

Lecture notes for Particle Physics MSc course

Gábor Takács

Budapest University of Technology and Economics

1st January 2025

Abstract

These notes are meant to accompany the course BMETE15MF43 “Particle Physics”, taught by the author at the Budapest University of Technology and Economics as part of the Physics MSc curriculum.

Note: measured values of particle properties are quoted from the Review of Particle Physics [K.A. Olive et al. (Particle Data Group), Chinese Phys. C38 (2014) No 9], except in cases where some other specific sources are given.

Disclaimer: the present version is preliminary. Some sections are not yet completed, and there could be typos and other mistakes in the text. I would be grateful for anyone bringing errors to my attention. - G.T.

Contents

1 Overview of distance and energy scales	6
2 Special relativity	8
2.1 Four-vectors and Lorentz invariance	8
2.2 Examples of Lorentz transformations	12
2.3 Parity and time reversal	14
2.4 Free relativistic particle	15
2.5 The velocity and energy-momentum four-vectors	17
3 Particles: classification by quantum numbers and interactions	18
3.1 Units $\hbar = c = k_B = 1$	18
3.1.1 Electromagnetic interactions	18
3.1.2 Gravity	20
3.2 Symmetries, conserved quantities and quantum numbers	20
3.2.1 Discrete symmetries	22
3.2.2 Continuous symmetries: energy and momentum	22
3.2.3 Continuous symmetries: example of electric charge	23
3.2.4 Charge conjugation	23
3.3 The particle zoo	24
3.3.1 Leptons	24
3.3.2 Hadrons	24
4 Relativistic wave equations 1: the Klein-Gordon equation	27
4.1 Quantum mechanics of a non-relativistic point particle	27
4.2 The Klein-Gordon equation	30
4.3 Peculiarities of the Klein-Gordon equation	32
4.3.1 No spin	33
4.3.2 Negative energy solutions	33
4.3.3 No probability density	34
4.3.4 Klein paradox and completeness	35
5 Relativistic wave equations 2: the Dirac equation	39
5.1 Deriving the Dirac equation	39
5.2 Covariance and spinors	41
5.3 Properties of spinor transformations	43
5.3.1 Group property	43
5.3.2 Infinitesimal form of spinor transformations	43
5.3.3 Skew unitarity and the conjugate Dirac spinor	45
5.4 Spin, parity and chirality	47
5.4.1 Parity, Weyl spinors and chirality	47
5.4.2 Weyl representation of Dirac matrices and spin	50
5.4.3 Dirac and Majorana representations	51
5.5 Conserved current and probabilistic interpretation	53
5.6 Solutions of the free Dirac equation	54
5.6.1 Plane waves	54
5.6.2 Physical interpretation of the free Dirac equation	58

5.7	Weyl spinors	61
5.8	Dirac equation in an external electromagnetic field	63
5.9	Charge conjugation	66
5.10	Majorana fermions and Majorana mass term	68
6	Weak interactions	71
6.1	Beta decay and the neutrino	71
6.2	Parity violation	72
6.2.1	Parity symmetry	72
6.2.2	The $\theta - \tau$ puzzle and the proposal of parity violation	73
6.2.3	The Wu experiment and neutrino chirality	74
6.3	CP violation	78
6.3.1	CP symmetry	78
6.3.2	The Cronin-Fitch experiment	78
6.4	CPT, matter and antimatter	80
7	Strong interactions	83
7.1	Isospin symmetry	83
7.1.1	Isospin for nucleons	83
7.1.2	Representations of $SU(2)$ symmetry	84
7.1.3	$SU(2)$ Clebsh-Gordan coefficients and a simple application of isospin symmetry	87
7.1.4	Yukawa model and the π meson	89
7.2	Strangeness and hypercharge	91
7.3	The $SU(3)$ quark model	93
7.3.1	The $SU(3)$ group	93
7.3.2	Representation theory of $SU(3)$	96
7.3.3	The eightfold way: meson multiplets	100
7.3.4	The eightfold way: baryon multiplets and the GMO mass formula	104
7.3.5	The meson mass spectrum	109
7.3.6	Quark model	111
7.3.7	Experimental evidence for quarks	114
7.3.8	Quark masses	118
8	Lagrangian field theory	120
8.1	Locality and fields	120
8.2	Action and Euler-Lagrange equations	121
8.2.1	The Klein-Gordon action for real and complex scalar fields	123
8.2.2	The Dirac action	124
8.3	Symmetries and Noether's theorem	124
8.3.1	Energy-momentum tensor	126
8.3.2	Angular momentum conservation	130
8.3.3	The Belinfante-Rosenfeld energy momentum tensor	133
8.3.4	Other conserved charges	134

9 Electromagnetic field	138
9.1 Relativistic electrodynamics	138
9.1.1 Electromagnetic field in relativistic form	138
9.1.2 Relativistic particle in an external electromagnetic field	140
9.2 Action and energy-momentum tensor	144
9.3 Gauge invariance and minimal coupling	145
10 Elements of quantum field theory	148
10.1 Canonical quantisation of free fields	148
10.2 Free particles	153
10.3 Dyson series and Feynman diagrams	156
10.3.1 Interacting real scalar field	159
10.3.2 Quantum electrodynamics	160
11 Weak interactions revisited	162
11.1 Lepton number conservation	162
11.1.1 The Cowan-Reines neutrino experiment	163
11.1.2 Discovering the muon neutrino	165
11.2 Charged current interactions	166
11.3 FCNC, the GIM mechanism and the discovery of charm	170
11.3.1 The Cabibbo angle	170
11.3.2 FCNC and the discovery of charm	171
11.3.3 The third generation of leptons and quarks	175
11.4 Flavour mixing, CP violation and neutrino oscillations	177
11.4.1 The Cabibbo-Kobayashi-Maskawa matrix	177
11.4.2 CP violation in the neutral kaon system	179
11.4.3 Neutrino oscillations	182
12 Non-Abelian gauge fields and quantum chromodynamics	188
12.1 Non-Abelian gauge invariance	188
12.1.1 Global non-Abelian symmetry groups	188
12.1.2 Covariant derivative and the gauge fields	189
12.1.3 The gauge field action	193
12.1.4 Recovering the Abelian case	195
12.2 Discovery of quark colour	196
12.3 Quantum chromodynamics: the gauge theory of strong interactions	199
12.3.1 Confinement and asymptotic freedom	200
12.3.2 Running coupling, effective theories and the Theory of Everything	200
12.3.3 Asymptotic freedom	204
12.3.4 Confinement	209
12.4 Jet physics and evidence for gluons	213
13 Gauge theory of weak interactions	215
13.1 Intermediate vector bosons	215
13.1.1 Problems with the four-fermion interaction	215
13.1.2 The charged weak intermediate vector bosons	217
13.2 Spontaneous symmetry breaking	220

13.2.1 Spontaneous symmetry breaking and Goldstone's theorem	220
13.2.2 Chiral symmetry breaking in strong interactions	222
13.2.3 The Higgs mechanism	225
13.3 Elements of the Glashow-Salam-Weinberg model	228
13.3.1 Gauge symmetry and fermion multiplets	228
13.3.2 Higgs mechanism and vector boson masses	230
13.3.3 Fermion masses and mixing angles	234
13.3.4 The Higgs boson and present status of the Standard Model	237
14 Accelerators	239
14.1 Linac, cyclotron, synchrotron	239
14.2 Cross section and luminosity	239
15 Detectors	240
15.1 Anatomy of high energy particle detectors	240
15.2 Principles of particle identification	240
A Tensor method for Lie group representations	241
A.1 Tensors for $GL(n, \mathbb{C})$	241
A.2 Interesting subgroups of $GL(n, \mathbb{C})$	242
A.3 Relation to the permutation group	243
A.4 Tensors for $SU(2)$	244
A.4.1 Irreducible representations of $SU(2)$	246
A.4.2 Relation between $SU(2)$ and $SO(3)$	247
A.4.3 The Lie algebra $su(2)$ and its representations	250
A.4.4 $su(2)$ representations in the tensor product method	251
A.5 Tensor method for $SU(n)$	253
A.5.1 Reducing the problem to $\mathcal{T}^{(N,0)}$	253
A.5.2 Tensor symmetry classes and Young tableaux	254
A.5.3 The adjoint representation	256
A.6 $SU(2)$ examples	257
A.7 $SU(3)$ examples	260
A.7.1 $\mathbf{3} \otimes \mathbf{3} = \mathbf{3} \oplus \mathbf{6}$	261
A.7.2 $\mathbf{3} \otimes \mathbf{3} = \mathbf{1} \oplus \mathbf{8}$	261
A.7.3 $\mathbf{3} \otimes \mathbf{3} \otimes \mathbf{3} = \mathbf{1} \oplus \mathbf{8} \oplus \mathbf{8} \oplus \mathbf{10}$	262
B Running coupling and asymptotic freedom	264
B.1 The running coupling	264
B.2 Charged scalar: Landau diamagnetism	266
B.3 Dirac fermion: Landau paramagnetism and Pauli diamagnetism	267
B.4 Non-Abelian gauge fields	269

1 Overview of distance and energy scales

The aim of particle physics is to study the fundamental constituents of matter and their interactions. Presently, this mostly involves studying *sub-nuclear physics*, i.e. physics at scales that are smaller than the atomic nucleus. It is worthwhile to recall the scales of diverse natural phenomena, shown in the table below. The values just give a rough order of magnitude within factors of order unity.

Scale (m)	Phenomena
10^{-35}	Planck scale (quantum gravity)
10^{-32}	cosmological inflation
$< 10^{-19}$	the unknown...
$10^{-19} - 10^{-12}$	presently known particle physics
10^{-14}	atomic nucleus
$10^{-10} - 10^{-9}$	atoms and molecules
$10^{-6} - 10^{-4}$	cells
$10^{-3} - 10^6$	human activity
$10^6 - 10^8$	planets
$10^8 - 10^{12}$	stars
10^{14}	size of solar system
10^{16}	light year; typical interstellar distances
$10^{20} - 10^{22}$	galaxies
10^{26}	size of the observable universe

However, particle physics does directly influence phenomena at much larger scales, especially at the cosmological level. As an extreme example, *cosmological inflation* (which is a process of presently unknown but most likely particle physics origin) is estimated to produce a Universe which is homogeneous over scales around $10^7 - 10^8$ times the linear size of the observable part (the Universe itself may even be infinite in extent).

Now we would like to resolve phenomena below the scale 10^{-12} m. If we use light to probe the system, that is already hard γ -ray and it is not easy to build a microscope for that. For scales even below, there is no hope of probing them with light. So we want some other objects with wavelengths between 10^{-12} and 10^{-20} m, say. By the de Broglie relation this corresponds to momentum

$$p = \frac{h}{\lambda} \quad (1.1)$$

As usual in particle physics, we are going to measure energy in gigaelectronvolts:

$$1 \text{ GeV} = 10^9 \text{ eV} = 1.602176487(40) \cdot 10^{-10} \text{ J} \quad (1.2)$$

The product of Planck's constant and light speed is

$$hc = 1.98644568 \cdot 10^{-25} \text{ J}\cdot\text{m} = 1.23984199 \cdot 10^{-15} \text{ GeV}\cdot\text{m} \quad (1.3)$$

For example, the momentum corresponding to resolving a distance of 10^{-18} m is around

$$p = 1240 \text{ GeV}/c \quad (1.4)$$

Note that using the relativistic expressions for energy and momentum

$$\mathcal{E} = \frac{mc^2}{\sqrt{1 - \frac{v^2}{c^2}}} \quad p = \frac{mv}{\sqrt{1 - \frac{v^2}{c^2}}} \quad (1.5)$$

even a proton with

$$m_p c^2 = 0.938272046(21) \text{ GeV} \quad (1.6)$$

of rest mass would be highly relativistic at such energies. Indeed, its velocity in units of the light speed would be

$$\frac{v}{c} = \frac{pc}{\mathcal{E}} = 0.9999997 \dots \quad (1.7)$$

For an electron with

$$m_e c^2 = 0.5109989461(31) \text{ MeV} \quad (1.8)$$

this is even more extreme with

$$\frac{v}{c} = 0.999999999999915 \quad (1.9)$$

So concerning the relativistic energy-momentum relation

$$\mathcal{E} = \sqrt{p^2 c^2 + m^2 c^4} = \begin{cases} mc^2 + \frac{p^2}{2m} + O(p^4) & p \ll mc \\ pc & p \gg mc \end{cases} \quad (1.10)$$

the particles we consider are almost always in the ultra-relativistic regime, corresponding to the second option shown. To characterise such particles, one often uses the relativistic γ factor

$$\gamma = \frac{\mathcal{E}}{mc^2} = \left(1 - \frac{v^2}{c^2}\right)^{-1/2} \quad (1.11)$$

Non-relativistic physics is only valid when

$$\gamma - 1 \ll 1 \quad (1.12)$$

In contrast, the case

$$\gamma \gg 1 \quad (1.13)$$

is called ultra-relativistic. In fact, in such a case one can easily compute

$$\frac{v}{c} = 1 - \frac{1}{2\gamma^2} + O(\gamma^{-4}) \quad (1.14)$$

which shows that even for a moderate value $\gamma = 10$ the velocity already agrees with the speed of light within half percent. As a result, *description of particle physics phenomena requires special relativity*, mostly in the extreme relativistic regime. Therefore we summarise the elements of relativity in the following Section.

2 Special relativity

2.1 Four-vectors and Lorentz invariance

In Minkowski's four-vector formalism, the *space-time* coordinates are

$$\begin{aligned} x^\mu &= (x^0, \vec{x}) \\ x^0 &= ct \end{aligned} \quad (2.1)$$

The invariant Minkowski 'distance' is

$$\begin{aligned} ds^2 &= c^2 dt^2 - d\vec{x}^2 = \eta_{\mu\nu} dx^\mu dx^\nu \\ \eta_{\mu\nu} &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \end{aligned} \quad (2.2)$$

and the trajectory of light rays corresponds to four-dimensional curves (*world-lines*) that satisfy

$$ds^2 = 0 \Rightarrow \left| \frac{d\vec{x}}{dt} \right| = c \quad (2.3)$$

The vectors we introduced so far are all of the variety called *contravariant*, which is signalled by upper four-indices. Given a contravariant vector v^μ , it is convenient to introduce the *covariant* form

$$v_\mu = \eta_{\mu\nu} v^\nu \quad (2.4)$$

i.e. we have the following relations which allow for lowering and raising indices whenever necessary:

$$\begin{aligned} v_0 &= v^0 \\ v_i &= -v^i \end{aligned} \quad (2.5)$$

The inverse of this relation can be written as

$$v^\mu = \eta^{\mu\nu} v_\nu \quad (2.6)$$

where $\eta^{\mu\nu}$ is the matrix inverse of $\eta_{\mu\nu}$; in special relativity, the two matrices are numerically equal (but the corresponding objects are conceptually different: one is the metric tensor and the other is its inverse).

We demand that the above relations must be preserved by transformation between inertial systems. As a result, we can make transformations of the form

$$x'^\mu = \Lambda^\mu_\nu x^\nu + a^\mu \quad (2.7)$$

such that

$$ds'^2 = ds^2 \quad (2.8)$$

Such transformations are called *Poincaré transformations*. The above condition can be written

$$\eta_{\mu\nu} \Lambda^\mu_\sigma \Lambda^\nu_\tau dx^\sigma dx^\tau = \eta_{\sigma\tau} dx^\sigma dx^\tau \quad (2.9)$$

which should hold for any dx^μ , leading to

$$\eta_{\mu\nu}\Lambda^\mu_\sigma\Lambda^\nu_\tau = \eta_{\sigma\tau} \quad (2.10)$$

These look like 16 equations, but in fact they are symmetric in σ and τ , so only 10 are independent. Therefore, the 16 unknowns Λ^μ_ν are determined in terms of 6 arbitrary parameters. Before we solve these equations, let us consider the properties of the solutions. Writing these into matrix notations we have

$$\Lambda^T \eta \Lambda = \eta \quad (2.11)$$

The following properties follow immediately:

1. The unit matrix $\Lambda^\mu_\nu = \delta^\mu_\nu$ is a solution.
2. Any solution is invertible, since $\det \Lambda = \pm 1 \neq 0$.
3. The inverse of any solution is a solution, since

$$\begin{aligned} \Lambda^T \eta \Lambda = \eta &\Rightarrow \eta = (\Lambda^T)^{-1} \eta \Lambda^{-1} \\ \text{and} &\quad (\Lambda^T)^{-1} = (\Lambda^{-1})^T \end{aligned} \quad (2.12)$$

4. If Λ_1 and Λ_2 are solutions, then $\Lambda_1 \Lambda_2$ is a solution:

$$(\Lambda_1 \Lambda_2)^T \eta \Lambda_1 \Lambda_2 = \Lambda_2^T \Lambda_1^T \eta \Lambda_1 \Lambda_2 = \Lambda_2^T \eta \Lambda_2 = \eta \quad (2.13)$$

Therefore the set of real matrices

$$\mathcal{L} = \{\Lambda : \Lambda^T \eta \Lambda = \eta\} \quad (2.14)$$

forms a group, using the matrix product as group multiplication. This is called the *Lorentz group*, and a Poincaré transformation with no translation ($a^\mu = 0$) is called a *Lorentz transformation*. Note that Poincaré transformations also form a group.

Exercise 2.1. Show that the inverse of a Lorentz matrix can be obtained by lowering the first index and raising the second one and then transposing them

$$(\Lambda^{-1})^\tau_\nu = \Lambda^\tau_\nu = \eta_{\nu\nu'} \eta^{\tau\tau'} \Lambda^{\nu'}_{\tau'} \quad (2.15)$$

Solution.

$$\eta_{\mu\nu} \Lambda^\mu_\sigma \Lambda^\nu_\tau = \eta_{\sigma\tau} \Rightarrow (\Lambda^{-1})^\tau_\nu = \eta_{\mu\nu} \Lambda^\mu_\sigma \eta^{\sigma\tau} \quad (2.16)$$

Another way to see this is that one must have

$$(\Lambda^{-1})^\tau_\nu \Lambda^\nu_\rho = \delta^\tau_\rho \quad (2.17)$$

and indeed

$$\eta_{\mu\nu} \Lambda^\mu_\sigma \eta^{\sigma\tau} \Lambda^\nu_\rho = \eta_{\mu\nu} \Lambda^\mu_\sigma \Lambda^\nu_\rho \eta^{\sigma\tau} = \eta_{\sigma\rho} \eta^{\sigma\tau} = \delta^\tau_\rho \quad (2.18)$$

□

Exercise 2.2. For a contravariant vector v transforming as

$$v'^{\mu} = \Lambda^{\mu}_{\nu} v^{\nu} \quad (2.19)$$

compute the matrix $\bar{\Lambda}$ transforming the covariant vector:

$$v'_{\mu} = \bar{\Lambda}_{\mu}^{\nu} v_{\nu} \quad (2.20)$$

Solution. Writing the contravariant transformation in the form

$$\eta^{\mu\tau} v'_{\tau} = \Lambda^{\mu}_{\sigma} \eta^{\sigma\nu} v_{\nu} \quad (2.21)$$

we get

$$v'_{\mu} = \eta_{\mu\tau} \Lambda^{\tau}_{\sigma} \eta^{\sigma\nu} v_{\nu} \quad (2.22)$$

so

$$\bar{\Lambda}_{\mu}^{\nu} = \eta_{\mu\tau} \Lambda^{\tau}_{\sigma} \eta^{\sigma\nu} \quad (2.23)$$

or, using the previous result

$$\bar{\Lambda} = (\Lambda^{-1})^T \quad (2.24)$$

□

Let us comment on the transposition involved in equation (2.24). Note that the indices of a (contravariant) Lorentz transformation have the order

$$\Lambda^{\mu}_{\nu}$$

i.e.. the first one is upper and the second one is lower. This is also true for the inverse matrix. However, the covariant matrix must have its first index down and the second up, which is achieved by transposing the indices of the inverse matrix. This also ensures that the group property of the Lorentz matrices is preserved in the covariant form:

Exercise 2.3. Show that for any two Lorentz matrices Λ_1 and Λ_2 with the product

$$\Lambda_{12} = \Lambda_1 \Lambda_2$$

it is also true that

$$\bar{\Lambda}_{12} = \bar{\Lambda}_1 \bar{\Lambda}_2$$

Solution. It is easiest to do this in the matrix form

$$\bar{\Lambda}_{12} = (\Lambda_{12}^{-1})^T = (\Lambda_2^{-1} \Lambda_1^{-1})^T = (\Lambda_1^{-1})^T (\Lambda_2^{-1})^T = \bar{\Lambda}_1 \bar{\Lambda}_2$$

□

This relation means that the matrices $\bar{\Lambda}$ form a representation of the Lorentz group, which is just obtained by taking the inverse and the transpose, called the *contragredient*. One can also

visualize the contravariant-covariant distinction by writing the contravariant vectors as column vectors:

$$\begin{pmatrix} v^0 \\ v^1 \\ v^2 \\ v^3 \end{pmatrix}$$

Action of a 4×4 matrix Λ on such a column vector leads to another column vector by the standard rules of matrix multiplications. Covariant vectors are, on the other hand, row vectors

$$(w_0 \ w_1 \ w_2 \ w_3)$$

Multiplying a row vector with a column vector leads to a 1×1 matrix, i.e. a number:

$$\begin{aligned} (w_0 \ w_1 \ w_2 \ w_3) \begin{pmatrix} v^0 \\ v^1 \\ v^2 \\ v^3 \end{pmatrix} &= w_0 v^0 + w_1 v^1 + w_2 v^2 + w_3 v^3 = w_\mu v^\mu \\ &= w^0 v_0 - w^1 v_1 - w^2 v_2 - w^3 v_3 = \eta_{\mu\nu} v^\mu w^\nu \\ &= w_0 v_0 - w_1 v_1 - w_2 v_2 - w_3 v_3 = \eta^{\mu\nu} v_\mu w_\nu \end{aligned}$$

The fact that the covariant transformation is the contragredient of the contravariant one ensures that the above so-called Minkowski product is a scalar quantity:

Exercise 2.4. Show that the Minkowski product

$$v^\mu w^\nu \eta_{\mu\nu} = v^\mu w_\mu = v_\mu w^\mu \quad (2.25)$$

is Lorentz invariant.

Solution.

$$v'_\mu w'^\mu = \Lambda_\mu^\nu v_\nu \Lambda^\mu_\sigma w^\sigma = (\Lambda^{-1})^\mu_\sigma \Lambda_\mu^\nu v_\nu w^\sigma = \delta_\sigma^\nu v_\nu w^\sigma = v_\sigma w^\sigma \quad (2.26)$$

□

Exercise 2.5. Show that the vector of coordinate derivatives

$$\partial_\mu = \frac{\partial}{\partial x^\mu} \quad (2.27)$$

transforms as a covariant four-vector.

Solution. This can be shown by the following consideration: supposing that

$$\partial'_\mu = \bar{\Lambda}_\mu^\nu \partial_\nu \quad (2.28)$$

we obtain

$$\begin{aligned} \partial'_\mu x'^\nu &= \delta_\mu^\nu = \partial_\mu x^\nu \\ &\Downarrow \\ \bar{\Lambda}_\mu^\tau \partial_\tau \Lambda^\nu_\sigma x^\sigma &= \delta_\mu^\nu \\ &\Downarrow \\ \bar{\Lambda}_\mu^\sigma \Lambda^\nu_\sigma &= \delta_\mu^\nu \end{aligned} \quad (2.29)$$

so indeed

$$\bar{\Lambda} = (\Lambda^{-1})^T \quad (2.30)$$

□

A quantity transforming like a product of vectors is a *tensor*

$$(T')^{\mu_1 \dots \mu_m}_{\nu_1 \dots \nu_n} = \Lambda^{\mu_1}_{\sigma_1} \dots \Lambda^{\mu_m}_{\sigma_m} \Lambda^{\tau_1}_{\nu_1} \dots \Lambda^{\tau_n}_{\nu_n} T^{\sigma_1 \dots \sigma_m}_{\tau_1 \dots \tau_n} \quad (2.31)$$

of m contravariant and n covariant indices, called an (m, n) -tensor. Relation

$$\eta_{\mu\nu} \Lambda^\mu_\sigma \Lambda^\nu_\tau = \eta_{\sigma\tau} \quad (2.32)$$

shows that $\eta_{\mu\nu}$ is an invariant $(0, 2)$ -tensor; similarly, $\eta^{\mu\nu}$ is an invariant $(2, 0)$ -tensor. The $(1, 1)$ -tensor version of η is just identical to Kronecker's delta:

$$\eta^\nu_\mu = \eta_{\mu\sigma} \eta^{\sigma\nu} = \delta^\nu_\mu \quad (2.33)$$

Note that Lorentz invariant tensor operations always contain index sums running over covariant and contravariant indices in pairs which is called *contracting the indices*.

2.2 Examples of Lorentz transformations

Orthogonal transformations We can look for special solutions in the form

$$\Lambda = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & & & \\ 0 & \mathcal{O} & & \\ 0 & & & \end{pmatrix} \quad (2.34)$$

where $\mathcal{O}$ is a real 3×3 matrix.

Exercise 2.6. Show that for Λ being a Lorentz matrix is equivalent to

$$\mathcal{O}^T \mathcal{O} = \mathbf{1}_{3 \times 3} \quad (2.35)$$

Solution. The defining property

$$\Lambda^T \eta \Lambda = \eta \quad (2.36)$$

of Lorentz matrices can be evaluated explicitly as follows:

$$\begin{aligned} & \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & & & \\ 0 & \mathcal{O}^T & & \\ 0 & & & \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & & & \\ 0 & \mathcal{O} & & \\ 0 & & & \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & & & \\ 0 & -\mathcal{O}^T \mathcal{O} & & \\ 0 & & & \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \end{aligned} \quad (2.37)$$

□

So in this case $\mathcal{O}$ can be an orthogonal transformation, i.e. any rotation or reflection of the three spatial coordinates x^1, x^2, x^3 . Rotations are those transformations which preserve orientation, i.e.

$$\det \mathcal{O} = +1 \quad (2.38)$$

Any other orthogonal matrix can be generated from rotations by multiplication with spatial parity transformation

$$P = -\mathbf{1}_{3 \times 3} \quad (2.39)$$

Rotations themselves have three continuous angular parameters; one can use the Euler angles familiar from the motion of rigid bodies, or parametrise rotations by giving a vector

$$\vec{\varphi} \quad (2.40)$$

whose direction gives the axis of rotation and length gives the angle.

Boosts Let us look at a transformation that acts on time and a spatial dimension, say $x^0 = ct$ and $x^1 = x$. We can write the Lorentz transformation as

$$\begin{aligned} x'^0 &= \alpha x^0 + \beta x^1 \\ x'^1 &= \gamma x^0 + \delta x^1 \end{aligned} \quad (2.41)$$

Exercise 2.7. Find the general solution for the coefficients $\alpha, \beta, \gamma, \delta$.

Solution. Keeping only the first two rows and columns of the equation □

$$\Lambda^T \eta \Lambda = \eta \quad (2.42)$$

the coefficients satisfy the equations

$$\begin{pmatrix} \alpha & \gamma \\ \beta & \delta \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad (2.43)$$

or, in components

$$\begin{aligned} \alpha^2 - \gamma^2 &= 1 \\ \beta^2 - \delta^2 &= -1 \\ \alpha\beta - \gamma\delta &= 0 \end{aligned} \quad (2.44)$$

The first two equations are solved by

$$\begin{aligned} \alpha &= \cosh \eta \\ \gamma &= \sinh \eta \\ \delta &= \cosh \eta' \\ \beta &= \sinh \eta' \end{aligned} \quad (2.45)$$

while the third one gives

$$\tanh \eta = \tanh \eta' \Rightarrow \eta = \eta' \quad (2.46)$$

So we get the transformation rules

$$\begin{aligned}x'^0 &= x^0 \cosh \eta + x^1 \sinh \eta \\x'^1 &= x^0 \sinh \eta + x^1 \cosh \eta\end{aligned}\tag{2.47}$$

Reintroducing the usual time variable, we get

$$\begin{aligned}ct' &= ct \cosh \eta + x \sinh \eta \\x' &= ct \sinh \eta + x \cosh \eta\end{aligned}\tag{2.48}$$

What is the physical meaning of this transformation? Let's see how the origin of the system K' moves in K :

$$\begin{aligned}0 &= x' = ct \sinh \eta + x \cosh \eta \\&\Downarrow \\x &= -ct \tanh \eta\end{aligned}\tag{2.49}$$

So we get

$$\tanh \eta = -\frac{v}{c}\tag{2.50}$$

where v is the velocity of K' which moves along the x direction with respect to K . Using the identities

$$\sinh \eta = \frac{\tanh \eta}{\sqrt{1 - \tanh^2 \eta}} \quad \cosh \eta = \frac{1}{\sqrt{1 - \tanh^2 \eta}}\tag{2.51}$$

we get the so-called Lorentz boost in the x direction

$$t' = \frac{t - vx/c^2}{\sqrt{1 - v^2/c^2}} \quad x' = \frac{x - vt}{\sqrt{1 - v^2/c^2}}\tag{2.52}$$

while also $y' = y$ and $z' = z$. There are two other independent boost transformations along y and z .

It is clear that for a general relative velocity vector $\vec{v}$ of K' with respect to K the proper generalisation is

$$t' = \frac{t - \vec{v} \cdot \vec{x}/c^2}{\sqrt{1 - v^2/c^2}} \quad \vec{x}' = \frac{\vec{x} - \vec{v}t}{\sqrt{1 - v^2/c^2}}\tag{2.53}$$

since one can always take the first axis parallel to $\vec{v}$, in which case these expressions reduce to the former case. So the boosts depend on three real parameters, which are the components of $\vec{v}$. Together with the three angular parameters for rotations, this indeed makes 6 continuous parameters altogether for the general Lorentz transformations.

2.3 Parity and time reversal

Note also that it would also have been possible to take

$$\begin{aligned}\alpha &= \delta = -\cosh \eta \\ \beta &= \gamma = \sinh \eta\end{aligned}\tag{2.54}$$

Such a choice would make t' flow backwards with respect to t . Besides the rotations and boosts, there are therefore two *discrete Lorentz transformations*, namely

$$\text{parity} : \mathcal{P} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \quad \text{time reversal} : \mathcal{T} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (2.55)$$

In general, it is possible to prove that

$$|\Lambda_0^0| \geq 1 \quad (2.56)$$

Transformations with $\Lambda_0^0 \geq 1$ preserve, while those with $\Lambda_0^0 \leq -1$ reverse the arrow of time. The Lorentz group is made of four components

$$\mathcal{L} = \mathcal{L}_+^\uparrow \cup \mathcal{L}_+^\downarrow \cup \mathcal{L}_-^\uparrow \cup \mathcal{L}_-^\downarrow \quad (2.57)$$

where the arrow shows the action on the direction of time, and the index $\pm$ corresponds to the determinant of the matrix being ± 1 . All rotations and boosts, together with the identity are in the component $\mathcal{L}_+^\uparrow$, while the rest of the components can be generated from the first by multiplication with a particular discrete transformation, namely:

$$\begin{aligned} \mathcal{L}_+^\downarrow &: \text{by } \mathcal{PT} \\ \mathcal{L}_-^\uparrow &: \text{by } \mathcal{P} \\ \mathcal{L}_-^\downarrow &: \text{by } \mathcal{T} \end{aligned} \quad (2.58)$$

2.4 Free relativistic particle

Take an object in inertial motion. In its proper system of reference we have

$$d\tau^2 = \frac{1}{c^2} ds^2 \quad (2.59)$$

where τ is the object's proper time. The proper time elapsed along any trajectory $\vec{x}(t)$ is

$$\tau = \int_{t_1}^{t_2} dt \sqrt{1 - \vec{\beta}^2(t)} \quad (2.60)$$

where the square root factor corresponds to time dilatation, with

$$\vec{\beta}(t) = \frac{\vec{v}}{c} = \frac{1}{c} \frac{d\vec{x}}{dt} \quad (2.61)$$

being the velocity in units of c . For a point particle, we define its motion by Fermat's principle. The particle takes the path along which the elapsed proper time is extremal (in this case, the extremum is a maximum). So *the action for a free relativistic point particle* is

$$S = -A \int_{t_1}^{t_2} dt \sqrt{1 - \dot{\vec{x}}(t)^2/c^2} \quad (2.62)$$

with the Lagrangian

$$L = -A \sqrt{1 - \dot{\vec{x}}(t)^2/c^2} \quad (2.63)$$

Note that

$$\dot{\vec{x}}(t)^2 = \dot{\vec{x}}(t) \cdot \dot{\vec{x}}(t) = \dot{x}^i(t)\dot{x}^i(t) \quad (2.64)$$

since we use here the standard Euclidean three-vector scalar product.

Exercise 2.8. Compute the canonical momentum, the Hamiltonian and the equations of motion from the above Lagrangian.

Solution. The canonical momentum is

$$p_i = \frac{\partial L}{\partial \dot{x}^i} = A \frac{\dot{x}_i(t)/c^2}{\sqrt{1 - \dot{\vec{x}}(t)^2/c^2}} \quad (2.65)$$

and the Hamiltonian is

$$H = p_i \dot{x}^i(t) - L = \frac{A}{\sqrt{1 - \dot{\vec{x}}(t)^2/c^2}} = \sqrt{A^2 + \vec{p}^2 c^2} \quad (2.66)$$

The Euler-Lagrange equations are

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{x}^i} - \frac{\partial L}{\partial x^i} = 0 \quad \Rightarrow \quad \frac{dp_i}{dt} = 0 \quad (2.67)$$

which is just the correct equation of motion for a free particle. $\square$

Now let us introduce a new notation for A as

$$A = mc^2 \quad (2.68)$$

so that we obtain the relativistic expressions for the *action*, *momentum* and *energy* of a free relativistic particle of mass m :

$$\begin{aligned} S &= -mc^2 \int_{t_1}^{t_2} dt \sqrt{1 - \dot{\vec{x}}(t)^2/c^2} \\ p^i &= -p_i = \frac{m\dot{x}^i(t)}{\sqrt{1 - \dot{\vec{x}}(t)^2/c^2}} \\ \mathcal{E} &= \frac{mc^2}{\sqrt{1 - \dot{\vec{x}}(t)^2/c^2}} \end{aligned} \quad (2.69)$$

Note that for $|\dot{\vec{x}}| \ll c$ the Lagrangian becomes

$$L = -mc^2 \sqrt{1 - \dot{\vec{x}}(t)^2/c^2} = -mc^2 + \frac{1}{2}m\dot{\vec{x}}(t)^2 + O(\dot{\vec{x}}(t)^4/c^2) \quad (2.70)$$

so, apart from a constant term it reduces to the known non-relativistic Lagrangian.

2.5 The velocity and energy-momentum four-vectors

We can write the energy and momentum in four vector notation by introducing the *four-velocity*

$$u^\mu = \frac{dx^\mu}{d\tau} = \frac{1}{\sqrt{1 - \dot{\vec{x}}(t)^2/c^2}} \left(c, \dot{\vec{x}}(t) \right) \quad (2.71)$$

Since τ is an invariant, this is a four-vector, which always satisfies

$$u^\mu u^\nu \eta_{\mu\nu} = c^2 \quad (2.72)$$

so that the energy and momentum can be unified into the *energy-momentum four-vector*

$$p^\mu = mu^\mu = (\mathcal{E}/c, \vec{p}) \quad (2.73)$$

Note also that

$$p^\mu p^\nu \eta_{\mu\nu} = m^2 c^2 \quad (2.74)$$

or explicitly

$$\mathcal{E}^2 - \vec{p}^2 c^2 = m^2 c^4 \quad (2.75)$$

Now one can see that it is possible to have a *zero-mass object*, which satisfies

$$\mathcal{E} = |\vec{p}|c \quad (2.76)$$

and, since

$$\mathcal{E} \sqrt{1 - \dot{\vec{x}}(t)^2/c^2} = mc^2 = 0 \quad (2.77)$$

it will always move with the speed of light

$$\left| \frac{d\vec{x}}{dt} \right| = c \quad (2.78)$$

We can deduce from this that *photons must have zero mass*, since for in other case finite and real energy implies

$$\left| \frac{d\vec{x}}{dt} \right| < c \quad (2.79)$$

i.e. massive objects always move slower than c .

3 Particles: classification by quantum numbers and interactions

3.1 Units $\hbar = c = k_B = 1$

In particle physics there are *natural units* which can be expressed in terms of the unit of energy. This is shown in the following table:

Quantity	Natural unit	Short notation	SI
Energy	eV	eV	$1.6021766 \cdot 10^{-19}$ J
Mass	eV/c^2	eV	$1.782662 \cdot 10^{-36}$ kg
Distance	$\hbar c/\text{eV}$	eV^{-1}	$1.97327 \cdot 10^{-7}$ m
Time	$\hbar/\text{eV}$	eV^{-1}	$6.582119 \cdot 10^{-16}$ s
Momentum	eV/c	eV	$5.344286 \cdot 10^{-28}$ kg·m/s
Temperature	eV/k_B	eV	11604.5 K
Velocity	c	1	299792458 m/s

In these units, the mass of the electron is

$$m_e = 0.510998910 \text{ MeV} \quad (3.1)$$

while that of the proton is

$$m_p = 938.27205 \text{ MeV} = 0.93827205 \text{ GeV} \quad (3.2)$$

The atomic mass unit corresponds to

$$1 \text{ a.m.u.} = 931.49432 \text{ MeV} \quad (3.3)$$

Depending on the phenomenon, one can also use MeV or GeV as the basis of units. 1 GeV^{-1} is equal to $1.97327 \cdot 10^{-16}$ m as a distance and $6.582119 \cdot 10^{-25}$ s as a time unit. The present energy frontier corresponds to the operation of the LHC, which runs at a collision energy of 13 TeV that corresponds to distance scales a few times 10^{-20} m.

From now on we mainly use the natural units and put $\hbar = c = k_B = 1$ (with some exceptions for convenience or didactic purposes).

3.1.1 Electromagnetic interactions

Let us consider the magnitude of electromagnetic interactions in these units. The potential between two elementary charges is

$$V = \frac{e^2}{4\pi\epsilon_0 r} \quad (3.4)$$

where r is distance in meters. We can introduce the distance in the natural units by

$$r [\text{m}] = \hbar c R [\text{eV}^{-1}] \quad (3.5)$$

and write

$$V = \frac{\alpha}{R} \quad (3.6)$$

where

$$\alpha = \frac{e^2}{4\pi\epsilon_0\hbar c} \approx \frac{1}{137.0359992} \quad (3.7)$$

is the *fine structure constant* which gives the strength of the electromagnetic interaction. Since $\alpha \ll 1$ the electromagnetic interaction is usually weak; the exception is when there are very strong electric or magnetic fields.

We can get a better handle on this by looking at the hydrogen atom. The non-relativistic Schrödinger equation in natural units is

$$-\frac{1}{2m}\Delta\Psi - \frac{\alpha}{R}\Psi = E\Psi \quad (3.8)$$

Suppose the electron is localised at a typical distance d , which then gives the characteristic scale over which the wave function Ψ varies. Then the size of the derivative term can be estimated as

$$|\Delta\Psi| \sim \frac{1}{d^2}|\Psi| \quad (3.9)$$

In a bound state the kinetic and potential energies are similar in magnitude, so

$$\frac{1}{2m}|\Delta\Psi| \sim \frac{1}{2md^2}|\Psi| \sim \frac{\alpha}{d}|\Psi| \quad (3.10)$$

so the size of the hydrogen atom is

$$d \sim \frac{1}{\alpha m} \approx 268 \text{ MeV}^{-1} \approx 5.3 \cdot 10^{-11} \text{ m} \quad (3.11)$$

This exactly coincides with the Bohr radius. For the energy we get the order of magnitude estimate

$$E \sim -\frac{\alpha}{d} = -\alpha^2 m \approx -27.2 \text{ eV} \quad (3.12)$$

while the actual value for the ground state energy is just half of this number:

$$E_0 \approx -13.6 \text{ eV} \quad (3.13)$$

This shows that our order-of-magnitude estimate is correct. The scales for which the electron behaves as a relativistic quantum particle, on the other hand, correspond to distances of the order of the (reduced) *Compton wavelength*

$$\frac{1}{m} \approx \frac{1}{0.511 \text{ MeV}} \approx 3.86 \cdot 10^{-13} \text{ m} \quad (3.14)$$

The hydrogen atom is two orders of magnitude larger, corresponding to the small value of α . At the same time, the typical energy scale (c.f. E_0) is more than four orders of magnitude smaller than the rest mass of the electron. As a result, the non-relativistic approximation to the spectrum works very well, and the hydrogen atom can be described to a very good precision as a *non-relativistic* quantum system. In fact, with a few notable exceptions , *phenomena in atomic physics and chemistry can be understood using standard non-relativistic quantum mechanics*; relativistic effects give only small corrections that are suppressed by α (such as the fine structure of the hydrogen atom, whence the name *fine structure constant*).

