are needed as input parameters, have been obtained by fitting the proton and Kaon masses, respectively. The lattice simulations from the Budapest-Marseille-Wuppertal collaboration have given the baryon

masses for N , Λ , Σ , Δ , Σ^* , Ξ , Ω , and the masses of the mesons ρ and K^* . The results are within the experimental errors.

There have been many other experiments that support QCD theory. Let us mention one more example: Calculating the anomalous magnetic moment of the muon in QED in the same way, as it was done for the electron, the result does not agree with the extremely precise experimental value. However, taking hadronic contributions from QCD into account, the agreement is perfect.

5 Electroweak Theory

5.1 Weak Interactions

The weak interactions of elementary particles are distinguished from other interactions by some characteristic properties like lifetimes, strength of coupling, cross-sections, and violation of symmetries. We refer to the introductory lectures on particle physics for details.

Some typical processes of weak interactions are the following.

a) Leptonic processes

Muon decay: $\mu^- \longrightarrow e^- + \bar{\nu}_e + \nu_\mu$

$e\nu$ -scattering: $e^- \nu_\mu \longrightarrow \mu^- + \nu_e$

b) Semi-leptonic processes. They involve hadrons.

β -decay $n \longrightarrow p^+ + e^- + \bar{\nu}_e$

β -decay in quark picture

$d \longrightarrow u + e^- + \bar{\nu}_e$

Pion decay²⁴ $\pi^- \longrightarrow \mu^- + \bar{\nu}_\mu$ ($d\bar{u} \longrightarrow \mu^- + \bar{\nu}_\mu$)

c) Non-leptonic weak interaction processes

Λ decay $\Lambda^0 \longrightarrow p^+ + \pi^-$ ($uds \longrightarrow uud + d\bar{u}$)

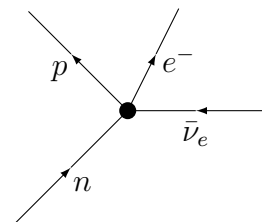
Kaon decay $K^- \longrightarrow \pi^- + \pi^0$ ($s\bar{u} \longrightarrow d\bar{u} + \frac{1}{\sqrt{2}}(-u\bar{u} + d\bar{d})$)

These weak interaction processes violate isospin symmetry.

5.1.1 Fermi theory of weak interaction

As mentioned in the introduction, in 1932 Enrico Fermi formulated a theory for the β -decay of the neutron as a four fermion process. The Lagrangian consists of the free parts for n , p^+ , ν , and e^- , plus the four-fermion interaction term

$$\mathcal{L}_w = G (\bar{e}(x)\gamma^\mu\nu_e(x)) (\bar{p}(x)\gamma_\mu n(x)). \quad (5.1)$$



$\mathcal{L}_w$ is invariant under space reflections. (Parity operator P.) In Fermi's theory parity is conserved.

5.1.2 Parity violation

In 1957 parity violation, occurring in the β -decay of ^{60}Co , was experimentally established by Wu. At that time, when the neutrinos were regarded to be

²⁴There also is a dominant hadronic decay mode giving the same particles. The leptonic mode is distinguished by its timescale.

massless particles, the ν_e was always considered to be left-handed, and the $\bar{\nu}_e$ right handed. Madame Wu found that the e^- arising in this decay is dominantly left-handed, violating parity conservation. Parity violation in Kaon-decay was proposed by Lee and Yang in 1966.

In order to describe parity violation and handedness of neutrinos, let us recall some facts about chirality from Sec. 4.1.3. The projection operators on left- and right-handed spinors are, respectively

$$P_L = \frac{1}{2}(1 - \gamma_5), \quad P_R = \frac{1}{2}(1 + \gamma_5). \quad (5.2)$$

With $\psi_L = P_L\psi$, $\psi_R = P_R\psi$, the Dirac Lagrangian is

$$\mathcal{L} = \bar{\psi}(i\gamma^\mu\partial_\mu - m)\psi \quad (5.3)$$

$$= \bar{\psi}_L i\gamma^\mu\partial_\mu\psi_L + \bar{\psi}_R i\gamma^\mu\partial_\mu\psi_R - m(\bar{\psi}_L\psi_R + \bar{\psi}_R\psi_L). \quad (5.4)$$

If $m = 0$, $\mathcal{L}$ separates into a left-handed and a right-handed part.

Consider a left-handed Dirac spinor plane wave solution

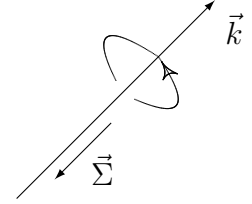
$$u(k) = u_L(k) = \frac{1}{2}(1 - \gamma_5)u(k), \quad (5.5)$$

$$\gamma_5 u_L(k) = \gamma_5 \frac{1}{2}(1 - \gamma_5)u(k) = \frac{1}{2}(\gamma_5 - 1)u(k) = -u_L(k), \quad (5.6)$$

since $(\gamma_5)^2 = 1$.

With the spin-operator $\vec{S} = (1/2)\vec{\Sigma}$ we define helicity λ to be the projection of spin onto the direction of momentum. In the massless case one can show that helicity and chirality are equal:

$$\vec{\Sigma} \cdot \frac{\vec{k}}{|\vec{k}|} u_L(k) = \gamma_5 u_L(k) = -u_L(k). \quad (5.7)$$



The helicity in this case is $\lambda = -\frac{1}{2}$.

Similarly $\gamma_5 u_R(\vec{k}) = +u_R(\vec{k})$, the particle is right-handed, and the helicity is $\lambda = +\frac{1}{2}$.

Since all neutrinos were found to be left-handed and all antineutrinos to be right-handed, we have

$$\nu = \nu_L = \frac{1}{2}(1 - \gamma_5)\nu, \quad (5.8)$$

$$\bar{\nu} = \bar{\nu}_R = \frac{1}{2}(1 + \gamma_5)\bar{\nu}. \quad (5.9)$$

The non-existence of right-handed neutrinos implies that parity is violated. Nowadays, experiments have shown that neutrinos are massive particles. Due to their very small mass, they almost always move with a speed very close to the speed of light. Nevertheless, in a moving frame that moves even faster, they appear to be right-handed. So, helicity (and chirality) can take both values.

5.1.3 V-A theory

The V-A theory was developed by Feynman and Gell-Mann in 1958 and independently by Marshak and Sudarshan in the same year. It assumes massless neutrinos and takes chirality and parity violation into account.

The neutrino part in the Lagrangian is replaced by²⁵

$$\bar{e}(x)\gamma_\mu\nu_e(x) \rightarrow \bar{e}(x)\gamma_\mu\frac{1}{2}(1-\gamma_5)\nu_e(x) \quad (5.10)$$

$$\begin{aligned} &= \bar{e}_L(x)\gamma_\mu\nu_{e,L}(x) \\ &= \underbrace{\frac{1}{2}\bar{e}(x)\gamma_\mu\nu_e(x)}_{\text{vector current}} - \underbrace{\frac{1}{2}\bar{e}(x)\gamma_\mu\gamma_5\nu_e(x)}_{\text{axial vector current}} \end{aligned} \quad (5.11)$$

$$= \frac{1}{2} \left(V_\mu^{(e)}(x) - A_\mu^{(e)}(x) \right). \quad (5.12)$$

Thus, the V-A theory modifies the Fermi theory, represented by the vector current V , by subtracting the axial vector current term A .

There are further contributions, which come from the muon and tau lepton. The total weak leptonic current then reads

$$J_\mu^{(l)} = 2\bar{e}_L(x)\gamma_\mu\nu_{e,L}(x) + 2\bar{\mu}_L(x)\gamma_\mu\nu_{\mu,L}(x) + 2\bar{\tau}_L(x)\gamma_\mu\nu_{\tau,L}(x). \quad (5.13)$$

Hadronic Currents

Hadronic currents enter in the same way as leptonic currents. For the process that a down quark is converted to an up quark, the current is

$$J_\mu^{(h)} = 2\bar{u}_L(x)\gamma_\mu d_L(x). \quad (5.14)$$

(For nucleons one would write $2\bar{p}_L\gamma_\mu n_L$.)

²⁵Compare the calculations for vector currents on page 94.

In total we have

$$J_\mu(x) = J_\mu^{(l)}(x) + J_\mu^{(h)}(x). \quad (5.15)$$

$$\begin{aligned} \mathcal{L}_w &= \frac{G_F}{\sqrt{2}} J_\mu(x) J^\mu(x)^+ \\ &= \frac{G_F}{\sqrt{2}} \left(J_\mu^{(l)}(x) J^{\mu(l)}(x)^+ \right. && \text{leptonic interactions} \\ &\quad + J_\mu^{(l)}(x) J^{\mu(h)}(x)^+ + J_\mu^{(h)}(x) J^{\mu(l)}(x)^+ && \text{semi-leptonic interactions} \\ &\quad \left. + J_\mu^{(h)}(x) J^{\mu(h)}(x)^+ \right). && \text{hadronic interactions} \end{aligned} \quad (5.16)$$

The Fermi coupling constant G_F is the same for all weak interaction processes. Its value is

$$G_F = 1.03 \cdot 10^{-5} \text{ m}_p^{-2} = 4.51 \cdot 10^{-33} \text{ cm}^2 = 1.17 \cdot 10^{-5} \text{ GeV}^{-2}. \quad (5.17)$$

Modifications of the V-A theory

a) For nucleons (and other hadrons) the hadronic current is replaced by

$$J_\mu^{(h)} = g_V V^{(h)} - g_A A^{(h)}. \quad (5.18)$$

While $g_V = 1$ always, the value of the other parameter is in the case of nucleons

$$g_A \approx 1.24, \quad (5.19)$$

whereas for quarks $g_A = 1$.

b) For the u, d, s and c-quarks, forming the doublets

$$\begin{pmatrix} u \\ d \end{pmatrix}, \quad \begin{pmatrix} c \\ s \end{pmatrix}, \quad (5.20)$$

the Lagrangian of V-A theory only allows transitions $u \leftrightarrow d$ and $c \leftrightarrow s$ between quarks. At times, when the charmed quark was not yet known, weak processes had been observed that change strangeness, like the weak Kaon decay

$$K^+ \longrightarrow \mu^+ + \bar{\nu}_\mu. \quad (5.21)$$

Numerical relations between strangeness-conserving and strangeness-changing processes motivate to incorporate strangeness-changing processes by postulating that mixtures of s and d states

$$d' = d \cos \theta_C + s \sin \theta_C \quad (5.22)$$

$$s' = s \cos \theta_C - d \sin \theta_C. \quad (5.23)$$

take part in the weak interactions. Including charmed particles, the hadronic part of the current is

$$J_\mu^{(h)} = 2\bar{u}_L\gamma_\mu d'_L + 2\bar{c}_L\gamma_\mu s'_L. \quad (5.24)$$

The angle θ_C is called Cabibbo angle, its value is

$$\theta_C \approx 13^\circ. \quad (5.25)$$

The V-A theory very successfully describes many weak processes at relatively low energies. Owing to the structure of the currents J_μ , in the processes described by the V-A theory, the electric charge does always change in the leptonic sector. One speaks of “charged currents”.