¹The few exceptions include the colour of copper and gold, which is caused by relativistic corrections, and also the liquidity of mercury at room temperature. Relativistic corrections are also important for inner electrons of atoms with a large atomic number Z , which results in an enhanced effective interaction constant $Z\alpha$.

3.1.2 Gravity

Gravity behaves in an entirely different way. Consider the Newtonian potential between two equal masses

$$V_G = \frac{Gm^2}{r} = \frac{m^2}{M_{Pl}^2 R} \quad (3.15)$$

where $G = 6.674 \cdot 10^{-11} \text{Nm}^2/\text{kg}^2$ is Newton's gravitational constants and

$$M_{Pl} = \sqrt{\frac{\hbar c}{G}} \approx 2.1765 \cdot 10^{-8} \text{kg} \approx 1.2209 \cdot 10^{19} \text{ GeV} \quad (3.16)$$

is the Planck mass. For two protons at a given distance, gravitational interaction is much weaker than the electromagnetic one, indeed, it is smaller by a factor

$$\frac{m_p^2}{M_{Pl}^2} / \alpha \sim 10^{-36} \quad (3.17)$$

As a result, *gravity can be neglected in all high-energy particle phenomena that we can presently study by experiments*^[2] However, we can expect quantum gravitational phenomena to appear at length and time scales M_{Pl}^{-1} which correspond to the

$$\begin{aligned} \text{Planck length:} \quad l_{Pl} &= \sqrt{\frac{\hbar G}{c^3}} \approx 1.6162 \cdot 10^{-35} \text{ m} \\ \text{Planck time:} \quad t_{Pl} &= \sqrt{\frac{\hbar G}{c^5}} \approx 5.391 \cdot 10^{-44} \text{ s} \end{aligned} \quad (3.18)$$

3.2 Symmetries, conserved quantities and quantum numbers

We consider a quantum system with a Hilbert space $\mathcal{H}$ and a Hamiltonian operator H , so that the time evolution is described by the Schrödinger equation

$$i \frac{\partial}{\partial t} |\Psi(t)\rangle = H |\Psi(t)\rangle \quad (3.19)$$

We recall that observables correspond to self-adjoint operators $\mathcal{O} : \mathcal{H} \rightarrow \mathcal{H}$. Given such an operator, its eigenvalues and eigenstates satisfy^[3]

$$\mathcal{O} |\Phi_\alpha\rangle = \lambda_\alpha |\Phi_\alpha\rangle \quad (3.20)$$

and the eigenstates can be taken to form a full orthonormal basis:

$$\langle \Phi_\alpha | \Phi_\beta \rangle = \delta_{\alpha\beta} \quad (3.21)$$

²We remark that the effect of gravity on quantum mechanical systems can already be studied directly using trapped cold neutrons; the measurements agree with the predictions obtained by including the gravitational potential in the non-relativistic Schrödinger equation.

³We ignore here the issues related to operators with continuous spectra, such as position and momentum. They can be treated using the spectral theorem of functional analysis, which we can consider to be a technical refinement of the simple-minded formalism used here, without any novel physical content.

When undertaking a measurement of the observable $\mathcal{O}$, the *possible outcomes* are given by the *eigenvalues* λ_α , while the *probability of obtaining a given outcome* in a state $|\Psi\rangle$ is given by the *Born rule*

$$p_\alpha = |\langle\Psi|\Phi_\alpha\rangle|^2 \quad (3.22)$$

A *symmetry* is nothing else then a map

$$\mathcal{U} : \mathcal{H} \rightarrow \mathcal{H} \quad (3.23)$$

that preserves the probabilities of any possible measurement in any possible state. E. Wigner has proven the following result:

Theorem 3.1. (*Wigner's theorem*) Any symmetry $\mathcal{U}$ can be described by either a unitary or an anti-unitary operator, i.e. it must satisfy

$$\mathcal{U}\mathcal{U}^\dagger = \mathcal{U}^\dagger\mathcal{U} = \mathbf{1} \quad (3.24)$$

and be either

$$\begin{aligned} \text{linear:} \quad & \mathcal{U}(\alpha|\Psi_1\rangle + \beta|\Psi_2\rangle) = \alpha\mathcal{U}|\Psi_1\rangle + \beta\mathcal{U}|\Psi_2\rangle \\ \text{or antilinear:} \quad & \mathcal{U}(\alpha|\Psi_1\rangle + \beta|\Psi_2\rangle) = \alpha^*\mathcal{U}|\Psi_1\rangle + \beta^*\mathcal{U}|\Psi_2\rangle \end{aligned} \quad (3.25)$$

where $|\Psi_{1,2}\rangle \in \mathcal{H}$ and $\alpha, \beta \in \mathbb{C}$.

In addition, a *symmetry must preserve the time evolution*, i.e. if we transform the state and then evolve it in time, it must be the same as first evolving and then transforming it. Therefore

$$\mathcal{U}H\mathcal{U}^{-1} = H \quad (3.26)$$

i.e. *the symmetry operator must commute with the Hamiltonian*

$$\mathcal{U}H = H\mathcal{U} \Leftrightarrow [H, \mathcal{U}] = 0 \quad (3.27)$$

Therefore H and $\mathcal{U}$ have a common eigenstate basis. Let us suppose now that $\mathcal{U}$ is unitary. The above property implies that taking an eigenstate of $\mathcal{U}$ as initial condition

$$\mathcal{U}|\Psi(0)\rangle = \eta|\Psi(0)\rangle \quad (3.28)$$

the time evolved state will be an eigenstate of $\mathcal{U}$ with the same eigenvalue, independent of time

$$\mathcal{U}|\Psi(t)\rangle = \eta|\Psi(t)\rangle \quad (3.29)$$

This is easy to prove

$$\begin{aligned} \partial_t (\mathcal{U}|\Psi(t)\rangle - \eta|\Psi(t)\rangle) &= \mathcal{U}\partial_t|\Psi(t)\rangle - \eta\partial_t|\Psi(t)\rangle = -\mathcal{U}iH|\Psi(t)\rangle + \eta iH|\Psi(t)\rangle \\ &= -iH\mathcal{U}|\Psi(t)\rangle + \eta iH|\Psi(t)\rangle = -iH\eta|\Psi(t)\rangle + \eta iH|\Psi(t)\rangle = 0 \end{aligned} \quad (3.30)$$

So energy eigenstates of the Hamiltonian can be characterised by the eigenvalue of $\mathcal{U}$. These eigenvalues are called *quantum numbers*. There are two distinct cases according to the nature of $\mathcal{U}$: discrete and continuous symmetries.

3.2.1 Discrete symmetries

The prime example of discrete symmetries is *parity* (spatial reflection), represented by a unitary operator $\mathcal{P}$ on the Hilbert space. The eigenvalue of $\mathcal{P}$ is called the *parity of the state*.

Another example of discrete symmetry is *time reversal*, which can be represented by an operator $\mathcal{T}$. However, under time reversal one expects that while momentum changes sign (particle propagating back in time), energy does not:

$$\mathcal{T}\vec{P}\mathcal{T}^{-1} = -\vec{P} \quad \mathcal{T}H\mathcal{T}^{-1} = H \quad (3.31)$$

Should energy change sign, the spectrum would not be bounded from below, leading to an unstable physical system without a ground state. On the other hand, $\mathcal{T}$ must change time evolution to run backwards:

$$\mathcal{T}e^{-itH}\mathcal{T}^{-1} = e^{itH} \quad (3.32)$$

So $\mathcal{T}$ must be an anti-unitary operator. We now show that the eigenvalue of such an operator has no physical meaning. First of all, due to $\mathcal{T}^\dagger = \mathcal{T}^{-1}$, any eigenvalue must be a phase. Second, if we find an eigenstate

$$\mathcal{T}|\psi\rangle = \xi|\psi\rangle \quad (3.33)$$

then redefining the phase of the state

$$|\psi'\rangle = e^{i\alpha}|\psi\rangle \quad (3.34)$$

it satisfies

$$\mathcal{T}|\psi'\rangle = \xi e^{-2i\alpha}|\psi'\rangle \quad (3.35)$$

However, states related by a phase redefinition are physically equivalent, since they do not differ in any observable properties. Therefore, *the eigenvalue of an anti-unitary symmetry operator cannot have any physical significance*.

3.2.2 Continuous symmetries: energy and momentum

An example of a continuous symmetry is *time translation invariance*. In fact, the Schrödinger equation

$$i\frac{\partial}{\partial t}|\Psi(t)\rangle = H|\Psi(t)\rangle \quad (3.36)$$

tells us that for a Hamiltonian that does not depend explicitly on t we have

$$|\Psi(t + \Delta t)\rangle = e^{-iH\Delta t}|\Psi(t)\rangle \quad (3.37)$$

So *energy is nothing else than the generator of time translations*.

Recalling that in one-particle quantum mechanics the momentum operator is

$$\vec{p} = -i\vec{\nabla} \quad (3.38)$$

one can see that *momentum is the generator of spatial translations*

$$\Psi(\vec{x} + \Delta\vec{x}, t) = e^{i\Delta\vec{x}\cdot\vec{p}}\Psi(\vec{x}, t) \quad (3.39)$$

In fact, the general definition of momentum in a quantum theory is exactly via its identification with the generator of spatial translations. Let us denote the unitary operator implementing spatial translations by a vector $\vec{a}$ as

$$\mathcal{U}_s(\vec{a}) \quad (3.40)$$

Since by the group property of translations

$$\mathcal{U}_s(\vec{a})\mathcal{U}_s(\vec{a}') = \mathcal{U}_s(\vec{a} + \vec{a}') \quad (3.41)$$

it can always be written as an exponential

$$\mathcal{U}_s(\vec{a}) = e^{i\vec{a} \cdot \vec{P}} \quad (3.42)$$

Unitarity of $\mathcal{U}_s(\vec{a})$ implies that P^i , $i = 1, 2, 3$ are Hermitian operators, which are defined to be the spatial components of the physical momentum of the quantum system. This can be joined with the energy to form the relativistic energy-momentum 4-vector P^μ by defining

$$P^0 = H \quad (3.43)$$

and the unitary operator implementing the 4-translation $x^\mu \rightarrow x^\mu + a^\mu$ can be written as

$$e^{-ia^\mu P_\mu} \quad (3.44)$$

3.2.3 Continuous symmetries: example of electric charge

Let us consider now the operator $\mathcal{Q}$ which corresponds to electric charge. We know that electric charge is conserved, therefore it must commute with the Hamiltonian

$$[H, \mathcal{Q}] = 0 \quad (3.45)$$

The transformations

$$\mathcal{U}_q(\alpha) = e^{i\alpha\mathcal{Q}} \quad (3.46)$$

form a *representation of the group* $U(1)$ composed of the complex numbers of unit modulus

$$U(1) = \{e^{i\alpha} \mid -\pi \leq \alpha < \pi\} \quad (3.47)$$

provided all eigenvalues of $\mathcal{Q}$ are quantised in integers. This is satisfied for all observed particles if $\mathcal{Q}$ is measured in units of the elementary charge.

Other examples of similar charges include *leptonic charge*, *baryonic charge* and *strangeness*, all of which will be encountered in the sequel.

3.2.4 Charge conjugation

Most of the particles in nature have an *antiparticle*, which has the same mass, but all of its conserved $U(1)$ charges (electric, leptonic, baryonic etc.) have the opposite sign. Some particles, such as the photon γ or the neutral pion π^0 are identical to their antiparticle, so they cannot carry any conserved $U(1)$ charge. In fact, it is a consequence of basic principles of relativistic quantum theory that *any particle carrying a non-trivial conserved $U(1)$ charge must have its antiparticle*. One can then define a new discrete symmetry operation called *charge conjugation* and denoted by $\mathcal{C}$, that simply swaps every particle for its antiparticle and vice versa, without changing any quantum numbers of the state apart from reversing the signs of all conserved $U(1)$ charges. A consequence of relativistic quantum theory is that the *masses of particles and antiparticles must be exactly equal* (see the discussion of CPT invariance in Section [6.4](#)).

3.3 The particle zoo

The classification of particles relies on the following properties:

- *the interaction they participate in;*
- *mass;*
- *spin and statistics* (boson/fermion)
- *electric charge;*
- *additional conserved quantum numbers.*

3.3.1 Leptons

Leptons are fermionic, spin 1/2 particles that do not participate in the strong interactions⁴. According to their electromagnetic properties, there are *charged leptons* and neutral ones; the latter are called the *neutrinos*. Their most basic properties are summarised in the table below.

Name	Symbol	Electric charge	Mass (MeV)	Spin	Mean lifetime (s)
Electron/positron	e^-/e^+	$-1/+1$	0.510998910(13)	1/2	stable
Muon/anti-muon	μ^-/μ^+	$-1/+1$	105.6583668(38)	1/2	$2.197019(21) \cdot 10^{-6}$
Tau/anti-tau	τ^-/τ^+	$-1/+1$	1776.84(17)	1/2	$2.906(10) \cdot 10^{-13}$
Electron (anti-)neutrino	$\nu_e/\bar{\nu}_e$	0	< 0.0000022	1/2	unknown
Muon (anti-)neutrino	$\nu_\mu/\bar{\nu}_\mu$	0	< 0.17	1/2	unknown
Tau (anti-)neutrino	$\nu_\tau/\bar{\nu}_\tau$	0	< 15.5	1/2	unknown

3.3.2 Hadrons

Hadrons are particles that participate in strong interactions. Those with integer spins are bosonic and are called *mesons*, while the ones with half-integer spins obey Fermi-Dirac statistics and are called *baryons*⁵. The most important hadrons discovered before the 1970s⁶ are listed in the tables below.

⁴Strong interactions are the origin for the nuclear force responsible for binding atomic nuclei together, but their nature is eventually more involved as we shall see later

⁵These names originate from the masses of the particles. Before the discovery of the τ and the heavier mesons, the rule was that leptons are light, mesons have an intermediate mass and baryons are heavy. Their names are from the Greek words λεπτός (leptós) = "fine, small, thin"; μέσος (mesos) = "intermediate"; βαρύς (barys) = "heavy".

⁶These are the ones which will later play a role in the $SU(3)$ quark model. Since then, a much larger number of hadronic particles have also been discovered, as will be discussed later.

Spin-0 mesons					
Name	Particle	Anti-particle	Electric charge	Rest mass (MeV)	Mean lifetime (s)
Charged pion	π^+	π^-	± 1	139.57018 ± 0.00035	$(2.6033 \pm 0.0005) \cdot 10^{-8}$
Neutral pion	π^0	Self	0	134.9766 ± 0.0006	$(8.52 \pm 0.18) \cdot 10^{-17}$
Eta	η	Self	0	547.862 ± 0.018	$(5.02 \pm 0.19) \cdot 10^{-19}$
Eta prime	η'	Self	0	957.78 ± 0.06	$(3.32 \pm 0.15) \cdot 10^{-21}$
Charged kaon	K^+	K^-	± 1	493.677 ± 0.016	$(1.2380 \pm 0.0021) \cdot 10^{-8}$
Neutral kaon (short)	K_S^0	Self	0	497.614 ± 0.024	$(8.954 \pm 0.004) \cdot 10^{-11}$
Neutral kaon (long)	K_L^0	Self	0	497.614 ± 0.024	$(5.116 \pm 0.021) \cdot 10^{-8}$
Spin-1 mesons					
Charged rho meson	ρ^+	ρ^-	± 1	775.11 ± 0.34	$(4.41 \pm 0.02) \cdot 10^{-24}$
Neutral rho meson	ρ^0	Self	0	775.26 ± 0.25	$(4.45 \pm 0.03) \cdot 10^{-24}$
Omega meson	ω	Self	0	782.65 ± 0.12	$(7.75 \pm 0.07) \cdot 10^{-23}$
Phi meson	ϕ	Self	0	1019.461 ± 0.019	$(1.54 \pm 0.01) \cdot 10^{-22}$
Kaon	K^{*+}	K^{*-}	± 1	891.66 ± 0.026	$(3.26 \pm 0.06) \cdot 10^{-23}$
Kaon	K^{*0}	K^{*0}	0	895.81 ± 0.19	$(1.39 \pm 0.02) \cdot 10^{-23}$

Spin-1/2 baryons					
Name	Particle	Antiparticle	Electric charge	Rest mass (MeV)	Mean lifetime (s)
Proton	p	$\bar{p}$	± 1	$938.272046(21)$	Stable
Neutron	n	$\bar{n}$	0	$939.565379(21)$	880.0 ± 0.9 (free n)
Lambda	Λ^0	$\bar{\Lambda}^0$	0	1115.683 ± 0.006	$(2.632 \pm 0.020) \cdot 10^{-10}$
Sigma	Σ^+	Σ^-	± 1	1189.37 ± 0.07	$(8.018 \pm 0.026) \cdot 10^{-11}$
Sigma	Σ^0	$\bar{\Sigma}^0$	0	1192.642 ± 0.024	$(7.4 \pm 0.7) \cdot 10^{-20}$
Sigma	Σ^-	Σ^+	∓ 1	1197.449 ± 0.030	$(1.479 \pm 0.011) \cdot 10^{-10}$
Xi	Ξ^0	$\bar{\Xi}^0$	0	1314.86 ± 0.20	$(2.90 \pm 0.09) \cdot 10^{-10}$
Xi	Ξ^-	$\bar{\Xi}^+$	∓ 1	1321.71 ± 0.07	$(1.639 \pm 0.015) \cdot 10^{-10}$
Spin-3/2 baryons					
Delta	Δ^{++}	Δ^{--}	± 2	1232 ± 2	$(5.63 \pm 0.14) \cdot 10^{-24}$
Delta	Δ^+	Δ^-	± 1	1232 ± 2	$(5.63 \pm 0.14) \cdot 10^{-24}$
Delta	Δ^0	$\bar{\Delta}^0$	0	1232 ± 2	$(5.63 \pm 0.14) \cdot 10^{-24}$
Delta	Δ^-	Δ^+	∓ 1	1232 ± 2	$(5.63 \pm 0.14) \cdot 10^{-24}$
Sigma	Σ^{*+}	Σ^{*-}	± 1	1382.8 ± 0.4	$(1.839 \pm 0.0041) \cdot 10^{-23}$
Sigma	Σ^{*0}	$\bar{\Sigma}^{*0}$	0	1383.7 ± 1.0	$(1.83 \pm 0.25) \cdot 10^{-23}$
Sigma	Σ^{*-}	Σ^{*+}	∓ 1	1387.2 ± 0.5	$(1.671 \pm 0.089) \cdot 10^{-23}$
Xi	Ξ^{*0}	$\bar{\Xi}^{*0}$	0	1531.80 ± 0.32	$(7.23 \pm 0.40) \cdot 10^{-23}$
Xi	Ξ^{*-}	$\bar{\Xi}^{*+}$	∓ 1	1535.0 ± 0.6	$(6.6 \pm 1.3) \cdot 10^{-23}$
Omega	Ω^-	$\bar{\Omega}^+$	∓ 1	1672.45 ± 0.29	$(8.21 \pm 0.11) \cdot 10^{-11}$

We remark that in contrast to the case of mesons, anti-baryons form a set disjoint from the baryons. For example, Σ^- is not the antiparticle of Σ^+ as evidenced by their different mass; the antiparticle of Σ^+ is $\bar{\Sigma}^-$ which is different from Σ^- . For the mesons, the antiparticle of π^+ is π^- and that of K^+ is K^- . This is explained by the quark model considered in Section [7.3](#).

Note also that there are clearly three different time-scales for the decays. The shortest lifetimes are of order $10^{-24} - 10^{-22}$ seconds and correspond to decay processes mediated by the strong

interactions (*strong decays*). In these processes the decay products are hadrons, e.g.

$$\begin{aligned}\Delta^+ &\rightarrow p + \pi^0 \\ \Delta^+ &\rightarrow n + \pi^+\end{aligned}\tag{3.48}$$

We remark that there by the second half of the 1960s there was also known a huge number of very short-lived *resonances* (with strong decays), corresponding to excited states of the particles listed in the above tables.

The second class of decays has a lifetime between 10^{-20} and 10^{-17} seconds; these are mediated by electromagnetic interactions (*electromagnetic decays*), such as e.g. the decay of the neutral pion

$$\pi^0 \rightarrow 2\gamma\tag{3.49}$$

or of the neutral sigma

$$\Sigma^0 \rightarrow \Lambda^0 + \gamma\tag{3.50}$$

Finally, there are decays mediated by the weak interactions with a time scale typically longer than 10^{-11} seconds, which may involve leptons in the final state such as

$$\begin{aligned}\pi^+ &\rightarrow \mu^+ + \nu_\mu \\ n &\rightarrow p + e^- + \bar{\nu}_e\end{aligned}\tag{3.51}$$

called *weak leptonic/semi-leptonic decays*, or may lead to purely hadronic decay products

$$\begin{aligned}\Lambda^0 &\rightarrow p + \pi^- \\ \Lambda^0 &\rightarrow n + \pi^0\end{aligned}\tag{3.52}$$

called *weak non-leptonic decays*.

4 Relativistic wave equations 1: the Klein-Gordon equation

The considerations in this chapter and the next one are based on the idea of extending quantum mechanics in a relativistic way. Ultimately this simple idea fails and relativistic quantum theory must be formulated in terms of local quantum fields. However, the wave equations we obtain still play an important role, albeit not as relativistic versions of the Schrödinger equation but rather as building blocks for the relativistic field theories of particle physics. To keep the relation to quantum mechanical ideas transparent, in this section and the next one we restore the explicit dependence on $\hbar$ and c .

Klein-Gordon equation is the simplest relativistic field equation, that was originally obtained as a tentative quantum mechanical description of a relativistic point particle. As we shall see there are profound reasons why this quantum mechanical interpretation is not ultimately satisfactory. However, the quantum field theory version of this equation does describe a field with quanta that are spin zero bosonic particles. An important example is the Higgs field in the Standard Model of particle physics.

4.1 Quantum mechanics of a non-relativistic point particle

We recall first the quantum mechanics of a non-relativistic point particle. The Hamilton operator has the form

$$H = \frac{\vec{P}^2}{2m} + V(\vec{X}) \quad (4.1)$$

where X^i and P^i , $i = 1, 2, 3$ are Hermitian operators satisfying the canonical commutation relations

$$[X^k, P^l] = i\hbar\delta^{kl} \quad (4.2)$$

According to von Neumann's theorem, any two ways of representing this algebra are unitarily equivalent to each other. The system is therefore uniquely specified by H , which gives the time-evolution of the state $|\Psi(t)\rangle$ represented by an element of a (separable) Hilbert space $\mathcal{H}$ on which X and P act, as dictated by the Schrödinger equation

$$i\hbar\frac{\partial}{\partial t}|\Psi(t)\rangle = H|\Psi(t)\rangle \quad (4.3)$$

The form of H is dictated by the fact that *the classical limit should give back the dynamics of a point particle of mass m in an external potential V* . The validity of this quantum-classical correspondence can eventually be proven, but the choice of H is *not unique*; it can be modified by any term that vanishes as $\hbar \rightarrow 0$ and so has no effect on the classical limit. Eventually, *specifying the Hamilton operator H as above amounts to defining the dynamics of the system at the quantum level*. There is no unique procedure of “quantisation” that can define a quantum system from its classical limit; all quantisation procedures used in physics are in fact *heuristic arguments* justified by the end result, i.e. the description of observable phenomena (such as e.g. the spectrum of hydrogen atom or molecular vibrations) by the resulting quantum model.

Observables are Hermitian operators on the Hilbert space. Given such an operator $\mathcal{O}$, we can consider its eigenstates⁷

$$\mathcal{O}|\psi_k\rangle = \lambda_k|\psi_k\rangle \quad (4.4)$$

⁷More precisely, to include cases with continuous spectra such as X and P , the following should be formulated using the spectral theorem of functional analysis. However, we shall follow usual habits in treating only the case with discrete spectra and normalisable eigenvectors.

where the $|\psi_k\rangle$ form a complete orthonormal basis of the Hilbert space. Then the fundamental postulate of quantum mechanics states that a measurement of the observable $\mathcal{O}$ at some time t can yield as a result one of the eigenvalues λ_k , with the following probability (*Born's rule*)

$$P_k(t : \mathcal{O} = \lambda_k) = |\langle \Psi(t) | \psi_k \rangle|^2 \quad (4.5)$$

We can now introduce the (improper, i.e. not normalisable in the Hilbert sense) eigenvectors of the position operator satisfying

$$\begin{aligned} X^i |\vec{x}\rangle &= x^i |\vec{x}\rangle \\ \langle \vec{x}' | \vec{x} \rangle &= \delta(\vec{x} - \vec{x}') \end{aligned} \quad (4.6)$$

Exercise 4.1. Prove that $\langle \vec{x}' | \vec{x} \rangle = 0$ for $\vec{x}' \neq \vec{x}$.

Solution. We use the eigenstate equation of X^i , once on the left and once on the right state to write

$$\langle \vec{x}' | X^i | \vec{x} \rangle = x'^i \langle \vec{x}' | \vec{x} \rangle = x^i \langle \vec{x}' | \vec{x} \rangle \quad (4.7)$$

so

$$(x'^i - x^i) \langle \vec{x}' | \vec{x} \rangle = 0 \quad (4.8)$$

□

The state $|\Psi(t)\rangle$ is completely characterised by the wave function

$$\Psi(t, \vec{x}) = \langle \vec{x} | \Psi(t) \rangle \quad (4.9)$$

and the fundamental operators are represented by the derivative and multiplication operators

$$P^k = \frac{\hbar}{i} \frac{\partial}{\partial x^k} \quad X^k = x^k \times \quad (4.10)$$

so that the Schrödinger equation can be written as

$$i\hbar \frac{\partial \Psi(t, \vec{x})}{\partial t} = \left(-\frac{\hbar^2}{2m} \Delta + V(\vec{x}) \right) \Psi(t, \vec{x}) \quad (4.11)$$

The wave function has an interesting interpretation. Let us consider an observable operator which can be expressed with the coordinates alone:

$$\mathcal{O}_f = f(\vec{X}) \quad (4.12)$$

then its expectation value can be written as

$$\langle \Psi(t) | \mathcal{O} | \Psi(t) \rangle = \int d^3x f(\vec{x}) |\Psi(\vec{x}, t)|^2 \quad (4.13)$$

Exercise 4.2. Prove that

$$\langle \Psi(t) | f(\vec{X}) | \Psi(t) \rangle = \int d^3x f(\vec{x}) |\Psi(\vec{x}, t)|^2 \quad (4.14)$$

Solution. The statement that the wave-function completely characterises the state can be written as the identity

$$|\Psi(t)\rangle = \int d^3x |\vec{x}\rangle \Psi(t, \vec{x}) \quad (4.15)$$

The equality can be seen from the fact that the right hand side has the same wave function as the left

$$\langle \vec{y} | \left(\int d^3x |\vec{x}\rangle \Psi(t, \vec{x}) \right) = \int d^3x \langle \vec{y} | \vec{x} \rangle \Psi(t, \vec{x}) = \int d^3x \delta(\vec{y} - \vec{x}) \Psi(t, \vec{x}) = \Psi(t, \vec{y}) \quad (4.16)$$

So we can compute

$$\begin{aligned} \langle \Psi(t) | f(\vec{X}) | \Psi(t) \rangle &= \int d^3x \int d^3y \Psi^\dagger(t, \vec{y}) \langle \vec{y} | f(\vec{X}) | \vec{x} \rangle \Psi(t, \vec{x}) \\ \text{using } f(\vec{X}) | \vec{x} \rangle &= f(\vec{x}) | \vec{x} \rangle \\ &= \int d^3x \int d^3y \Psi^\dagger(t, \vec{y}) f(\vec{x}) \delta(\vec{y} - \vec{x}) \Psi(t, \vec{x}) \\ &= \int d^3x f(\vec{x}) |\Psi(\vec{x}, t)|^2 \end{aligned} \quad (4.17)$$

□

As a result,

$$\rho(\vec{x}, t) = |\Psi(\vec{x}, t)|^2 \quad (4.18)$$

can be interpreted as the probability density of finding the particle in space, and it must be normalised as

$$\int d^3x |\Psi(\vec{x}, t)|^2 = 1 \quad (4.19)$$

In fact, since the total sum of probabilities for any observable must be

$$\sum_k P_k(t : \mathcal{O} = \lambda_k) = 1 \quad (4.20)$$

we always get

$$\sum_k |\langle \Psi(t) | \psi_k \rangle|^2 = \langle \Psi(t) | \Psi(t) \rangle = 1 \quad (4.21)$$

There are two important properties that make this interpretation consistent:

1. As a probability density, ρ must be *positive*, which trivially holds.
2. There must exist a *current density* $\vec{j}(\vec{x}, t)$ such that ρ and $\vec{j}$ satisfy a *continuity equation*:

$$\frac{\partial \rho}{\partial t} + \vec{\nabla} \cdot \vec{j} = 0 \quad (4.22)$$

This is important because this implies that the integral of ρ over all space is independent of time

$$\frac{\partial}{\partial t} \int d^3x \rho(\vec{x}, t) = 0 \quad (4.23)$$

so Schrödinger time evolution *conserves the total sum of probabilities*, i.e. the norm of the wave function, once it is normalised at any particular instant of time.

Exercise 4.3. Prove that the proper current is

$$\vec{j} = -\frac{i\hbar}{2m} \left(\Psi^\dagger \vec{\nabla} \Psi - \Psi \vec{\nabla} \Psi^\dagger \right) \quad (4.24)$$

Solution. We have

$$\begin{aligned} i\hbar \frac{\partial \Psi}{\partial t} &= \left(-\frac{\hbar^2}{2m} \Delta + V \right) \Psi \\ -i\hbar \frac{\partial \Psi^\dagger}{\partial t} &= \left(-\frac{\hbar^2}{2m} \Delta + V \right) \Psi^\dagger \end{aligned} \quad (4.25)$$

so

$$i\hbar \frac{\partial}{\partial t} (\Psi \Psi^\dagger) = -\frac{\hbar^2}{2m} (\Psi^\dagger \Delta \Psi - \Psi \Delta \Psi^\dagger) \quad (4.26)$$

Writing

$$(\Psi^\dagger \Delta \Psi - \Psi \Delta \Psi^\dagger) = \vec{\nabla} \cdot (\Psi^\dagger \vec{\nabla} \Psi - \Psi \vec{\nabla} \Psi^\dagger) \quad (4.27)$$

we can rearrange the above equation as

$$\frac{\partial \rho}{\partial t} + \vec{\nabla} \cdot \vec{j} = 0 \quad (4.28)$$

with $\vec{j}$ as above. □

4.2 The Klein-Gordon equation

For a free particle, the elementary solution of the Schrödinger equation

$$i\hbar \frac{\partial \Psi(t, \vec{x})}{\partial t} = -\frac{\hbar^2}{2m} \Delta \Psi(t, \vec{x}) \quad (4.29)$$

is a plane wave

$$\begin{aligned} \Psi_{\vec{p}}(t, \vec{x}) &= A e^{-i\mathcal{E}t/\hbar + i\vec{p} \cdot \vec{x}/\hbar} \\ \mathcal{E} &= \frac{\vec{p}^2}{2m} \end{aligned} \quad (4.30)$$

which has definite momentum and energy

$$P^i \Psi_{\vec{p}}(t, \vec{x}) = p^i \Psi_{\vec{p}}(t, \vec{x}) \quad H \Psi_{\vec{p}}(t, \vec{x}) = \mathcal{E} \Psi_{\vec{p}}(t, \vec{x}) \quad (4.31)$$

Eventually, the normalisable states can be constructed as wave packets

$$\Psi(t, \vec{x}) = \int \frac{d^3p}{(2\pi\hbar)^{3/2}} A(\vec{p}) e^{-i\mathcal{E}t/\hbar + i\vec{p} \cdot \vec{x}/\hbar} \quad (4.32)$$

with

$$\int d^3p |A(\vec{p})|^2 < \infty \quad (4.33)$$

Exercise 4.4. Compute the normalisation condition for the wave packet amplitude $A(\vec{p})$.

Solution.

$$\begin{aligned}
\int d^3x |\Psi(\vec{x}, t)|^2 &= \int d^3x \int \frac{d^3p}{(2\pi\hbar)^{3/2}} A(\vec{p}) e^{-i\mathcal{E}t/\hbar + i\vec{p}\cdot\vec{x}/\hbar} \int \frac{d^3p'}{(2\pi\hbar)^{3/2}} A^\dagger(\vec{p}') e^{i\mathcal{E}'t/\hbar - i\vec{p}'\cdot\vec{x}/\hbar} \\
&= \int \frac{d^3p d^3p'}{(2\pi)^3} A(\vec{p}) A^\dagger(\vec{p}') e^{-i(\mathcal{E}-\mathcal{E}')t/\hbar} \int \frac{d^3x}{\hbar^3} e^{i(\vec{p}-\vec{p}')\cdot\vec{x}/\hbar} \\
\text{using } &\int \frac{d^3x}{\hbar^3} e^{i(\vec{p}-\vec{p}')\cdot\vec{x}/\hbar} = \int d^3z e^{i(\vec{p}-\vec{p}')\cdot\vec{z}} = (2\pi)^3 \delta(\vec{p}-\vec{p}') \\
&= \int d^3p |A(\vec{p})|^2
\end{aligned} \tag{4.34}$$

so the condition is

$$\int d^3p |A(\vec{p})|^2 = 1 \tag{4.35}$$

□

Exercise 4.5. Show that for a free particle of momentum $\vec{p}$

$$\vec{j} = \frac{\vec{p}}{m} \rho \tag{4.36}$$

which is consistent with a motion of velocity $\vec{p}/m$.

Solution. Using

$$\Psi_{\vec{p}}(t, \vec{x}) = A e^{-i\mathcal{E}t/\hbar + i\vec{p}\cdot\vec{x}/\hbar} \tag{4.37}$$

we can compute directly

$$\begin{aligned}
\rho &= |\Psi_{\vec{p}}|^2 = |A|^2 \\
\vec{j} &= -\frac{i\hbar}{2m} \left(\Psi_{\vec{p}}^\dagger \vec{\nabla} \Psi_{\vec{p}} - \Psi_{\vec{p}} \vec{\nabla} \Psi_{\vec{p}}^\dagger \right) = \frac{\vec{p}}{m} |A|^2
\end{aligned} \tag{4.38}$$

□

However, this describes particles with the *non-relativistic* energy-momentum (“dispersion”) relation

$$\mathcal{E} = \frac{\vec{p}^2}{2m} \tag{4.39}$$

To get *relativistic* particles, we need a dispersion relation of the form

$$\mathcal{E}^2 - \vec{p}^2 c^2 = m^2 c^4 \tag{4.40}$$

The simplest wave equation with this property is

$$\left(\partial_t^2 - c^2 \Delta + \frac{m^2 c^4}{\hbar^2} \right) \Psi(t, \vec{x}) = 0 \tag{4.41}$$

Indeed, substituting a plane wave

$$\Psi_{\vec{p}}(t, \vec{x}) = A e^{-i\mathcal{E}t/\hbar + i\vec{p}\cdot\vec{x}/\hbar} \tag{4.42}$$

and demanding that the solution is non-trivial, i.e. $A \neq 0$, leads to the relation

$$-\frac{\mathcal{E}^2}{\hbar^2} + \frac{\vec{p}^2 c^2}{\hbar^2} + \frac{m^2 c^4}{\hbar^2} = 0 \quad (4.43)$$

which is just right! So our first guess for a relativistic Schrödinger equation is the *Klein-Gordon equation*:

$$\left(\partial_t^2 - c^2 \Delta + \frac{m^2 c^4}{\hbar^2} \right) \Psi(t, \vec{x}) = 0 \quad (4.44)$$

Exercise 4.6. Write the Klein-Gordon equation and the plane wave solution in a manifestly covariant form.

Solution. Dividing by c^2

$$\left(\frac{1}{c^2} \partial_t^2 - \Delta + \frac{m^2 c^2}{\hbar^2} \right) \Psi(t, \vec{x}) = 0 \quad (4.45)$$

and using

$$\begin{aligned} \partial_\mu &= \left(\frac{1}{c} \partial_t, \vec{\nabla} \right) \\ \partial^\mu &= \left(\frac{1}{c} \partial_t, -\vec{\nabla} \right) \end{aligned} \quad (4.46)$$

the Klein-Gordon equation can be rewritten as

$$\begin{aligned} (\partial_\mu \partial^\mu + M^2) \Psi(t, \vec{x}) &= 0 \\ \text{with } M &= \frac{mc}{\hbar} \end{aligned} \quad (4.47)$$

while using the four-vectors

$$\begin{aligned} x^\mu &= (ct, \vec{x}) \\ p^\mu &= (\mathcal{E}/c, \vec{p}) \Rightarrow p_\mu = (\mathcal{E}/c, -\vec{p}) \end{aligned} \quad (4.48)$$

allows the plane wave to be recast in the form

$$\Psi_{\vec{p}}(t, \vec{x}) = A e^{-ip_\mu x^\mu / \hbar} \quad (4.49)$$

□

4.3 Peculiarities of the Klein-Gordon equation

In fact, a closer inspection of the Klein-Gordon equation establishes that this is not exactly what we bargained for... In the following we go through the problems we face when trying to interpret this equation as a quantum mechanical model of a relativistic particle.

4.3.1 No spin

This is not a problem per se, just a fact: the Klein-Gordon equation describes *spinless* (i.e. zero spin) particles. The reason is that Ψ is a Lorentz scalar field, so when changing the four-coordinates with a Lorentz transformation

$$x' = \Lambda x \quad (4.50)$$

the field in the new coordinates takes the form

$$\Psi'(x') = \Psi(\Lambda^{-1}x') \quad (4.51)$$

i.e. for every point with new coordinate x' we look up its old coordinate $x = \Lambda^{-1}x'$ and read off the field value at that point.

Now the spin is the particle's internal angular momentum, so to examine it we need to choose the particle's own reference system, called the rest frame where the orbital angular momentum resulting from the particle's motion is absent. In this frame the Klein-Gordon plane wave is

$$\Psi_{\vec{0}}(t, \vec{x}) = Ae^{-imc^2t/\hbar} \quad (4.52)$$

This function is invariant under the special Lorentz transformations called spatial rotations, i.e.

$$\Psi_{\vec{0}}(t, \mathcal{O}\vec{x}) = \Psi_{\vec{0}}(t, \vec{x}) \quad (4.53)$$

However this is impossible if the particle has a non-zero spin. If the angular momentum in the rest frame is non-zero, the wave function cannot be rotationally invariant since the angular momentum is a *vector*, and a rotation changes its direction. Therefore in the case of a non-zero spin a rotation must take a state with some angular momentum to a state with the rotated angular momentum, which is a different state. In fact, as we shall see in the case of the Dirac equation, the transformation of the wave function under rotations in the rest frame tells us exactly what the spin of the particle is.

So the Klein-Gordon equation can only describe a particle with a zero spin. Such particles do exist in nature. Most of them are not elementary, such as the pions or kaons, but they could still be described as a point particle as long as their internal structure is irrelevant for the physical processes we are trying to capture. On the other hand, the Higgs particle is an example (in fact, the only known one) of a zero-spin boson which is elementary according to our present knowledge.