Problems of the V-A theory

- It is not renormalisable. We have outlined the theory based on tree level diagrams. Adding loop-corrections will give non-renormalisable infinities.
- It shows bad behaviour at high energies. Scattering cross-section behave like

$$\sigma \sim G_F^2 \cdot s, \quad s = (p + q)^2 = E_{CMS}^2 \quad (5.26)$$

At high energies this violates the rigorous “unitary bound”

$$\sigma \leq \frac{4\pi}{s}. \quad (5.27)$$

This happens at $s \approx 1 \text{ TeV}^2$.

- The discovery of neutral currents at CERN, 1973. The following processes have been found. (X stands for a hadron.)

$$\begin{aligned} \bar{\nu}_\mu + e^- &\longrightarrow \bar{\nu}_\mu + e^-, \\ \nu_\mu + N &\longrightarrow \nu_\mu + X, \\ \bar{\nu}_\mu + N &\longrightarrow \bar{\nu}_\mu + X. \end{aligned}$$

Intermediate vector bosons

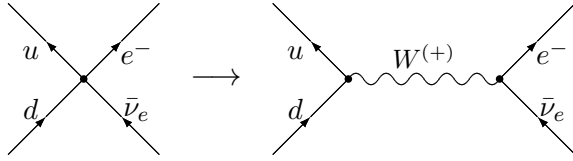
In order to improve the high-energy behaviour, it was proposed that the weak interactions are mediated by bosons, analogous to the photon in QED. For the charged currents two charged spin 1 particles, the vector bosons $W_\mu^{(+)}$ and $W_\mu^{(-)}$ have been postulated. If these particles were massless like photons,

their range would be infinite in contrast to the short range nature of weak interactions. Therefore, the W-bosons are massive.

The coupling to other fields are given by

$$\mathcal{L}_w = g_W J_\mu W^{(-)\mu} + g_W J_\mu^+ W^{(+)\mu}. \quad (5.28)$$

In β -decay, for example, the 4-fermion-vertex of the Fermi theory is replaced by a graph containing a $W^{(+)}$ propagator.



For low energies the V-A theory should be reproduced. This is the case if

$$\frac{g_W^2}{m_W^2} = \frac{G_F}{\sqrt{2}}, \quad (5.29)$$

and consequently the mass m_W has to be large.

To describe the neutral currents within this approach, there should in addition exist a neutral intermediate vector boson, denoted $Z^{(0)}$.

In fact, in 1983 the W and Z bosons have been found with the Super-Proton-Synchrotron (SPS) at CERN. Nevertheless, there remains a serious problem: The theory of intermediate vector bosons is not renormalisable. This fact is related to the mass of the intermediate particles. (The renormalisability of QED is connected with the fact that photons are massless.)

So we are faced with the question:

Is there a theory with intermediate vector bosons,
which is renormalisable ?

One can think of two different approaches to answer this question. On the one hand, one could start with the most general Lagrangian, having all imaginable kinds of intermediate particles and appropriate symmetries, and then sort out all those theories, which are non-renormalisable. The other path, which we shall follow, is to try to find a gauge theory that describes the weak interactions. The immediate problem with this attempt is that no mass term is allowed for gauge bosons, the gauge bosons are massless.

The solution to this problem is the “Higgs mechanism” ²⁶

²⁶Here, the word “mechanism” does not mean a physical process but a theoretical construction.

5.2 Higgs Mechanism

5.2.1 Spontaneous breakdown of a global symmetry

We shall now consider a concept in field theory, which is also known in other areas of physics like phase transitions, e.g. ferromagnetism, spin waves or even laser theory. Consider a complex scalar field with Lagrangian

$$\mathcal{L} = \partial_\mu \phi^* \partial^\mu \phi - m^2 \phi^* \phi - \lambda (\phi^* \phi)^2. \quad (5.30)$$

The coefficient m^2 should be thought of as a real parameter, which can be negative, too. $\mathcal{L}$ has a global U(1) symmetry

$$\phi(x) \longrightarrow e^{i\alpha} \phi(x). \quad (5.31)$$

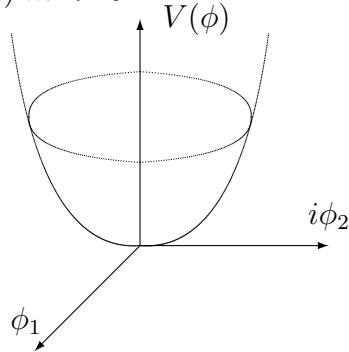
Let us look at the potential part of the Lagrangian

$$V(\phi) = m^2 \phi^* \phi + \lambda (\phi^* \phi)^2 \quad (5.32)$$

as a real valued function over the complex plane with

$$\phi = \frac{1}{\sqrt{2}}(\phi_1 + i\phi_2). \quad (5.33)$$

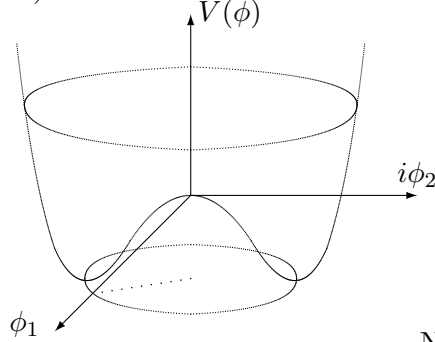
Case a) $m^2 > 0$



The symmetry is unbroken, the minimum of the potential is not degenerate, and the mean value for the field is

$$\langle \phi \rangle = 0. \quad (5.34)$$

Case b) $m^2 < 0$



Now the minima of the potential are located on a circle with radius

$$|\phi| = \sqrt{\frac{-m^2}{2\lambda}} =: \frac{v}{\sqrt{2}}. \quad (5.35)$$

In the classical theory, a groundstate corresponds to a field configuration with minimal energy, which in this case is a constant field ϕ_0 with a value in one of the potential minima. Without loss of generality we choose it to be real,

$$\phi_0 = \frac{v}{\sqrt{2}}. \quad (5.36)$$

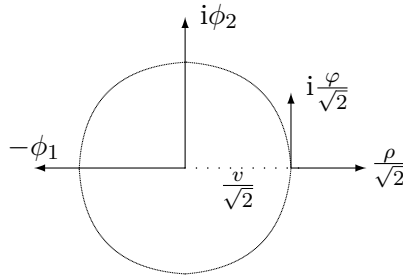
Let us assume that in the quantum theory the corresponding situation holds and the mean value of the field is

$$\langle \phi \rangle = \frac{v}{\sqrt{2}}. \quad (5.37)$$

(More precisely there will be corrections to this value.) The symmetry is then spontaneously broken.

We decompose $\phi(x)$ into its mean value and a complex remainder,

$$\phi(x) = \frac{1}{\sqrt{2}}(v + \rho(x) + i\varphi(x)). \quad (5.38)$$



The Lagrangian density, written in terms of the real fields $\rho(x)$ and $\varphi(x)$, becomes up to a constant

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \rho)^2 + \frac{1}{2}(\partial_\mu \varphi)^2 - \lambda v^2 \rho^2 - \lambda v(\rho^3 + \rho \varphi^2) - \frac{\lambda}{4}(\rho^2 + \varphi^2)^2. \quad (5.39)$$

We see that $\rho(x)$ is a massive field coupled to the massless field $\varphi(x)$. The mass m_ρ is given by

$$m_\rho^2 = 2\lambda v^2 = 2|m^2|. \quad (5.40)$$

$\rho(x)$ describes radial excitations of the original field. The massless field $\varphi(x)$ corresponds to tangential excitations. These are called *Goldstone bosons*.

A general statement about the existence of Goldstone bosons is made by the

Goldstone theorem:

Spontaneous breakdown of a continuous, global symmetry leads to one massless spin 0 particle for each generator of the group that is spontaneously broken.

In the above situation, the U(1) group has one generator, and one Goldstone boson field arises.

5.2.2 Higgs mechanism

Now we turn to a seemingly similar situation, where the symmetry is, however, a local gauge symmetry. Will there be massless Goldstone bosons?

Let us consider the abelian Higgs model, describing a complex scalar field coupled to gauge field.

$$\mathcal{L} = (D_\mu \phi)^*(D^\mu \phi) - V(\phi) - \frac{1}{4}F_{\mu\nu}F^{\mu\nu}, \quad (5.41)$$

$$D_\mu = \partial_\mu + iqA_\mu. \quad (5.42)$$

In this “scalar QED”, $A_\mu(x)$ is the gauge field and the gauge group is U(1). As before the potential is a quartic one,

$$V(\phi) = m^2 \phi^* \phi + \lambda(\phi^* \phi)^2, \quad (5.43)$$

and we consider the case $m^2 < 0$.

As in the previous section, we expand the field $\phi(x)$ near the real minimum of $V(\phi)$, but this time we choose a different parameterisation in terms of a

radial variable $\rho(x)$ and an angular variable $\xi(x)$:

$$\phi(x) = \frac{1}{\sqrt{2}}(v + \rho(x))e^{i\xi(x)/v} \quad (5.44)$$

$$= \frac{1}{\sqrt{2}}(v + \rho(x) + i\xi(x) + \dots). \quad (5.45)$$

After a few lines of algebra the Lagrangian becomes

$$\begin{aligned} \mathcal{L} = & \frac{1}{2}(\partial_\mu \rho)^2 - \lambda v^2 \rho^2 \\ & + \frac{1}{2}(\partial_\mu \xi)^2 + qv A_\mu \partial^\mu \xi \\ & + \frac{1}{2}q^2 v^2 A_\mu A^\mu - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} \\ & + \dots \text{ (interaction terms).} \end{aligned} \quad (5.46)$$

As before, the mass belonging to the radial excitations is

$$m_\rho^2 = 2\lambda v^2. \quad (5.47)$$

There is also term quadratic in the gauge field, namely $\frac{1}{2}q^2 v^2 A_\mu A^\mu$, from which we read off a mass for the field A_μ , which is

$$m_A = qv. \quad (5.48)$$

There is no mass term for the field $\xi(x)$ and it looks like a massless Goldstone field. However, there is a mixing term $qv A_\mu \partial^\mu \xi$, quadratic in the fields, and we have to clarify the situation.

Looking at the parameterisation of $\phi(x)$, we recognise that the factor

$$\exp\left(\frac{i}{v}\xi(x)\right), \quad (5.49)$$

has precisely the form of a local gauge transformation. Therefore the field $\xi(x)$ can be transformed away by means of a compensating gauge transformation:

$$\phi(x) \longrightarrow \phi'(x) := \exp\left(-\frac{i}{v}\xi(x)\right)\phi(x) = \frac{1}{\sqrt{2}}(v + \rho(x)). \quad (5.50)$$

At the same time the gauge field A_μ is transformed as

$$A_\mu \longrightarrow A'_\mu(x) = A_\mu(x) + \frac{1}{qv}\partial_\mu \xi(x) =: B_\mu(x). \quad (5.51)$$

This gauge is called “unitary gauge”. The Lagrangian now becomes

$$\begin{aligned}\mathcal{L} = & \frac{1}{2}(\partial_\mu \rho)^2 - \frac{m_\rho^2}{2}\rho^2 - \lambda v \rho^3 - \frac{\lambda}{4}\rho^4 \\ & - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \frac{1}{2}q^2 v^2 B_\mu B^\mu \\ & + v q^2 \rho B_\mu B^\mu + \frac{1}{2}q^2 \rho^2 B_\mu B^\mu,\end{aligned}\tag{5.52}$$

where $F_{\mu\nu}$ is now formed from $B_\mu(x)$. The mixing term has disappeared. The remaining physical fields are the following.