4.3.2 Negative energy solutions

A problem with the equation is that the dispersion relation has two solutions

$$\mathcal{E} = \pm \sqrt{m^2c^4 + \vec{p}^2c^2} \quad (4.54)$$

Therefore there are two solutions for any value of $\vec{p}$ that can be chosen as

$$\Psi_{\vec{p}}^{(+)}(x) = Ae^{-ip_{\mu}x^{\mu}/\hbar} \quad \Psi_{\vec{p}}^{(-)}(x) = Ae^{+ip_{\mu}x^{\mu}/\hbar} \quad (4.55)$$

where we always choose

$$p_0 = +\sqrt{\vec{p}^2 + m^2c^2} \quad (4.56)$$

This is not necessarily a problem as long as the particle moves in a slowly varying background field. However, in a rapidly varying background a problem known as the Klein paradox arises (cf. Subsection [4.3.4](#)), which is in conflict with the single-particle interpretation.

Another problem is that the system is unstable in the presence of a dynamical field such as quantum electromagnetic field (with photons as quanta): the problem is even worse than for the Klein paradox as the particle can loose energy by endlessly emitting photons and tumbling down along the negative energy levels into the infinite recesses of the DRQH (*Dark Relativistic Quantum Hell*)...

The solution to this problem is in fact that the field Ψ is *not a wave function describing the state of a single particle, but a quantum variable itself*. That leads to quantum field theory, in which case the proper interpretation is a second quantised description of many-particle states of a spinless bosonic particle.

4.3.3 No probability density

It is not very hard to construct a conserved four-current for the Klein-Gordon equation, which is of the form

$$j^\mu = \frac{i\hbar}{2m} (\Psi^\dagger \partial^\mu \Psi - \Psi \partial^\mu \Psi^\dagger) \quad (4.57)$$

Exercise 4.7. Prove that

$$\partial_\mu j^\mu = 0 \quad (4.58)$$

Solution. We have

$$\begin{aligned} (\partial_\mu \partial^\mu + M^2) \Psi &= 0 \\ (\partial_\mu \partial^\mu + M^2) \Psi^\dagger &= 0 \end{aligned} \quad (4.59)$$

so

$$\begin{aligned} 0 &= \Psi^\dagger (\partial_\mu \partial^\mu + M^2) \Psi - \Psi (\partial_\mu \partial^\mu + M^2) \Psi^\dagger \\ &= \partial_\mu (\Psi^\dagger \partial^\mu \Psi - \Psi \partial^\mu \Psi^\dagger) \end{aligned} \quad (4.60)$$

□

There is a slight problem here: the Klein-Gordon equation, in contrast to the Schrödinger equation, is real, hence nothing excludes taking Ψ to be real. In that case, however, the current j^μ is trivial. We can get over this by declaring that we consider Ψ to be complex, but this was just the first warning sign. There is a much bigger fly in the ointment: the putative probability density

$$\rho = j^0/c = \frac{i\hbar}{2mc} (\Psi^\dagger \partial_t \Psi - \Psi \partial_t \Psi^\dagger) \quad (4.61)$$

although it satisfies the continuity equation

$$\frac{\partial \rho}{\partial t} + \vec{\nabla} \cdot \vec{j} = 0 \quad (4.62)$$

is *not positive definite*! This alone seems to exclude the probabilistic interpretation and with that a sensible quantum mechanics built upon the Klein-Gordon equation.

Exercise 4.8. Compute the current j^μ for a positive and negative energy plane wave. Verify that

$$\vec{j} = \rho \vec{v} \quad (4.63)$$

as expected for a motion of velocity $\vec{v}$.

Solution. For

$$\Psi_{\vec{p}}^{(+)}(x) = A e^{-ip_\mu x^\mu / \hbar} \quad (4.64)$$

we get

$$j^{(+)\mu} = \frac{p^\mu}{m} |A|^2 \quad (4.65)$$

while for

$$\Psi_{\vec{p}}^{(-)}(x) = A e^{+ip_\mu x^\mu / \hbar} \quad (4.66)$$

the result is

$$j^{(-)\mu} = -\frac{p^\mu}{m} |A|^2 \quad (4.67)$$

In both cases

$$\frac{\vec{j}}{\rho} = \frac{\vec{p}c}{p^0} = \vec{v} \quad (4.68)$$

since

$$p^0 = \frac{\mathcal{E}}{c} = \frac{mc}{\sqrt{1 - \vec{v}^2/c^2}} \quad \vec{p} = \frac{m\vec{v}}{\sqrt{1 - \vec{v}^2/c^2}} \quad (4.69)$$

Note that

$$\begin{aligned} \rho^{(+)} &= \frac{p^0}{m} |A|^2 > 0 \\ \rho^{(-)} &= -\frac{p^0}{m} |A|^2 < 0 \end{aligned} \quad (4.70)$$

so the density is negative for negative energy waves. $\square$

There still seems to be a way out. Maybe the negative energy solutions can just be discarded and the dynamics can be restricted to the subspace of positive energy solutions, which have a positive density. This would also solve the stability problem discussed in Subsection [4.3.2](#). However, this turns out not to be the case: to get a complete basis for the field one needs to retain both types of plane waves, as shown below.

4.3.4 Klein paradox and completeness

We consider the scattering of one-dimensional plane waves on a potential step

$$V(x) = \begin{cases} V_0 & x \geq 0 \\ 0 & x < 0 \end{cases} \quad (4.71)$$

Now for $x < 0$ we still have the ordinary Klein-Gordon equation

$$(\partial_0^2 - \partial_1^2 + M^2) \Psi = 0 \quad (4.72)$$

For $x > 0$, however, the proper dispersion relation is

$$(\mathcal{E} - V_0)^2 - \vec{p}^2 c^2 = m^2 c^4 \quad (4.73)$$

leading to the equation

$$((\partial_0 - \alpha)^2 - \partial_1^2 + M^2) \Psi = 0 \quad \alpha = \frac{V_0}{i\hbar c} \quad (4.74)$$

Let us consider a scattering where there is an incident and reflected wave in the domain $x > 0$ and an outgoing wave for $x < 0$, all of which are assumed to be of positive energy:

$$\begin{aligned} \Psi_{<} &= A e^{-i\mathcal{E}t/\hbar} (e^{ipx/\hbar} + R e^{-ipx/\hbar}) \\ \Psi_{>} &= A e^{-i\mathcal{E}t/\hbar} T e^{iqx/\hbar} \end{aligned} \quad (4.75)$$

with

$$p = \frac{1}{c} \sqrt{\mathcal{E}^2 - m^2 c^4} \quad q = \frac{1}{c} \sqrt{(\mathcal{E} - V_0)^2 - m^2 c^4} \quad (4.76)$$

where R and T are the reflection and transmission amplitudes. Since the Klein-Gordon equation is second order in spatial derivatives, Ψ and $\partial_x \Psi$ must be continuous (similarly to the Schrödinger equation).

Exercise 4.9. Determine R and T .

Solution.

$$\begin{aligned} \Psi_{<}(x=0) &= \Psi_{>}(x=0) &: 1 + R &= T \\ \partial_x \Psi_{<}(x=0) &= \partial_x \Psi_{>}(x=0) &: p(1 - R) &= qT \end{aligned} \quad (4.77)$$

leading to

$$R = \frac{p - q}{p + q} \quad T = \frac{2p}{p + q} \quad (4.78)$$

□

Now we can determine the incident, transmitted and reflected fluxes:

$$j_{in} = |A|^2 \frac{p}{m} \quad j_R = |A|^2 |R|^2 \frac{p}{m} \quad j_T = |A|^2 |T|^2 \frac{q}{m} \quad (4.79)$$

and so the reflection and transmission probabilities are

$$r = \frac{j_R}{j_{in}} = \left(\frac{p - q}{p + q} \right)^2 \quad t = \frac{j_T}{j_{in}} = \frac{4pq}{(p + q)^2} \quad (4.80)$$

which satisfy

$$r + t = 1 \quad (4.81)$$

However, this is only valid as long as

$$\mathcal{E} - V_0 \geq mc^2 \quad (4.82)$$

For

$$-mc^2 < \mathcal{E} - V_0 < mc^2 \quad (4.83)$$

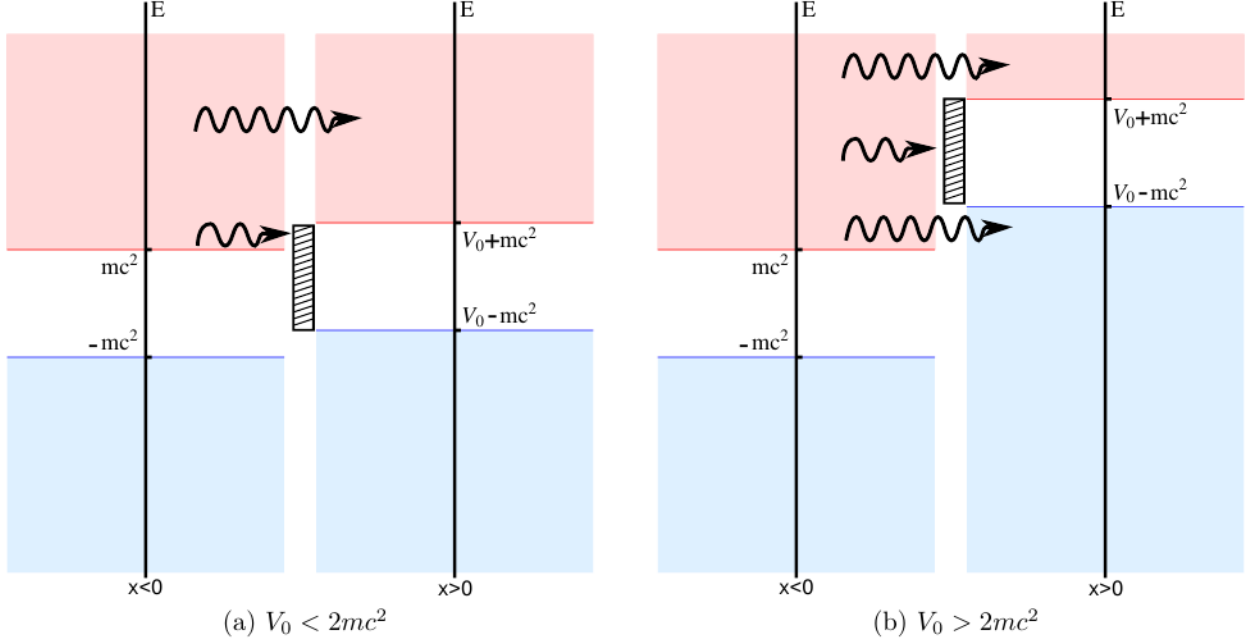


Figure 4.1: Illustration of the Klein paradox

the value of q is imaginary, and so the wave-function decays exponentially for $x > 0$. This is still fine, and is interpreted as a total reflection since for $q = i\kappa$ we get

$$j_{in} = |A|^2 \frac{p}{m} \quad j_R = |A|^2 |R|^2 \frac{p}{m} = j_{in} \quad j_T = 0 \quad (4.84)$$

so there is a total reflection as the wave cannot propagate beyond $x = 0$.

The real surprise is that for

$$V_0 > \mathcal{E} + mc^2 \quad (4.85)$$

although the energy is way below the necessary amount for penetration, q becomes real again! *The wave does propagate into the region behind the barrier.* To understand that, consider the spectrum of the theory in the two regions far from $x = 0$. For $x < 0$ we get the standard spectrum consisting of positive and negative energy states, while for $x > 0$ the whole spectrum is shifted up by V_0 . As illustrated in Figure 4.1, nothing strange happens as long as $V_0 < 2mc^2$: particles with insufficient energy are simply forbidden to propagate on the other side. However, for $V_0 > 2mc^2$ there is a range of energies where particles with energies naively forbidden to pass over the barrier find a negative energy level to travel into, making it possible for them to cross the barrier.

The conclusion is that one cannot exclude the negative energy levels from the Hilbert space, otherwise for the states with

$$V_0 > \mathcal{E} + mc^2 \quad (4.86)$$

the tentative probability current would not be conserved, since they would have $r < 1$, but no level available for transmission on the other side. On the other hand, the presence of these states leads to negative probabilities.

Note that the above results exclude the existence of a fully consistent probabilistic interpretation for the Klein-Gordon equation: excluding the negative energy states leads to positive probabilities, but due to incompleteness of the wave function basis violates probability conservation, while

including the negative energy states leads to a conserved, but indefinite density which cannot be interpreted as probability.

Another reason for which the negative energy states cannot be excluded is the nature of the initial value problem. Note that the Klein-Gordon equation is second order in derivatives, so in addition to Ψ its first time derivative must also be specified as an initial condition. Consider a generic solution of the free Klein-Gordon equation:

$$\begin{aligned}\Psi(t, \vec{x}) &= \underbrace{\int \frac{d^3p}{(2\pi)^3} A(\vec{p}) e^{-i\mathcal{E}(\vec{p})t/\hbar + i\vec{p}\vec{x}/\hbar}}_{\text{positive energy modes}} + \underbrace{\int \frac{d^3p}{(2\pi)^3} B(\vec{p}) e^{i\mathcal{E}(\vec{p})t/\hbar - i\vec{p}\vec{x}/\hbar}}_{\text{negative energy modes}} \\ \mathcal{E}(\vec{p}) &= \sqrt{\vec{p}^2 c^2 + m^2 c^4}\end{aligned}\tag{4.87}$$

One can then evaluate

$$\begin{aligned}\Psi(0, \vec{x}) &= \int \frac{d^3p}{(2\pi)^3} A(\vec{p}) e^{i\vec{p}\vec{x}/\hbar} + \int \frac{d^3p}{(2\pi)^3} B(\vec{p}) e^{-i\vec{p}\vec{x}/\hbar} = \int \frac{d^3p}{(2\pi)^3} [A(\vec{p}) + B(-\vec{p})] e^{i\vec{p}\vec{x}/\hbar} \\ \partial_t \Psi(0, \vec{x}) &= - \int \frac{d^3p}{(2\pi)^3} \frac{i\mathcal{E}(\vec{p})}{\hbar} A(\vec{p}) e^{i\vec{p}\vec{x}/\hbar} + \int \frac{d^3p}{(2\pi)^3} \frac{i\mathcal{E}(\vec{p})}{\hbar} B(\vec{p}) e^{-i\vec{p}\vec{x}/\hbar} \\ &= \int \frac{d^3p}{(2\pi)^3} \left(-\frac{i\mathcal{E}(\vec{p})}{\hbar} \right) [A(\vec{p}) - B(-\vec{p})] e^{i\vec{p}\vec{x}/\hbar}\end{aligned}\tag{4.88}$$

The two initial conditions together specify all the coefficients A and B by Fourier transform. Therefore, the positive energy modes alone do not form a complete basis for the time evolution of the Klein-Gordon equation.

5 Relativistic wave equations 2: the Dirac equation

5.1 Deriving the Dirac equation

We have seen that the Klein-Gordon equation suffers from several problems and limitations:

- it can only account for excitations of spin 0;
- there is no probability interpretation;
- negative energy state signal a potential instability of the system, and there is no easy way to solve this in the framework of quantum mechanics;
- it does not describe the fine structure of hydrogen-like atoms correctly.

The second and the third problem are in fact related, and are also tied with the fact that the equation is second order in derivatives. So in contrast to the Schrödinger equation, it is necessary to specify both the function Φ and its derivative for a well-defined initial value problem.

The idea behind the Dirac equation is simply to find a first order relativistic wave equation. Such an equation can be written in the form

$$(i\Gamma^\mu \hbar \partial_\mu + M) \Psi = 0 \quad (5.1)$$

where Γ^μ is some object that must transform as a four-vector under Lorentz transformations, while M is a Lorentz scalar. On the other hand, we expect the relativistic dispersion relation to be valid, so Ψ must also satisfy the Klein-Gordon equation

$$(\hbar^2 \partial_\mu \partial^\mu + m^2 c^2) \Psi = 0 \quad (5.2)$$

where m is the mass of the particle described by the equation, and this must be a consequence of the first-order equation. Let us evaluate

$$\begin{aligned} (i\Gamma^\mu \hbar \partial_\mu - M) (i\Gamma^\mu \hbar \partial_\mu + M) \Psi &= 0 \\ \Downarrow \\ (-\hbar^2 \Gamma^\mu \Gamma^\nu \partial_\mu \partial_\nu + i\hbar [\Gamma^\mu, M] \partial_\mu - M^2) \Psi &= 0 \end{aligned} \quad (5.3)$$

Using the symmetry of the second partial derivatives this implies that the following must hold

$$\begin{aligned} \Gamma^\mu \Gamma^\nu + \Gamma^\nu \Gamma^\mu &= 2\eta^{\mu\nu} \mathbf{1} \\ [\Gamma^\mu, M] &= 0 \\ M &= mc \end{aligned} \quad (5.4)$$

where $\mathbf{1}$ is the identity operator. The first relation defines the so-called Clifford algebra associated to the Lorentz group of four-dimensional Minkowski space time. The second states that operator M commutes with the Clifford generators Γ^μ . while the third actually states that M is nothing else but the mass operator, i.e. its eigenvalues are equal to the particle mass (times c). We can now make use of the following mathematical theorem:

Theorem 5.1. *The Clifford algebra*

$$\Gamma^\mu \Gamma^\nu + \Gamma^\nu \Gamma^\mu = 2\eta^{\mu\nu} \mathbf{1} \quad (5.5)$$

has a single 4-dimensional irreducible representation up to similarity transformation. All other representations of this algebra are simply direct sums of the 4-dimensional one.

In other words, if we find operators satisfying the Clifford relation, they can always be represented as block diagonal matrices made up of 4-dimensional blocks in some basis:

$$\Gamma^\mu = \begin{pmatrix} \gamma_1^\mu & 0 & \dots & 0 \\ 0 & \gamma_2^\mu & \dots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \dots & \gamma_N^\mu \end{pmatrix} \quad (5.6)$$

where the γ_k^μ are 4×4 matrices. They are also related to each other by a similarity transformation, so we can find 4×4 matrices A_k such that all of them are transformed into some canonical form⁸

$$\gamma_k^\mu = A_k \gamma^\mu A_k^{-1} \quad (5.7)$$

Such a transformation amounts to a change of basis that preserves the block diagonal form, so we can write

$$\Gamma^\mu = \begin{pmatrix} \gamma^\mu & 0 & \dots & 0 \\ 0 & \gamma^\mu & \dots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \dots & \gamma^\mu \end{pmatrix} \quad (5.8)$$

by some appropriate choice of basis, where the 4×4 matrices satisfy

$$\{\gamma^\mu, \gamma^\nu\} = 2\eta^{\mu\nu} \mathbf{1}_{4 \times 4} \quad (5.9)$$

where

$$\{A, B\} = AB + BA \quad (5.10)$$

is called the anti-commutator of operators A and B .

Irreducibility means that there is no basis in which the 4×4 γ^μ can be decomposed simultaneously into a block-diagonal form in terms of smaller than 4-dimensional blocks. That implies that the only matrices commuting with all the γ^μ are proportional to the 4×4 identity matrix (this is the consequence of Schur's lemma):

$$[\gamma^\mu, A] = 0 \Rightarrow A = \alpha \mathbf{1}_{4 \times 4} \quad (5.11)$$

As a result

$$M = \begin{pmatrix} m_1 c \mathbf{1} & 0 & \dots & 0 \\ 0 & m_2 c \mathbf{1} & \dots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & \dots & m_N c \mathbf{1} \end{pmatrix} \quad (5.12)$$

where the labels 4×4 were omitted for brevity. From now on, $\mathbf{1}$ will always denote the identity matrix of appropriate dimension for the given relation it occurs.

Therefore the general solution is just N copies of the same equation for different masses. Consider a single 4-dimensional piece, which then has a definite mass m ; this is expected to describe the relativistic quantum mechanics of some particle of mass m . This is the celebrated *Dirac equation*:

$$(i\hbar\gamma^\mu\partial_\mu - mc)\Psi = 0 \quad (5.13)$$

Now let us understand the physical content of this equation.

⁸Here there is no summation meant for the repeated index k !

5.2 Covariance and spinors

The field in the Dirac equation is a 4-component object, which is called a *Dirac spinor*. Its index will be denoted by Greek letters; this will not cause confusion with 4-vector indices as the type of the object will be clear from the context it appears:

$$\Psi_\xi(t, \vec{x}) \quad , \quad \xi = 1, 2, 3, 4 \quad (5.14)$$

We need to determine the *transformation law for the spinors*. Under a Lorentz transformation

$$x'^\mu = \Lambda^\mu_\nu x^\nu \quad (5.15)$$

one can write

$$\Psi'(x') = S(\Lambda)\Psi(x) \quad (5.16)$$

where $S(\Lambda)$ is some transformation acting on the Dirac spinor indices. It can be obtained by stipulating that the Dirac equation is covariant, i.e.

$$(i\hbar\gamma^\mu\partial_\mu - mc)\Psi(x) = 0 \Leftrightarrow (i\hbar\gamma^\mu\partial'_\mu - mc)\Psi'(x') = 0 \quad (5.17)$$

Now using

$$\partial'_\mu = \bar{\Lambda}_\mu^\nu \partial_\nu \quad (5.18)$$

the Dirac equation transforms according to

$$(i\hbar\gamma^\mu\partial'_\mu - mc)\Psi'(x') = (i\hbar\gamma^\mu\bar{\Lambda}_\mu^\nu\partial_\nu - mc)S(\Lambda)\Psi(x) = 0 \quad (5.19)$$

Multiplying from the left by $S(\Lambda)^{-1}$ gives

$$(i\hbar S(\Lambda)^{-1}\gamma^\mu\bar{\Lambda}_\mu^\nu S(\Lambda)\partial_\nu - mc)\Psi(x) = 0 \quad (5.20)$$

This is identical to the Dirac equation in coordinates x when

$$S(\Lambda)^{-1}\gamma^\mu\bar{\Lambda}_\mu^\nu S(\Lambda) = \gamma^\nu \quad (5.21)$$

Exercise 5.2. Show that from the above equation it follows that γ^μ transforms as a contravariant 4-vector.

Solution. The above equation can be multiplied by $S(\Lambda)$ from the left and $S(\Lambda)^{-1}$ from the right to give

$$S(\Lambda)\gamma^\nu S(\Lambda)^{-1} = \gamma^\mu\bar{\Lambda}_\mu^\nu \quad (5.22)$$

Now using

$$\bar{\Lambda} = (\Lambda^{-1})^T \quad (5.23)$$

the left-hand side can be written as

$$S(\Lambda)\gamma^\nu S(\Lambda)^{-1} = (\Lambda^{-1})^\nu_\mu \gamma^\mu \quad (5.24)$$

□

An alternative way of deriving the result is to consider the following condition of covariance

$$S(\Lambda)^{-1} \gamma^\mu \frac{\partial}{\partial x'^\mu} S(\Lambda) = \gamma^\nu \frac{\partial}{\partial x^\nu} \quad (5.25)$$

Using

$$\begin{aligned} \frac{\partial}{\partial x'^\mu} &= \frac{\partial x^\nu}{\partial x'^\mu} \frac{\partial}{\partial x^\nu} = (\Lambda^{-1})^\nu_\mu \frac{\partial}{\partial x^\nu} \\ &\text{since } x^\nu = (\Lambda^{-1})^\nu_\mu x'^\mu \end{aligned} \quad (5.26)$$

implies that

$$S(\Lambda) \gamma^\nu S(\Lambda)^{-1} = (\Lambda^{-1})^\nu_\mu \gamma^\mu \quad (5.27)$$

This is our fundamental equation defining $S(\Lambda)$ for any given Lorentz matrix Λ . The solution for $S(\Lambda)$ is however, not unique, since any scalar factor multiplying $S(\Lambda)$ eventually cancels. This can be partially fixed by choosing solutions with $\det S(\Lambda) = +1$. Even with this additional condition the solution is not unique since for any particular solution $S(\Lambda)$ the matrices

$$\zeta S(\Lambda) \quad (5.28)$$

with $\zeta^4 = 1$ also solve the fundamental equation and have unit determinant. In fact, the ambiguity can be fixed further up to an overall sign, as shown later using the Weyl decomposition discussed in Subsection [5.4.1](#)

Exercise 5.3. Show that Clifford's theorem implies that there always exists a solution for $S(\Lambda)$.

Solution. The Lorentz transformed Dirac matrices

$$\gamma'^\nu = (\Lambda^{-1})^\nu_\mu \gamma^\mu \quad (5.29)$$

satisfy the Clifford algebra:

$$\begin{aligned} \{\gamma'^\mu, \gamma'^\nu\} &= (\Lambda^{-1})^\mu_\rho (\Lambda^{-1})^\nu_\sigma \{\gamma^\rho, \gamma^\sigma\} \\ &= (\Lambda^{-1})^\mu_\rho (\Lambda^{-1})^\nu_\sigma 2\eta^{\rho\sigma} \mathbf{1} \\ &= 2\eta^{\mu\nu} \mathbf{1} \end{aligned} \quad (5.30)$$

So by Clifford's theorem there exists a similarity transformation $S(\Lambda)$ such that

$$S(\Lambda) \gamma^\nu S(\Lambda)^{-1} = (\Lambda^{-1})^\nu_\mu \gamma^\mu \quad (5.31)$$

holds, and its determinant can always be chosen 1 by multiplying it with an appropriate scalar factor. $\square$

5.3 Properties of spinor transformations

5.3.1 Group property

Exercise 5.4. Show that $S(\Lambda_1)S(\Lambda_2)$ solves the fundamental relation for $\Lambda_1\Lambda_2$.

Solution. Defining

$$\begin{aligned}
 S(\Lambda_1)S(\Lambda_2)\gamma^\mu S(\Lambda_2)^{-1}S(\Lambda_1)^{-1} &= S(\Lambda_1) [(\Lambda_2^{-1})^\mu{}_\nu \gamma^\nu] S(\Lambda_1)^{-1} \\
 &= (\Lambda_2^{-1})^\mu{}_\nu (\Lambda_1^{-1})^\nu{}_\sigma \gamma^\sigma \\
 &= (\Lambda_2^{-1}\Lambda_1^{-1})^\mu{}_\sigma \gamma^\sigma \\
 &= ((\Lambda_1\Lambda_2)^{-1})^\mu{}_\sigma \gamma^\sigma
 \end{aligned} \tag{5.32}$$

□

Therefore, taking into account the sign ambiguity in $S(\Lambda)$ we obtain

$$S(\Lambda_1\Lambda_2) = \pm S(\Lambda_1)S(\Lambda_2) \tag{5.33}$$

The $\pm$ sign ambiguity shows that $S(\Lambda)$ is not a proper representation of the Lorentz group. Instead the map

$$\Lambda \mapsto S(\Lambda) \tag{5.34}$$

is a so-called *projective representation*. Such a representation of a symmetry only keeps the multiplication property of the transformations up to a phase, and is *allowed in quantum theory* since such phases cancel from all transition probabilities and expectation values of the observables.

5.3.2 Infinitesimal form of spinor transformations

Let us take an infinitesimal Lorentz transformation

$$\Lambda^\mu{}_\nu = \delta^\mu{}_\nu + \omega^\mu{}_\nu + O(\omega^2) \quad , \quad \omega_{\mu\nu} = -\omega_{\nu\mu} \tag{5.35}$$

so that

$$(\Lambda^{-1})^\mu{}_\nu = \delta^\mu{}_\nu - \omega^\mu{}_\nu + O(\omega^2) \tag{5.36}$$

Looking for the spinor transformation in the form

$$\begin{aligned}
 S(\Lambda) &= \mathbf{1} - \frac{i}{4}\omega_{\mu\nu}\sigma^{\mu\nu} + O(\omega^2) \\
 S(\Lambda)^{-1} &= \mathbf{1} + \frac{i}{4}\omega_{\mu\nu}\sigma^{\mu\nu} + O(\omega^2)
 \end{aligned} \tag{5.37}$$

it can be assumed that

$$\sigma^{\mu\nu} = -\sigma^{\nu\mu} \tag{5.38}$$

due to the antisymmetry of $\omega^{\mu\nu}$. Using the fundamental relation

$$S(\Lambda)\gamma^\nu S(\Lambda)^{-1} = (\Lambda^{-1})^\nu{}_\mu \gamma^\mu \tag{5.39}$$

leads to

$$[\gamma^\lambda, \sigma^{\mu\nu}] = 2i (\eta^{\lambda\mu}\gamma^\nu - \eta^{\lambda\nu}\gamma^\mu) \tag{5.40}$$

Exercise 5.5. Derive the above relation.

Solution. Substituting our starting formulae into

$$S(\Lambda)\gamma^\lambda S(\Lambda)^{-1} = (\Lambda^{-1})^\lambda_\nu \gamma^\nu \quad (5.41)$$

leads to

$$\left(1 - \frac{i}{4}\omega_{\mu\nu}\sigma^{\mu\nu}\right) \gamma^\lambda \left(1 + \frac{i}{4}\omega_{\mu\nu}\sigma^{\mu\nu}\right) = (\delta^\lambda_\nu - \omega^\lambda_\nu) \gamma^\nu + O(\omega^2) \quad (5.42)$$

i.e.

$$\gamma^\lambda + \frac{i}{4}\omega_{\mu\nu}[\gamma^\lambda, \sigma^{\mu\nu}] = \gamma^\lambda - \omega_{\mu\nu}\eta^{\lambda\mu}\gamma^\nu \quad (5.43)$$

Using the antisymmetry of $\omega_{\mu\nu}$ this can be rewritten as

$$\frac{i}{4}\omega_{\mu\nu}[\gamma^\lambda, \sigma^{\mu\nu}] = \frac{1}{2}\omega_{\mu\nu}(\eta^{\lambda\nu}\gamma^\mu - \eta^{\lambda\mu}\gamma^\nu) \quad (5.44)$$

and since this is valid for any ω we obtain

$$\frac{i}{4}[\gamma^\lambda, \sigma^{\mu\nu}] = \frac{1}{2}(\eta^{\lambda\nu}\gamma^\mu - \eta^{\lambda\mu}\gamma^\nu) \quad (5.45)$$

from which relation (5.40) follows. $\square$

A solution to (5.40) is given by the matrices

$$\sigma^{\mu\nu} = \frac{i}{2}[\gamma^\mu, \gamma^\nu] \quad (5.46)$$

Exercise 5.6. Verify the above solution using the Jacobi identity

$$[A, [B, C]] = \{\{A, B\}, C\} - \{B, \{A, C\}\} \quad (5.47)$$

Solution. One can easily compute

$$\begin{aligned} -2i[\gamma^\lambda, \sigma^{\mu\nu}] &= [\gamma^\lambda, [\gamma^\mu, \gamma^\nu]] = \{\{\gamma^\lambda, \gamma^\mu\}, \gamma^\nu\} - \{\gamma^\mu, \{\gamma^\lambda, \gamma^\nu\}\} \\ &= 2\eta^{\lambda\mu}\{\mathbf{1}, \gamma^\nu\} - 2\eta^{\lambda\nu}\{\gamma^\mu, \mathbf{1}\} \\ &= 4(\eta^{\lambda\mu}\gamma^\nu - \eta^{\lambda\nu}\gamma^\mu) \end{aligned} \quad (5.48)$$

from which

$$[\gamma^\lambda, \sigma^{\mu\nu}] = 2i(\eta^{\lambda\mu}\gamma^\nu - \eta^{\lambda\nu}\gamma^\mu) \quad (5.49)$$

$\square$

Exercise 5.7. Verify that if there are two solutions $\sigma_1^{\mu\nu}$ and $\sigma_2^{\mu\nu}$ to

$$[\gamma^\lambda, \sigma^{\mu\nu}] = 2i(\eta^{\lambda\mu}\gamma^\nu - \eta^{\lambda\nu}\gamma^\mu) \quad (5.50)$$

then there exist numbers $\kappa^{\mu\nu}$ such that

$$\sigma_2^{\mu\nu} = \sigma_1^{\mu\nu} + \kappa^{\mu\nu}\mathbf{1} \quad (5.51)$$

Solution. This follows from

$$[\gamma^\lambda, \sigma_2^{\mu\nu} - \sigma_1^{\mu\nu}] = 0 \quad (5.52)$$

and Clifford's theorem (irreducibility of the 4 dimensional Dirac spinor) + Schur's lemma. $\square$

A finite Lorentz transformation can always be divided up into infinitesimal ones as

$$\Lambda = \lim_{N \rightarrow \infty} \left(1 + \frac{\omega}{N}\right)^N = e^\omega \quad (5.53)$$

which leads to the expression

$$S(\Lambda) = \lim_{N \rightarrow \infty} \left(1 - \frac{i}{4} \frac{\omega_{\mu\nu}}{N} \sigma^{\mu\nu}\right)^N = \exp\left(-\frac{i}{4} \omega_{\mu\nu} \sigma^{\mu\nu}\right) \quad (5.54)$$

The remaining ambiguity (5.51) leads a redefinition by an overall phase:

$$S(\Lambda) = \exp\left(-\frac{i}{4} \omega_{\mu\nu} \sigma^{\mu\nu}\right) e^{-\frac{i}{4} \omega_{\mu\nu} \kappa^{\mu\nu}} \quad (5.55)$$

which can always be omitted by the rules of quantum theory. So one can take

$$\sigma^{\mu\nu} = \frac{i}{2} [\gamma^\mu, \gamma^\nu] \quad (5.56)$$

as the generators of infinitesimal spinor transformations and

$$S(\Lambda) = \exp\left(-\frac{i}{4} \omega_{\mu\nu} \sigma^{\mu\nu}\right) \quad (5.57)$$

as the form of finite spinor transformations.

5.3.3 Skew unitarity and the conjugate Dirac spinor

To proceed, it is necessary to specify the conjugation properties of the γ matrices first. Note that

$$(\gamma^0)^2 = \mathbf{1} \quad (\gamma^i)^2 = -\mathbf{1} \quad i = 1, 2, 3 \quad (5.58)$$

so that γ^0 has eigenvalues ± 1 , while the eigenvalues of γ^i are $\pm i$. As a result, the γ matrices can be chosen to have the following Hermiticity properties

$$\gamma^{0\dagger} = \gamma^0 \quad \gamma^{i\dagger} = -\gamma^i \quad (5.59)$$

From now on we consider only representations in which properties are satisfied. Using $\gamma^0 \gamma^i = -\gamma^i \gamma^0$ leads to

$$\gamma^{\mu\dagger} = \gamma^0 \gamma^\mu \gamma^0 \quad (5.60)$$

Using $(\gamma^0)^{-1} = \gamma^0$ implies that $\gamma^{\mu\dagger}$ and γ^μ are similar, which is again a consequence of Clifford's theorem since the matrices $\gamma^{\mu\dagger}$ satisfy the same anti-commutation relations as the γ^μ .

Exercise 5.8. Show that

$$\sigma^{\mu\nu\dagger} = \gamma^0 \sigma^{\mu\nu} \gamma^0 \quad (5.61)$$

Solution.

$$\begin{aligned}
\sigma^{\mu\nu\dagger} &= -\frac{i}{2} (\gamma^\mu \gamma^\nu - \gamma^\nu \gamma^\mu)^\dagger = -\frac{i}{2} (\gamma^{\nu\dagger} \gamma^{\mu\dagger} - \gamma^{\mu\dagger} \gamma^{\nu\dagger}) \\
&= \frac{i}{2} \left(\gamma^0 \gamma^\mu \underbrace{\gamma^0 \gamma^0}_1 \gamma^\nu \gamma^0 - \gamma^0 \gamma^\nu \underbrace{\gamma^0 \gamma^0}_1 \gamma^\mu \gamma^0 \right) = \gamma^0 \left[\frac{i}{2} (\gamma^\mu \gamma^\nu - \gamma^\nu \gamma^\mu) \right] \gamma^0
\end{aligned} \tag{5.62}$$

□

Exercise 5.9. Verify that $R(\Lambda) = \gamma^0 S(\Lambda)^\dagger \gamma^0$ solves the same equation as $S(\Lambda)^{-1}$, i.e.

$$R(\Lambda)^{-1} \gamma^\nu R(\Lambda) = (\Lambda^{-1})^\nu{}_\mu \gamma^\mu \tag{5.63}$$

Solution.

$$\begin{aligned}
S(\Lambda) \gamma^\nu S(\Lambda)^{-1} &= (\Lambda^{-1})^\nu{}_\mu \gamma^\mu \\
&\Downarrow \\
S(\Lambda)^{\dagger-1} \gamma^{\nu\dagger} S(\Lambda)^\dagger &= (\Lambda^{-1})^\nu{}_\mu \gamma^{\mu\dagger} \\
&\Downarrow \\
(\gamma^0 S(\Lambda)^\dagger \gamma^0)^{-1} \gamma^\nu (\gamma^0 S(\Lambda)^\dagger \gamma^0) &= (\Lambda^{-1})^\nu{}_\mu \gamma^\mu
\end{aligned} \tag{5.64}$$

□

Now the following *skew unitarity* property of $S(\Lambda)$ can be proven:

$$S(\Lambda)^{-1} = \gamma^0 S(\Lambda)^\dagger \gamma^0 \tag{5.65}$$

This is in fact quite straightforward:

$$\gamma^0 S(\Lambda)^\dagger \gamma^0 = \gamma^0 \exp \left(\frac{i}{4} \omega_{\mu\nu} \sigma^{\mu\nu\dagger} \right) \gamma^0 = \gamma^0 \gamma^0 \exp \left(\frac{i}{4} \omega_{\mu\nu} \sigma^{\mu\nu} \right) \gamma^0 \gamma^0 = \exp \left(\frac{i}{4} \omega_{\mu\nu} \sigma^{\mu\nu} \right) = S(\Lambda)^{-1} \tag{5.66}$$

If γ^0 was the unit matrix, then this would be just standard unitarity. However, that is not possible; in fact we shall see shortly that γ^0 has eigenvalues ± 1 , both doubly degenerate. γ^0 can be considered a sort of “metric” which defines an invariant inner product in spinor space. In fact, it is easy to show that given two spinors ψ and χ , the product

$$\psi^\dagger \gamma^0 \chi \tag{5.67}$$

is invariant under any Lorentz transformation Λ :

$$\begin{aligned}
\psi^\dagger \gamma^0 \chi &\rightarrow_\Lambda \psi^\dagger S(\Lambda)^\dagger \gamma^0 S(\Lambda) \chi \\
&= \psi^\dagger \gamma^0 \gamma^0 S(\Lambda)^\dagger \gamma^0 S(\Lambda) \chi \\
&= \psi^\dagger \gamma^0 S(\Lambda)^{-1} S(\Lambda) \chi \\
&= \psi^\dagger \gamma^0 \chi
\end{aligned} \tag{5.68}$$

Therefore, the natural conjugate of any Dirac spinor ψ can be defined as

$$\bar{\psi} = \psi^\dagger \gamma^0 \tag{5.69}$$

and the invariant product of spinors ψ and χ is then

$$\bar{\psi}\chi \quad (5.70)$$

In component notations, ψ is usually represented as a column vector

$$\psi = \begin{pmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \psi_4 \end{pmatrix} \quad (5.71)$$

while

$$\psi^\dagger = (\psi_1 \quad \psi_2 \quad \psi_3 \quad \psi_4) \quad (5.72)$$

is a row vector. Then by the usual rules of matrix multiplication it follows that $\bar{\psi}$ is a row vector and $\bar{\psi}\chi$ is a 1×1 matrix, which is correct for a scalar.

From the Dirac equation for ψ it follows that the *conjugate Dirac spinor* $\bar{\psi}$ satisfies the equation

$$\bar{\psi}(x) \left(i\hbar\gamma^\mu \overleftarrow{\partial}_\mu + mc \right) = 0 \quad (5.73)$$

where the arrow indicated that the derivative acts to the left, i.e.

$$\bar{\psi}(x) \left(i\hbar\gamma^\mu \overleftarrow{\partial}_\mu + mc \right) = i\hbar\partial_\mu \bar{\psi}(x) \gamma^\mu + mc\bar{\psi}(x) \quad (5.74)$$

Exercise 5.10. Derive the conjugate Dirac equation

$$\bar{\psi}(x) \left(i\hbar\gamma^\mu \overleftarrow{\partial}_\mu + mc \right) = 0 \quad (5.75)$$

Solution. Taking the Hermitian conjugate of the Dirac equation

$$0 = (i\hbar\gamma^\mu \partial_\mu - mc) \psi(x) \quad (5.76)$$

leads to

$$\begin{aligned} 0 &= \psi^\dagger(x) \left(-i\hbar\gamma^{\mu\dagger} \overleftarrow{\partial}_\mu - mc \right) \\ &= \psi^\dagger(x) \left(-i\hbar\gamma^0\gamma^\mu\gamma^0 \overleftarrow{\partial}_\mu - mc \right) \end{aligned} \quad (5.77)$$

and multiplying by $-\gamma^0$ to the right results in the conjugate Dirac equation. $\square$

5.4 Spin, parity and chirality

5.4.1 Parity, Weyl spinors and chirality

It is easy to see that

$$S(\Lambda)\gamma^\nu S(\Lambda)^{-1} = (\Lambda^{-1})^\nu{}_\mu \gamma^\mu \quad (5.78)$$

implies that the spinor transformation corresponding to parity can be represented as

$$S(\mathcal{P}) = \gamma^0 \quad (5.79)$$

since

$$\gamma^0 \gamma^0 \gamma^0 = \gamma^0 \quad , \quad \gamma^0 \gamma^i \gamma^0 = -\gamma^i \quad (5.80)$$

Now let us introduce a special element of the Clifford algebra:

$$\gamma^5 = i\gamma^0 \gamma^1 \gamma^2 \gamma^3 = \frac{i}{4!} \epsilon_{\mu\nu\rho\sigma} \gamma^\mu \gamma^\nu \gamma^\rho \gamma^\sigma \quad (5.81)$$

which has very interesting properties. The second representation is valid due to

$$\epsilon_{0123} = +1 \quad (5.82)$$

and the complete antisymmetry of ϵ together with the anti-commutation relations of the γ^μ .

Exercise 5.11. Show that

$$\begin{aligned} (\gamma^5)^2 &= \mathbf{1} \quad , \quad \gamma^{5\dagger} = \gamma^5 \\ \{\gamma^5, \gamma^\mu\} &= 0 \quad , \quad [\gamma^5, \sigma^{\mu\nu}] = 0 \end{aligned} \quad (5.83)$$

Solution. By direct calculation

$$\begin{aligned} (\gamma^5)^2 &= -\gamma^0 \gamma^1 \gamma^2 \gamma^3 \gamma^0 \gamma^1 \gamma^2 \gamma^3 = -(\gamma^0)^2 (\gamma^1)^2 (\gamma^2)^2 (\gamma^3)^2 = \mathbf{1} \\ \gamma^{5\dagger} &= -i\gamma^{3\dagger} \gamma^{2\dagger} \gamma^{1\dagger} \gamma^{0\dagger} = i\gamma^3 \gamma^2 \gamma^1 \gamma^0 = i\gamma^0 \gamma^1 \gamma^2 \gamma^3 = \gamma^5 \\ \gamma^5 \gamma^\mu &= i\gamma^0 \gamma^1 \gamma^2 \gamma^3 \gamma^\mu = -i\gamma^\mu \gamma^0 \gamma^1 \gamma^2 \gamma^3 = -\gamma^\mu \gamma^5 \\ \gamma^5 \sigma^{\mu\nu} &= \frac{i}{2} \gamma^5 (\gamma^\mu \gamma^\nu - \gamma^\nu \gamma^\mu) = \frac{i}{2} (\gamma^\mu \gamma^\nu - \gamma^\nu \gamma^\mu) \gamma^5 = \sigma^{\mu\nu} \gamma^5 \end{aligned} \quad (5.84)$$

□

From these properties it follows that the 4-dimensional space $\mathcal{S}$ of spinors can be split into the eigenspaces of γ^5 of eigenvalues ± 1 :

$$\begin{aligned} \mathcal{S} &= \mathcal{S}_+ \oplus \mathcal{S}_- \\ \mathcal{S}_\pm &= \{\psi \in \mathcal{S} | \gamma^5 \psi = \pm \psi\} \end{aligned} \quad (5.85)$$

Since the generators $\sigma^{\mu\nu}$ commute with γ^5 , so do the spinor transformations $S(\Lambda)$:

$$[\gamma^5, S(\Lambda)] = 0 \quad (5.86)$$

and therefore the spinor transformations are block-diagonal in a basis that is put together from two bases in the spaces $\mathcal{S}_\pm$

$$S(\Lambda) = \begin{pmatrix} s_-(\Lambda) & 0 \\ 0 & s_+(\Lambda) \end{pmatrix} \quad (5.87)$$

This can be used to reduce the ambiguity (5.28) by demanding that both matrices $s_\pm(\Lambda)$ have a unit determinant, which due to the size 2×2 of $s_\pm(\Lambda)$ only allows for an overall sign ambiguity in

both. Note that this still allows for more freedom than an overall sign in $S(\Lambda)$ but this can be fixed by continuity of the dependence of $s_{\pm}(\Lambda)$ on Λ and choosing them equal (i.e. $s_+ = s_- = \pm \mathbf{1}_{2 \times 2}$) for the case when Λ is the unit Lorentz transformation.

However, there is a sign ambiguity that cannot be eliminated for the following reason: a 2π rotation changes the sign of the spinor field. Suppose we take the z axis as the axis of rotation and recall that a rotation by angle φ around the axis z can be obtained as

$$S(\varphi) = e^{-i\varphi \frac{1}{2} \sigma^{12}} \quad (5.88)$$

where

$$\sigma^{12} = \begin{pmatrix} \sigma^3 & 0 \\ 0 & \sigma^3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \quad (5.89)$$

(in the Weyl representation; the result for $S(2\pi)$, being proportional to the identity is independent of the choice of the spinor basis, so our conclusion is independent of the choice of the representation. It can be easily seen that

$$S(2\pi) = -\mathbf{1}_{4 \times 4} \quad (5.90)$$

Since a rotation by 2π is an identity transformation in physics which must leave all observables invariant, the Dirac field cannot be an observable. In fact, all local observables are expressed in terms of even products of the Dirac field.

As a result of the residual sign ambiguity, the spinorial transformations satisfy

$$S(\Lambda_1)S(\Lambda_2) = \pm S(\Lambda_1\Lambda_2) \quad (5.91)$$

(cf. the discussion of the group property in Subsection [5.3](#)).

In such a basis, the Dirac spinor can be written as a composite of two so-called Weyl spinors

$$\psi = \begin{pmatrix} \psi_- \\ \psi_+ \end{pmatrix} \quad (5.92)$$

with

$$\gamma^5 = \begin{pmatrix} -\mathbf{1}_{2 \times 2} & 0 \\ 0 & +\mathbf{1}_{2 \times 2} \end{pmatrix} \quad (5.93)$$

and the projectors on $\mathcal{S}_{\pm}$ are

$$P_{\pm} = \frac{1 \pm \gamma^5}{2} \quad (5.94)$$

Because of

$$\gamma^0 \gamma^5 = -\gamma^5 \gamma^0 \quad (5.95)$$

parity reflection as represented by

$$S(\mathcal{P}) = \gamma^0 \quad (5.96)$$

interchanges the two subspaces:

$$S(\mathcal{P}) : \mathcal{S}_{\pm} \mapsto \mathcal{S}_{\mp} \quad (5.97)$$

As a result, *Weyl spinors are chiral objects*, with $\mathcal{S}_+$ taken as right-handed and $\mathcal{S}_-$ taken as left-handed, which are spatial reflections of each other.

5.4.2 Weyl representation of Dirac matrices and spin

Any basis choice for the Dirac spinor space in which γ^5 is diagonal and

$$\gamma^{0\dagger} = \gamma^0 \quad \gamma^{i\dagger} = -\gamma^i \quad (5.98)$$

leads to a so-called Weyl representation of Dirac matrices. A simple example is

$$\gamma^0 = \begin{pmatrix} 0 & \mathbf{1}_{2 \times 2} \\ \mathbf{1}_{2 \times 2} & 0 \end{pmatrix} \quad , \quad \gamma^i = \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix} \quad (5.99)$$

where

$$\sigma^1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad , \quad \sigma^2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad , \quad \sigma^3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad (5.100)$$

are the Pauli matrices. We recall here the most important relations satisfied by the Pauli matrices

$$\begin{aligned} [\sigma^j, \sigma^k] &= 2i\epsilon^{jkl}\sigma^l \\ \{\sigma^j, \sigma^k\} &= 2\delta^{jk}\mathbf{1}_{2 \times 2} \end{aligned} \quad (5.101)$$

which enable fast computations when using the block formalism. One can also write these relations as

$$\sigma^j \sigma^k = \delta^{jk}\mathbf{1}_{2 \times 2} + i\epsilon^{jkl}\sigma^l \quad (5.102)$$

or as the Pauli vector identity

$$(\vec{\sigma} \cdot \vec{a})(\vec{\sigma} \cdot \vec{b}) = \vec{a} \cdot \vec{b}\mathbf{1} + i\vec{\sigma} \cdot (\vec{a} \times \vec{b}) \quad (5.103)$$

In the Weyl representation

$$\gamma^5 = \begin{pmatrix} -\mathbf{1}_{2 \times 2} & 0 \\ 0 & +\mathbf{1}_{2 \times 2} \end{pmatrix} \quad (5.104)$$

and so the upper two components of the Dirac spinor form the left-handed, while the lower two form the right-handed Weyl spinor.

Exercise 5.12. Compute the Lorentz generators $\sigma^{\mu\nu}$ in Weyl representation.

Solution. By direct calculation

$$\begin{aligned} \sigma^{0j} &= \frac{i}{2}[\gamma^0, \gamma^j] = i \begin{pmatrix} -\sigma^j & 0 \\ 0 & \sigma^j \end{pmatrix} \\ \sigma^{jk} &= \frac{i}{2}[\gamma^j, \gamma^k] = \epsilon^{jkl} \begin{pmatrix} \sigma^l & 0 \\ 0 & \sigma^l \end{pmatrix} \end{aligned} \quad (5.105)$$

□

Recalling the infinitesimal form of Lorentz and the corresponding spinor transformations

$$\begin{aligned} \Lambda &= \mathbf{1} - \frac{i}{2}\omega_{\mu\nu}J^{\mu\nu} \\ S(\Lambda) &= \mathbf{1} - \frac{i}{4}\omega_{\mu\nu}\sigma^{\mu\nu} \end{aligned} \quad (5.106)$$

leads to the following correspondence between the generators

$$J^{\mu\nu} \leftrightarrow \frac{1}{2}\sigma^{\mu\nu} \quad (5.107)$$

In particular, for the spatial rotations

$$J^{jk} = \epsilon_{jkl}J^l \quad (5.108)$$

where J^l , $l = 1, 2, 3$ are the generators of usual three-dimensional rotations. So J^l has the spinor representation

$$J^l \leftrightarrow \frac{1}{2} \begin{pmatrix} \sigma^l & 0 \\ 0 & \sigma^l \end{pmatrix} \quad (5.109)$$

As a result, the Dirac-spinor contains two spin-1/2 Pauli spinors and therefore describe two spin-1/2 objects. The conclusion is that

the Dirac equation describes spin-1/2 particles

thereby solving one of our problems with the Klein-Gordon equation by incorporating the spin.

5.4.3 Dirac and Majorana representations

A Dirac representation for the γ matrices is one in which γ^0 is diagonal. Since γ^5 does not commute with γ^0 , it is not diagonal in a Dirac representation.

Exercise 5.13. Show that the matrices

$$\gamma_D^\mu = A\gamma^\mu A^{-1} \quad (5.110)$$

where

$$A = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ -1 & 0 & 1 & 0 \\ 0 & -1 & 0 & 1 \end{pmatrix} \quad (5.111)$$

form a Dirac representation.

Solution. By direct calculation

$$\begin{aligned} \gamma_D^0 &= \begin{pmatrix} \mathbf{1}_{2 \times 2} & 0 \\ 0 & -\mathbf{1}_{2 \times 2} \end{pmatrix} \\ \gamma_D^l &= \begin{pmatrix} 0 & \sigma^l \\ -\sigma^l & 0 \end{pmatrix} \end{aligned} \quad (5.112)$$

□

Exercise 5.14. Compute γ_D^5 and $\sigma_D^{\mu\nu}$.

Solution. By direct calculation

$$\begin{aligned}\gamma_D^5 &= \begin{pmatrix} 0 & \mathbf{1}_{2 \times 2} \\ \mathbf{1}_{2 \times 2} & 0 \end{pmatrix} \\ \sigma_D^{0l} &= i \begin{pmatrix} 0 & \sigma^l \\ \sigma^l & 0 \end{pmatrix} \\ \sigma_D^{jk} &= \epsilon_{jkl} \begin{pmatrix} \sigma^l & 0 \\ 0 & \sigma^l \end{pmatrix}\end{aligned}\tag{5.113}$$

Note that γ_D^5 is indeed non-diagonal, and the spatial rotation generators are unchanged from the Weyl representation. $\square$

Another interesting class of representations is the one in which the Dirac equation is real, so that all γ -matrices are purely imaginary. This is called a Majorana representation, and in a Majorana basis the solutions of the Dirac equation can be restricted to be either purely real or purely imaginary.

Exercise 5.15. Show that the matrices

$$\gamma_M^\mu = A \gamma^\mu A^{-1}\tag{5.114}$$

where

$$A = \frac{1}{2} \begin{pmatrix} 1 & i & 1 & -i \\ -i & 1 & i & 1 \\ 1 & -i & -1 & -i \\ i & 1 & i & -1 \end{pmatrix}\tag{5.115}$$

form a Majorana representation.

Solution. By direct calculation

$$\begin{aligned}\gamma_M^0 &= \begin{pmatrix} 0 & \sigma^2 \\ \sigma^2 & 0 \end{pmatrix} & , & \quad \gamma_M^1 = \begin{pmatrix} i\sigma^3 & 0 \\ 0 & i\sigma^3 \end{pmatrix} \\ \gamma_M^2 &= \begin{pmatrix} 0 & -\sigma^2 \\ \sigma^2 & 0 \end{pmatrix} & , & \quad \gamma_M^3 = \begin{pmatrix} -i\sigma^1 & 0 \\ 0 & -i\sigma^1 \end{pmatrix}\end{aligned}\tag{5.116}$$

$\square$

Exercise 5.16. Compute γ_M^5 .

Solution. By direct calculation

$$\gamma_M^5 = \begin{pmatrix} \sigma^2 & 0 \\ 0 & -\sigma^2 \end{pmatrix}\tag{5.117}$$

Note that γ_M^5 is again non-diagonal. $\square$

5.5 Conserved current and probabilistic interpretation

Let us introduce the following current

$$j^\mu = \bar{\psi}\gamma^\mu\psi \quad (5.118)$$

which transforms as a 4-vector and satisfies the continuity equation

$$\partial_\mu j^\mu = 0 \quad (5.119)$$

as long as ψ satisfies the Dirac equation

$$(i\hbar\gamma^\mu\partial_\mu - mc)\psi(x) = 0 \quad (5.120)$$

Exercise 5.17. Show that j^μ transforms as a (contravariant) 4-vector.

Solution. Under a Lorentz transformation Λ

$$\begin{aligned} j^\mu &= \bar{\psi}\gamma^\mu\psi \xrightarrow{\Lambda} \bar{\psi}S(\Lambda)^{-1}\gamma^\mu S(\Lambda)\psi \\ &\quad \text{using } S(\Lambda)\gamma^\mu S(\Lambda)^{-1} = (\Lambda^{-1})^\mu{}_\nu\gamma^\nu \\ &= \Lambda^\mu{}_\nu\bar{\psi}\gamma^\nu\psi \\ &= \Lambda^\mu{}_\nu j^\nu \end{aligned} \quad (5.121)$$

□

Exercise 5.18. Show that

$$\partial_\mu j^\mu = 0 \quad (5.122)$$

follows from the Dirac equation

Solution. Writing the Dirac equation and its conjugate

$$\begin{aligned} (i\hbar\gamma^\mu\partial_\mu - mc)\psi(x) &= 0 \\ \bar{\psi}(x)\left(i\hbar\gamma^\mu\overleftarrow{\partial}_\mu + mc\right) &= 0 \end{aligned} \quad (5.123)$$

Multiplying the first by $-i\bar{\psi}/\hbar$ from the left and the second by $-i\psi/\hbar$ from the right and adding the results leads to

$$\begin{aligned} \bar{\psi}(x)\gamma^\mu\partial_\mu\psi(x) + \bar{\psi}(x)\gamma^\mu\overleftarrow{\partial}_\mu\psi(x) &= 0 \\ \Downarrow \\ \partial_\mu(\bar{\psi}(x)\gamma^\mu\psi(x)) &= 0 \end{aligned} \quad (5.124)$$

□

The conserved density

$$j^0 = \psi^\dagger\psi \quad (5.125)$$

is *positive definite*, which shows that the Dirac equation provides a solution to the other problem plaguing the Klein-Gordon equation and makes a probabilistic interpretation possible. Note that

the conserved current j^μ is bilinear in the Dirac field and so it is an allowed observable (cf. Subsection [5.4.1](#)).

Note that the Dirac equation can be written in the Schrödinger form by the following consideration

$$\begin{aligned} i\hbar\gamma^0\partial_0\psi &= i\hbar\gamma^0\frac{1}{c}\partial_t\psi = \left(i\hbar\vec{\gamma} \cdot \vec{\nabla} + mc\right)\psi(x) \\ \Downarrow \\ i\hbar\partial_t\psi &= H_D\psi \end{aligned} \tag{5.126}$$

where the Dirac Hamiltonian is

$$\begin{aligned} H_D &= i\hbar c\vec{\alpha} \cdot \vec{\nabla} + mc^2\beta \\ \alpha^i &= \gamma^0\gamma^i, \quad \beta = \gamma^0 \end{aligned} \tag{5.127}$$

This makes it manifest that similarly to the Schrödinger equation, the initial value problem is specified by giving only the state at the initial time.

5.6 Solutions of the free Dirac equation

5.6.1 Plane waves

The easiest way to solve the free Dirac equation

$$(i\hbar\gamma^\mu\partial_\mu - mc)\psi(x) = 0 \tag{5.128}$$

is to use the Dirac basis in which

$$\gamma^0 = \begin{pmatrix} \mathbf{1}_{2\times 2} & 0 \\ 0 & -\mathbf{1}_{2\times 2} \end{pmatrix}, \quad \gamma^l = \begin{pmatrix} 0 & \sigma^l \\ -\sigma^l & 0 \end{pmatrix} \tag{5.129}$$

Since $\psi(x)$ satisfies the Klein-Gordon equation, there are both positive and negative energy solutions. These can be written in the form

$$\psi_+(x) = u(\vec{p})e^{-ip_\mu x^\mu/\hbar}, \quad \psi_-(x) = v(\vec{p})e^{ip_\mu x^\mu/\hbar} \tag{5.130}$$

with

$$p^0 = p_0 = +\sqrt{\vec{p}^2 + m^2c^2} \tag{5.131}$$

Substituting these into the Dirac equation leads to the following equations for the spinor amplitudes:

$$\begin{aligned} (\gamma^\mu p_\mu - mc)u(\vec{p}) &= 0 \\ (\gamma^\mu p_\mu + mc)v(\vec{p}) &= 0 \end{aligned} \tag{5.132}$$

To simplify notations it is useful to introduce the Dirac slash. For any 4-vector a let us denote

$$\not{a} = \gamma^\mu a_\mu \tag{5.133}$$

so the Dirac equation in momentum space can be written as

$$\begin{aligned} (\not{p} - mc)u(\vec{p}) &= 0 \\ (\not{p} + mc)v(\vec{p}) &= 0 \end{aligned} \tag{5.134}$$

For later use, we note some interesting properties of the Dirac slash:

$$\begin{aligned} \not{a} \not{b} + \not{b} \not{a} &= 2a_\mu b^\mu \\ \gamma_5 \not{a} &= -\not{a} \gamma_5 \end{aligned} \quad (5.135)$$

the first follows from the Clifford algebra, while the second from the anti-commutation of γ_5 with all the γ^μ .

Let us first consider the solutions in the rest frame, i.e.

$$\vec{p} = 0 \Rightarrow p_0 = mc \quad (5.136)$$

leading to

$$\begin{aligned} \gamma^0 u(\vec{0}) &= +u(\vec{0}) \\ \gamma^0 v(\vec{0}) &= -v(\vec{0}) \end{aligned} \quad (5.137)$$

so the independent solutions are

$$\begin{aligned} u(\vec{p}=0, s=+1/2) &= \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} , & u(\vec{p}=0, s=-1/2) &= \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \\ v(\vec{p}=0, s=+1/2) &= \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} , & v(\vec{p}=0, s=-1/2) &= \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} \end{aligned} \quad (5.138)$$

The notation reflects the fact that due to the identification of the rotation generator around the z axis

$$J^3 = \frac{1}{2} \begin{pmatrix} \sigma^3 & 0 \\ 0 & \sigma^3 \end{pmatrix} \quad (5.139)$$

$u(\vec{0}, s)$ and $v(\vec{0}, s)$ can be interpreted to correspond to states with the z projection of the spin equal to s .

Moving positive energy solutions can be obtained by writing

$$u = \begin{pmatrix} \chi \\ \varphi \end{pmatrix} \quad (5.140)$$

with χ and φ 2-component spinors, which satisfy

$$\begin{aligned} (p_0 - mc)\chi - \vec{\sigma} \cdot \vec{p} \varphi &= 0 \\ -(p_0 + mc)\varphi + \vec{\sigma} \cdot \vec{p} \chi &= 0 \end{aligned} \quad (5.141)$$

here it is important to note that $p^i = -p_i$, i.e. $p_\mu = (p_0, -\vec{p})$. Using the second relation, the general positive energy solution is

$$u(\vec{p}, s) = \mathcal{N} \begin{pmatrix} \chi_s \\ \frac{\vec{\sigma} \cdot \vec{p}}{p_0 + mc} \chi_s \end{pmatrix} \quad (5.142)$$

with the following choice of basis:

$$\chi_{+1/2} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad , \quad \chi_{-1/2} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad (5.143)$$

and where $\mathcal{N}$ is a suitable normalisation factor.

Exercise 5.19. Show that the other relation

$$(p_0 - mc)\chi - \vec{\sigma} \cdot \vec{p}\varphi = 0 \quad (5.144)$$

automatically holds by virtue of the dispersion relation.

Solution. Substituting the solution

$$\begin{aligned} & (p_0 - mc)\chi - \vec{\sigma} \cdot \vec{p} \frac{\vec{\sigma} \cdot \vec{p}}{p_0 + mc} \chi \\ &= (p_0 - mc)\chi - \frac{(\vec{\sigma} \cdot \vec{p})^2}{p_0 + mc} \chi \\ &= \frac{1}{p_0 + mc} (p_0^2 - m^2 c^2 - \vec{p}^2) \chi = 0 \end{aligned} \quad (5.145)$$

where we used that

$$(\vec{\sigma} \cdot \vec{p})^2 = \sigma^i \sigma^j p^i p^j = \frac{1}{2} \{ \sigma^i, \sigma^j \} p^i p^j = p^i p^i \mathbf{1}_{2 \times 2} \quad (5.146)$$

□

Exercise 5.20. Compute the normalisation factor $\mathcal{N}$ from the (Lorentz-invariant) condition

$$\bar{u}(\vec{p}, s) u(\vec{p}, s) = 1 \quad (5.147)$$

supposing that

$$\chi^\dagger \chi = 1 \quad (5.148)$$

Solution. Using the Hermiticity of the Pauli matrices one can write

$$\begin{aligned} \bar{u}(\vec{p}, s) u(\vec{p}, s) &= |\mathcal{N}|^2 \left(\chi^\dagger \chi^\dagger \frac{\vec{\sigma} \cdot \vec{p}}{p_0 + mc} \right) \begin{pmatrix} \mathbf{1}_{2 \times 2} & 0 \\ 0 & -\mathbf{1}_{2 \times 2} \end{pmatrix} \begin{pmatrix} \chi \\ \frac{\vec{\sigma} \cdot \vec{p}}{p_0 + mc} \chi \end{pmatrix} \\ &= |\mathcal{N}|^2 \left(\chi^\dagger \chi - \chi^\dagger \left(\frac{\vec{\sigma} \cdot \vec{p}}{p_0 + mc} \right)^2 \chi \right) = |\mathcal{N}|^2 \left(1 - \frac{\vec{p}^2}{(p_0 + mc)^2} \right) \\ &= |\mathcal{N}|^2 \left(1 - \frac{(p_0)^2 - m^2 c^2}{(p_0 + mc)^2} \right) = |\mathcal{N}|^2 \frac{2mc}{p_0 + mc} = 1 \end{aligned} \quad (5.149)$$

The usual choice is

$$\mathcal{N} = \sqrt{\frac{p_0 + mc}{2mc}} \quad (5.150)$$

□

Note that using the above normalisation

$$\bar{u}(\vec{p}, s)u(\vec{p}, s') = \chi_s^\dagger \chi_{s'} \quad (5.151)$$

or, using the basis states $s, s' = \pm 1/2$

$$\bar{u}(\vec{p}, s)u(\vec{p}, s') = \delta_{ss'} \quad (5.152)$$

One can treat the negative energy case

$$(\not{p} + mc)v(\vec{p}) = 0 \quad (5.153)$$

in a similar way by writing

$$v = \begin{pmatrix} \chi \\ \varphi \end{pmatrix} \quad (5.154)$$

and obtaining

$$\begin{aligned} (p_0 + mc)\chi - \vec{\sigma} \cdot \vec{p}\varphi &= 0 \\ -(p_0 - mc)\varphi + \vec{\sigma} \cdot \vec{p}\chi &= 0 \end{aligned} \quad (5.155)$$

The solutions are

$$v(\vec{p}, s) = \mathcal{N} \begin{pmatrix} \frac{\vec{\sigma} \cdot \vec{p}}{p_0 + mc} \varphi_s \\ \varphi_s \end{pmatrix} \quad (5.156)$$

with the same normalisation factor, but this time satisfying the condition

$$\bar{v}(\vec{p}, s)v(\vec{p}, s') = -\varphi_s^\dagger \varphi_{s'} \quad (5.157)$$

or, again using the basis states $s, s' = \pm 1/2$

$$\bar{v}(\vec{p}, s)v(\vec{p}, s') = -\delta_{ss'} \quad (5.158)$$

Our results about the plane wave solutions are collected in the table below.

Energy	Plane wave form	Spinor amplitude	Basis	Invariant product
+	$\psi_+(x) = u(\vec{p})e^{-ip_\mu x^\mu/\hbar}$	$u(\vec{p}, s) = \sqrt{\frac{p_0 + mc}{2mc}} \begin{pmatrix} \chi_s \\ \frac{\vec{\sigma} \cdot \vec{p}}{p_0 + mc} \chi_s \end{pmatrix}$	$s = \pm \frac{1}{2}$	$\bar{u}(\vec{p}, s)u(\vec{p}, s') = \delta_{ss'}$
-	$\psi_-(x) = v(\vec{p})e^{+ip_\mu x^\mu/\hbar}$	$v(\vec{p}, s) = \sqrt{\frac{p_0 + mc}{2mc}} \begin{pmatrix} \frac{\vec{\sigma} \cdot \vec{p}}{p_0 + mc} \varphi_s \\ \varphi_s \end{pmatrix}$		$\bar{v}(\vec{p}, s)v(\vec{p}, s') = -\delta_{ss'}$
where $p_0 = +\sqrt{\vec{p}^2 + m^2 c^2}$, $\chi_{+1/2} = \varphi_{+1/2} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $\chi_{-1/2} = \varphi_{-1/2} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$				

Exercise 5.21. Compute the negative energy solutions and their properties, repeating the steps for the positive energy ones.

Solution. Explicit calculation. □

The Hermitian norms of the plane waves can also be computed with the result

$$\begin{aligned} u^\dagger(\vec{p}, s)u(\vec{p}, s') &= \frac{p_0}{mc}\delta_{ss'} \\ v^\dagger(\vec{p}, s)v(\vec{p}, s') &= \frac{p_0}{mc}\delta_{ss'} \quad s, s' = \pm 1/2 \end{aligned} \quad (5.159)$$

Note that these are not Lorentz invariant.

Exercise 5.22. Compute the Hermitian norms.

Solution. Starting from

$$u(\vec{p}, s) = \sqrt{\frac{p_0 + mc}{2mc}} \begin{pmatrix} \chi_s \\ \frac{\vec{\sigma} \cdot \vec{p}}{p_0 + mc} \chi_s \end{pmatrix} \quad (5.160)$$

we get

$$\begin{aligned} u^\dagger(\vec{p}, s)u(\vec{p}, s') &= \frac{p_0 + mc}{2mc} \left(1 + \frac{\vec{p}^2}{(p_0 + mc)^2} \right) \chi_s^\dagger \chi_{s'} \\ &= \frac{(p_0 + mc)^2 + p_0^2 - m^2 c^2}{2mc(p_0 + mc)} \chi_s^\dagger \chi_{s'} \\ &= \frac{p_0}{mc} \chi_s^\dagger \chi_{s'} \end{aligned} \quad (5.161)$$

Similar calculation gives the Hermitian norms of the $v(\vec{p}, s)$. □

5.6.2 Physical interpretation of the free Dirac equation

The presence of the negative-energy particles still leads to the same instabilities in the presence of interactions. They cannot be omitted from the space of states; the Klein paradox works the same way as for the Klein-Gordon equation. This is clear from the fact that the physical picture as shown in Fig. 4.1 is unchanged.

However, the particles obeying the Dirac equation have spin 1/2 and therefore are fermions according to Pauli's principle⁹. So there is a simple way to stabilise the system by assuming that in the ground state, the negative energy states are filled with particles. This is called the *Dirac sea*.

The Dirac sea, however, has serious consequences for the physical behaviour of the system as shown by the analysis below.

Energy and momentum; antiparticles The energy and momentum of the plane wave modes can be obtained using the explicit expression for the energy and momentum operators:

$$\hat{E} = i\hbar\partial_t = i\hbar c\partial_0 \quad , \quad \hat{P}^k = -i\hbar\partial_k \quad (5.162)$$

⁹In relativistic quantum field theory it can be proven rigorously that integer spin particles must obey Bose-Einstein, while half-integer ones must be subject to Fermi-Dirac statistics; this is called the spin-statistics theorem.

$$\begin{aligned}
\psi_+(x) &= u(p)e^{-ip_\mu x^\mu/\hbar} &: \hat{E}\psi_+ &= p_0 c \psi_+ & \hat{P}^k \psi_+ &= -p_k \psi_+ = p^k \psi_+ \\
\psi_-(x) &= u(p)e^{+ip_\mu x^\mu/\hbar} &: \hat{E}\psi_- &= -p_0 c \psi_- & \hat{P}^k \psi_- &= p_k \psi_- = -p^k \psi_-
\end{aligned} \tag{5.163}$$

Excitations of the systems correspond to either

- particles in positive energy states. They carry momentum $\vec{p}$ and energy

$$E = p_0 c = \sqrt{\vec{p}^2 c^2 + m^2 c^4} \tag{5.164}$$

, and can be interpreted as relativistic particles of mass m and spin $1/2$, which is represented by the 2-spinor amplitude χ_s in the plane wave solution.

- holes in negative energy states. Since they correspond to missing negative energy particles, they again carry momentum $\vec{p}$ and energy

$$E = p_0 c = \sqrt{\vec{p}^2 c^2 + m^2 c^4} \tag{5.165}$$

and can be interpreted as relativistic particles of mass m and spin $1/2$, which is represented by the 2-spinor amplitude φ_s in the plane wave solution; however, their spin projection is rather $-s$ in their rest frame. In addition, if the particles carry some charge (electric or other), holes carry the opposite charge.

As a result, *the holes correspond to anti-particles*, leading to the Dirac's famous prediction of the positron when the Dirac equation is applied to describe the electron, which was confirmed by Anderson in 1931, earning him the Nobel prize in 1936. Dirac shared the Nobel prize with Schrödinger in 1933.

When taking interaction with photons into account, photons with energy larger than $2mc^2$ can excite a negative energy particle to a positive energy state, leading to the creation of electron-positron pairs. Due to energy and momentum conservation, this can only proceed if there is some additional partner present, for example in the electric field of an atomic nucleus that can absorb some amount of energy and momentum.

Exercise 5.23. Show that conservation of energy and momentum forbids the creation of an electron-positron pair from a single photon without any additional interaction present.

Solution. Suppose the photon has 4-momentum k^μ , while the created pair has 4-momenta p_1^μ and p_2^μ . We then have the following constraint equations

$$\begin{aligned}
\text{Momentum conservation:} & \quad k^\mu = p_1^\mu + p_2^\mu \\
\text{Electron/positron rest mass:} & \quad p_{1\mu} p_1^\mu = p_{2\mu} p_2^\mu = m^2 c^2 \\
\text{Photon rest mass:} & \quad k^\mu k_\mu = 0
\end{aligned} \tag{5.166}$$

But this is impossible to satisfy because p_1 and p_2 are time-like, so their sum is also time-like, while k is light-like. In more detail, the following must hold

$$0 = k^\mu k_\mu = 2m^2 c^2 + 2p_{1\mu} p_2^\mu \tag{5.167}$$

Now one can write

$$p_{1\mu} p_2^\mu = \sqrt{\vec{p}_1^2 + m^2 c^2} \sqrt{\vec{p}_2^2 + m^2 c^2} - \vec{p}_1 \cdot \vec{p}_2 \geq |\vec{p}_1| |\vec{p}_2| - \vec{p}_1 \cdot \vec{p}_2 \geq 0 \tag{5.168}$$

so it is impossible to fulfil energy and momentum conservation. $\square$

It is also possible that a particle and an anti-particle described by a hole annihilate, which in the Dirac sea picture corresponds to a transition to the empty negative energy level corresponding to the hole. This is accompanied by the emission of a photon pair (which allows to satisfy energy and momentum conservation). The creation and annihilation processes signal the breakdown of one-particle quantum mechanics; whenever the negative energy levels come into play the system must be described with multi-particle states allowing for transition between states with different particle/antiparticle numbers, which forces the introduction of quantum field theory and second quantisation.

Probability current and velocity One can compute the probability density and current for plane wave solutions with the result

$$\begin{aligned}\rho_{\pm}(x) &= j_{\pm}^0(x) = \psi_{\pm}^{\dagger}(x)\psi_{\pm}(x) = \frac{p_0}{mc} \\ \vec{j}_{\pm}(x) &= \psi_{\pm}^{\dagger}(x)\gamma^0\vec{\gamma}\psi_{\pm}(x) = \pm\frac{\vec{p}}{mc}\end{aligned}\tag{5.169}$$

The covariant form of the continuity equation

$$\partial_{\mu}j^{\mu} = 0\tag{5.170}$$

gets rewritten in the usual form when changing the variable $x^0 = ct$:

$$\frac{\partial\rho}{\partial t} + \vec{\nabla} \cdot c\vec{j} = 0\tag{5.171}$$

so the probability current is eventually $c\vec{j}$. This is perfectly consistent with the particle interpretation, since these densities satisfy

$$c\vec{j}_{\pm} = \pm\rho_{\pm}\vec{v}\tag{5.172}$$

where

$$\vec{v} = \frac{\vec{p}c}{p_0}\tag{5.173}$$

is indeed the relativistic expression for the 3-velocity in terms of the energy and momentum. The minus sign in the expression for the negative energy plane waves is explained by the fact that their momentum is $-\vec{p}$; it is the holes which carry momentum $\vec{p}$.

Exercise 5.24. Compute the above formulae for the densities and currents.

Solution. For the positive energy solutions we obtain

$$\begin{aligned}\rho_{+}(x) &= \psi_{+}^{\dagger}(x)\psi_{+}(x) = \frac{p_0 + mc}{2mc} \left(\chi^{\dagger} \chi^{\dagger} \frac{\vec{\sigma} \cdot \vec{p}}{p_0 + mc} \right) \left(\frac{\chi}{p_0 + mc} \right) \\ &= \frac{p_0 + mc}{2mc} \left(1 + \frac{\vec{p}^2}{(p_0 + mc)^2} \right) = \frac{p_0 + mc}{2mc} \left(1 + \frac{p_0^2 - m^2c^2}{(p_0 + mc)^2} \right) \\ &= \frac{p_0}{mc}\end{aligned}\tag{5.174}$$

$$\begin{aligned}
j_+^k(x) &= \psi_+^\dagger(x) \gamma^0 \gamma^k \psi_+(x) = \frac{p_0 + mc}{2mc} \left(\chi^\dagger \chi^\dagger \frac{\vec{\sigma} \cdot \vec{p}}{p_0 + mc} \right) \begin{pmatrix} \mathbf{1}_{2 \times 2} & 0 \\ 0 & -\mathbf{1}_{2 \times 2} \end{pmatrix} \begin{pmatrix} 0 & \sigma^k \\ -\sigma^k & 0 \end{pmatrix} \begin{pmatrix} \chi \\ \frac{\vec{\sigma} \cdot \vec{p}}{p_0 + mc} \chi \end{pmatrix} \\
&= \frac{p_0 + mc}{2mc} \left(\chi^\dagger \chi^\dagger \frac{\vec{\sigma} \cdot \vec{p}}{p_0 + mc} \right) \begin{pmatrix} 0 & \sigma^k \\ \sigma^k & 0 \end{pmatrix} \begin{pmatrix} \chi \\ \frac{\vec{\sigma} \cdot \vec{p}}{p_0 + mc} \chi \end{pmatrix} \\
&= \frac{p_0 + mc}{2mc} \chi^\dagger \left(\sigma^k \frac{\vec{\sigma} \cdot \vec{p}}{p_0 + mc} + \frac{\vec{\sigma} \cdot \vec{p}}{p_0 + mc} \sigma^k \right) \chi = \frac{1}{2mc} \chi^\dagger \{ \sigma^k, \sigma^l \} p^l \chi \\
&= \frac{1}{2mc} \chi^\dagger 2 \delta^{kl} p^l \chi = \frac{p^k}{mc}
\end{aligned} \tag{5.175}$$

Similar computation gives the results for the negative energy solutions. $\square$

Plane waves can be normalised when putting the particle in the box of volume V by requiring

$$\int_V d^3x \rho_{\pm(V)}(x) = \int_V d^3x \psi_{(V)\pm}^\dagger(x) \psi_{(V)\pm}(x) = 1 \tag{5.176}$$

However, our plane wave solutions satisfy instead

$$\int_V d^3x \psi_\pm^\dagger(x) \psi_\pm(x) = \frac{p_0}{mc} V \tag{5.177}$$

So the plane waves normalised in a box can be written as

$$\psi_{(V)\pm}(x) = \sqrt{\frac{mc}{p_0 V}} \psi_\pm(x) \tag{5.178}$$

5.7 Weyl spinors

In the Weyl representation with diagonal γ_5

$$\gamma_5 = \begin{pmatrix} -\mathbf{1}_{2 \times 2} & 0 \\ 0 & +\mathbf{1}_{2 \times 2} \end{pmatrix} \tag{5.179}$$

the Dirac spinor can be written in terms of chiral components

$$\psi = \begin{pmatrix} \phi_- \\ \phi_+ \end{pmatrix} \tag{5.180}$$

where $\phi_\pm$ are two-dimensional so-called *Weyl spinors*. The two *chiral components* of the Dirac spinor can also be written as

$$\psi_- = \frac{1 - \gamma_5}{2} \psi = \begin{pmatrix} \phi_- \\ 0 \end{pmatrix} \quad \psi_+ = \frac{1 + \gamma_5}{2} \psi = \begin{pmatrix} 0 \\ \phi_+ \end{pmatrix} \tag{5.181}$$

Chirality is nothing else than the eigenvalue of γ_5 ; therefore ϕ_- has negative and ϕ_+ has positive chirality. Using

$$\gamma^0 = \begin{pmatrix} 0 & \mathbf{1}_{2 \times 2} \\ \mathbf{1}_{2 \times 2} & 0 \end{pmatrix} \quad , \quad \gamma^k = \begin{pmatrix} 0 & \sigma^k \\ -\sigma^k & 0 \end{pmatrix} \tag{5.182}$$

the Dirac equation can be written as

$$\begin{aligned} i\hbar\partial_0\phi_+ + i\hbar\sigma^k\partial_k\phi_+ - mc\phi_- &= 0 \\ i\hbar\partial_0\phi_- - i\hbar\sigma^k\partial_k\phi_- - mc\phi_+ &= 0 \end{aligned} \quad (5.183)$$

For zero mass, the two components decouple:

$$i\partial_0\phi_\pm = \mp i\sigma^k\partial_k\phi_\pm \quad (5.184)$$

Let us consider positive and negative energy plane wave solutions of these equations

$$\begin{aligned} \phi_\pm^{(+)}(x) &= u_\pm(p)e^{-ip_\mu x^\mu/\hbar} \\ \phi_\pm^{(-)}(x) &= v_\pm(p)e^{+ip_\mu x^\mu/\hbar} \end{aligned} \quad (5.185)$$

with $p_0 = |\vec{p}|$. Writing

$$e^{-ip_\mu x^\mu/\hbar} = e^{-ip_0 x^0/\hbar} e^{+ip^k x^k/\hbar} \quad (5.186)$$

the positive and negative chirality Weyl equations become

$$\begin{aligned} \frac{\vec{\sigma} \cdot \vec{p}}{|\vec{p}|} u_+(p) &= +u_+(p) \\ \frac{\vec{\sigma} \cdot \vec{p}}{|\vec{p}|} v_+(p) &= +v_+(p) \\ \frac{\vec{\sigma} \cdot \vec{p}}{|\vec{p}|} u_-(p) &= -u_-(p) \\ \frac{\vec{\sigma} \cdot \vec{p}}{|\vec{p}|} v_-(p) &= -v_-(p) \end{aligned} \quad (5.187)$$

These equations mean that *for a negative chirality Weyl particle, the spin points opposite to the direction of momentum, so the particle is left-handed*. For the negative energy mode, we already noted that the negative energy particle has momentum $-\vec{p}$ and spin $\vec{\sigma}$. So in the case of the anti-particle, which is a hole in the Dirac sea, the momentum is $\vec{p}$ and the spin is $-\vec{\sigma}$, therefore the anti-particle is right-handed. For the positive chirality Weyl particle, the situation is just the opposite. This is summarised in the table below:

Chirality	Particle	Antiparticle
+1	right-handed (spin parallel with motion)	left-handed (spin opposite to motion)
-1	left-handed (spin opposite to motion)	right-handed (spin parallel with motion)

This is relevant for weak interactions. It turns out that only the left-handed fermions and right-handed anti-fermions couple to the $W^\pm$ bosons that mediate weak decays. In the approximation when the neutrino mass is taken to be zero¹⁰, the neutrino can be modelled as a negative chirality Weyl particle as discussed later in Subsection 6.2.3. Note that since parity is represented by

$$\gamma^0 = \begin{pmatrix} 0 & \mathbf{1}_{2 \times 2} \\ \mathbf{1}_{2 \times 2} & 0 \end{pmatrix} \quad (5.188)$$

¹⁰Indeed there is as yet no direct evidence for a non-zero neutrino mass, although the observed phenomenon of neutrino mixing implies that all three mass differences must be non-zero, so at least two neutrinos should have a tiny, but non-zero mass.

it takes the positive and negative chirality components into each other, so it changes the handedness of the particles. This can also be seen from the fact that the spin is a pseudo-vector, while the momentum is a polar vector, so their scalar product changes sign under spatial reflection P .