- $\rho(x)$ is a massive scalar field with $m_\rho^2 = 2\lambda v^2$.
- $B_\mu(x)$ is a *massive* vector field with mass $m_v = qv$. It describes “massive photons”.

The usual massless photons have two transverse degrees of freedom (components), whereas the massive field $B_\mu(x)$ has three degrees of freedom. Its additional “longitudinal” degree of freedom originates from the alleged Goldstone field $\xi(x)$. This is sometimes expressed by saying, “the Goldstone boson is eaten by the vector boson.”

The last two terms in the Lagrangian are interaction terms between the two fields.

In the unitary gauge the local gauge symmetry is hidden. In the literature one sometimes finds the formulation of a “spontaneous breaking of local gauge symmetry”. This is not quite correct. Local gauge symmetries cannot be broken spontaneously (“Elitzur’s theorem”). Instead, the fields $\rho(x)$ and $B_\mu(x)$ are gauge invariant combinations of the original fields.

5.3 Glashow-Weinberg-Salam Model

The Glashow-Weinberg-Salam model is a gauge theory describing the electroweak interactions. The gauge bosons, mediating the weak and electromagnetic interactions, are the massive W^+ , W^- , Z^0 , and the massless photon A_μ . The crucial piece in the formulation of a gauge theory with massive vector bosons is the Higgs mechanism. The gauge group is

$$\text{gauge group} = \underbrace{\text{SU}(2)}_{\text{weak isospin}} \otimes \underbrace{\text{U}(1)}_{\text{weak hypercharge}}_Y \tag{5.53}$$

and the model uses the Higgs mechanism to create the masses m_W and m_Z .

Historically, the basic structure was formulated by Sheldon Glashow (1961), but without the Higgs mechanism, and the complete form was found by Steven Weinberg (1967) and Abdus Salam (1968).

Let us start by concentrating on leptons. In the weak interactions only the left handed components couple to the SU(2) gauge field. Therefore the group is denoted SU(2)_L. It acts on the doublets

$$\begin{pmatrix} \nu_e \\ e^- \end{pmatrix}_L = \frac{1}{2}(1 - \gamma_5) \begin{pmatrix} \nu_e \\ e^- \end{pmatrix}, \quad (5.54)$$

and on the doublets of the remaining two families

$$\begin{pmatrix} \nu_\mu \\ \mu^- \end{pmatrix}_L, \quad \begin{pmatrix} \nu_\tau \\ \tau^- \end{pmatrix}_L. \quad (5.55)$$

This SU(2) symmetry is called weak isospin, its generators are denoted T_1 , T_2 , T_3 , and the corresponding quantum numbers of the lepton doublets are

$$t = \frac{1}{2}, \quad t_3 = \pm \frac{1}{2}. \quad (5.56)$$

The right-handed parts of the leptons, $(e^-)_R$, $(\mu^-)_R$, $(\tau^-)_R$, do not have right-handed neutrino-partners, and are singlets ($t = 0$) under weak isospin.

The other symmetry group U(1)_Y is associated with the weak hypercharge Y , which is related to the electric charge Q by

$$Q = t_3 + \frac{y}{2}. \quad (5.57)$$

		ν	e_L^-, μ_L^-, τ_L^-	e_R^-, μ_R^-, τ_R^-
weak isospin	t_3	$+\frac{1}{2}$	$-\frac{1}{2}$	0
hypercharge	y	-1	-1	-2
electric charge	Q	0	-1	-1

Gauge fields

The symmetry groups are now gauged by introducing corresponding gauge fields.

symmetry group	generators	gauge field	field strength
SU(2)	$T_a = \frac{1}{2}\tau_a$	$W_\mu^a(x)$	$W_{\mu\nu}^a(x)$
U(1)	$Y = y \mathbf{1}$	$B_\mu(x)$	$B_{\mu\nu}(x)$

The gauge field dynamics is given by the gauge part Lagrangian

$$\mathcal{L}_g = -\frac{1}{2}W_{\mu\nu}^a W^{a\mu\nu} - \frac{1}{4}B_{\mu\nu}B^{\mu\nu}. \quad (5.58)$$

The fermionic part of the Lagrangian is

$$\begin{aligned} \mathcal{L}_f = & \overline{(\nu_e, e^-)_L} i\gamma^\mu \left(\partial_\mu + i\frac{g'}{2}YB_\mu + i\frac{g}{2}\tau_a W_\mu^a \right) \begin{pmatrix} \nu_e \\ e^- \end{pmatrix}_L \\ & + \overline{e_R^-} i\gamma^\mu \left(\partial_\mu + i\frac{g'}{2}YB_\mu \right) e_R^- \\ & + \text{the same terms for muon and tauon fields.} \end{aligned} \quad (5.59)$$

Higgs field

Finally, we have to add the Higgs field part of the Lagrangian. For the Higgs field we take a complex SU(2) doublet

$$\phi(x) = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix} \quad (5.60)$$