On the other hand, as seen from the above table charge conjugation C also changes handedness, so the combination CP in fact preserves chirality and maps $\psi_{\pm}$ to $\psi_{\pm}$. *Therefore the handedness of weak interactions in itself does not break CP .* Although CP is broken in nature, the origin of this breaking is thought to come from a mixing between the three generations of quarks (cf. Subsection 11.4). The experimental observation of neutrino oscillations (cf. Subsection 11.4.3) opened another possibility for CP breaking whose source is the neutrino mixing matrix.

5.8 Dirac equation in an external electromagnetic field

The electromagnetic field described by a four-vector potential A_{μ} , can be included in the Dirac equation

$$(i\hbar\gamma^{\mu}\partial_{\mu} - mc)\psi(x) = 0 \quad (5.189)$$

using the prescription of *minimal subtraction*

$$p^{\mu} \rightarrow p^{\mu} - qA^{\mu} \quad (5.190)$$

This is the relativistic extension of the non-relativistic recipe

$$\vec{p} \rightarrow \vec{p} - q\vec{A} \quad (5.191)$$

and is related to the concept of *gauge invariance* introduced later in Section 9.3. This procedure leads to

$$[\gamma^{\mu}(i\hbar\partial_{\mu} - qA_{\mu}(x)) - mc]\psi(x) = 0 \quad (5.192)$$

or, using Dirac's notation

$$(i\hbar\not{\partial} - q\not{A} - mc)\psi = 0 \quad (5.193)$$

This equation can be decomposed into upper and lower components

$$\Psi = \begin{pmatrix} \chi \\ \varphi \end{pmatrix} \quad (5.194)$$

with the following result (in the Dirac representation for the γ matrices)

$$\begin{aligned} (i\hbar\partial_t - qV - mc^2)\chi + c\vec{\sigma} \cdot (i\hbar\vec{\nabla} + q\vec{A})\varphi &= 0 \\ (-i\hbar\partial_t + qV - mc^2)\varphi - c\vec{\sigma} \cdot (i\hbar\vec{\nabla} + q\vec{A})\chi &= 0 \end{aligned} \quad (5.195)$$

where V and $\vec{A}$ are the electromagnetic scalar and vector potentials, respectively.

Exercise 5.25. Derive the above equations.

Solution. Using

$$\gamma_D^0 = \begin{pmatrix} \mathbf{1}_{2 \times 2} & 0 \\ 0 & -\mathbf{1}_{2 \times 2} \end{pmatrix}, \quad \gamma_D^l = \begin{pmatrix} 0 & \sigma^l \\ -\sigma^l & 0 \end{pmatrix} \quad (5.196)$$

we obtain

$$\begin{aligned}(i\hbar\partial_0 - qA_0 - mc)\chi + \sigma^l(i\hbar\partial_l - qA_l)\varphi &= 0 \\ (-i\hbar\partial_0 + qA_0 - mc)\varphi - \sigma^l(i\hbar\partial_l - qA_l)\chi &= 0\end{aligned}\tag{5.197}$$

Substituting

$$x^0 = ct \quad , \quad \partial_0 = \frac{1}{c}\partial_t \quad , \quad A_0 = \frac{1}{c}V \quad , \quad A^l = -A_l\tag{5.198}$$

we get

$$\begin{aligned}(i\hbar\partial_t - qV - mc^2)\chi + c\vec{\sigma} \cdot (i\hbar\vec{\nabla} + q\vec{A})\varphi &= 0 \\ (-i\hbar\partial_t + qV - mc^2)\varphi - c\vec{\sigma} \cdot (i\hbar\vec{\nabla} + q\vec{A})\chi &= 0\end{aligned}\tag{5.199}$$

□

To obtain the non-relativistic limit, consider relatively small momenta and weak fields, so the dominant terms become

$$\begin{aligned}(i\hbar\partial_t - mc^2)\chi + \dots &= 0 \\ (-i\hbar\partial_t - mc^2)\varphi + \dots &= 0\end{aligned}\tag{5.200}$$

As a result, most of the time dependence can be factored out as

$$\begin{pmatrix} \chi \\ \varphi \end{pmatrix} = e^{-imc^2t/\hbar} \begin{pmatrix} \Psi \\ \Phi \end{pmatrix}\tag{5.201}$$

leading to

$$\begin{aligned}(i\hbar\partial_t - qV)\Psi + c\vec{\sigma} \cdot (i\hbar\vec{\nabla} + q\vec{A})\Phi &= 0 \\ (-i\hbar\partial_t + qV - 2mc^2)\Phi - c\vec{\sigma} \cdot (i\hbar\vec{\nabla} + q\vec{A})\Psi &= 0\end{aligned}\tag{5.202}$$

Assuming that Ψ and Φ are slowly varying in time, and also that the potential is small ($|qV| \ll 2mc^2$), the second equation can be solved for Φ :

$$\Phi = -\frac{1}{2mc}\vec{\sigma} \cdot (i\hbar\vec{\nabla} + q\vec{A})\Psi\tag{5.203}$$

which can be substituted into the first equation to give

$$(i\hbar\partial_t - qV)\Psi - \frac{1}{2m}\left[\vec{\sigma} \cdot (i\hbar\vec{\nabla} + q\vec{A})\right]^2\Psi = 0\tag{5.204}$$

This leads to the Pauli equation

$$\begin{aligned}i\hbar\frac{\partial\Psi}{\partial t} &= H\Psi \\ \text{where } H &= \frac{1}{2m}\left[\vec{\sigma} \cdot (\vec{p} - q\vec{A})\right]^2 + qV \\ \text{and } \vec{p} &= \frac{\hbar}{i}\vec{\nabla}\end{aligned}\tag{5.205}$$

The Pauli Hamiltonian can also be written as

$$H = \underbrace{\frac{1}{2m} (\vec{p} - q\vec{A})^2 + qV}_{\text{Schrödinger Hamiltonian}} - \underbrace{\frac{q\hbar}{2m} \vec{\sigma} \cdot \vec{B}}_{\text{Pauli term}} \quad (5.206)$$

Exercise 5.26. Compute the above form of the Hamiltonian.

Solution.

$$\begin{aligned} \left[\vec{\sigma} \cdot (\vec{p} - q\vec{A}) \right]^2 &= \sigma^j \sigma^k (p^j - qA^j)(p^k - qA^k) \\ &= (\delta^{jk} + i\epsilon_{jkl}\sigma^l)(p^j - qA^j)(p^k - qA^k) \\ &= (\vec{p} - q\vec{A})^2 + \frac{1}{2}i\epsilon_{jkl}\sigma^l [p^j - qA^j, p^k - qA^k] \end{aligned} \quad (5.207)$$

Now we can write

$$[p^j - qA^j, p^k - qA^k] = -q([p^j, A^k] + [A^j, p^k]) = -q(-i\hbar\partial_j A^k + i\hbar\partial_k A^j) \quad (5.208)$$

so

$$\frac{1}{2}i\epsilon_{jkl}\sigma^l [p^j - qA^j, p^k - qA^k] = \frac{1}{2}i\sigma^l(-q)(-i\hbar)2\epsilon_{ljk}\frac{\partial}{\partial x^j}A^k = -q\hbar\vec{\sigma} \cdot \vec{B} \quad (5.209)$$

from which the result follows. $\square$

Note that the non-relativistic spin operator is

$$\vec{S} = \frac{\hbar}{2}\vec{\sigma} \quad (5.210)$$

so the magnetic moment of the Dirac particle can be written as

$$\vec{\mu} = g\frac{q}{2m}\vec{S} \quad (5.211)$$

where g is the so-called *gyromagnetic factor*, and the Dirac equation (with minimal coupling to the electromagnetic field) predicts its value to be exactly $g = 2$. In experiment, the measured values of gyromagnetic factors are (the numbers in the parentheses are the errors of the last two digits)

$$\begin{aligned} \text{electron: } g_e &= 2.0023193043617(15) \\ \text{muon: } g_\mu &= 2.0023318414(12) \\ \text{proton: } g_p &= 5.585694713(46) \end{aligned} \quad (5.212)$$

The deviation for the electron can be explained taking into account the effects of quantum electrodynamics; for the muon, the effects of weak and strong interactions are also needed. The big deviation in proton g -factor from 2 is due to the proton being a composite particle of quarks and gluons; due to the strong interactions the proton spin and magnetic moment is not well understood at present. The same goes for the neutron; despite being neutral, it's magnetic moment is

$$\vec{\mu}_n = g_n \frac{e}{2m_n} \vec{S} \quad (5.213)$$

where $e = 1.6021717662 \cdot 10^{-19}$ coulombs is the elementary charge, and

$$g_n = -3.82608545(90) \quad (5.214)$$

5.9 Charge conjugation

It is interesting to consider what is transformation corresponding to charge conjugation which maps particles into antiparticles. Using the example of quantum electrodynamics, it is clear that taking the complex conjugate of the covariant derivative

$$(\hbar\partial_\mu + iqA_\mu)^* = \hbar\partial_\mu - iqA_\mu \quad (5.215)$$

corresponds to changing q to $-q$. For the scalar field, therefore the charge conjugate of the field is

$$(C\Phi)(x) = \xi\Phi^*(x) \quad (5.216)$$

where ξ is the *charge conjugation parity* of the field which must satisfy

$$\xi^*\xi = 1 \quad (5.217)$$

since repeating C should give the identity.

For the Dirac field charge conjugation is a little more involved due to spinor structure. Charge conjugation is a map $\psi \rightarrow \psi^c$ interchanging the particle with the antiparticle; therefore it should reverse the charge $q \rightarrow -q$:

$$(i\hbar\not{\partial} - q\not{A} - mc)\psi = 0 \Leftrightarrow (i\hbar\not{\partial} + q\not{A} - mc)\psi^c = 0 \quad (5.218)$$

To construct this transformation, take the conjugate Dirac equation

$$\bar{\psi}(x) \left(-i\hbar\overleftarrow{\not{\partial}} - q\not{A} - mc \right) = 0 \quad (5.219)$$

and transpose it

$$\gamma^{\mu T} (-i\hbar\partial_\mu - qA_\mu - mc) \bar{\psi}^T = 0 \quad (5.220)$$

Suppose that there exists a *charge conjugation matrix* $\mathcal{C}$ satisfying

$$-\mathcal{C}\gamma^{\mu T}\mathcal{C}^{-1} = \gamma^\mu \quad (5.221)$$

then one can take

$$\psi^c = \xi\mathcal{C}\bar{\psi}^T \quad (5.222)$$

where ξ is the charge conjugation parity of the field Ψ .

In fact, Clifford's theorem guarantees the existence of such a matrix since

$$\{\gamma^\mu, \gamma^\nu\} = 2\eta^{\mu\nu}\mathbf{1} \Rightarrow \{-\gamma^{\mu T}, -\gamma^{\nu T}\} = 2\eta^{\mu\nu}\mathbf{1} \quad (5.223)$$

Its explicit form, however, is representation dependent. In the Dirac representation

$$\begin{aligned} \gamma_D^0 &= \begin{pmatrix} \mathbf{1}_{2\times 2} & 0 \\ 0 & -\mathbf{1}_{2\times 2} \end{pmatrix} \\ \gamma_D^l &= \begin{pmatrix} 0 & \sigma^l \\ -\sigma^l & 0 \end{pmatrix} \end{aligned} \quad (5.224)$$

the γ matrices satisfy

$$\gamma_D^0 = \gamma_D^{0T} \quad \gamma_D^1 = -\gamma_D^{1T} \quad \gamma_D^2 = \gamma_D^{2T} \quad \gamma_D^3 = -\gamma_D^{3T} \quad (5.225)$$

so a suitable choice is

$$\mathcal{C}_D = i\gamma^2\gamma^0 = -i \begin{pmatrix} 0 & \sigma^2 \\ \sigma^2 & 0 \end{pmatrix} \quad (5.226)$$

which satisfies

$$-\mathcal{C}_D = \mathcal{C}_D^{-1} = \mathcal{C}_D^T = \mathcal{C}_D^\dagger \quad (5.227)$$

Exercise 5.27. Find the charge conjugation matrix in the Weyl and Majorana representations.

Solution. In the Weyl representation

$$\gamma^0 = \begin{pmatrix} 0 & \mathbf{1}_{2 \times 2} \\ \mathbf{1}_{2 \times 2} & 0 \end{pmatrix} \quad , \quad \gamma^i = \begin{pmatrix} 0 & -\sigma^i \\ \sigma^i & 0 \end{pmatrix} \quad (5.228)$$

we have

$$\gamma^0 = \gamma^{0T} \quad \gamma^1 = -\gamma^{1T} \quad \gamma^2 = \gamma^{2T} \quad \gamma^3 = -\gamma^{3T} \quad (5.229)$$

so the correct choice is

$$\mathcal{C} = i\gamma^2\gamma^0 = i \begin{pmatrix} -\sigma^2 & 0 \\ 0 & \sigma^2 \end{pmatrix} \quad (5.230)$$

and again

$$-\mathcal{C} = \mathcal{C}^{-1} = \mathcal{C}^T = \mathcal{C}^\dagger \quad (5.231)$$

In the Majorana representation

$$\begin{aligned} \gamma_M^0 &= \begin{pmatrix} 0 & \sigma^2 \\ \sigma^2 & 0 \end{pmatrix} \quad , \quad \gamma_M^1 = \begin{pmatrix} i\sigma^3 & 0 \\ 0 & i\sigma^3 \end{pmatrix} \\ \gamma_M^2 &= \begin{pmatrix} 0 & -\sigma^2 \\ \sigma^2 & 0 \end{pmatrix} \quad , \quad \gamma_M^3 = \begin{pmatrix} -i\sigma^1 & 0 \\ 0 & -i\sigma^1 \end{pmatrix} \end{aligned} \quad (5.232)$$

we have

$$\gamma_M^0 = -\gamma_M^{0T} \quad \gamma_M^1 = \gamma_M^{1T} \quad \gamma_M^2 = \gamma_M^{2T} \quad \gamma_M^3 = \gamma_M^{3T} \quad (5.233)$$

so a suitable choice is

$$\mathcal{C}_M = -\gamma_M^0 = - \begin{pmatrix} 0 & \sigma^2 \\ \sigma^2 & 0 \end{pmatrix} \quad (5.234)$$

With this choice

$$\mathcal{C}_M = \mathcal{C}_M^{-1} = -\mathcal{C}_M^T = \mathcal{C}_M^\dagger \quad (5.235)$$

□

5.10 Majorana fermions and Majorana mass term

Note that in the Majorana representation

$$\begin{aligned}\psi^c &= \mathcal{C}_M \bar{\psi}^T = \mathcal{C}_M \gamma_M^{0T} \psi^{\dagger T} = (-\gamma_M^0)^2 \psi^* \\ &= \psi^*\end{aligned}\tag{5.236}$$

so charge conjugation is equivalent to complex conjugation for the Dirac field. Additionally, in this representation the Dirac equation

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

is purely real since the γ matrices [5.116](#) are purely imaginary, so the Dirac equation can be restricted to real fields ψ . This observation leads to the possibility of neutral spin-1/2 particles, called Majorana fermions. These correspond to a purely real field ψ in Majorana representation. In a general representation, a Majorana field satisfies

$$\psi^c = \psi\tag{5.237}$$

Exercise 5.28. Compute the components of ψ^c in Weyl and Dirac representations.

Solution. Writing

$$\psi = \begin{pmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \psi_4 \end{pmatrix}\tag{5.238}$$

in both representations the result is

$$\psi^c = \begin{pmatrix} \psi_4^* \\ -\psi_3^* \\ -\psi_2^* \\ \psi_1^* \end{pmatrix}\tag{5.239}$$

□

In the Weyl representation, charge conjugation swaps the upper/lower two components of the Dirac spinor, which is equivalent to swapping the left/right chiral components; this latter observation is true independently of the representation.

Exercise 5.29. Verify explicitly that $(\psi^c)_R = (\psi_L)^c$ and $(\psi^c)_L = (\psi_R)^c$.

Solution. The simplest way is in Weyl representation, where charge conjugation is

$$\psi^c = \begin{pmatrix} \psi_4^* \\ -\psi_3^* \\ -\psi_2^* \\ \psi_1^* \end{pmatrix}\tag{5.240}$$

and

$$\psi_L = \frac{1 - \gamma_5}{2} \psi = \begin{pmatrix} \psi_1 \\ \psi_2 \\ 0 \\ 0 \end{pmatrix} \quad \psi_R = \frac{1 + \gamma_5}{2} \psi = \begin{pmatrix} 0 \\ 0 \\ \psi_3 \\ \psi_4 \end{pmatrix} \quad (5.241)$$

□

so

$$(\psi^c)_R = \begin{pmatrix} 0 \\ 0 \\ -\psi_2^* \\ \psi_1^* \end{pmatrix} \equiv (\psi_L)^c \quad \text{and} \quad (\psi^c)_L = \begin{pmatrix} \psi_4^* \\ -\psi_3^* \\ 0 \\ 0 \end{pmatrix} \equiv (\psi_R)^c.$$

In the following we shall use the convention that

$$\psi_R^c = (\psi_R)^c \quad \text{and} \quad \psi_L^c = (\psi_L)^c$$

so ψ_L^c is a right-handed, while ψ_R^c is a left-handed spinor.

This observation makes it possible to introduce another type of mass term, called the Majorana mass term. Note that Lorentz invariance requires that the mass couples left-handed components to right-handed and vice versa, as manifest in writing the Dirac mass term as

$$m_D \bar{\psi} \psi = m_D (\bar{\psi}_L \psi_R + \bar{\psi}_R \psi_L) \quad (5.242)$$

in terms of the chiral components

$$\psi_{L/R} = \frac{1 \mp \gamma_5}{2} \psi.$$

Since the charge conjugate of a left-handed component is right-handed and vice-versa, Lorentz invariance also allows for the *Majorana mass terms*

$$\begin{aligned} & \frac{1}{2} m_L (\bar{\psi}_L^c \psi_L + \bar{\psi}_L \psi_L^c) \\ & \frac{1}{2} m_R (\bar{\psi}_R^c \psi_R + \bar{\psi}_R \psi_R^c) \end{aligned} \quad (5.243)$$

Exercise 5.30. Compute the terms $\bar{\psi}_L^c \psi_L$ and $\bar{\psi}_R^c \psi_R$ explicitly in Weyl representation.

Solution. We have

$$\psi_L = \frac{1 - \gamma_5}{2} \psi = \begin{pmatrix} \psi_1 \\ \psi_2 \\ 0 \\ 0 \end{pmatrix} \quad \psi_L^c = \begin{pmatrix} 0 \\ 0 \\ -\psi_2^* \\ \psi_1^* \end{pmatrix} \quad (5.244)$$

from which we can write

$$\begin{aligned} \bar{\psi}_L^c \psi_L &= (\psi_L^c)^\dagger \gamma_0 \psi_L = \begin{pmatrix} 0 & 0 & -\psi_2 & \psi_1 \end{pmatrix} \begin{pmatrix} 0 & \mathbf{1}_{2 \times 2} \\ \mathbf{1}_{2 \times 2} & 0 \end{pmatrix} \begin{pmatrix} \psi_1 \\ \psi_2 \\ 0 \\ 0 \end{pmatrix} \\ &= \begin{pmatrix} 0 & 0 & -\psi_2 & \psi_1 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ \psi_1 \\ \psi_2 \end{pmatrix} = -\psi_2 \psi_1 + \psi_1 \psi_2. \end{aligned} \quad (5.245)$$

While this naively seems to vanish, it must be remembered that fermions obey anticommutation relations, so the two terms are actually equal! Similarly

$$\psi_R = \frac{1 + \gamma_5}{2} \psi = \begin{pmatrix} 0 \\ 0 \\ \psi_3 \\ \psi_4 \end{pmatrix} \quad \psi_R^c = \begin{pmatrix} \psi_4^* \\ -\psi_3^* \\ 0 \\ 0 \end{pmatrix} \quad (5.246)$$

leads to

$$\overline{\psi_R^c} \psi_R = \psi_4 \psi_3 - \psi_3 \psi_4. \quad (5.247)$$

We note that the Dirac mass term has the form

$$\bar{\psi} \psi = \psi_3^\dagger \psi_1 + \psi_4^\dagger \psi_2 + \psi_1^\dagger \psi_3 + \psi_2^\dagger \psi_4 \quad (5.248)$$

□

Note that under a $U(1)$ transformation $\psi \rightarrow e^{-i\alpha} \psi$ the Dirac mass term is invariant, while the Weyl mass terms transform as

$$\overline{\psi_L^c} \psi_L \rightarrow e^{-2i\alpha} \overline{\psi_L^c} \psi_L \quad \overline{\psi_R^c} \psi_R \rightarrow e^{-2i\alpha} \overline{\psi_R^c} \psi_R \quad (5.249)$$

with their Hermitian conjugates transforming as

$$\overline{\psi_L} \psi_L^c \rightarrow e^{+2i\alpha} \overline{\psi_L} \psi_L^c \quad \overline{\psi_R} \psi_R^c \rightarrow e^{+2i\alpha} \overline{\psi_R} \psi_R^c \quad (5.250)$$

For a Majorana fermion we have

$$\psi_L^c = \psi_R \quad \text{and} \quad \psi_R^c = \psi_L$$

so for a Majorana fermion the Dirac mass term is just a special case of (5.243) with $m_D = m_L = m_R$.

6 Weak interactions

Since the electromagnetic interaction is already familiar from undergraduate studies, we first introduce the weak interactions which led to the discovery of the neutrino.

6.1 Beta decay and the neutrino

In 1911, Lise Meitner and Otto Hahn measured the energy spectrum of electrons emitted in a β decay, and found a wide distribution. This was unexpected as it seemed to contradict energy and momentum conservation. Imagine a nucleus N at rest decaying as

$$N \rightarrow N' + e^- \quad (6.1)$$

By momentum conservation, the daughter nucleus and the electron would have opposite momenta $\vec{p}$ and $-\vec{p}$, and energy conservation implies

$$m_N = \sqrt{m_e^2 + \vec{p}^2} + \sqrt{m_{N'}^2 + \vec{p}^2} \quad (6.2)$$

which has a unique solution for the magnitude p . Therefore the energy $\sqrt{m_e^2 + \vec{p}^2}$ of the emitted electron should be sharply defined. Instead, the distribution of the electron kinetic energy looks like this (taken from a later historical measurement of the decay of ^{210}Bi):

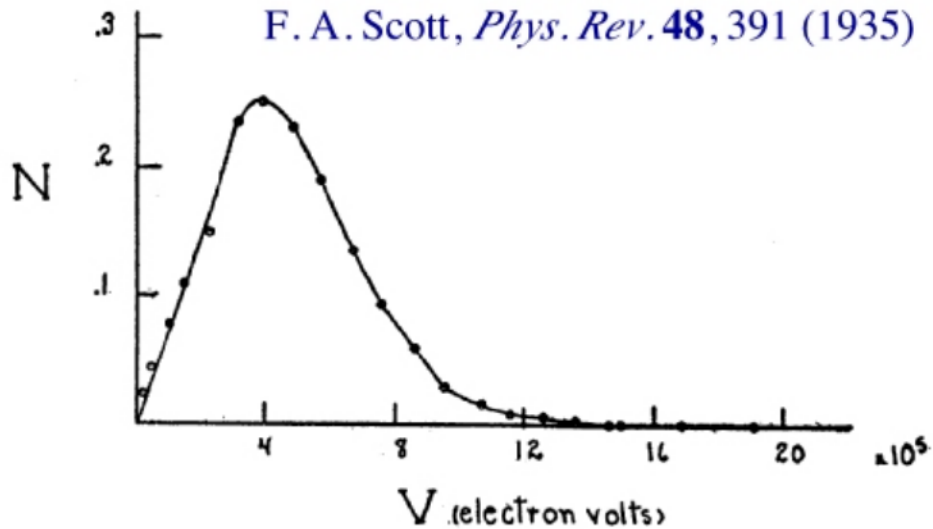


FIG. 5. Energy distribution curve of the beta-rays.

The spectrum is obviously very wide. There is a maximum value for the electron kinetic energy, but otherwise it has a continuous distribution. The explanation for this was provided by Fermi in 1933, who developed an earlier suggestion by Pauli that *there should be a third decay product*. Fermi called this new particle the *neutrino* (Italian expression for a small neutral particle). The mass of this particle can be determined from the way the distribution approaches zero at the maximum energy, i.e. from the behaviour of the spectrum at its endpoint; at present the direct searches for the neutrino mass in beta decay put an upper limit of about 2.2 eV, but there are several proposed experiments planning to improve upon this.

The particle hypothesised by Fermi is today called the anti electron neutrino, and the full equation for the beta decay takes the form:

$$N \rightarrow N' + e^- + \bar{\nu}_e \quad (6.3)$$

The simplest assumption at the time was that this was a zero mass, neutral fermionic particle. The fact that the spin of the neutrino was half-integer followed from considering the spins of the initial and daughter nucleus, and imposing angular momentum conservation. The simplest assumption consistent with all empirical data was that this new particle has spin 1/2, and that it does not participate in the electromagnetic and strong interactions. For a free neutron, the decay

$$n \rightarrow p^+ + e^- + \bar{\nu}_e \quad (6.4)$$

has a lifetime of around 15 minutes, from which one can derive the strength of the interaction responsible for the process. Such a long lifetime as opposed to typical electromagnetic transitions or strong interaction processes points to a very weak coupling, hence the name “weak interaction”.

6.2 Parity violation

6.2.1 Parity symmetry

One of the greatest surprises of modern physics is the discovery that *Nature distinguishes right and left handed objects*, or putting otherwise, a physical process does not necessarily occur with the same probability as its mirror image. In any given reference system, parity corresponds to the spatial reflection

$$\vec{x} \rightarrow -\vec{x} \quad (6.5)$$

The corresponding element of the Lorentz group is

$$P = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \quad (6.6)$$

and (provided it is a symmetry) it is represented on Hilbert space by a unitary operator $\mathcal{P}$. Spatial reflection reverses momentum, but keeps angular momentum, so a particle state with momentum $\vec{p}$ and spin z -projection σ satisfies

$$\mathcal{P}|\vec{p}\rangle_\sigma = \eta|-\vec{p}\rangle_\sigma \quad (6.7)$$

where $\eta = \pm 1$ is called the internal parity (in principle it could depend on σ , but using Lorentz invariance it can be shown to be independent of both $\vec{p}$ and σ). Under the composition rules of group representations, the transformation of multi-particle states is then

$$\mathcal{P}|\vec{p}_1, \dots, \vec{p}_n\rangle_{\sigma_1 \dots \sigma_n} = \eta_1 \dots \eta_n |-\vec{p}_1, \dots, -\vec{p}_n\rangle_{\sigma_1 \dots \sigma_n} \quad (6.8)$$

Now let us assume the initial state is an eigenstate of parity i.e.

$$\mathcal{P}|\Psi(0)\rangle = \pm|\Psi(0)\rangle \quad (6.9)$$

If the dynamics is parity invariant, i.e. the Hamiltonian H commutes with $\mathcal{P}$

$$[H, \mathcal{P}] = 0 \quad (6.10)$$

the time-evolved state

$$|\Psi(t)\rangle = e^{-iHt}|\Psi(0)\rangle \quad (6.11)$$

has the same parity at every instant t :

$$\mathcal{P}|\Psi(t)\rangle = \pm|\Psi(t)\rangle \quad (6.12)$$

As a result, parity even/odd states evolve into parity even/odd states, respectively. It had long been assumed that parity is conserved in Nature; electromagnetic and strong interactions indeed conserve parity.

For some particles, η must be chosen as a matter of convention. It turns out to be enough to choose the internal parities of electrons, protons and neutrons: the convention is $\eta = +1$ for all of them. The rest follows from considering electromagnetic and strong interaction processes. For example, to determine the parity of π^- one can take the strong interaction process

$$\pi^- + D \rightarrow 2n^0 \quad (6.13)$$

where D is the deuteron. It is known that the deuteron has orbital angular momentum $L = 0$ and total spin $S_D = 1$. The parity of the orbital wave function is $(-1)^L = 1$ while the spin wave function is a triplet made up of two $1/2$ spins, which is also symmetric and therefore even. As a result the deuteron's total parity is

$$\eta_D = \eta_p \eta_n (-1)^L (+1)_{\text{spin}} = +1 \quad (6.14)$$

The two neutrons in the outgoing state have total spin $S = 1$ due to spin conservation (which is valid because of the reaction being essentially non-relativistic), i.e. their spin wave function is symmetric. However, Pauli principle implies that this is only possible if the spatial wave-function is odd i.e. $(-1)^L = -1$. So the total parity of the outgoing two-neutron state is odd

$$\eta_{2n} = -1 \quad (6.15)$$

From the conservation of parity

$$\eta_{\pi^-} \cdot \eta_D = \eta_{2n} \quad (6.16)$$

and so

$$\eta_{\pi^-} = -1 \quad (6.17)$$

In fact, isospin symmetry (cf. Section [7.1](#)) then implies that all the pions are odd under parity i.e. they all have $\eta = -1$. Similarly, all kaons have $\eta = -1$, a result consistent with flavour $SU(3)$ symmetry discussed in Section [7.3](#).

Another interesting case is that of the photon. Since it is possible to have atomic electromagnetic transitions emitting/absorbing a single photon that change the total orbital momentum L by one unit without changing the spin, the photon must be odd under parity:

$$\eta_\gamma = -1 \quad (6.18)$$

6.2.2 The $\theta - \tau$ puzzle and the proposal of parity violation

As data on hadrons started piling up in the 1950s, a curious phenomenon was observed by Dalitz and Fabri in 1953. It was found that there existed two particles with exactly the same mass, charge and other measurable physical parameters. However, one of them decayed as

$$\theta \rightarrow \pi^+ + \pi^0 \quad (6.19)$$

and the other one as

$$\tau \rightarrow \pi^+ + \pi^0 + \pi^0 \quad (6.20)$$

The two end states have opposite parity, so if parity was conserved these two particles had to be different. However, all their other physical properties were found to be the same. It was also clear from their lifetimes that the decay proceeded via weak interactions.

In 1956 Lee and Yang suggested that *the two particles were identical to each other*; in fact, both of them was the same as the K^+ meson. However this implies that weak interaction cannot conserve parity; Lee and Yang therefore suggested an experiment to test this directly in radioactive β decay, which was performed in 1957 by Wu.

6.2.3 The Wu experiment and neutrino chirality

The experiment performed by Wu was based on the decay sequence of cobalt-60:



The result of the β -decay is an excited state of the nickel nucleus, which then decays by emitting two gamma rays. Cooling the cobalt sample to near absolute zero and applying a strong magnetic field the nucleus spins can be aligned. The experiment is shown in Figure 6.1¹¹. The presence of γ photons is important as a cross-check since they are produced by electromagnetic interaction which preserves parity. Wu's experiment compared the distribution of gamma and electron emissions with the nuclear spins in opposite orientations. If the electrons were always found to be emitted in the same direction and in the same proportion as the gamma rays that would mean that weak interactions conserved parity just as electromagnetism. A difference in the direction of decays would establish that weak interaction violated parity.

Wu found that the γ rays were emitted asymmetrically, 60% in the vertical while 40% in the horizontal direction. This establishes that the cobalt sample is spin polarised, and it is invariant under switching the direction of current in the solenoid which flips the spin direction; this is due to the parity invariant nature of electromagnetic interactions. On the other hand, switching the direction of the nuclear spin flips the distribution of electrons which are always emitted preferentially opposite to the nuclear spin, with a 70%-30% asymmetry.

Figure 6.2 illustrates that parity symmetry would imply that the electrons should have no preference to be emitted either to the half space in the direction or the half space opposite to the spin direction. This is a simple consequence of the electron momenta changing sign with parity, in contrast to the nucleus spin which is invariant under spatial reflection. *So the fact that electrons are preferentially emitted opposite to the nuclear spin violates parity!*

Lee and Yang showed that the violation observed by Wu is *maximal*, and it is connected to *neutrino chirality*. Those times neutrinos were believed to have zero mass. For a zero mass particle, the projection of its spin on the direction of its momentum (called *helicity*) is Lorentz invariant. For neutrinos it means that they could come in two *chiral* versions, called Weyl fermions (cf. Section 5.7). However, it turns out that anti-neutrinos emitted in beta decay are all right-handed i.e. they are spinning in the direction of their momentum, corresponding to *negative chirality* for the Weyl fermion.

¹¹Drawing from Wikipedia By Pen88, with English translation by Stigmatella aurantiaca - Derived from Wu-Experiment_wikipedia by Pen88 on the German Wikipedia, substituting the German legends with English legends., CC BY-SA 3.0, <https://commons.wikimedia.org/w/index.php?curid=29895892>

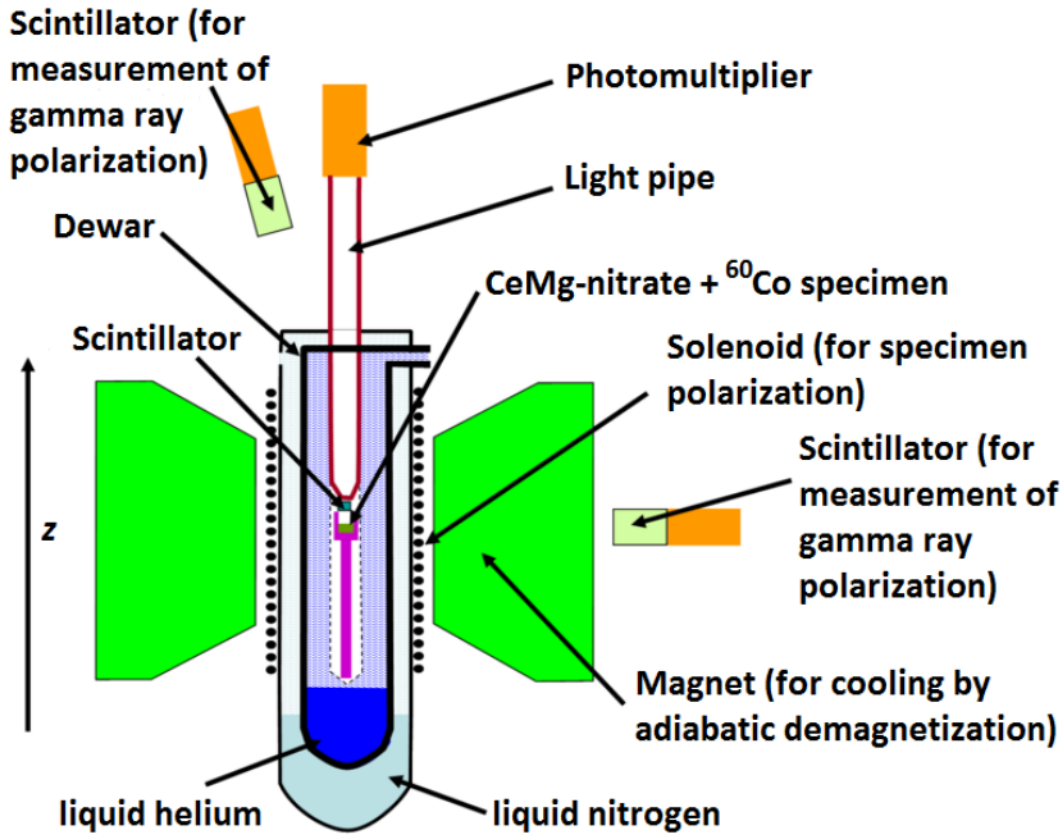


Figure 6.1: The Wu experiment (picture from Wikimedia Commons, available at [https://en.wikipedia.org/wiki/File:Wu-Experiment_\(English\).png](https://en.wikipedia.org/wiki/File:Wu-Experiment_(English).png))

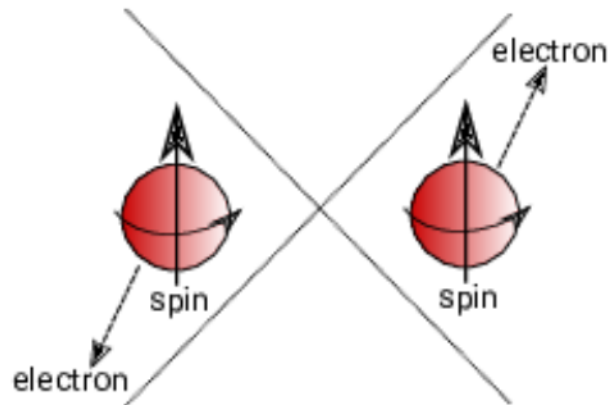


Figure 6.2: The left hand side shows an electron emitted in the half space opposite to the spin. On the right hand side the situation is the parity image of the left one with the centre of transformation being the intersection point of the dotted lines. Note that the direction of rotation, and therefore of the spin, is the same, while the electron is now emitted in the other half space. If weak interaction was parity invariant, the probability for both processes would be the same, therefore the total probability of electrons emitted in both half spaces should agree.