The Higgs field has weak isospin $t = 1/2$ and weak hypercharge $y_\phi = 1$. The upper component ϕ^+ has $t_3 = +1/2$, $y_\phi = 1$ and therefore charge $Q = +1$, whereas the lower component ϕ^0 has $t_3 = -1/2$, $y_\phi = 1$ and charge $Q = 0$. In matrix form this reads

$$\begin{aligned} Q\phi &= \left(T_3 + \frac{Y}{2} \right) \phi = \left(\frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} + \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \right) \phi \\ &= \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \phi. \end{aligned} \quad (5.61)$$

The covariant derivative of the Higgs field is

$$D_\mu \phi = \left(\partial_\mu + i\frac{g'}{2}B_\mu + i\frac{g}{2}\tau_a W_\mu^a \right) \phi. \quad (5.62)$$

Its Lagrangian is

$$\mathcal{L}_h = (D_\mu \phi)^\dagger (D^\mu \phi) - \mu^2 \phi^\dagger \phi - \lambda (\phi^\dagger \phi)^2, \quad (\mu^2 < 0, \lambda > 0). \quad (5.63)$$

The potential has minima at

$$\phi^\dagger \phi = \frac{-\mu^2}{2\lambda}. \quad (5.64)$$

The vacuum state is chosen to be

$$\phi_0 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix}, \quad v^2 = -\frac{\mu^2}{\lambda}. \quad (5.65)$$

It is chosen such that $Q\phi_0 = 0$. This guarantees that the electromagnetic gauge group $U(1)$, which is generated by Q , is unaffected by the Higgs mechanism (it is “unbroken”), so that the photon remains massless. We shall see this explicitly below.

Without the gauge interactions we would have a spontaneously broken global symmetry, and there would be

- 3 massless Goldstone bosons,
- 1 massive Higgs field.

In the presence of the gauge fields, the fields are transformed into the unitary gauge as in Sec. 5.2.2, so that we have

$$\phi(x) \longrightarrow \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + \rho(x) \end{pmatrix}. \quad (5.66)$$

The Lagrangian of the Higgs and gauge fields then becomes

$$\begin{aligned} \mathcal{L}_h = & \frac{1}{2}(\partial_\mu \rho)^2 - \frac{1}{2}m_\rho^2 \rho^2 - \lambda v \rho^3 - \frac{\lambda}{4} \rho^4 \\ & + \frac{1}{2}m_W^2(W_\mu^1 W^{\mu 1} + W_\mu^2 W^{\mu 2}) \\ & + \frac{v^2}{8}(g' B_\mu - g W_\mu^3)(g' B^\mu - g W^{\mu 3}) \\ & + \text{interaction terms.} \end{aligned} \quad (5.67)$$

As a result one gets a massive neutral Higgs field ρ with $m_\rho^2 = 2\lambda v^2$, and massive vector fields W^1 and W^2 with $m_W = \frac{1}{2}gv$. The charged bosons $W^\pm$ are defined as the complex linear combinations

$$W^\pm = \frac{1}{\sqrt{2}}(W_\mu^1 \mp iW_\mu^2) \quad (5.68)$$

and we have

$$W_\mu^1 W^{\mu 1} + W_\mu^2 W^{\mu 2} = 2W_\mu^+ W^{\mu -}. \quad (5.69)$$

The remaining fields B_μ and W_μ^3 appear mixed in the quadratic terms. They can be diagonalised by means of

$$Z_\mu := \frac{gW_\mu^3 - g'B_\mu}{\sqrt{g^2 + g'^2}}, \quad (5.70)$$

$$A_\mu := \frac{g'W_\mu^3 + gB_\mu}{\sqrt{g^2 + g'^2}}. \quad (5.71)$$

The quadratic term then reads

$$\frac{1}{2}m_Z^2 Z_\mu Z^\mu, \quad (5.72)$$

and we see that the masses belonging to the fields Z_μ and A_μ are

$$m_Z = \frac{1}{2}v\sqrt{g^2 + g'^2}, \quad m_A = 0. \quad (5.73)$$

In this way fields for the massive Z^0 boson and the massless photon emerge. Defining an angle θ_W through

$$\tan \theta_W = \frac{g'}{g}, \quad \text{or} \quad \frac{m_W}{m_Z} = \cos \theta_W, \quad (5.74)$$

the fields can be represented as

$$\begin{pmatrix} Z_\mu \\ A_\mu \end{pmatrix} = \begin{pmatrix} \cos \theta_W & -\sin \theta_W \\ \sin \theta_W & \cos \theta_W \end{pmatrix} \begin{pmatrix} W_\mu^3 \\ B_\mu \end{pmatrix}. \quad (5.75)$$

θ_W is called weak mixing angle or “Weinberg angle”. We also have

$$\cos \theta_W = \frac{g}{\sqrt{g^2 + g'^2}}, \quad \sin \theta_W = \frac{g'}{\sqrt{g^2 + g'^2}}. \quad (5.76)$$

To identify the charges of the fields, we consider the covariant derivative

$$D_\mu = \partial_\mu + i\frac{g'}{2}YB_\mu + igT_aW_\mu^a \quad (5.77)$$

and recall

$$Q = T_3 + \frac{1}{2}Y. \quad (5.78)$$

With

$$T_\pm := T_1 \pm iT_2 \quad (5.79)$$

we have

$$[Q, T_\pm] = [T_3, T_\pm] = \pm T_\pm, \quad [Q, Y] = 0. \quad (5.80)$$

Then we get the charged field components as parts of

$$gT_aW_\mu^a = \frac{g}{\sqrt{2}} \underbrace{(T_+W_\mu^+ + T_-W_\mu^-)}_{\pm \text{ charged fields}} + gT_3W_\mu^3. \quad (5.81)$$