This fact alone violates parity, since under spatial reflection a right-handed anti-neutrino would turn into a left-handed one (note that parity P reverses momentum, while leaving spin invariant). Therefore the reflected version of the beta decay simply does not exist: this is the reason why parity violation is maximal (a non-maximal case would be when the reflected process exists, but proceeds with a different probability). Conservation of momentum and angular momentum results in this asymmetry to be reflected in the electron distribution, and the asymmetry Lee and Yang obtained from this model matched the one observed by Wu. For the description of violation of parity conservation in weak interactions Lee and Yang were awarded the Nobel Prize in physics in 1957.

Direct evidence for neutrino chirality was obtained by Goldhaber, Grodzins and Sunyar in 1957. They considered the radioactive decay

$$^{152}\text{Eu} + e^- \longrightarrow ^{152m}\text{Sm} + \nu_e \longrightarrow ^{152}\text{Sm} + \gamma + \nu_e \quad (6.22)$$

which proceeds by capturing a K shell electron. Now it turns out that angular momentum conservation makes the helicities of the outgoing neutrino and gamma photon correlated. Helicity is defined as the projection of the spin on the direction of the momentum of the particle:

$$\chi = \frac{\vec{S} \cdot \vec{p}}{|\vec{p}|} \quad (6.23)$$

Recalling the result (5.187) and its discussion, for a positive chirality Weyl fermion the particle has positive helicity ($\chi = +1/2$) while the antiparticle has negative helicity ($\chi = -1/2$), with the opposite assignment for negative chirality. From nuclear physics it is known that the ^{152}Eu and ^{152}Sm nuclei have spin 0, while the ^{152m}Sm nucleus has spin 1. As a result, there are the following possibilities for spin projections along a reference directions:

$$\begin{array}{ccccccc} ^{152}\text{Eu} & + & e^- & \longrightarrow & ^{152m}\text{Sm} & + & \nu_e & \longrightarrow & ^{152}\text{Sm} & + & \gamma & + & \nu_e \\ 0 & & +1/2 & & +1 & & -1/2 & & 0 & & +1 & & -1/2 \\ 0 & & -1/2 & & -1 & & +1/2 & & 0 & & -1 & & +1/2 \end{array} \quad (6.24)$$

which shows that *the direction of the photon and neutrino spins are always antiparallel to each other*. If one can select the cases where the neutrino and photon are emitted back-to-back, the helicity of the photon would be the same as the helicity of the neutrino. Selection of these cases is possible due to the fact that the Sm nucleus suffers a recoil from neutrino emission, and subsequently the photon suffers a Doppler shift from the moving nucleus, in addition to losing some of its momentum due to the nuclear recoil. Now the energy of the neutrino is 840 keV, while the photon energy is 961 keV in the laboratory frame; when the photon and neutrino are emitted back-to-back, the Doppler shift and the recoil effect compensate each other sufficiently well so that the outgoing photon can be absorbed by a separate ^{152}Sm scatterer. Selecting those events in which photon reabsorption happens, one can make sure the neutrino was emitted in the direction opposite to the photon.

The photon helicity corresponds to light polarisation and can be measured using Compton scattering on polarised electrons, which were provided by magnetised Fe surrounding the Eu source. The photons themselves were detected using a NaI scintillation counter; the experimental setup is shown in Fig 6.3

The photon polarisation measured this way is in one-to-one correspondence with the neutrino helicity, and the result for the latter was:

$$\chi_\nu = -1.0 \pm 0.3 \quad (6.25)$$

concluding that neutrinos are indeed left-handed.

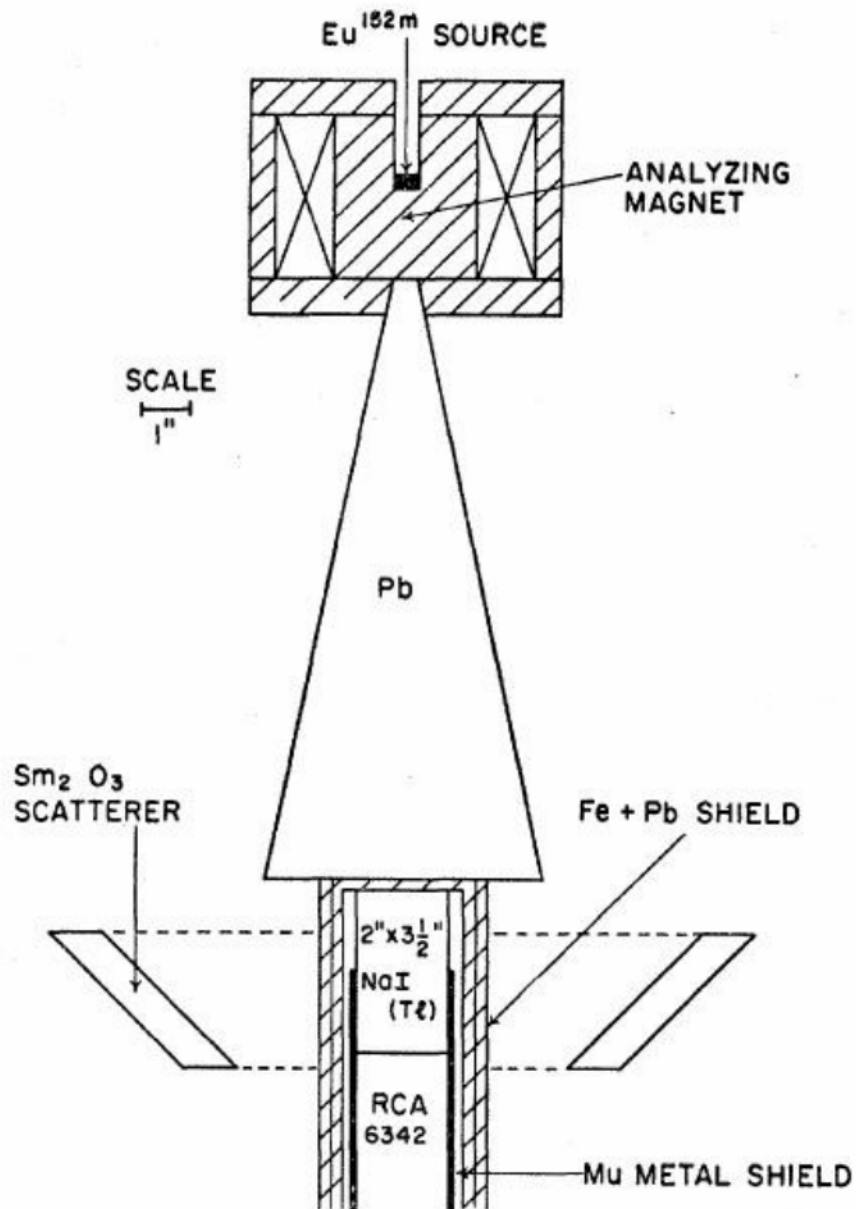


Figure 6.3: The setup of the Goldhaber-Grodzins-Sunyar experiment.

From: M. Goldhaber, L. Grodzins and A. W. Sunyar, Phys. Rev. 109 (1958) 1015.

6.3 CP violation

6.3.1 CP symmetry

Once parity P is violated, it is an interesting question whether some other combination is preserved. Indeed, the weak interaction model built upon four-Fermi interactions that is discussed in Subsection [11.2](#) preserves the combination CP where C is charge conjugation that swaps particles with their anti-particles while leaving their momenta and spin invariant:

$$\mathcal{C}|\vec{p}, n\rangle_\sigma = \xi_n|\vec{p}, n^c\rangle_\sigma \quad (6.26)$$

where n denotes the species of the particle and n^c is the corresponding anti-particle species. As a result, antimatter has opposite chirality compared to matter, and an “antimatter Wu experiment” would result in just the opposite result: the positrons produced by the anti-cobalt nuclei would fly preferentially in the direction of the nuclear spin.

CP symmetry works even for chiral neutrinos, since the CP image of a left-handed neutrino is a right-handed antineutrino, exactly as it should be for a Weyl fermion of negative chirality. Indeed, after the discovery of parity violation CP symmetry was still consistent with all observed weak (and also electromagnetic and strong) processes, until a famous experiment performed by Cronin and Fitch in 1964.

6.3.2 The Cronin-Fitch experiment

The first observation of CP violation happened for neutral kaons that are charge neutral. However, they have a charge called strangeness that is conserved under strong interactions, and therefore *the particle and anti-particle are different*. Denoting them K^0 and $\bar{K}^0$, these two states are *related by CP symmetry*:

$$|\bar{K}^0\rangle = \mathcal{CP}|K^0\rangle \quad (6.27)$$

Now if CP is a symmetry of the Hamiltonian, it is expected that energy eigenstates are also eigenstates of CP . Indeed, there exist two neutral kaon states in Nature, one of them has a short lifetime $\tau_S \approx 9 \cdot 10^{-11}$ s, while the other is noted for its “longevity”: $\tau_L \approx 5.2 \cdot 10^{-8}$ s. They have the following main decay routes:

$$\begin{aligned} K_S^0 &\longrightarrow \begin{cases} \pi^+ + \pi^- \\ \pi^0 + \pi^0 \end{cases} \\ K_L^0 &\longrightarrow \begin{cases} \pi^\pm + e^\mp + \bar{\nu}_e/\nu_e \\ \pi^\pm + \mu^\mp + \bar{\nu}_\mu/\nu_\mu \\ \pi^+ + \pi^- + \pi^0 \\ \pi^0 + \pi^0 + \pi^0 \end{cases} \end{aligned} \quad (6.28)$$

The pionic decay routes imply that K_S^0 has $CP = +1$, while K_L^0 has $CP = -1$, so one can write them as

$$\begin{aligned} |K_S^0\rangle &= \frac{1}{\sqrt{2}} (|K^0\rangle + |\bar{K}^0\rangle) \\ |K_L^0\rangle &= \frac{1}{\sqrt{2}} (|K^0\rangle - |\bar{K}^0\rangle) \end{aligned} \quad (6.29)$$

Cronin and Fitch prepared a beam of hadrons by colliding 30 GeV protons from the AGS (Alternating Gradient Synchrotron, Brookhaven) with a fixed target, obtaining a beam of neutral particles at angle 30° removing charged particles with a sweeping magnet and led collimators. This leaves a beam consisting of neutral kaons: other neutral particles, such as π^0 have a very short lifetime and decay almost immediately, and not even the Λ^0 would make it through the beamline with its lifetime of about $2.6 \cdot 10^{-10}$ s and decay products $n + \pi^0$ or $p + \pi^-$. After propagating around 17.6 meters (57 feet), the beam consists purely of K_L^0 , the K_S^0 component having long decayed even after accounting for relativistic time dilation effects. To their surprise, they observed that *for one in every 500 decays, the K_L^0 ended up in a two pion end state, clearly violating CP symmetry.*

Exercise 6.1. Compute the fraction of K_S^0 remaining in the beam using the lifetime τ_S and assuming a speed $v = 0.997c$.

Solution. At the speed given above the time dilation factor is

$$\gamma = (1 - v^2/c^2)^{-1/2} \approx 12.92 \quad (6.30)$$

corresponding to kaons of energy

$$E = m_K c^2 \gamma \approx 6 \text{ GeV} \quad (6.31)$$

The time dependence of the number of K_S^0 is then given by

$$N_S(t) = N_S(0) e^{-t/\gamma\tau_S} \quad (6.32)$$

which translates into a spatial dependence

$$N_S(x) = N_S(0) e^{-x/l_S} \quad (6.33)$$

with the attenuation length

$$l_S = v\gamma\tau_S \approx 0.35 \text{ m} \quad (6.34)$$

So, after propagating 17.6 meters, the number of K_S^0 is depleted by a factor

$$e^{-17.6/0.35} \approx 1.4 \cdot 10^{-22} \quad (6.35)$$

therefore there are no K_S^0 at the end of the beamline. For K_L^0 the attenuation length is

$$l_L = v\gamma\tau_L \approx 200 \text{ m} \quad (6.36)$$

which is an order of magnitude longer than the beamline. These numbers show that the observed 1/500 branching ratio of two pion decays cannot be explained by a residual fraction of K_S^0 present at the end of the beamline¹² □

¹²Here we remark that *CP* violation was so unexpected that while the measurement result themselves were not questioned as the experiment itself was considered accurate and professionally done, alternative explanations were sought for the experimental results for quite some time. An important confirmation for the interpretation of the experimental results as *CP* violation came from the observation of quantum interference between the 2-pion decays of K_S^0 and K_L^0 , using the fact that K_S^0 can be regenerated from K_L^0 by scattering on a material (Cronin and Fitch used beryllium for this purpose). Such interference can also be used to measure the tiny difference between the masses of K_S^0 and K_L^0 , resulting in a mass difference 10^{-15} times smaller than the masses themselves. The interested reader is invited to consult e.g. M.S. Sozzi: *Discrete Symmetries and CP Violation: From Experiment to Theory*, Oxford University press, 2008, pp. 244-249.

There is also a way to see CP violation *directly* in the decays of the K_L^0 meson by comparing the probabilities for the decays

$$K_L^0 \longrightarrow \begin{cases} \pi^+ + e^- + \bar{\nu}_e \\ \pi^- + e^+ + \nu_e \end{cases} \quad (6.37)$$

CP symmetry transforms an end state with an electron into one with a positron and vice versa; it also changes the signs of momenta, but after summing up the end states the total probabilities of the two decays should be the same if CP holds. However, the experimental result (measured in 1999) is that the decay to the positronic state is more frequent by an amount of 3.3×10^{-3} . This tiny difference, however, has a huge implication: *matter and antimatter are really different*, since they are distinguished e.g. by K_L^0 decays!

By now, CP violation was observed in the decays of other mesons as well; a more detailed discussion of the mechanism of CP violation can be found in Section 11.4. For the original discovery of CP violation, Cronin and Fitch received the Nobel prize in 1980.

6.4 CPT, matter and antimatter

There is another important discrete symmetry corresponding to the direction of time, which is called time reversal and denoted by T . It is represented by an operator $\mathcal{T}$ on the Hilbert space which acts on particle states as follows:

$$\mathcal{T}|\vec{p}, n\rangle_\sigma = \zeta_n(-1)^{s-\sigma} |-\vec{p}, n\rangle_{-\sigma} \quad (6.38)$$

where ζ_n is a phase factor, s is the spin of the particle, and σ denotes its projection on some chosen direction in which we label the spin states in our basis of the Hilbert space. Part of the above formula is obvious since time reversal changes the signs of both the momentum and angular momentum (therefore also of the spin), while the factor $(-1)^{s-\sigma}$ can be obtained by considering constraints resulting from representation of spatial rotations on the state.

An interesting property of the operator $\mathcal{T}$ is that while it keeps the scalar product invariant (as any symmetry should), it is not a proper linear operator. Indeed, let us suppose that the Hamiltonian is time reversal invariant, i.e. it commutes with $\mathcal{T}$. Then if $\mathcal{T}$ were a linear operator it would map the time evolution operator onto itself:

$$\mathcal{T}e^{-iHt}\mathcal{T}^{-1} = e^{-iHt} \quad (6.39)$$

which is not correct since it should change the sign of t . However Wigner's theorem (cf. Thm. 3.1) allows for symmetries represented by anti-unitary operators satisfying

$$\mathcal{T}(\alpha|\Psi_1\rangle + \beta|\Psi_2\rangle) = \alpha^*\mathcal{T}|\Psi_1\rangle + \beta^*\mathcal{T}|\Psi_2\rangle \quad (6.40)$$

In this case

$$\mathcal{T}i = -i\mathcal{T} \quad (6.41)$$

so

$$\mathcal{T}e^{-iHt}\mathcal{T}^{-1} = e^{iHt} \quad (6.42)$$

which is the correct relation. Note that due to anti-linearity of $\mathcal{T}$ the phase ζ_n has no physical meaning, and so there is no “time parity”. This can be seen by redefining the phase of the state by introducing new one-particle states as

$$|\vec{p}, n\rangle'_\sigma = \lambda_n|\vec{p}, n\rangle_\sigma \quad (6.43)$$

Now

$$\mathcal{T}|\vec{p}, n\rangle'_\sigma = \lambda_n^* \mathcal{T}|\vec{p}, n\rangle_\sigma = \lambda_n^* \zeta_n (-1)^{s-\sigma} |-\vec{p}, n\rangle_{-\sigma} = \lambda_n^* \lambda_n^{-1} \zeta_n (-1)^{s-\sigma} |-\vec{p}, n\rangle'_{-\sigma} \quad (6.44)$$

So in the new basis the time reversal phase for the particle species n is

$$\zeta'_n = \lambda_n^* \lambda_n^{-1} \zeta_n \quad (6.45)$$

which can be set to unity by choosing the phase redefinition as

$$\lambda_n = \zeta_n^{1/2} \quad (6.46)$$

(the choice of the sign of the square root does not matter).

It turns out that the following so-called *CPT* theorem holds (cf. also Section [10.9](#)):

Theorem 6.2. (*CPT*) *In any local relativistic quantum theory with a self-adjoint Hamiltonian, CPT is an exact symmetry.*

For the notion of locality, which is necessary for consistency between special relativity and quantum theory, cf. Section [8.1](#). The proof of this theorem is quite lengthy and involves a deep understanding of representation theory of the Poincaré group. Since according to our present state of knowledge, the fundamental principles involved in the assumptions of the theorem are known to hold to a very high degree of accuracy, *CPT is expected to be a symmetry of all particle physics phenomena.*

A fundamental consequence of *CPT* symmetry is that *the mass of a particle is exactly equal to that of the corresponding antiparticle.* This is easy to see from

$$\mathcal{CPT}|\vec{p}, n\rangle_\sigma = \eta_n \xi_n \zeta_n (-1)^{s-\sigma} |\vec{p}, n^c\rangle_{-\sigma} \quad (6.47)$$

Since $\mathcal{CPT}$ commutes with the Hamiltonian, the energies of the states $|\vec{p}, n\rangle_\sigma$ and $|\vec{p}, n^c\rangle_{-\sigma}$ are equal for all values of $\vec{p}$ which implies that the masses of the particle and the anti-particle are equal. *CPT* symmetry was tested to a very high accuracy in the neutral kaon system by measuring the mass difference between particle and anti-particle with the result

$$\frac{|m_K - m_{\bar{K}}|}{m_K} \lesssim 10^{-18} \quad (6.48)$$

It has also been tested by measuring the proton-antiproton mass and charge difference, which has found no difference to a relative accuracy of a few times 10^{-9} . The CERN ASACUSA experiment measured the antiproton-electron mass ratio

$$\frac{m_{\bar{p}}}{m_e} = 1836.1526736(23) \quad (6.49)$$

using antiprotonic helium atoms, and the result agrees with the proton-electron mass ratio which is known to a similar precision:

$$\frac{m_p}{m_e} = 1836.15267389(17) \quad (6.50)$$

(the numbers of parentheses give the error as the magnitude of one standard deviation in the last two digits).

Another interesting consequence is that since CP is violated, so is T , so the fundamental physical processes are not time reversal invariant. Note, however, that this has no direct relation to the irreversibility formulated in the second law of thermodynamics.

Furthermore, CP violation implies that matter and antimatter are truly different, albeit the violation is small. The difference between matter and antimatter is related to the cosmological matter-antimatter asymmetry, which is the observation that *our Universe seems to consist exclusively of ordinary matter*, with antimatter arising only temporarily in high energy processes. Although it is in principle possible that the Universe is simply “born that way”, it is more plausible that the asymmetry is generated in a hypothetical process called *baryogenesis* in the high energy environment right after the Big Bang. Andrei Sakharov, renown Soviet physicist (who was a political dissident and winner of the Nobel peace prize in 1975) showed that this is possible under three conditions:

- There must clearly exist *baryon number violating processes*, otherwise every reaction would produce equal quantities of matter and antimatter. The Standard Model does allow for such processes to proceed via a high-energy configuration called the “sphaleron”, and most models beyond the Standard Model involve explicit baryon number violation.
- Baryogenesis must take place *out of thermal equilibrium*; it is not that uncommon for processes in the early Universe to fall out of equilibrium due to the rapid expansion of the Universe during early times. This is necessary so that reactions generating baryon number do not cancel with the reverse processes.
- It is *necessary to have CP violation*, since the whole process requires an asymmetry between matter and antimatter.

The discovery of CP violation indicates that a baryogenesis scenario is indeed possible. Although CP violation is small, it is the case that the asymmetry which must have existed or been generated at early times is around a part in one billion, making such a scenario not so far fetched. There are several proposed mechanisms, but as yet no experimentally validated consensus theory, leaving baryogenesis a fundamental open problem in today’s physics.

7 Strong interactions

7.1 Isospin symmetry

7.1.1 Isospin for nucleons

The proton and neutron masses are astonishingly close

$$m_p = 938.28 \text{ MeV} \quad m_n = 939.57 \text{ MeV} \quad (7.1)$$

On this basis Heisenberg proposed that the proton and neutron are actually two states of the same particle, the nucleon and they are no different under the strong interactions. The small mass difference can be due to the charge of the proton, reflecting energy stored in its electric field. However, this would suggest that the proton should be heavier, so *there must be some additional source for the mass difference*. In any case, one can consider an $SU(2)$ symmetry that acts on the flavour labels p and n , similarly to the usual spin components of a spin-1/2 particle. In the flavour basis

$$p = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad n = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad (7.2)$$

By the analogy to the case of the spin, this is called an *isospin symmetry*, denoted by $SU(2)_I$. An $SU(2)$ transformation is a 2×2 unitary matrix U with unit determinant

$$UU^\dagger = U^\dagger U = \mathbf{1}_{2 \times 2} \quad \det U = 1 \quad (7.3)$$

Exercise 7.1. Writing

$$U = e^{-iT} \quad (7.4)$$

show that

$$T^\dagger = T \quad \text{Tr } T = 0 \quad (7.5)$$

Solution.

$$U^\dagger U = e^{iT^\dagger} e^{-iT} = \mathbf{1} \quad (7.6)$$

i.e.

$$e^{iT^\dagger} = e^{iT} \quad (7.7)$$

so $T = T^\dagger$. In addition

$$\det U = \det e^{-iT} = e^{-i\text{Tr } T} = 1 \quad (7.8)$$

so

$$\text{Tr } T = 0 \quad (7.9)$$

□

Therefore T is a 2×2 traceless Hermitian matrix, and any such matrix depends on three real parameters, and can be written as a linear combination of the Pauli matrices

$$T = \sum_{k=1}^3 \epsilon_k \frac{\sigma_k}{2} \quad \sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad (7.10)$$

which satisfy the Lie algebra

$$\left[\frac{\sigma_k}{2}, \frac{\sigma_l}{2}\right] = i\epsilon_{klm} \frac{\sigma_m}{2} \quad (7.11)$$

A general group element can be parametrised as

$$U = e^{-i\epsilon_k \frac{\sigma_k}{2}} \quad (7.12)$$

while small transformations (close to the identity) can be written as

$$U = 1 - i\epsilon_k \frac{\sigma_k}{2} + O(\epsilon^2) \quad (7.13)$$

This is in fact a concrete realisation of $SU(2)$ as transformations acting in a two-dimensional Hilbert space. In an abstract (representation independent) notation, the infinitesimal isospin transformations are generated by three generators I_k , $k = 1, 2, 3$ satisfying the algebra

$$[I_k, I_l] = i\epsilon_{klm} I_m \quad (7.14)$$

The nucleon transforms as an $I = 1/2$ multiplet, and the two independent states correspond to $I_3 = \pm 1/2$

$$p = |1/2, +1/2\rangle \quad n = |1/2, -1/2\rangle \quad (7.15)$$

In this basis, the generators are represented by the Pauli matrices

$$I_k \rightarrow \frac{\sigma_k}{2} \quad (7.16)$$

It is assumed that isospin is a symmetry of the strong interactions, described by a hamiltonian H_s :

$$[H_s, I_k] = 0 \quad (7.17)$$

while the total Hamiltonian is

$$H = H_s + H_b \quad (7.18)$$

with H_b containing isospin breaking terms.

7.1.2 Representations of $SU(2)$ symmetry

In order to understand the consequences of the isospin symmetry, it is necessary to learn how an $SU(2)$ symmetry works. It is convenient to introduce operators

$$I_{\pm} = I_1 \pm iI_2 \quad (7.19)$$

which satisfy the commutation relations

$$\begin{aligned} [I_3, I_{\pm}] &= \pm I_{\pm} \\ [I_+, I_-] &= 2I_3 \end{aligned} \quad (7.20)$$

For any multiplet transforming under $SU(2)$ symmetry, one can introduce a basis of I_3 eigenstates

$$I_3|m\rangle = m|m\rangle \quad (7.21)$$

Exercise 7.2. Show that

$$I_{\pm}|m\rangle = C_{\pm}(m)|m \pm 1\rangle \quad (7.22)$$

with some coefficients $C_{\pm}(m)$.

Solution. The eigenvalue corresponding to $I_+|m\rangle$ can be computed by acting with I_3 and using the commutation relations

$$I_3 I_+ |m\rangle = (I_+ I_3 + I_+) |m\rangle = (m+1) I_+ |m\rangle \quad (7.23)$$

so

$$I_+ |m\rangle = C_+(m) |m+1\rangle \quad (7.24)$$

The eigenvalue corresponding to $I_-|m\rangle$ can be computed as

$$I_3 I_- |m\rangle = (I_- I_3 - I_-) |m\rangle = (m-1) I_- |m\rangle \quad (7.25)$$

so

$$I_- |m\rangle = C_-(m) |m-1\rangle \quad (7.26)$$

□

According to Wigner's theorem, *continuous symmetries must be represented by unitary transformations*, therefore the infinitesimal generators are Hermitian, leading to

$$I_k^\dagger = I_k \quad I_{\pm}^\dagger = I_{\mp} \quad (7.27)$$

The latter, in particular, leads to the relation

$$C_+(m) = \langle m+1 | I_+ |m\rangle = (\langle m+1 | I_-) |m\rangle = C_-(m+1)^* \quad (7.28)$$

Choosing the relative phases of the states $|m\rangle$ properly, one can ensure that all $C_+(m)$ and therefore also $C_-(m)$ are real and positive.

As a result of the Hermiticity of I_3 , the vectors $|m\rangle$ are mutually orthogonal and can be normalised as

$$\langle m | m' \rangle = \delta_{mm'} \quad (7.29)$$

This can be used to determine the $C_+(m)$ writing the length squared of $I_+|m\rangle$ as

$$\langle m | I_- I_+ |m\rangle = C_+(m)^2 \quad (7.30)$$

on the one hand, and as

$$\langle m | I_- I_+ |m\rangle = \langle m | (-2I_3 + I_+ I_-) |m\rangle = -2m + C_-(m)^2 = -2m + C_+(m-1)^2 \quad (7.31)$$

so the following recursion relation holds:

$$C_+(m)^2 = C_+(m-1)^2 - 2m \quad (7.32)$$

Now this must stop at some maximum m , otherwise at some point $C_+(m)^2$ turns negative, which is impossible. Therefore there must be a maximum for m which we denote as j , such that

$$I_+ |j\rangle = 0 \quad (7.33)$$

i.e.

$$C_+(j) = 0 \quad (7.34)$$

Starting from this it is easy to compute

$$C_+(m)^2 = (j - m)(j + m + 1) \quad (7.35)$$

and so

$$C_-(m)^2 = (j + m)(j - m + 1) \quad (7.36)$$

On the other hand, the same argument can be repeated for I_- , leading to a minimum for m as well. Using the expression for $C_-(m)$ this corresponds to $m = -j$. However, the difference between j and $-j$ must be an integer, leading to

$$j \in \frac{1}{2}\mathbb{N} \quad \text{i.e.} \quad j = 0, 1/2, 1, 3/2, \dots \quad (7.37)$$

In conclusion, $SU(2)$ multiplets are given by

$$\begin{aligned} |j, m\rangle & : \quad m = -j, -j + 1, \dots, j \\ I_3 |j, m\rangle & = m |j, m\rangle \\ I_{\pm} |j, m\rangle & = \sqrt{(j \mp m)(j \pm m + 1)} |j, m \pm 1\rangle \end{aligned} \quad (7.38)$$

In group theory jargon these are called the *irreducible representations* of $SU(2)$. They correspond to $2j + 1$ dimensional vector spaces on which an action of the generators and therefore the group is given, and are irreducible in the sense that there is no proper subspace which is closed under the group action.

One can also compute the so-called quadratic Casimir

$$I^2 = I_1^2 + I_2^2 + I_3^2 = I_3^2 + \frac{1}{2}I_+I_- + \frac{1}{2}I_-I_+ \quad (7.39)$$

It is easy to verify that

$$[I^2, I_k] = 0 \quad (7.40)$$

so, in particular I^2 can be diagonalised simultaneously with I_3 . Since it commutes with $I_{\pm}$, its eigenvalue is the same for all $|m\rangle$ in the same multiplet. Therefore

$$I^2 |j, m\rangle = C(j) |j, m\rangle \quad (7.41)$$

and it can be evaluated by considering $m = j$

$$I^2 |j, m\rangle = \left(I_3^2 + \frac{1}{2}I_+I_- + \frac{1}{2}I_-I_+ \right) |j, m\rangle = \left(j^2 + \frac{1}{2}\sqrt{2j}\sqrt{2j} + 0 \right) |j, m\rangle = j(j + 1) |j, m\rangle \quad (7.42)$$

so

$$C(j) = j(j + 1) \quad (7.43)$$

The value of the quadratic Casimir specifies j and therefore characterises the multiplet uniquely.

7.1.3 $SU(2)$ Clebsh-Gordan coefficients and a simple application of isospin symmetry

Let us consider now states composed of two particles with isospins j_1 and j_2 . The question is what multiplets of $SU(2)_I$ occur in this two-particle state? This is answered by the composition rule

$$j_1 \otimes j_2 = |j_1 - j_2| \oplus |j_1 - j_2| + 1 \oplus \cdots \oplus j_1 + j_2 \quad (7.44)$$

which is known from the quantum theory of angular momentum. To illustrate this, consider the case $j_1 = 1/2$ and $j_2 = 1/2$. The basis of the full isospin space is

$$\begin{aligned} |++\rangle &= |1/2, +1/2\rangle \otimes |1/2, +1/2\rangle \\ |+-\rangle &= |1/2, +1/2\rangle \otimes |1/2, -1/2\rangle \\ |-+\rangle &= |1/2, -1/2\rangle \otimes |1/2, +1/2\rangle \\ |--\rangle &= |1/2, -1/2\rangle \otimes |1/2, -1/2\rangle \end{aligned} \quad (7.45)$$

On these states an $SU(2)$ transformation acts as product

$$e^{-i\epsilon_k \frac{\sigma_k}{2}} \otimes e^{-i\epsilon_k \frac{\sigma_k}{2}} \quad (7.46)$$

i.e. both spins must be transformed at the same time. Expanding in ϵ

$$e^{-i\epsilon_k \frac{\sigma_k}{2}} \otimes e^{-i\epsilon_k \frac{\sigma_k}{2}} = 1 - i\epsilon_k \left(\frac{\sigma_k}{2} \otimes \mathbf{1} + \mathbf{1} \otimes \frac{\sigma_k}{2} \right) + O(\epsilon^2) \quad (7.47)$$

so generators act as

$$\frac{\sigma_k}{2} \otimes \mathbf{1} + \mathbf{1} \otimes \frac{\sigma_k}{2} \quad (7.48)$$

This is valid for the product of any two representations, i.e. the action of the generators I_k on a two-fold product is

$$I_k \otimes \mathbf{1} + \mathbf{1} \otimes I_k \quad (7.49)$$

where $\mathbf{1}$ is the unit, i.e. one must act on both components and sum the result. For the product of n multiplets one gets the following action of the generators

$$I_k \otimes \mathbf{1} \otimes \cdots \otimes \mathbf{1} + \mathbf{1} \otimes I_k \otimes \cdots \otimes \mathbf{1} + \cdots + \mathbf{1} \otimes \cdots \otimes \mathbf{1} \otimes I_k \quad (7.50)$$

where in every term the generator acts on precisely one factor in the product and then these are summed over.

Applying these results to our case leads to

$$\begin{aligned} I_3 |++\rangle &= +1 \cdot |++\rangle \\ I_3 |+-\rangle &= 0 \cdot |+-\rangle \\ I_3 |-+\rangle &= 0 \cdot |-+\rangle \\ I_3 |--\rangle &= -1 \cdot |--\rangle \end{aligned} \quad (7.51)$$

It is clear that this space contains a $j = 1$ and a $j = 0$ representation, but these must be disentangled. It is easy to do this by taking the top vector $|++\rangle$ and generating the components of the triplet by action of I_- :

$$\begin{aligned} |1, +1\rangle &= |++\rangle \\ I_- |1, +1\rangle &= \sqrt{2} |1, 0\rangle = |+-\rangle + |-+\rangle \\ I_-^2 |1, +1\rangle &= 2 |1, -1\rangle = 2 |--\rangle \end{aligned} \quad (7.52)$$

Now the singlet can be found as the state orthogonal to $|1, 0\rangle$:

$$|0, 0\rangle = \frac{1}{2} (|+-\rangle - |-+\rangle) \quad (7.53)$$

The end result is that

$$\begin{aligned} |1, +1\rangle &= |1/2, +1/2\rangle \otimes |1/2, +1/2\rangle \\ |1, 0\rangle &= \frac{1}{\sqrt{2}} |1/2, +1/2\rangle \otimes |1/2, -1/2\rangle + \frac{1}{\sqrt{2}} |1/2, -1/2\rangle \otimes |1/2, +1/2\rangle \\ |1, -1\rangle &= |1/2, -1/2\rangle \otimes |1/2, -1/2\rangle \\ |0, 0\rangle &= \frac{1}{\sqrt{2}} |1/2, +1/2\rangle \otimes |1/2, -1/2\rangle - \frac{1}{\sqrt{2}} |1/2, -1/2\rangle \otimes |1/2, +1/2\rangle \end{aligned} \quad (7.54)$$

The coefficients appearing in these composition rules are called *Clebsh-Gordan coefficients*. They are defined from the expansion

$$|j, m\rangle = \sum_{m_1, m_2} \begin{pmatrix} j_1 & j_2 & j \\ m_1 & m_2 & m \end{pmatrix} |j_1, m_1\rangle \otimes |j_2, m_2\rangle \quad (7.55)$$

Obviously they satisfy that

$$\begin{aligned} \begin{pmatrix} j_1 & j_2 & j \\ m_1 & m_2 & m \end{pmatrix} &= 0 \quad j \notin \{|j_1 - j_2|, |j_1 - j_2| + 1, \dots, j_1 + j_2\} \\ \text{or } m &\neq m_1 + m_2 \end{aligned} \quad (7.56)$$

Note that we have just computed the following non-zero Clebsh-Gordan coefficients

$$\begin{aligned} \begin{pmatrix} 1/2 & 1/2 & 1 \\ 1/2 & 1/2 & 1 \end{pmatrix} &= 1 & \begin{pmatrix} 1/2 & 1/2 & 1 \\ -1/2 & -1/2 & -1 \end{pmatrix} &= 1 \\ \begin{pmatrix} 1/2 & 1/2 & 1 \\ 1/2 & -1/2 & 0 \end{pmatrix} &= \frac{1}{\sqrt{2}} & \begin{pmatrix} 1/2 & 1/2 & 1 \\ -1/2 & 1/2 & 0 \end{pmatrix} &= \frac{1}{\sqrt{2}} \\ \begin{pmatrix} 1/2 & 1/2 & 0 \\ 1/2 & -1/2 & 0 \end{pmatrix} &= \frac{1}{\sqrt{2}} & \begin{pmatrix} 1/2 & 1/2 & 0 \\ -1/2 & 1/2 & 0 \end{pmatrix} &= -\frac{1}{\sqrt{2}} \end{aligned} \quad (7.57)$$

Exercise 7.3. Compute the Clebsh-Gordan coefficients for $\mathbf{1} \otimes \frac{1}{2}$

Solution. Applying the above procedure leads to the results

$$\begin{aligned} |3/2, +3/2\rangle &= |1, +1\rangle \otimes |1/2, +1/2\rangle \\ |3/2, +1/2\rangle &= \frac{1}{\sqrt{3}} |1, +1\rangle \otimes |1/2, -1/2\rangle + \sqrt{\frac{2}{3}} |1, 0\rangle \otimes |1/2, +1/2\rangle \\ |3/2, -1/2\rangle &= \frac{1}{\sqrt{3}} |1, -1\rangle \otimes |1/2, +1/2\rangle + \sqrt{\frac{2}{3}} |1, 0\rangle \otimes |1/2, -1/2\rangle \\ |3/2, -3/2\rangle &= |1, -1\rangle \otimes |1/2, -1/2\rangle \\ |1/2, +1/2\rangle &= \sqrt{\frac{2}{3}} |1, +1\rangle \otimes |1/2, -1/2\rangle - \frac{1}{\sqrt{3}} |1, 0\rangle \otimes |1/2, +1/2\rangle \\ |1/2, -1/2\rangle &= \sqrt{\frac{2}{3}} |1, -1\rangle \otimes |1/2, +1/2\rangle - \frac{1}{\sqrt{3}} |1, 0\rangle \otimes |1/2, -1/2\rangle \end{aligned} \quad (7.58)$$

from which the Clebsh-Gordan coefficients can be read off. $\square$

7.1.4 Yukawa model and the π meson

From the study of nuclei it is known that strong interactions only act at short distances. A characteristic distance can be taken as the size¹³ of the simplest nucleus, the deuteron which is about 2 fm, or the helium nucleus, which is about 1.7 fm. A short range interaction cannot be mediated by a massless particle like the photon, since the potential field satisfies the wave equation

$$\square V(x) = f(x) \quad (7.59)$$

where f is the source. The static potential of a point source then satisfies the Poisson equation

$$\Delta V(\vec{x}) = -\delta(\vec{x}) \quad (7.60)$$

which has the solution

$$V(\vec{x}) = \frac{1}{4\pi|\vec{x}|} \quad (7.61)$$

giving a long-distance interaction. However, one may consider an interaction mediated by a massive field with the equation

$$(\square + \mu^2) V(x) = f(x) \quad (7.62)$$

or for the static potential

$$(\Delta - \mu^2) V(\vec{x}) = -\delta(\vec{x}) \quad (7.63)$$

which has the solution

$$V(\vec{x}) = \frac{e^{-\mu|\vec{x}|}}{4\pi|\vec{x}|} \quad (7.64)$$

Matching the expected range gives

$$\mu = \frac{1}{2 \text{ fm}} \approx 100 \text{ MeV}/\hbar c \quad (7.65)$$

Based on this, H. Yukawa predicted a *mediating particle with a mass around 100 MeV*. The muon (μ) discovered in 1936 seemed to fit the bill with its mass of 106 MeV; however, it was observed to penetrate solid material easily, excluding its participation in the strong interactions. High altitude atmospheric observations found the pion (π) in 1947, which turned out to be a true strongly interacting particle with a mass of around 140 MeV and spin 0, which is created by cosmic radiation and decays to the muons observed at the surface of earth. For his prediction, Yukawa earned the Nobel prize in 1949, while C. Powell was awarded a Nobel prize in 1950 for developing and applying the technique of particle detection using photographic emulsions which lead to the discovery of the pion. As the tables in Section 3.3.2 show, the pion comes as a triplet composed of π^+ , π^0 and π^- , and so it is natural to guess that it is an isospin triplet with $I = 1$, the three states being

$$\pi^+ = |1, +1\rangle \quad \pi^0 = |1, 0\rangle \quad \pi^- = |1, -1\rangle \quad (7.66)$$

Now the interaction proceeds by pion exchange, which makes sense from the isospin point of view, since the 1/2 isospin representation of the nucleon can be found in the product of a 1 representation of a pion with a 1/2 of the nucleon (cf. Exercise 7.3).