For the fields W_μ^3 and B_μ we find

$$\begin{aligned}
gT_3W_\mu^3 + g'\frac{Y}{2}B_\mu &= (g\sin\theta_W T_3 + g'\cos\theta_W \frac{Y}{2})A_\mu \\
&\quad + (g\cos\theta_W T_3 - g'\sin\theta_W \frac{Y}{2})Z_\mu \\
&= \frac{gg'}{\sqrt{g^2 + g'^2}}(T_3 + \frac{Y}{2})A_\mu \\
&\quad + \sqrt{g^2 + g'^2} \left(T_3 - \sin^2\theta_W (T_3 + \frac{Y}{2}) \right) Z_\mu \\
&= \frac{gg'}{\sqrt{g^2 + g'^2}}QA_\mu \\
&\quad + \sqrt{g^2 + g'^2} (T_3 - \sin^2\theta_W Q) Z_\mu.
\end{aligned} \tag{5.82}$$

The photon field A_μ couples to the charge Q , as it should. From the expression we can identify the coupling constant of the photon, the electric charge unit

$$e_0 = \frac{gg'}{\sqrt{g^2 + g'^2}} = g\sin\theta_W = g'\cos\theta_W. \tag{5.83}$$

Altogether we can write the fields appearing in the covariant derivative in the form

$$\begin{aligned}
g'\frac{Y}{2}B_\mu + gT_aW_\mu^a &= \frac{e_0}{\sqrt{2}\sin\theta_W}(T_+W_\mu^+ + T_-W_\mu^-) \\
&\quad + e_0QA_\mu \\
&\quad + \frac{e_0}{\sin\theta_W\cos\theta_W}(T_3 - \sin^2\theta_W Q)Z_\mu.
\end{aligned} \tag{5.84}$$

From the discussion above it is already clear that the masses of the W and Z bosons stem from the non-vanishing vacuum value of the Higgs field. To make it manifest, let us nevertheless show this explicitly again. The covariant derivative acts on the constant vacuum value of the Higgs field as

$$\begin{aligned}
D_\mu\phi_0 &= \frac{ig}{\sqrt{2}}(W_\mu^+T_+ + W_\mu^-T_-)\phi_0 + \frac{ig}{\cos\theta_W}Z_\mu T_3\phi_0 \\
&= \frac{ig}{2}W_\mu^+ \begin{pmatrix} v \\ 0 \end{pmatrix} - \frac{ig}{2\sqrt{2}\cos\theta_w}Z_\mu \begin{pmatrix} 0 \\ v \end{pmatrix}.
\end{aligned} \tag{5.85}$$

In the vacuum state the kinetic term of the Higgs field then yields

$$(D_\mu\phi_0)^\dagger(D^\mu\phi_0) = -\frac{g^2v^2}{8}(2W_\mu^+W^{\mu-} + \frac{1}{\cos^2\theta_W}Z_\mu Z^\mu), \tag{5.86}$$

and we get the masses of the gauge bosons

$$m_W = \frac{1}{2}gv = \frac{e_0 v}{2 \cos \theta_W}, \quad m_Z = \frac{m_w}{\cos \theta_W}. \quad (5.87)$$

Fermion masses

Under the local gauge symmetry group

$$\text{SU}(2)_L \otimes \text{U}(1)_Y \quad (5.88)$$

the right- and left-handed fields e_R and e_L transform differently. Therefore, the mass term

$$m \bar{\psi} \psi = m(\bar{\psi}_L \psi_R + \bar{\psi}_R \psi_L) \quad (5.89)$$

cannot be invariant, and the bare fermion masses have to be zero. How can we get fermion masses?

The solution is that the fermion masses come from the Higgs field, like the masses of the vector bosons. This can be achieved by coupling the Higgs field to the fermion fields through a Yukawa coupling

$$\mathcal{L}_Y = -G_e \overline{(\nu_e, e^-)}_L \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} e_R^- + \text{Hermitian conjugate}. \quad (5.90)$$

The same terms are added for muon and tauon fields. These Yukawa terms are invariant under $\text{SU}(2)_L$. They are also invariant under $\text{U}(1)_Y$ because the hypercharges add to zero:

$$\underbrace{\overline{(\nu_e, e^-)}_L \phi e_R^-}_{Y_\phi = 1 \quad +1 -2 = 0} \quad (5.91)$$

In the unitary gauge

$$\begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + \rho(x) \end{pmatrix} \quad (5.92)$$

the Lagrangian becomes

$$\mathcal{L}_Y = - \underbrace{\frac{G_e v}{\sqrt{2}} \bar{e}^- e^-}_{\text{mass}} - \underbrace{\frac{G_e}{\sqrt{2}} \rho \bar{e}^- e^-}_{\text{remaining Yukawa coupling}} \quad (5.93)$$

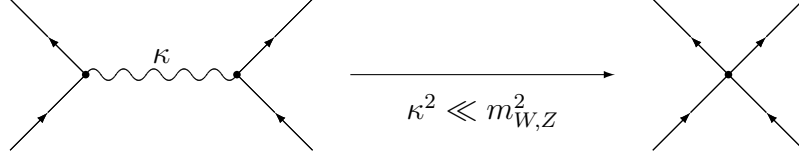
The electron mass is

$$m_e = \frac{G_e v}{\sqrt{2}}. \quad (5.94)$$

In this way all masses (of W , Z , leptons, quarks) are generated by the Higgs mechanism.