Since the elementary process consists of pion emission with the nucleon preserved (except for a possible change of its state between a proton and a neutron), nucleons do not disappear or created

¹³Here size is taken as the rms charge radius.

during this interaction. In addition, nucleon number is also preserved during the β decay. This leads to the concept of a conserved charge called *baryon number*¹⁴ or *baryon charge* B : baryons (like proton and neutron) have baryon charge $+1$, antibaryons (like antiproton and antineutron) have baryon charge -1 , and mesons (like the pion) have baryon charge zero. Notice that the electric charge of the hadrons considered so far (nucleon and pions) can be written as

$$Q = I_3 + \frac{B}{2} \quad (7.67)$$

We now consider a simple application of isospin symmetry. Consider the isospin space of the two-nucleon system, which can be decomposed as

$$\begin{aligned} |1, +1\rangle &= |pp\rangle \\ |1, 0\rangle &= \frac{1}{\sqrt{2}}|pn\rangle + \frac{1}{\sqrt{2}}|np\rangle \\ |1, -1\rangle &= |nn\rangle \\ |0, 0\rangle &= \frac{1}{\sqrt{2}}|pn\rangle - \frac{1}{\sqrt{2}}|np\rangle \end{aligned} \quad (7.68)$$

It is important to note that the above decomposition implies that the deuteron can only be an isosinglet state, otherwise there would exist a pp and an nn bound state as well.

Next, consider the following three scattering processes:

$$\begin{aligned} (a) : p + p &\rightarrow d + \pi^+ \\ (b) : p + n &\rightarrow d + \pi^0 \\ (c) : n + n &\rightarrow d + \pi^- \end{aligned} \quad (7.69)$$

The strong Hamiltonian preserves the isospin, therefore

$$\langle I, I_3 | H_s | I', I'_3 \rangle = 0 \quad \text{whenever either } I \neq I' \text{ or } I_3 \neq I'_3 \quad (7.70)$$

Now the end state is an isotriplet, so the only component of the initial state contributing to the process is the isotriplet. Therefore the amplitudes for these processes have the ratio

$$\mathcal{M}_a : \mathcal{M}_b : \mathcal{M}_c = 1 : \frac{1}{\sqrt{2}} : 1 \quad (7.71)$$

Neglecting isospin breaking effects (such as the proton-neutron mass difference) the processes are otherwise the same, and since transition probabilities are squares of amplitudes, their cross-sections have the ratio

$$\sigma_a : \sigma_b : \sigma_c = 2 : 1 : 2 \quad (7.72)$$

Process (c) is hard to perform experimentally, but (a) and (b) have been measured, and when corrections are made for electromagnetic effects, they are indeed found to satisfy the above relation.

We remark that the decomposition studied in exercise 7.3 can be applied to pion-nucleon processes; the interested reader can find the calculation in Section 4.3 of D. Griffiths: *Introduction to Elementary Particles*, Wiley-VCH, 2008.

¹⁴Strictly speaking, so far it would only be “nucleon number”, but by the time the principle was proposed it had also been observed in processes involving other baryons.

7.2 Strangeness and hypercharge

With the pion found it seemed that the theory of strong interactions was complete. However, this did not last long: in December of the same year new strongly interacting particles were found, and the list of such particles kept extending during the next decade. In addition, many of these particles were “strange” in the sense that they were only produced together, but they decayed to the good old nucleons and pions. For example, these new particles, called *K mesons* and *hyperons* were produced in reactions like



but never in reactions like



Since these processes had cross sections typical of strong interaction, it became clear that strong interaction conserved an *additional quantum number* called *strangeness* and denoted by S . To fix conventions, the strangeness of K^+ was chosen to be +1, and the quantum number of all other particles can be assigned by considering the allowed processes. This resulted in the following tables:

Spin-0 mesons							
Particle	Q	S	B	I	I_3	Rest mass (MeV)	Mean lifetime (s)
π^+	+1	0	0	1	1	139.57018 ± 0.00035	$(2.6033 \pm 0.0005) \cdot 10^{-8}$
π^-	-1	0	0	1	-1	139.57018 ± 0.00035	$(2.6033 \pm 0.0005) \cdot 10^{-8}$
π^0	0	0	0	1	0	134.9766 ± 0.0006	$(8.52 \pm 0.18) \cdot 10^{-17}$
η	0	0	0	0	0	547.862 ± 0.018	$(5.02 \pm 0.19) \cdot 10^{-19}$
K^+	+1	+1	0	1/2	+1/2	493.677 ± 0.016	$(1.2380 \pm 0.0021) \cdot 10^{-8}$
K^-	-1	-1	0	1/2	-1/2	493.677 ± 0.016	$(1.2380 \pm 0.0021) \cdot 10^{-8}$
K^0	0	+1	0	1/2	-1/2	497.614 ± 0.024	$(8.954 \pm 0.004) \cdot 10^{-11}$ (K_S^0)*
K^0	0	-1	0	1/2	+1/2	497.614 ± 0.024	$(5.116 \pm 0.021) \cdot 10^{-8}$ (K_L^0)
η'	0	0	0	0	0	957.78 ± 0.06	$(3.32 \pm 0.15) \cdot 10^{-21}$
*Note: for the neutral kaon, the states with definite lifetime are the S and L combinations of K^0 and $\bar{K}^0$.							
Spin-1 mesons							
ρ^+	+1	0	0	1	1	775.11 ± 0.34	$(4.41 \pm 0.02) \cdot 10^{-24}$
ρ^-	-1	0	0	1	-1	775.11 ± 0.34	$(4.41 \pm 0.02) \cdot 10^{-24}$
ρ^0	0	0	0	1	0	775.26 ± 0.25	$(4.45 \pm 0.03) \cdot 10^{-24}$
ω	0	0	0	0	0	782.65 ± 0.12	$(7.75 \pm 0.07) \cdot 10^{-23}$
K^{*+}	+1	+1	0	1/2	+1/2	891.66 ± 0.026	$(3.26 \pm 0.06) \cdot 10^{-23}$
K^{*-}	-1	-1	0	1/2	-1/2	891.66 ± 0.026	$(3.26 \pm 0.06) \cdot 10^{-23}$
K^{*0}	0	+1	0	1/2	-1/2	895.81 ± 0.19	$(1.39 \pm 0.02) \cdot 10^{-23}$
K^{*0}	0	-1	0	1/2	+1/2	895.81 ± 0.19	$(1.39 \pm 0.02) \cdot 10^{-23}$
ϕ	0	0	0	0	0	1019.461 ± 0.019	$(1.54 \pm 0.01) \cdot 10^{-22}$

Spin-1/2 baryons							
Particle	Q	S	B	I	I_3	Rest mass (MeV)	Mean lifetime (s)
p	+1	0	1	1/2	+1/2	938.272046(21)	Stable
n	0	0	1	1/2	-1/2	939.565379(21)	880.0 ± 0.9 (free n)
Λ^0	0	-1	1	0	0	1115.683 ± 0.006	$(2.632 \pm 0.020) \cdot 10^{-10}$
Σ^+	+1	-1	1	1	+1	1189.37 ± 0.07	$(8.018 \pm 0.026) \cdot 10^{-11}$
Σ^0	0	-1	1	1	0	1192.642 ± 0.024	$(7.4 \pm 0.7) \cdot 10^{-20}$
Σ^-	-1	-1	1	1	-1	1197.449 ± 0.030	$(1.479 \pm 0.011) \cdot 10^{-10}$
Ξ^0	0	-2	1	1/2	1/2	1314.86 ± 0.20	$(2.90 \pm 0.09) \cdot 10^{-10}$
Ξ^-	-1	-2	1	1/2	-1/2	1321.71 ± 0.07	$(1.639 \pm 0.015) \cdot 10^{-10}$
Spin-3/2 baryons							
Δ^{++}	+2	0	1	3/2	+3/2	1232 ± 2	$(5.63 \pm 0.14) \cdot 10^{-24}$
Δ^+	+1	0	1	3/2	+1/2	1232 ± 2	$(5.63 \pm 0.14) \cdot 10^{-24}$
Δ^0	0	0	1	3/2	-1/2	1232 ± 2	$(5.63 \pm 0.14) \cdot 10^{-24}$
Δ^-	-1	0	1	3/2	-3/2	1232 ± 2	$(5.63 \pm 0.14) \cdot 10^{-24}$
Σ^{*+}	+1	-1	1	1	+1	1382.8 ± 0.4	$(1.839 \pm 0.0041) \cdot 10^{-23}$
Σ^{*0}	0	-1	1	1	0	1383.7 ± 1.0	$(1.83 \pm 0.25) \cdot 10^{-23}$
Σ^{*-}	-1	-1	1	1	-1	1387.2 ± 0.5	$(1.671 \pm 0.089) \cdot 10^{-23}$
Ξ^{*0}	0	-2	1	1/2	+1/2	1531.80 ± 0.32	$(7.23 \pm 0.40) \cdot 10^{-23}$
Ξ^{*-}	-1	-2	1	1/2	-1/2	1535.0 ± 0.6	$(6.6 \pm 1.3) \cdot 10^{-23}$
Ω^-	-1	-3	1	0	0	1672.45 ± 0.29	$(8.21 \pm 0.11) \cdot 10^{-11}$

Note that antibaryons are not listed in these tables; their table can be obtained by simple charge conjugation (i.e. setting all their charges -1 times the charge of the corresponding baryon, and leaving their mass and lifetimes the same).

Strong decays of the lightest strange particle are forbidden by conservation of strangeness; therefore they decay via weak interaction, e.g.

$$\begin{aligned}
K^+ &\rightarrow \pi^+ + \pi^0 \\
K^+ &\rightarrow \mu^+ + \nu_\mu \\
\Lambda^0 &\rightarrow p^+ + \pi^- \\
\Sigma^+ &\rightarrow p^+ + \pi^0 \\
\Sigma^+ &\rightarrow n + \pi^+
\end{aligned} \tag{7.75}$$

and have notably longer lifetimes than particles for which strong decay channels are allowed, such as e.g. the Δ family of resonances that can decay into nucleons plus pions, or heavier strange particles which can decay to lighter strange ones plus pions. We remark that the lifetime of Σ^0 is neither typical of weak, nor of strong decays, which is explained by its decay to $\Lambda^0 + \gamma$ via electromagnetic interactions, which violates isospin conservation, but conserves strangeness, and is therefore allowed.

Note that the relation (7.67) can be extended to all the above hadrons as

$$Q = I_3 + \frac{Y}{2} \tag{7.76}$$

where

$$Y = B + S \tag{7.77}$$

is the so-called *hypercharge*.

7.3 The $SU(3)$ quark model

7.3.1 The $SU(3)$ group

With all these hadrons around, the natural question is whether there is any principle to classify them. It is natural to identify the following groups:

- 9 spin-0 (scalar) mesons,
- 9 spin-1 (vector) mesons,
- 8 spin-1/2 baryons,
- 10 spin-3/2 baryons.

Is there a symmetry which has exactly these irreducible representations (multiplets)? Note that it should contain isospin symmetry $SU(2)_I$ as a subgroup, and it must also have another generator corresponding to strangeness, or in view of (7.76), the hypercharge Y . The hypercharge generator must also commute with the isospin so that I^2 , I_3 and Y are simultaneously diagonalisable and can therefore be simultaneously assigned to hadrons as quantum numbers.

The simplest such extension is given by the special unitary group $SU(3)$ composed of 3×3 unitary matrices U satisfying $\det U = 1$. Following the same reasoning as for $SU(2)$ in Subsection 7.1.1, the generators of $SU(3)$ are traceless Hermitian 3×3 matrices of which there are exactly 8 independent, called the *Gell-Mann matrices*. They are named after M. Gell-Mann who proposed the $SU(3)$ classification of hadrons with Y. Neeman, modifying an earlier idea by S. Sakata and have the form:

$$\begin{aligned} \lambda_1 &= \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} & \lambda_2 &= \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} & \lambda_3 &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \\ \lambda_4 &= \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} & \lambda_5 &= \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix} & \lambda_6 &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \\ \lambda_7 &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix} & \lambda_8 &= \frac{1}{\sqrt{3}} \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{pmatrix} \end{aligned} \quad (7.78)$$

They satisfy

$$\text{Tr} \lambda_k \lambda_l = 2\delta_{kl} \quad (7.79)$$

By convention, the generators of the *fundamental representation* are taken as

$$\frac{\lambda_k}{2} \quad (7.80)$$

The isospin is identified with the two-dimensional subspace spanned by the first two basis vectors, so $k = 1, 2, 3$ correspond to $I_k, k = 1, 2, 3$. The third component is an isospin singlet, since $SU(2)_I$ acts only in the first two.

Any $SU(3)$ group element can be written as

$$U(\epsilon) = e^{-i\epsilon^a \frac{\lambda_a}{2}} \quad (7.81)$$

where ϵ^a , $a = 1, \dots, 8$ are real parameters; the identity element corresponds to $\epsilon^a = 0$.

The generators satisfy the *Lie algebra*

$$\left[\frac{\lambda_k}{2}, \frac{\lambda_l}{2} \right] = i f_3^{klm} \frac{\lambda_m}{2} \quad (7.82)$$

where f_3^{klm} are the *structure constants*, which are clearly antisymmetric in their first two indices:

$$f_3^{klm} = -f_3^{lkm} \quad (7.83)$$

However, even more is true due to the “orthonormal” choice of basis of generators according to (7.79). To see this, note that the structure constants can be expressed as

$$f_3^{klm} = -2i \text{Tr} \left(\left[\frac{\lambda_k}{2}, \frac{\lambda_l}{2} \right] \frac{\lambda_m}{2} \right) \quad (7.84)$$

Exercise 7.4. Using (7.84) show that

$$f_3^{klm} \quad (7.85)$$

is totally antisymmetric in all its indices.

Solution. This follows from the cyclic property of the trace:

$$\text{Tr } A_1 A_2 \dots A_n = \text{Tr } A_2 \dots A_n A_1 \quad (7.86)$$

Explicitly

$$\begin{aligned} f_3^{klm} &= -2i \text{Tr} \left(\left[\frac{\lambda_k}{2}, \frac{\lambda_l}{2} \right] \frac{\lambda_m}{2} \right) \\ &= -2i \text{Tr} \left(\frac{\lambda_k}{2} \frac{\lambda_l}{2} \frac{\lambda_m}{2} \right) + 2i \text{Tr} \left(\frac{\lambda_l}{2} \frac{\lambda_k}{2} \frac{\lambda_m}{2} \right) \\ &= -2i \text{Tr} \left(\frac{\lambda_m}{2} \frac{\lambda_k}{2} \frac{\lambda_l}{2} \right) + 2i \text{Tr} \left(\frac{\lambda_k}{2} \frac{\lambda_m}{2} \frac{\lambda_l}{2} \right) \\ &= -2i \text{Tr} \left(\left[\frac{\lambda_m}{2}, \frac{\lambda_k}{2} \right] \frac{\lambda_l}{2} \right) \\ &= f_3^{mkl} \end{aligned} \quad (7.87)$$

So f_3^{klm} is invariant under cyclic permutations of indices, and antisymmetric in the last two, from which it follows that it is totally antisymmetric. $\square$

Exercise 7.5. Compute f_3^{klm} .

Solution. Due to the complete antisymmetry, it is enough to compute only the case $k < l < m$, which can be performed using (7.84). The non-zero structure constants are

$$\begin{aligned} f_3^{123} &= 1 & f_3^{147} &= f_3^{246} = f_3^{257} = f_3^{345} = -f_3^{156} = -f_3^{367} = \frac{1}{2} \\ f_3^{458} &= f_3^{678} = \frac{\sqrt{3}}{2} \end{aligned} \quad (7.88)$$

$\square$

Denoting the abstract generators of $SU(3)$ by T_a , and the defining representation by $\mathbf{3}$, representation of the generator T_a in $\mathbf{3}$ is given by the Gell-Mann matrices:

$$R_{\mathbf{3}}(T_a) = \frac{\lambda_a}{2} \quad a = 1, \dots, 8 \quad (7.89)$$

If U is an element of $SU(3)$, then $R_{\bar{\mathbf{3}}}(U) = U^*$ (where the star denotes complex conjugation) forms a representation:

$$R_{\bar{\mathbf{3}}}(U_1)R_{\bar{\mathbf{3}}}(U_2) = R_{\bar{\mathbf{3}}}(U_1U_2) \quad (7.90)$$

and using (7.81) the generators are represented in $\bar{\mathbf{3}}$ as

$$R_{\bar{\mathbf{3}}}(T_a) = -\frac{\lambda_a^T}{2} \quad (7.91)$$

with the upper index T denoting matrix transposition.

Another important representation is the *adjoint representation* (denoted by $\mathbf{8}$) which is a representation of the Lie algebra on itself (as a vector space) given by the transformation of the generators following from the commutators

$$\text{adj}(T_a)X = [T_a, X] \quad (7.92)$$

where $X = c^a T^a$ is a generic element of the Lie algebra with $c^a \in \mathbb{C}$. For group elements U this representation is given by

$$x \rightarrow UxU^\dagger \quad (7.93)$$

where

$$x = c^a \frac{\lambda_a}{2} \quad (7.94)$$

is the matrix representing the Lie algebra element X in the defining representation $\mathbf{3}$.

Exercise 7.6. Show that adj is really a representation of the Lie algebra i.e.

$$[\text{adj}(T_a), \text{adj}(T_b)] = \text{adj}([T_a, T_b]) = if^{abc}\text{adj}(T_c) \quad (7.95)$$

and compute the matrix of the generator in the adjoint representation in the basis T^a .

Hint: use Jacobi's identity

$$[A, [B, C]] + [B, [C, A]] + [C, [A, B]] = 0 \quad (7.96)$$

Solution. Putting $A = T_a$, $B = T_b$ and $C = X$ gives

$$[T_a, [T_b, X]] + [T_b, [X, T_a]] + [X, [T_a, T_b]] = 0 \quad (7.97)$$

which can be rewritten as

$$[T_a, [T_b, X]] - [T_b, [T_a, X]] = [[T_a, T_b], X] \quad (7.98)$$

i.e. exactly

$$[\text{adj}(T_a), \text{adj}(T_b)](X) = \text{adj}([T_a, T_b])(X) \quad (7.99)$$

for every X . Using

$$[T_a, T_b] = if^{abc}T_c \quad (7.100)$$

and linearity of adj gives

$$[\text{adj}(T_a), \text{adj}(T_b)] = if^{abc}\text{adj}(T_c) \quad (7.101)$$

Now writing $X = c^b T_b$ one obtains

$$\text{adj}(T_a)X = [T_a, c^b T_b] = if^{abd}c^b T_d = c'^d T_d \quad (7.102)$$

i.e. $c'^d = if^{abd}c^b$, so the matrix representation of the adjoint generator corresponding to T_a is given by the Lie algebra structure constants:

$$\text{adj}(T_a)_{db} = if^{abd} \quad (7.103)$$

□

7.3.2 Representation theory of $SU(3)$

From Subsection [7.3.1](#), recall that one must fit

- 9 spin-0 (scalar) mesons,
- 9 spin-1 (vector) mesons,
- 8 spin-1/2 baryons,
- 10 spin-3/2 baryons

into multiplets of the proposed symmetry. Indeed, the adjoint octet **8** can fit the spin-1/2 baryons, while mesons can fit into a singlet **1** plus an octet **8**. The details of how this work is described the next Subsection, where we show that the quantum numbers do fit these assignments, and that several properties of these particles can be explained using the $SU(3)$ symmetry, and the way it is broken^{[15](#)}. In order to make this more precise, we must familiarise ourselves with the representation theory of $SU(3)$.

The defining representation **3** corresponds to a complex **3** component object q_i transforming under $U \in SU(3)$ as

$$q'_i = U_{ij}q_j \quad (7.104)$$

One can then define tensors transforming as

$$\mathcal{T}'_{i_1 \dots i_n} = U_{i_1 j_1} \dots U_{i_n j_n} \mathcal{T}_{j_1 \dots j_n} \quad (7.105)$$

which is just the representation

$$\underbrace{\mathbf{3} \otimes \dots \otimes \mathbf{3}}_{n \text{ factors}} \quad (7.106)$$

¹⁵It is clear that it cannot be an exact symmetry, since the members of proposed multiplets have different masses and isospin is broken anyway.

However, such a representation is reducible. For example, a second order tensor can be decomposed into its symmetric and antisymmetric part

$$\begin{aligned}\mathcal{T}_{ij} &= \mathcal{T}_{\{ij\}} + \mathcal{T}_{[ij]} \\ \mathcal{T}_{\{ij\}} &= \frac{1}{2}(\mathcal{T}_{ij} + \mathcal{T}_{ji}) \\ \mathcal{T}_{[ij]} &= \frac{1}{2}(\mathcal{T}_{ij} - \mathcal{T}_{ji})\end{aligned}\tag{7.107}$$

each transforming separately under $SU(3)$. These cannot be decomposed further, and the symmetric part is just a new representation **6** (it has six independent components). The antisymmetric part is three dimensional, and now we show that it is just $\bar{\mathbf{3}}$. Remember that $\det U = 1$, so

$$\epsilon_{ijk}U_{ii'}U_{jj'}U_{kk'} = \epsilon_{i'j'k'}\tag{7.108}$$

so

$$\epsilon_{ijk}U_{jj'}U_{kk'} = (U^{-1})_{i'i}\epsilon_{i'j'k'}\tag{7.109}$$

Unitarity implies that $U^{-1} = U^\dagger$ leading to

$$\epsilon_{ijk}U_{jj'}U_{kk'} = U_{ii'}^*\epsilon_{i'j'k'}\tag{7.110}$$

Consider now the tensor $\mathcal{T}_{[ij]}$ transforming in the antisymmetric part of $\mathbf{3} \otimes \mathbf{3}$. This is in one-to-one correspondence with a new $\bar{q}$ defined as

$$\bar{q}_i = \epsilon_{ijk}\mathcal{T}_{[jk]}$$

which transforms as

$$\epsilon_{ijk}\mathcal{T}_{[jk]} \rightarrow \epsilon_{ijk}U_{jj'}U_{kk'}\mathcal{T}_{[j'k']} = U_{ii'}^*\epsilon_{i'j'k'}\mathcal{T}_{[j'k']}\tag{7.111}$$

under $\bar{\mathbf{3}}$:

$$\bar{q}'_i = U_{ii'}^*\bar{q}_{i'}\tag{7.112}$$

showing that the antisymmetric part of $\mathbf{3} \otimes \mathbf{3}$ is indeed identical to $\bar{\mathbf{3}}$.

Exercise 7.7. Show for $SU(2)$ the representation $\bar{\mathbf{2}}$ obtained as the complex conjugate of $\mathbf{2}$ is equivalent to $\mathbf{2}$, with a change of basis (where $\mathbf{2}$ denotes the defining, i.e. $j = 1/2$ representation of $SU(2)$).

Solution. For $SU(2)$ the relation $\det U = +1$ can be written as

$$\epsilon_{ij}U_{ii'}U_{jj'} = \epsilon_{i'j'}\tag{7.113}$$

which implies that

$$\epsilon_{ij}U_{jj'} = \epsilon_{i'j'}(U^{-1})_{i'i} = U_{ii'}^*\epsilon_{i'j'}\tag{7.114}$$

which can also be written in matrix form as $\epsilon U = U^* \epsilon$ or in a more suggestive way as

$$U^* = \epsilon U \epsilon^{-1}\tag{7.115}$$

So if v_i ($i = 1, 2$) is an $SU(2)$ doublet, then

$$\bar{v}_i = \epsilon_{ij}v_j\tag{7.116}$$

transforms as an anti-doublet. Therefore the antidoublet $\bar{v}$ and the doublet v are related by the linear transformation (i.e. by a change of basis)

$$\bar{v}_1 = v_2 \quad \bar{v}_2 = -v_1 \quad (7.117)$$

This shows that there is no “antispin”: the spin of a particle and of an anti-particle is the same sort of quantum number, transforming the same way under rotations. $\square$

Therefore

$$\mathbf{3} \otimes \mathbf{3} = \bar{\mathbf{3}} \oplus \mathbf{6} \quad (7.118)$$

Similarly

$$\bar{\mathbf{3}} \otimes \bar{\mathbf{3}} = \mathbf{3} \oplus \bar{\mathbf{6}} \quad (7.119)$$

where $\bar{\mathbf{6}}$ is the complex conjugate of representation $\mathbf{6}$, identical to the symmetric part of $\bar{\mathbf{3}} \otimes \bar{\mathbf{3}}$.

One can also decompose $\mathbf{3} \otimes \bar{\mathbf{3}}$ by introducing a $\mathcal{V}$ tensor transforming as

$$\mathcal{V}_{i\bar{j}} \rightarrow U_{ii'} U_{\bar{j}\bar{j}'}^* \mathcal{V}_{i'\bar{j}'} \quad (7.120)$$

which can be written in matrix form as

$$\mathcal{V} \rightarrow U \mathcal{V} U^\dagger = U \mathcal{V} U^{-1} \quad (7.121)$$

Therefore

$$\text{Tr} \mathcal{V} = \mathcal{V}_{1\bar{1}} + \mathcal{V}_{2\bar{2}} + \mathcal{V}_{3\bar{3}} \quad (7.122)$$

is a singlet, while the rest

$$\tilde{\mathcal{V}}_{i\bar{j}} = \mathcal{V}_{i\bar{j}} - \frac{1}{3} (\text{Tr} \mathcal{V}) \delta_{i\bar{j}} \quad (7.123)$$

is an eight-dimensional representation, which can be shown to be equivalent to $\mathbf{8}$.

Exercise 7.8. Show that the traceless part of $\mathbf{3} \otimes \bar{\mathbf{3}}$ is just the adjoint representation $\mathbf{8}$.

Hint: write down the transformation under an infinitesimal group element

$$U_{kl} = \delta_{kl} - i\epsilon^a \frac{1}{2} (\lambda_a)_{kl} \quad (7.124)$$

Solution. $T_{i\bar{j}}$ transforms as

$$\mathcal{V}_{i\bar{j}} \rightarrow \mathcal{V}'_{i\bar{j}} = U_{ii'} U_{\bar{j}\bar{j}'}^* \mathcal{V}_{i'\bar{j}'} = \left[\delta_{ii'} - i\epsilon^a \frac{1}{2} (\lambda_a)_{ii'} \right] \left[\delta_{\bar{j}\bar{j}'} + i\epsilon^a \frac{1}{2} (\lambda_a)_{\bar{j}\bar{j}'}^* \right] \mathcal{V}_{i'\bar{j}'} \quad (7.125)$$

Keeping only terms up to first order in ϵ

$$\begin{aligned} \delta \mathcal{V}_{i\bar{j}} &= \mathcal{V}'_{i\bar{j}} - \mathcal{V}_{i\bar{j}} = -i\epsilon^a \left(\frac{1}{2} (\lambda_a)_{ii'} \mathcal{V}_{i'\bar{j}} - \frac{1}{2} (\lambda_a)_{\bar{j}\bar{j}'}^* \mathcal{V}_{i\bar{j}'} \right) \\ &= -i\epsilon^a \left(\frac{1}{2} (\lambda_a)_{ii'} \mathcal{V}_{i'\bar{j}} - \mathcal{V}_{i\bar{j}'} \frac{1}{2} (\lambda_a)_{\bar{j}\bar{j}} \right) \end{aligned} \quad (7.126)$$

where we used that $\lambda_{ij}^* = \lambda_{ji}$ due to Hermiticity. In matrix form, the infinitesimal transformation can be written as

$$\delta\mathcal{V} = -i\epsilon^a \left[\frac{\lambda^a}{2}, \mathcal{V} \right] \quad (7.127)$$

Decomposing

$$\mathcal{V} = \tilde{\mathcal{V}} + \frac{1}{3} (\text{Tr}\mathcal{V}) \mathbf{1}_{3 \times 3} \quad (7.128)$$

it is clear that the trace part drops from the transformation, while on the traceless part $\tilde{\mathcal{V}}$ the generators $\lambda_a/2$ are represented according to the adjoint representation (7.92). $\square$

As a more complicated example let us consider $\mathbf{3} \times \mathbf{3} \times \mathbf{3}$. That is a tensor of the form

$$\mathcal{W}_{ijk} \quad (7.129)$$

Now there are the following possible irreducible components:

- Antisymmetric in all indices:

$$\mathcal{W}_{[ijk]} = \frac{1}{6} (\mathcal{W}_{ijk} + \mathcal{W}_{jki} + \mathcal{W}_{kij} - \mathcal{W}_{jik} - \mathcal{W}_{kji} - \mathcal{W}_{ikj}) \quad (7.130)$$

Since the indices can take only three values for $SU(3)$, this is proportional to the Levi-Civita symbol

$$\mathcal{W}_{[ijk]} = W_1 \epsilon_{ijk} \quad (7.131)$$

and due to $\det U = 1$, W_1 is an invariant. So this is a singlet component $\mathbf{1}$.

- Symmetric in all indices

$$\mathcal{W}_{\{ijk\}} = \frac{1}{6} (\mathcal{W}_{ijk} + \mathcal{W}_{jki} + \mathcal{W}_{kij} + \mathcal{W}_{jik} + \mathcal{W}_{kji} + \mathcal{W}_{ikj}) \quad (7.132)$$

This tensor has 10 components; it is easy to count it since the number of independent components is given by the number of ways one can choose values for the three indices i, j, k from the set of three possibilities 1, 2, 3 with repetitions.

- Antisymmetric in i and j : that makes it a nine component object since $(\mathbf{3} \otimes \mathbf{3})_{\text{antisym}} \otimes \mathbf{3} = \bar{\mathbf{3}} \otimes \mathbf{3} = \mathbf{1} + \mathbf{8}$. The singlet here can be easily constructed since the basis of $(\mathbf{3} \otimes \mathbf{3})_{\text{antisym}} = \bar{\mathbf{3}}$ is

$$\begin{aligned} \mathcal{W}_{23k} - \mathcal{W}_{32k} &= \mathcal{W}_{\bar{1}k}^{(1)} \\ \mathcal{W}_{31k} - \mathcal{W}_{13k} &= \mathcal{W}_{\bar{2}k}^{(1)} \\ \mathcal{W}_{12k} - \mathcal{W}_{21k} &= \mathcal{W}_{\bar{3}k}^{(1)} \end{aligned} \quad (7.133)$$

and the singlet in $\bar{\mathbf{3}} \otimes \mathbf{3}$ is

$$\mathcal{W}_{\bar{1}1}^{(1)} + \mathcal{W}_{\bar{2}2}^{(1)} + \mathcal{W}_{\bar{3}3}^{(1)} = 6W_1 \quad (7.134)$$

so this was already included above as the fully antisymmetric part. The really independent components are the traceless part of $\mathcal{W}^{(1)}$, which we denote as $\tilde{\mathcal{W}}^{(1)}$ and is a representation $\mathbf{8}$.

Now it is possible to construct two more such combinations, taking it antisymmetric either in the first and third, or the second and third index, giving

$$\begin{aligned}\mathcal{W}_{1k}^{(2)} &= \mathcal{W}_{2k3} - \mathcal{W}_{3k2} & \mathcal{W}_{1k}^{(3)} &= \mathcal{W}_{k23} - \mathcal{W}_{k32} \\ \mathcal{W}_{2k}^{(2)} &= \mathcal{W}_{3k1} - \mathcal{W}_{1k3} & \mathcal{W}_{2k}^{(3)} &= \mathcal{W}_{k31} - \mathcal{W}_{k13} \\ \mathcal{W}_{3k}^{(2)} &= \mathcal{W}_{1k2} - \mathcal{W}_{2k1} & \mathcal{W}_{3k}^{(3)} &= \mathcal{W}_{k12} - \mathcal{W}_{k21}\end{aligned}\tag{7.135}$$

Just as for $\mathcal{W}^{(1)}$, the trace part of $\mathcal{W}^{(2,3)}$ is W_1 , while the traceless parts $\tilde{\mathcal{W}}^{(2,3)}$ transform as **8**.

This seems to give

$$\mathbf{3} \otimes \mathbf{3} \otimes \mathbf{3} \stackrel{?}{=} \mathbf{1} \oplus \mathbf{8} \oplus \mathbf{8} \oplus \mathbf{8} \oplus \mathbf{10}\tag{7.136}$$

But something is wrong: the total dimension on the left-hand side is 27, while on the right-hand side it is 35. It is easy to guess that one of the octets is probably not independent, and indeed this is the case:

$$\mathcal{W}_{ij}^{(1)} - \mathcal{W}_{ij}^{(2)} + \mathcal{W}_{ij}^{(3)} = 6W_1\delta_{ij}\tag{7.137}$$

which can be easily verified by hand. So

$$\tilde{\mathcal{W}}_{ij}^{(1)} - \tilde{\mathcal{W}}_{ij}^{(2)} + \tilde{\mathcal{W}}_{ij}^{(3)} = 0\tag{7.138}$$

and there are only two independent octets. Therefore the end result is

$$\mathbf{3} \otimes \mathbf{3} \otimes \mathbf{3} = \mathbf{1} \oplus \mathbf{8} \oplus \mathbf{8} \oplus \mathbf{10}\tag{7.139}$$

An easier way to get this is to consider

$$\mathbf{3} \otimes \mathbf{3} \otimes \mathbf{3} = \mathbf{3} \otimes \bar{\mathbf{3}} + \mathbf{3} \otimes \mathbf{6}\tag{7.140}$$

Now

$$\mathbf{3} \otimes \bar{\mathbf{3}} = \mathbf{1} \oplus \mathbf{8}\tag{7.141}$$

while

$$\mathbf{3} \otimes \mathbf{6} = \mathbf{3} \otimes (\mathbf{3} \otimes \mathbf{3})_{\text{symm}}\tag{7.142}$$

obviously contains **10** which is $(\mathbf{3} \otimes \mathbf{3} \otimes \mathbf{3})_{\text{sym}}$, and the rest is just a copy of **8**, obtained from $T^{(3)}$ after eliminating T_1 above. So $\mathbf{3} \otimes \mathbf{6} = \mathbf{8} \oplus \mathbf{10}$.

The above tensor method can be used to decompose any product representation, and even to find the Clebsh-Gordan coefficients; it can also be extended to $SU(N)$ with an arbitrary N . There are some systematic combinatorial methods to compute the decompositions of tensors using the method of Young tableaux, but we do not go into the details here.

7.3.3 The eightfold way: meson multiplets

OK, so we have the multiplets **1**, **8** and **10** that can potentially accommodate the observed hadrons, but how do they fit in there? The answer is Gell-Mann's "*eightfold way*"; the name itself is a reference to the Buddhist path to enlightenment, while at the same time eight is the number of independent symmetries encompassed by $SU(3)$.

Let us first consider the meson nonets, which can be fit into $\mathbf{3} \otimes \bar{\mathbf{3}} = \mathbf{1} \oplus \mathbf{8}$. Taking the generators T_a , $a = 1, \dots, 8$ of $SU(3)$, we can make the following identifications for isospin and hypercharge:

$$\begin{aligned} I_3 &= T_3 \\ Y &= \alpha T_8 \end{aligned} \quad (7.143)$$

where α is a normalisation to be determined, while the total isospin is embedded into $SU(3)$ as

$$I^2 = T_1^2 + T_2^2 + T_3^2 \quad (7.144)$$

The generators I_3 and Y in the representations $\mathbf{3}$ and $\bar{\mathbf{3}}$ can be written as

$$\begin{aligned} R_{\mathbf{3}}(I_3) &= \frac{1}{2}\lambda_3 = \frac{1}{2} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix} & R_{\bar{\mathbf{3}}}(I_3) &= -\frac{1}{2}\lambda_3 = \frac{1}{2} \begin{pmatrix} -1 & 0 & 0 \\ 0 & +1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \\ R_{\mathbf{3}}(Y) &= \frac{\alpha}{2}\lambda_8 = \frac{\alpha}{2\sqrt{3}} \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{pmatrix} & R_{\bar{\mathbf{3}}}(Y) &= -\frac{\alpha}{2}\lambda_8 = -\frac{\alpha}{2\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix} \end{aligned} \quad (7.145)$$

So the quantum number assignments for these representations are

$\mathbf{3}$			$\bar{\mathbf{3}}$		
i	I_3	Y	$\bar{i}$	I_3	Y
1	1/2	$-\frac{\alpha}{2\sqrt{3}}$	$\bar{1}$	-1/2	$\frac{\alpha}{2\sqrt{3}}$
2	-1/2	$-\frac{\alpha}{2\sqrt{3}}$	$\bar{2}$	+1/2	$\frac{\alpha}{2\sqrt{3}}$
3	0	$\frac{\alpha}{\sqrt{3}}$	$\bar{3}$	0	$-\frac{\alpha}{\sqrt{3}}$

The nonet is a tensor $\mathcal{V}_{i\bar{j}}$ of which the traceless part $\tilde{\mathcal{V}}_{i\bar{j}}$ is the octet. The pions must fit as an isospin triplet, so they must be the product of the two isospin doublets $i = 1, 2$ and $\bar{j} = 1, 2$. Based on the value of I_3 we can see that π^+ can only be the component $1\bar{2}$ and π^- can only be $\bar{1}2$. These states have $Y = 0$ which is also fulfilled. On the other hand K^+ and K^0 are an isospin doublet which can be assigned as $1\bar{3}$ and $2\bar{3}$. Since K^+ has strangeness +1 and baryon number 0, we have

$$Y(K^+) = 1 = -\frac{\alpha}{2\sqrt{3}} - \frac{\alpha}{\sqrt{3}} = -\frac{\sqrt{3}\alpha}{2} \quad (7.146)$$

so $\alpha = -2/\sqrt{3}$. Now the table can be extended with further details as

$\mathbf{3}$					$\bar{\mathbf{3}}$				
i	I_3	Y	I	Q	$\bar{i}$	I_3	Y	I	Q
1	1/2	$\frac{1}{3}$	1/2	$\frac{2}{3}$	$\bar{1}$	-1/2	$-\frac{1}{3}$	1/2	$-\frac{2}{3}$
2	-1/2	$\frac{1}{3}$	1/2	$-\frac{1}{3}$	$\bar{2}$	+1/2	$-\frac{1}{3}$	1/2	$\frac{1}{3}$
3	0	$-\frac{2}{3}$	0	$-\frac{1}{3}$	$\bar{3}$	0	$\frac{2}{3}$	0	$\frac{1}{3}$

where we used $Q = I_3 + Y/2$, and also put in I denoting the total isospin of the multiplet the given component belongs to.

Now π^0 is a little more complicated to obtain. Note that the $|1\rangle$ and $|2\rangle$ components form an isospin doublet, so in $SU(2)$ notations

$$\begin{aligned} |1\rangle &= |1/2, +1/2\rangle \\ |2\rangle &= |1/2, -1/2\rangle \\ |3\rangle &= |0, 0\rangle \end{aligned} \quad (7.147)$$

But by the rules of conjugation of isospin doublet (cf. Exercise 7.7) we obtain

$$\begin{aligned} |\bar{1}\rangle &= |1/2, -1/2\rangle \\ |\bar{2}\rangle &= -|1/2, +1/2\rangle \\ |\bar{3}\rangle &= |0, 0\rangle \end{aligned} \quad (7.148)$$

So the pion triplet can be assigned as

$$\begin{aligned} |\pi^+\rangle &= |1\bar{2}\rangle \\ |\pi^-\rangle &= |2\bar{1}\rangle \\ |\pi^0\rangle &= \frac{1}{\sqrt{2}} (|1\bar{1}\rangle - |2\bar{2}\rangle) \end{aligned} \quad (7.149)$$

The isospin singlet state in the **8** is given by

$$|\eta_8\rangle = \frac{1}{\sqrt{6}} (|1\bar{1}\rangle + |2\bar{2}\rangle - 2|3\bar{3}\rangle) \quad (7.150)$$

where the first two terms and the last one are both isospin singlets, and the relative coefficients are chosen so that the trace vanishes as required for the component **8**. The kaon isospin doublets can be identified with

$$\begin{aligned} |K^+\rangle &= |1\bar{3}\rangle & |K^0\rangle &= |2\bar{3}\rangle \\ |\bar{K}^0\rangle &= |3\bar{2}\rangle & |K^-\rangle &= |3\bar{1}\rangle \end{aligned} \quad (7.151)$$

(we omitted the $-$ sign in $|\bar{K}^0\rangle$ as the overall phase of a state has no physical meaning. Finally, the $SU(3)$ singlet **1** is

$$|\eta_0\rangle = \frac{1}{\sqrt{3}} (|1\bar{1}\rangle + |2\bar{2}\rangle + |3\bar{3}\rangle) \quad (7.152)$$

The physical states η and η' are eventually combinations of η_0 and η_8 due to $SU(3)$ breaking, although their mixing is small, so η is almost exclusively η_8 , while η' is η_0 .