Reduction to the Fermi theory

As discussed in Sec. 5.1.3, the reduction to the V-A theory at low energies follows the scheme



where the vector boson propagator is replaced by $ig_{\mu\nu}/m^2$. The resulting effective theory has a four-fermion vertex with coupling constant

$$\frac{G_F}{\sqrt{2}} = \frac{g^2}{8m_W^2} = \frac{1}{2v^2}. \quad (5.95)$$

From the known value of the Fermi coupling G_F one derives

$$v = (\sqrt{2} G_F)^{-\frac{1}{2}} = 246 \text{ GeV}. \quad (5.96)$$

From independent measurements the values

$$\sin^2 \theta_W = 0.231, \quad (5.97)$$

$$m_W = 80.42 \text{ GeV}, \quad (5.98)$$

$$m_Z = 91.19 \text{ GeV} \quad (5.99)$$

have been obtained. Recent experiments indicate the value

$$m_\rho \approx 125 \text{ GeV} \quad (5.100)$$

for the Higgs mass.

Quarks

The quarks couple to the gauge fields of $SU(2)_L \otimes U(1)_Y$ in a similar way as leptons. The left-handed quarks form weak iso-doublets, i.e. $t = 1/2$,

$$\begin{pmatrix} u \\ d' \end{pmatrix}_L, \quad \begin{pmatrix} c \\ s' \end{pmatrix}_L, \quad \begin{pmatrix} t \\ b' \end{pmatrix}_L. \quad (5.101)$$

Their weak hypercharge is $y = 1/3$ in order to yield the correct electric charges. The d' , s' and b' components have been denoted with a prime, because mixing among them will occur, generalising the Cabibbo angle. To remind you, for two generations of quarks the mixing of the d and s quarks is

$$d' = d \cos \theta_C + s \sin \theta_C \quad (5.102)$$

$$s' = -d \sin \theta_C + s \cos \theta_C. \quad (5.103)$$

The right-handed quarks

$$u_R, d'_R, c_R, s'_R, t_R, b'_R \quad (5.104)$$

form weak iso-singlets with $t = 0$ and $y = 4/3$ and $y = -2/3$, respectively. Quark masses are generated through Yukawa couplings to the Higgs field. The most general form of the resulting mass terms for the three generations is

$$\begin{pmatrix} \bar{u}_R & \bar{c}_R & \bar{t}_R \end{pmatrix} M_u \begin{pmatrix} u_L \\ c_L \\ t_L \end{pmatrix} + \begin{pmatrix} \bar{d}'_R & \bar{s}'_R & \bar{b}'_R \end{pmatrix} M_d \begin{pmatrix} d'_L \\ s'_L \\ b'_L \end{pmatrix} + h.c. \quad (5.105)$$

with two 3×3 mass matrices M_u and M_d . The mass matrices cannot be diagonalised simultaneously and one ends up with a diagonal M_u and

$$M_d = V \begin{pmatrix} m_d & 0 & 0 \\ 0 & m_s & 0 \\ 0 & 0 & m_b \end{pmatrix} V^\dagger. \quad (5.106)$$

The relation between the mass eigenstates d, s, b and the mixed states d', s', b' , entering the weak interactions, is therefore

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = V \begin{pmatrix} d \\ s \\ b \end{pmatrix} \quad (5.107)$$

with the unitary Cabibbo-Kobayashi-Maskawa (CKM) matrix. In a theory with only 2 generations one has the Cabibbo mixing

$$V = \begin{pmatrix} \cos \theta_c & \sin \theta_c \\ -\sin \theta_c & \cos \theta_c \end{pmatrix}. \quad (5.108)$$

The Glashow-Weinberg-Salam model gives a consistent theory of the weak and electromagnetic interactions. A detailed study of the interactions described by the theory shows

- there are charged currents as in the V-A theory,
- it predicts neutral currents coupling to left-handed and right-handed fermions: $J_\mu^{(\text{neutral})} = J_\mu^3 - \sin^2 \theta_W J_\mu^{(e.m.)}$,
- it contains further interactions of Higgs and gauge boson fields.

Properties of the GWS theory are

- Renormalisability.

Although the Lagrangian does not look renormalisable in the unitary gauge, it is! The proof (G. 't Hooft and others) is complicated. The basic idea is: without unitary gauge the theory is renormalisable, and it has been shown that this property is independent of the gauge.

- Phenomenologically the theory is extremely successful.
- It provides a unified description of the electromagnetic and weak interactions.

However, because the gauge group is the direct product of SU(2) for weak isospin and U(1) for electromagnetism, this is not really a full unification. This motivates to search for a unified theory with a simple Lie group, which cannot be decomposed into the direct product of subgroups.

- The theory is free of anomalies.

In field theory an anomaly occurs, when the classical theory has a symmetry and an associated conserved Noether current, $\partial_\mu j^\mu = 0$, which is not conserved in the quantum theory: $\partial_\mu j^\mu \neq 0$. The symmetry is then broken by a quantum effect. In the GWS theory there are possible anomalies, but they cancel each other. The condition for the absence of anomalies is that the electric charges of all left-handed particles and all right-handed particles in a generation cancel against each other:

$$\sum_L Q - \sum_R Q = 0. \quad (5.109)$$

Here the quark charges are counted with a factor of 3 due to the three colours. The condition is satisfied for leptons and quarks within each generation. As a consequence the number of colours is related to the cancellation of anomalies.

- The Glashow-Iliopoulos-Maiani (GIM) mechanism explains the suppression of flavour-changing neutral currents, if there is an even number N_f of flavours. It was suggested at a time, when only three quark flavours had been known, and led to the prediction of the charm quark.

Adding the strong interactions of quarks via QCD leads to the complete Standard Model. The total gauge group then is

$$\text{SU}(3)_c \otimes \text{SU}(2)_L \otimes \text{U}(1)_Y. \quad (5.110)$$