Exercise 7.9. Determine the assignments of the vector meson nonet following the above reasoning.

Solution. The end result is

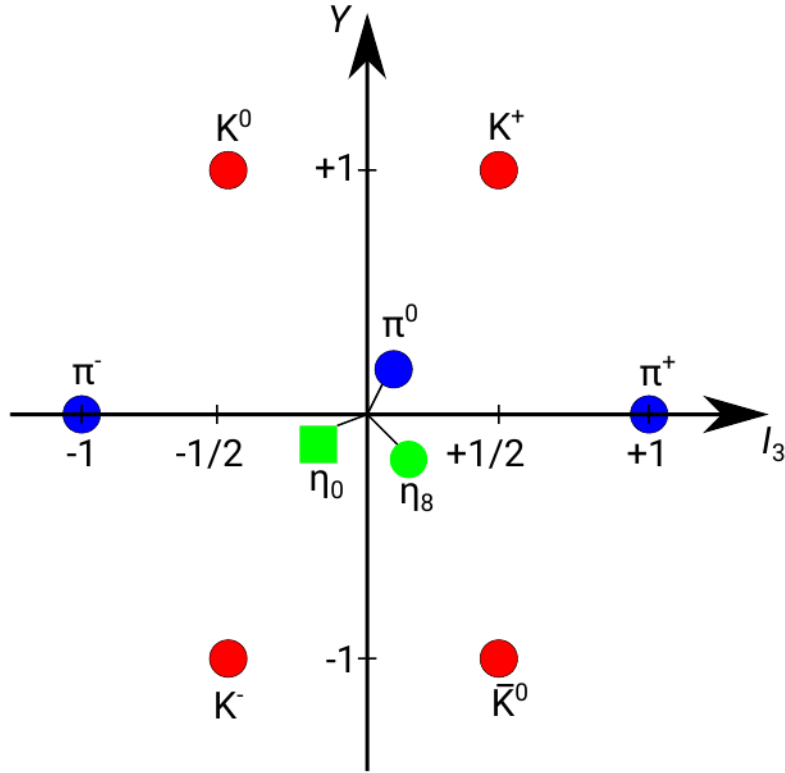
$$\begin{aligned} |\rho^+\rangle &= |1\bar{2}\rangle & |\rho^-\rangle &= |\bar{1}2\rangle & |\rho^0\rangle &= \frac{1}{\sqrt{2}} (|1\bar{1}\rangle - |2\bar{2}\rangle) \\ |K^{*+}\rangle &= |1\bar{3}\rangle & |K^{*0}\rangle &= |2\bar{3}\rangle \\ |\bar{K}^{*0}\rangle &= |3\bar{2}\rangle & |K^{*-}\rangle &= |3\bar{1}\rangle \end{aligned} \quad (7.153)$$

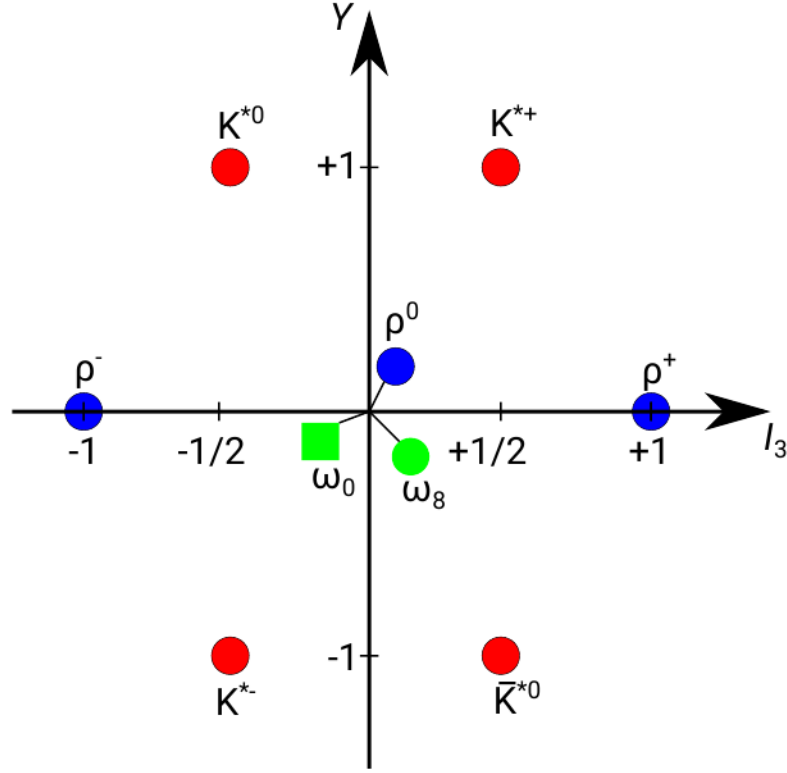
The isosinglet part of the octet and the $SU(3)$ singlet is

$$\begin{aligned} |\omega_8\rangle &= \frac{1}{\sqrt{6}} (|1\bar{1}\rangle + |2\bar{2}\rangle - 2|3\bar{3}\rangle) \\ |\omega_0\rangle &= \frac{1}{\sqrt{3}} (|1\bar{1}\rangle + |2\bar{2}\rangle + |3\bar{3}\rangle) \end{aligned} \quad (7.154)$$

which mix to form the state $|\omega\rangle$ and $|\phi\rangle$. More on this mixing can be found in Subsection 7.3.5. $\square$

The assignment of the scalar and vector mesons can be represented on a two-dimensional plane with axes I_3 and Y , leading to the diagrams (note that $Y = B + S = S$):





The colour codes are: green/red/blue denote isospin singlet/doublet/triplet states, while the shape codes are: circles for the $SU(3)$ octet ($\mathbf{8}$) states and boxes for $SU(3)$ singlet ($\mathbf{1}$) states.

7.3.4 The eightfold way: baryon multiplets and the GMO mass formula

To fit the baryons one can consider $\mathbf{3} \otimes \mathbf{3} \otimes \mathbf{3} = \mathbf{1} \oplus \mathbf{8} \oplus \mathbf{8} \oplus \mathbf{10}$; the spin-1/2 hadrons can fit into an $\mathbf{8}$, while the spin-3/2 can be fitted into $\mathbf{10}$. To get the correct baryon number, we must assign baryon number $B = 1/3$ to $\mathbf{3}$ and $B = -1/3$ to $\bar{\mathbf{3}}$, which allows us to extend the table with B and strangeness $S = Y - B$:

$\mathbf{3}$							$\bar{\mathbf{3}}$						
i	I_3	Y	I	Q	B	S	$\bar{i}$	I_3	Y	I	Q	B	S
1	1/2	$\frac{1}{3}$	1/2	$\frac{2}{3}$	$\frac{1}{3}$	0	1	-1/2	$-\frac{1}{3}$	1/2	$-\frac{2}{3}$	$-\frac{1}{3}$	0
2	-1/2	$\frac{1}{3}$	1/2	$-\frac{1}{3}$	$\frac{1}{3}$	0	2	+1/2	$-\frac{1}{3}$	1/2	$\frac{1}{3}$	$-\frac{1}{3}$	0
3	0	$-\frac{2}{3}$	0	$-\frac{1}{3}$	$\frac{1}{3}$	-1	3	0	$\frac{2}{3}$	0	$\frac{1}{3}$	$-\frac{1}{3}$	1

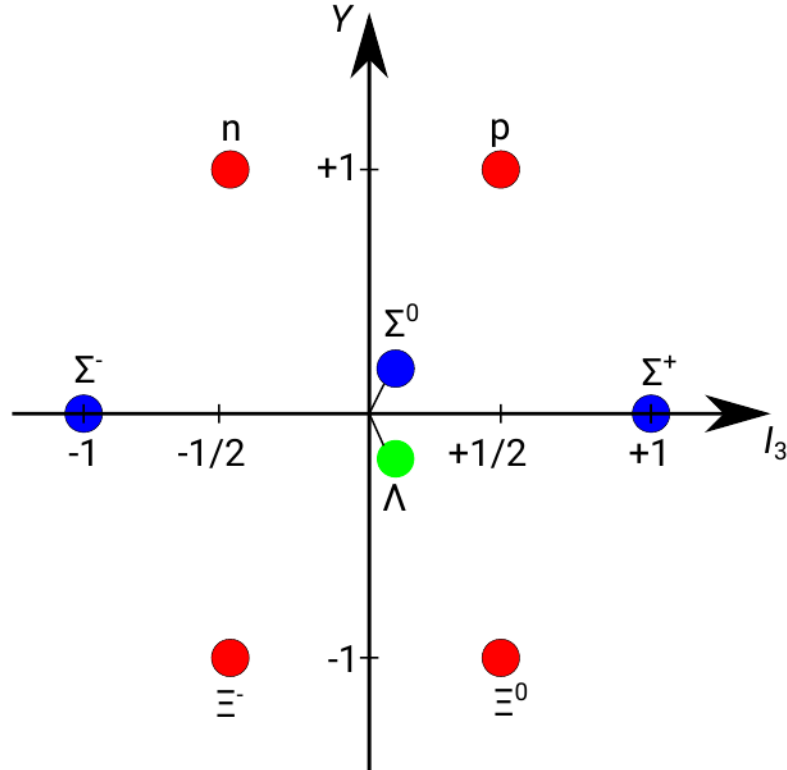
Recall that starting with the $\mathbf{3} \otimes \mathbf{3} \otimes \mathbf{3}$ tensor $\mathcal{W}_{ijk}$, the octet is the traceless part of $(\mathbf{3} \otimes \mathbf{3})_{\text{antisym}} \otimes \mathbf{3} = \bar{\mathbf{3}} \otimes \mathbf{3} = \mathbf{1} + \mathbf{8}$, where $(\mathbf{3} \otimes \mathbf{3})_{\text{antisym}} \otimes \mathbf{3}$ is

$$\begin{aligned}
\mathcal{W}_{23k} - \mathcal{W}_{32k} &= \mathcal{W}_{1k}^{(1)} \\
\mathcal{W}_{31k} - \mathcal{W}_{13k} &= \mathcal{W}_{2k}^{(1)} \\
\mathcal{W}_{12k} - \mathcal{W}_{21k} &= \mathcal{W}_{3k}^{(1)}
\end{aligned} \tag{7.155}$$

Considering the known quantum numbers of the baryon octet, this leads to the assignments¹⁶

$$\begin{aligned}
|p\rangle &= \frac{1}{\sqrt{2}} \mathcal{W}_{31}^{(1)} = \frac{1}{\sqrt{2}} (|121\rangle - |211\rangle) \\
|n\rangle &= \frac{1}{\sqrt{2}} \mathcal{W}_{32}^{(1)} = \frac{1}{\sqrt{2}} (|122\rangle - |212\rangle) \\
|\Sigma^+\rangle &= \frac{1}{\sqrt{2}} \mathcal{W}_{21}^{(1)} = \frac{1}{\sqrt{2}} (|311\rangle - |131\rangle) \\
|\Sigma^-\rangle &= \frac{1}{\sqrt{2}} \mathcal{W}_{12}^{(1)} = \frac{1}{\sqrt{2}} (|322\rangle - |232\rangle) \\
|\Sigma^0\rangle &= \frac{1}{2} (\mathcal{W}_{11}^{(1)} - \mathcal{W}_{22}^{(1)}) = \frac{1}{2} (|231\rangle - |321\rangle - |312\rangle + |132\rangle) \\
|\Lambda^0\rangle &= \frac{1}{2\sqrt{3}} (\mathcal{W}_{11}^{(1)} + \mathcal{W}_{22}^{(1)} - 2\mathcal{W}_{33}^{(1)}) = \frac{1}{2\sqrt{3}} (|231\rangle - |321\rangle + |312\rangle - |132\rangle - 2|123\rangle + 2|213\rangle) \\
|\Xi^0\rangle &= \frac{1}{\sqrt{2}} \mathcal{W}_{23}^{(1)} = \frac{1}{\sqrt{2}} (|313\rangle - |133\rangle) \\
|\Xi^-\rangle &= \frac{1}{\sqrt{2}} \mathcal{W}_{13}^{(1)} = \frac{1}{\sqrt{2}} (|323\rangle - |233\rangle)
\end{aligned} \tag{7.156}$$

which leads to the diagram (note that $Y = B + S = S + 1$):



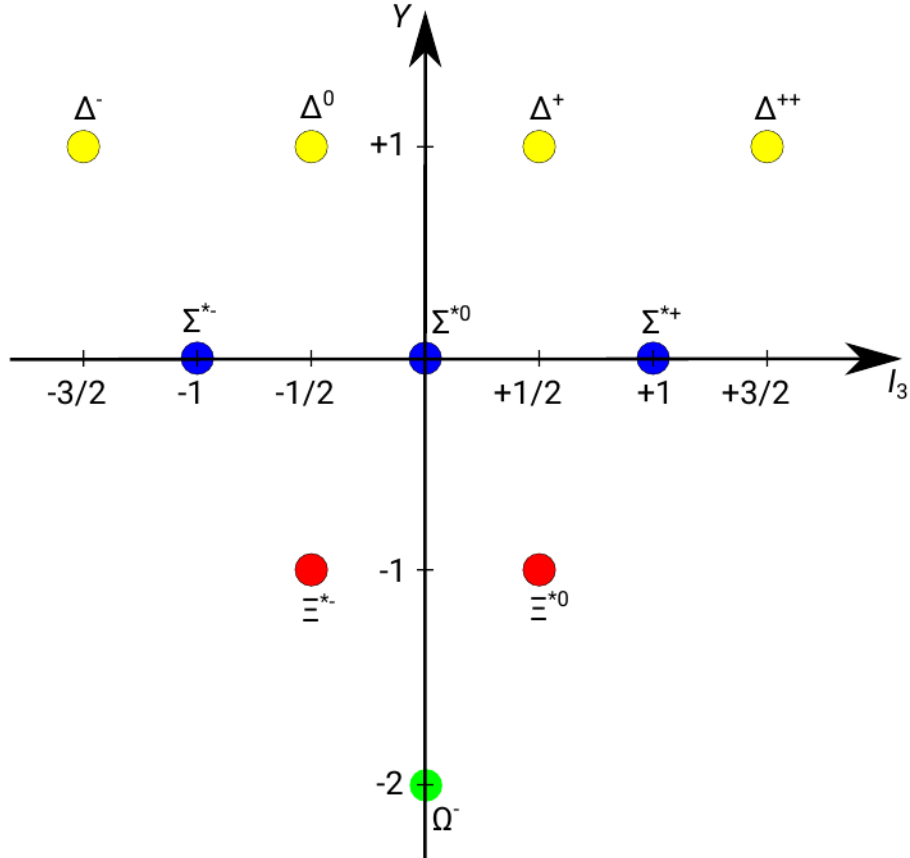
The colour codes are the same as for the mesons: green/red/blue denote isospin singlet/doublet/triplet states. Since all states are in **8**, there is no need to distinguish $SU(3)$ multiplets.

¹⁶In fact, the baryon wave functions are a little more complicated once the correct assignment of the spin part is taken into account; we return to this issue in Subsection 7.3.6

For the spin-3/2 hadrons, one considers $\mathbf{10} = (\mathbf{3} \otimes \mathbf{3} \otimes \mathbf{3})_{\text{symm}}$ and following the known quantum numbers leads to

$$\begin{aligned}
|\Delta^{++}\rangle &= |111\rangle \\
|\Delta^+\rangle &= \frac{1}{\sqrt{3}} (|112\rangle + |121\rangle + |211\rangle) \\
|\Delta^0\rangle &= \frac{1}{\sqrt{3}} (|122\rangle + |212\rangle + |221\rangle) \\
|\Delta^-\rangle &= |222\rangle \\
|\Sigma^{*+}\rangle &= \frac{1}{\sqrt{3}} (|113\rangle + |131\rangle + |311\rangle) \\
|\Sigma^{*0}\rangle &= \frac{1}{\sqrt{6}} (|123\rangle + |312\rangle + |231\rangle + |213\rangle + |132\rangle + |321\rangle) \\
|\Sigma^{*-}\rangle &= \frac{1}{\sqrt{3}} (|223\rangle + |232\rangle + |322\rangle) \\
|\Xi^{*0}\rangle &= \frac{1}{\sqrt{3}} (|133\rangle + |313\rangle + |331\rangle) \\
|\Xi^{*-}\rangle &= \frac{1}{\sqrt{3}} (|233\rangle + |323\rangle + |332\rangle) \\
|\Omega^-\rangle &= |333\rangle
\end{aligned} \tag{7.157}$$

which can be summarised in the diagram (note that $Y = B + S = S + 1$):



The colour codes are: green/red/blue/yellow denote isospin singlet/doublet/triplet/quadruplet states. Since all states are in **10**, there is no need to distinguish $SU(3)$ multiplets.

In fact, the hadron Ω^- was not known when Gell-Mann proposed the eightfold way; however, *it's existence follows from $SU(3)$ symmetry* since the multiplet must be complete. It was predicted in 1962 by Gell-Mann and independently, also by Y. Ne'eman. Its mass can be estimated observing the regular mass differences in the decuplet.

In case of an unbroken $SU(3)$ symmetry, all masses would be equal. However, the observed masses depend on the hypercharge, corresponding to $SU(3)$ symmetry breaking, and also slightly on I_3 which signals *isospin symmetry breaking*. Neglecting the latter for a moment, the strong interaction Hamiltonian can be written as

$$H_s = H_{su(3)} + H_{br} \quad (7.158)$$

where $H_{SU(3)}$ is $SU(3)$ symmetric, while H_{br} breaks $SU(3)$ leaving the isospin invariant. Looking at the Gell-Mann matrices (7.78) it is clear that the eighth component of the octet transforms represented by λ_8 in the right way, i.e. leaving $SU(2)_I$ invariant while breaking $SU(3)$. So we can propose that

$$H_{br} = hY \quad (7.159)$$

where the hypercharge generator Y carries the flavour labels and h acts on all other (momentum, spin etc.) degrees of freedom. This leads to the following prediction for the decuplet masses:

$$m = a + bY \quad (7.160)$$

The isospin averaged of the decuplet are

$$\begin{aligned} Y = 1 & : m_{\Delta} = \frac{m_{\Delta^{++}} + m_{\Delta^+} + m_{\Delta^0} + m_{\Delta^-}}{4} \approx 1232 \text{ MeV} \\ Y = 0 & : m_{\Sigma^*} = \frac{m_{\Sigma^{*+}} + m_{\Sigma^{*0}} + m_{\Sigma^{*-}}}{3} \approx 1384.5 \text{ MeV} \\ Y = -1 & : m_{\Xi^*} = \frac{m_{\Xi^{*0}} + m_{\Xi^{*-}}}{2} \approx 1533.4 \text{ MeV} \end{aligned} \quad (7.161)$$

which can be fitted with

$$a = 1383.3 \quad b = -150.7 \quad (7.162)$$

This leads to the prediction

$$m_{\Omega} = a - 2b \approx 1685 \text{ MeV} \quad (7.163)$$

More detailed considerations even predicted the expected decay products. The hadron Ω^- was indeed discovered in 1964, and Gell-Mann was awarded a Nobel Prize in physics in 1969 for his contributions and discoveries concerning the classification of elementary particles and their interactions.

The masses of the baryon octet do not follow the above simple rule. However this may be because second order corrections of $SU(3)$ breaking interactions may be relevant as well, leading to operators quadratic in the symmetry generators. The two quadratic operators consistent with the pattern of symmetry breaking are the squared isospin $\vec{I}^2 = I_1^2 + I_2^2 + I_3^2$ and Y^2 , leading to an improved mass formula

$$m = a' + b'Y + c'I(I+1) + d'Y^2 \quad (7.164)$$

However, for the decuplet all states satisfy the relation

$$I = 1 + Y/2 \quad (7.165)$$

so (7.164) can be rewritten as

$$m = a' + 2c' + \left(b' + \frac{3}{2}c'\right)Y + \left(\frac{1}{4}c' + d'\right)Y^2 \quad (7.166)$$

To keep the equation consistent with the linear mass splitting of the decuplet, we require

$$d' = -\frac{1}{4}c' \quad (7.167)$$

i.e. the mass formula is

$$m = a' + b'Y + c' \left(I(I+1) - \frac{Y^2}{4} \right) \quad (7.168)$$

In fact, it can be shown using a little more group theory that the combination of the generators

$$I(I+1) - \frac{Y^2}{4} \quad (7.169)$$

is the only quadratic one transforming as the 8th component with an octet, so it can be singled out by demanding that the symmetry breaking has this property. For the octet we have

$$\begin{aligned} I = \frac{1}{2}, Y = 1 & : m_N = \frac{m_p + m_n}{2} \approx 938.9 \text{ MeV} \\ I = 0, Y = 0 & : m_\Lambda \approx 1116 \text{ MeV} \\ I = 1, Y = 0 & : m_\Sigma = \frac{m_{\Sigma^+} + m_{\Sigma^-} + m_{\Sigma^0}}{3} \approx 1193.1 \text{ MeV} \\ I = \frac{1}{2}, Y = -1 & : m_\Xi = \frac{m_{\Xi^-} + m_{\Xi^0}}{2} \approx 1318.3 \text{ MeV} \end{aligned} \quad (7.170)$$

On the other hand, the formula predicts that

$$\begin{aligned} m_N &= a' + b' + \frac{c'}{2} \\ m_\Lambda &= a' \\ m_\Sigma &= a' + 2c' \\ m_\Xi &= a' - b' + \frac{c'}{2} \end{aligned} \quad (7.171)$$

which leads to the famous *Gell-Mann-Okubo mass relation*

$$m_N + m_\Xi = \frac{3m_\Lambda + m_\Sigma}{2} \quad (7.172)$$

which is satisfied with an accuracy of 0.6%. One can also determine the best fits of the parameters to be

$$a' = 1111.6 \text{ MeV} \quad b' = -189.7 \text{ MeV} \quad c' = 40.0 \text{ MeV} \quad (7.173)$$

These parameters fit the octet well, but for the decuplet they would give

$$\begin{aligned} a &= a' + 2c' = 1111.55 \text{ MeV} \\ b &= b' + \frac{3}{2}c' = -129.65 \text{ MeV} \end{aligned} \quad (7.174)$$

The discrepancy with (7.162) shows that the $SU(3)$ symmetric part a is very different for the octet and the decuplet. This can be solved, however, by noting that the octet has spin $s = 1/2$, while the decuplet has $s = 3/2$. Accounting for spin dependence leads to the *Gürsey-Radicati mass formula*

$$m = A + Bs(s + 1) + CY + D \left(I(I + 1) - \frac{Y^2}{4} \right) \quad (7.175)$$

Exercise 7.10. Determine parameters A, B, C, D by numerical fit and check the accuracy of (7.175).

Solution. Fitting this formula to the combined octet plus decuplet masses (7.161, 7.170) gives

$$A = 1066.7 \text{ MeV} \quad B = 66.8 \text{ MeV} \quad C = -195.0 \text{ MeV} \quad D = 33.0 \text{ MeV} \quad (7.176)$$

which compares to the measured masses as follows:

Masses (MeV)	m_N	m_Λ	m_Σ	m_Ξ	m_Δ	m_{Σ^*}	m_{Ξ^*}	m_Ω
Experiment	938.9	1115.7	1193.1	1318.3	1232	1384.5	1533.4	1672.5
GR formula	938.4	1116.8	1182.9	1328.3	1237.9	1383.3	1528.7	1674.1
Deviation (%)	0.05	0.1	0.9	0.8	0.5	0.09	0.3	0.1

so the formula is valid within 1% accuracy. □

The mass formula (7.175) can be derived by considering the combined flavour-spin symmetry $SU(6)$. The success in describing the hadron spectrum using the flavour symmetry and its pattern of breaking provided a strong evidence for Gell-Mann's eightfold way.

7.3.5 The meson mass spectrum

The GMO relation can also be tested for the meson spectrum. Naively, one would expect

$$2m_K = \frac{3m_\eta + m_\pi}{2} \quad (7.177)$$

Using

$$\begin{aligned} m_K &= \frac{493.677 + 497.614}{2} \approx 495.6 \text{ MeV} \\ m_\eta &\approx 547.9 \text{ MeV} \\ m_\pi &= \frac{2 \cdot 139.57018 + 134.9766}{3} \approx 138.0 \text{ MeV} \end{aligned} \quad (7.178)$$

leads to a 100 MeV, or about 10% discrepancy between the left and right hand sides. However, the relation holds much better for the squared masses:

$$2m_K^2 = \frac{3m_\eta^2 + m_\pi^2}{2} \quad (7.179)$$

which is true within about 6%. Construed as a prediction for the mass of η

$$m_\eta = \sqrt{\frac{4m_K^2 - m_\pi^2}{3}} \quad (7.180)$$

an substituting the values known for m_K and m_π gives $m_\eta \approx 567$ MeV, which is quite good. More advanced field theoretic modelling of the spin-0 mesons shows that the correct relation is indeed between the squared masses. Indeed, as shown later in Section 8.2 the Lagrangian (and therefore also the Hamiltonian) for boson fields depend on mass *quadratically*, while for fermions the dependence is linear, explaining the difference between the meson and baryon mass formulas.

However, the GMO relation seems to break down for the spin-1 meson octet. Naively, one would expect

$$3m_\omega^2 = 4m_{K^*}^2 - m_\rho^2 \quad (7.181)$$

Using $m_{K^*} = 890$ MeV and $m_\rho = 770$ MeV, this would predict $m_\omega = 926.5$ MeV, but experimentally $m_\omega = 783$ MeV. This discrepancy can be understood in terms of a *mixing between the octet and singlet mesons*. Namely, the predicted mass $m_8 = 926.5$ MeV corresponds to a state ω_8 with $Y = 0$ and $I_3 = 0$. On the other hand, there is also an $SU(3)$ singlet state ω_0 with $Y = 0$ and $I_3 = 0$ and a mass m_0 . The assumed $SU(3)$ breaking preserves $Y = 0$ and $I_3 = 0$ while it allows the singlet and octet to mix. The mass squared matrix is assumed to be the form

$$\begin{pmatrix} m_8^2 & \mu^2 \\ \mu^2 & m_0^2 \end{pmatrix} \quad (7.182)$$

This matrix can be diagonalised using a two-dimensional rotation to give the physical meson states ω and ϕ :

$$R \begin{pmatrix} m_8^2 & \mu^2 \\ \mu^2 & m_0^2 \end{pmatrix} R^T = \begin{pmatrix} m_\omega^2 & 0 \\ 0 & m_\phi^2 \end{pmatrix} \quad (7.183)$$

with

$$R = \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \quad (7.184)$$

so the physical states are

$$\begin{aligned} |\omega\rangle &= \cos \theta |\omega_8\rangle - \sin \theta |\omega_0\rangle \\ |\phi\rangle &= \sin \theta |\omega_8\rangle + \cos \theta |\omega_0\rangle \end{aligned} \quad (7.185)$$

Exercise 7.11. Compute the mixing angle θ using the mass $m_\phi = 1020$ MeV.

Solution. We can compute

$$\begin{aligned} R^T \begin{pmatrix} m_\omega^2 & 0 \\ 0 & m_\phi^2 \end{pmatrix} R &= \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} m_\omega^2 & 0 \\ 0 & m_\phi^2 \end{pmatrix} \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \\ &= \begin{pmatrix} m_\omega^2 \cos^2 \theta + m_\phi^2 \sin^2 \theta & (m_\omega - m_\phi) \cos \theta \sin \theta \\ (m_\omega - m_\phi) \cos \theta \sin \theta & m_\omega^2 \sin^2 \theta + m_\phi^2 \cos^2 \theta \end{pmatrix} \\ &= \begin{pmatrix} m_8^2 & \mu^2 \\ \mu^2 & m_0^2 \end{pmatrix} \end{aligned} \quad (7.186)$$

leads to

$$m_\omega^2 \cos^2 \theta + m_\phi^2 \sin^2 \theta = m_8^2 \quad (7.187)$$

which gives

$$\sin \theta = \sqrt{\frac{m_8^2 - m_\omega^2}{m_\phi^2 - m_\omega^2}} \approx 0.76 \quad (7.188)$$

□

7.3.6 Quark model

We saw that mesons can be classified by the product $\mathbf{3} \otimes \bar{\mathbf{3}} = \mathbf{1} \oplus \mathbf{8}$ and baryons by $\mathbf{3} \otimes \mathbf{3} \otimes \mathbf{3} = \mathbf{1} \oplus \mathbf{8} \oplus \mathbf{8} \oplus \mathbf{10}$ (antibaryons correspond to $\bar{\mathbf{3}} \otimes \bar{\mathbf{3}} \otimes \bar{\mathbf{3}} = \mathbf{1} \oplus \mathbf{8} \oplus \mathbf{8} \oplus \mathbf{10}$). It is natural to ask whether there are elementary building blocks corresponding to the triplet representation $\mathbf{3}$, with their antiparticles in $\bar{\mathbf{3}}$. Such particles were introduced in 1964 by Murray Gell-Mann who called them *quarks* after the following passage from James Joyce's *Finnegans Wake*:

*Three quarks for Muster Mark!
Sure he hasn't got much of a bark
And sure any he has it's all beside the mark.*

and also by George Zweig who called these constituents *aces*. The three constituents of the triplet are called up, down and strange; *up* (u) and *down* (d) got their name by being the two components of an isospin triplet, while *strange* (s) is an isospin singlet which carries the quantum number strangeness. The table at the beginning of Subsection can be rewritten as follows:

$\mathbf{3}$							$\bar{\mathbf{3}}$						
q	I_3	Y	I	Q	B	S	$\bar{q}$	I_3	Y	I	Q	B	S
u	$1/2$	$\frac{1}{3}$	$1/2$	$\frac{2}{3}$	$\frac{1}{3}$	0	$\bar{u}$	$-1/2$	$-\frac{1}{3}$	$1/2$	$-\frac{2}{3}$	$-\frac{1}{3}$	0
d	$-1/2$	$\frac{1}{3}$	$1/2$	$-\frac{1}{3}$	$\frac{1}{3}$	0	$\bar{d}$	$+1/2$	$-\frac{1}{3}$	$1/2$	$\frac{1}{3}$	$-\frac{1}{3}$	0
s	0	$-\frac{2}{3}$	0	$-\frac{1}{3}$	$\frac{1}{3}$	-1	$\bar{s}$	0	$\frac{2}{3}$	0	$\frac{1}{3}$	$-\frac{1}{3}$	1

Interestingly, these particles have *fractional charges*; such fundamental objects have never been observed in Nature. Indeed, no one has ever found a free quark, which is explained by quantum chromodynamics as explained later in Subsection 12.3.4 for the time being we must take this as just an observed fact. To be consistent with the observed spins of the hadrons, quarks must carry spin $1/2$ and are therefore *fermions* according to the spin-statistics theorem (cf. Section 10.2). Since they transform into each other under the $SU(3)$ symmetry, the three quarks are considered to be essentially three different states of the same object (similarly to proton and neutron being two states of nucleon according to isospin); the labels u , d and s are called *flavours*. Since $SU(3)$ is broken, the mass of the s quark is expected to be different from that of u and d :

$$m_s \gg m_{u,d} \quad (7.189)$$

In fact, isospin is also broken, although it is a much better symmetry than $SU(3)$; this seems to imply that $m_u \neq m_d$, but care must be taken to disentangle the effects of electromagnetic interaction which also contributes to isospin breaking effects such as e.g. the proton-neutron mass difference. Given that, as discussed in Subsection 7.1, the electromagnetic contribution would make the proton (with quark content uud) heavier, while in reality it is the neutron (udd) which has greater mass, we expect that $m_d > m_u$, with the difference being of an order of a few MeV, which is the characteristic magnitude for mass differences between members of the same isospin multiplet (cf. the hadron tables in Subsection 3.3.2). On the other hand, looking at the decuplet splittings s is expected to be about $130 - 140$ MeV heavier than either u or d which is true for the so-called constituent masses (cf. Table 11.2 in Subsection 11.3.3). The actual picture is somewhat more complicated due to strong interactions, and is discussed in some details in Subsection 7.3.8

The wave functions of the hadrons therefore correspond to bound states. The mesons are quark-antiquark bound states and their wave functions have the form

$$\int d^3p_1 \int d^3p_2 \sum_{f,\bar{f}} \sum_{\sigma_1,\sigma_2} \psi(\vec{p}_1, \vec{p}_2) \Phi_{f\bar{f}\Sigma_{\sigma_1\sigma_2}} |q_{f\sigma_1}(\vec{p}_1) \bar{q}_{\bar{f}\sigma_2}(\vec{p}_2)\rangle \quad (7.190)$$

where $\mathcal{F}$ describes the flavour $SU(3)$ structure and $\mathcal{S}$ the spin structure. In their ground states, the spatial wave function ψ must carry zero orbital momentum and therefore their total spins can be either 0 or 1. For the spin-0 octet, one obtains the following states (with the spatial part omitted)

$$\begin{aligned}
|K^+\rangle &= \frac{1}{\sqrt{2}} (|u_\uparrow \bar{s}_\downarrow\rangle - |u_\downarrow \bar{s}_\uparrow\rangle) \\
|K^0\rangle &= \frac{1}{\sqrt{2}} (|d_\uparrow \bar{s}_\downarrow\rangle - |d_\downarrow \bar{s}_\uparrow\rangle) \\
|K^-\rangle &= \frac{1}{\sqrt{2}} (|\bar{u}_\uparrow s_\downarrow\rangle - |\bar{u}_\downarrow s_\uparrow\rangle) \\
|\bar{K}^0\rangle &= \frac{1}{\sqrt{2}} (|\bar{d}_\uparrow s_\downarrow\rangle - |\bar{d}_\downarrow s_\uparrow\rangle) \\
|\pi^+\rangle &= \frac{1}{\sqrt{2}} (|u_\uparrow \bar{d}_\downarrow\rangle - |u_\downarrow \bar{d}_\uparrow\rangle) \\
|\pi^-\rangle &= \frac{1}{\sqrt{2}} (|\bar{u}_\uparrow d_\downarrow\rangle - |\bar{u}_\downarrow d_\uparrow\rangle) \\
|\pi^0\rangle &= \frac{1}{2} (|\bar{u}_\uparrow u_\downarrow\rangle - |\bar{u}_\downarrow u_\uparrow\rangle - |\bar{d}_\uparrow d_\downarrow\rangle + |\bar{d}_\downarrow d_\uparrow\rangle) \\
|\eta\rangle &= \frac{1}{\sqrt{12}} (|\bar{u}_\uparrow u_\downarrow\rangle + |\bar{d}_\uparrow d_\downarrow\rangle - 2|\bar{s}_\uparrow s_\downarrow\rangle - |\bar{u}_\downarrow u_\uparrow\rangle - |\bar{d}_\downarrow d_\uparrow\rangle + 2|\bar{s}_\downarrow s_\uparrow\rangle)
\end{aligned} \tag{7.191}$$

while the spin-0 singlet is

$$|\eta'\rangle = \frac{1}{\sqrt{6}} (|\bar{u}_\uparrow u_\downarrow\rangle + |\bar{d}_\uparrow d_\downarrow\rangle + |\bar{s}_\uparrow s_\downarrow\rangle - |\bar{u}_\downarrow u_\uparrow\rangle - |\bar{d}_\downarrow d_\uparrow\rangle - |\bar{s}_\downarrow s_\uparrow\rangle) \tag{7.192}$$

These states are nothing else than the flavour combinations constructed in Subsection [7.3.3](#) with the flavour assignments $1 = u$, $2 = d$, $3 = s$, combined with a singlet spin wave function

$$\frac{1}{\sqrt{2}} (|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle) \tag{7.193}$$

For the vector meson octet the spin part must be chosen to be the spin-1 triplet with components

$$|\uparrow\uparrow\rangle \quad \frac{1}{\sqrt{2}} (|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle) \quad |\downarrow\downarrow\rangle \tag{7.194}$$

and the interested reader is invited to write down the full wave functions. Here we only consider the following components

$$\begin{aligned}
|\omega_8\rangle &= \frac{1}{\sqrt{6}} (|\bar{u}u\rangle + |\bar{d}d\rangle - 2|\bar{s}s\rangle) \\
|\omega_0\rangle &= \frac{1}{\sqrt{3}} (|\bar{u}u\rangle + |\bar{d}d\rangle + |\bar{s}s\rangle)
\end{aligned} \tag{7.195}$$

where both spatial and spin components are suppressed. Recall that the physical states correspond to

$$\begin{aligned}
|\omega\rangle &= \cos\theta |\omega_8\rangle - \sin\theta |\omega_0\rangle \\
|\phi\rangle &= \sin\theta |\omega_8\rangle + \cos\theta |\omega_0\rangle
\end{aligned} \tag{7.196}$$

due to octet-singlet mixing. Now it is interesting to note that for the special case

$$\cos \theta = \frac{1}{\sqrt{3}} \approx 0.577 \quad \sin \theta = \sqrt{\frac{2}{3}} \approx 0.816 \quad (7.197)$$

one obtains

$$\begin{aligned} |\omega\rangle &= \frac{1}{\sqrt{2}} (|\bar{u}u\rangle + |\bar{d}d\rangle) \\ |\phi\rangle &= |s\bar{s}\rangle \end{aligned} \quad (7.198)$$

so that ω is purely made up of non-strange quarks, while ϕ contains only strange quarks. The mixing angle $\sin \theta \approx 0.76$ computed in Exercise 7.11 is indeed close to this so-called *maximal mixing*, so ϕ is dominated by strange quark content, which explains some of its properties, such as decay ratios as shown in Subsection 12.3.3

For the baryon decuplet, the total spin is $3/2$, which corresponds to a totally symmetric spin wave function which has the following 4 independent states:

$$\begin{aligned} |\uparrow\uparrow\uparrow\rangle &= \frac{1}{\sqrt{3}} (|\downarrow\uparrow\uparrow\rangle + |\uparrow\downarrow\uparrow\rangle + |\uparrow\uparrow\downarrow\rangle) \\ \frac{1}{\sqrt{3}} (|\downarrow\downarrow\uparrow\rangle + |\downarrow\uparrow\downarrow\rangle + |\uparrow\downarrow\downarrow\rangle) & \quad |\downarrow\downarrow\downarrow\rangle \end{aligned} \quad (7.199)$$

The flavour wave-function (7.157) is also completely symmetric and has 10 components. Since only these $4 \times 10 = 40$ states are found in Nature, the spatial wave function must also be completely symmetric, otherwise there would be more states corresponding to its orientation. *Now this is a problem since the quarks are fermions obeying Pauli principle, therefore the total state must be antisymmetric.* As discussed later in Section 12.2 this is resolved via introducing the *colour* degree of freedom, allowing the construction of quantum chromodynamics, which gives the fundamental description of strong interactions.

The wave-function for the baryon octet is more complicated; the interested reader is referred to Section 5.6 of the book D. Griffiths: *Introduction to Elementary Particles*. Here we only note that the baryon wave function constructed from the quark model gives a good description of magnetic momenta of the baryon octet, and the construction also explains while there is only a single baryon octet despite having two copies of **8** in the product $\mathbf{3} \otimes \mathbf{3} \otimes \mathbf{3}$ ¹⁷. This construction can also be described using an $SU(6)$ symmetry incorporating both spin and flavour indices, as sketched below.

$SU(6)$ is obtained by extending flavour symmetry $SU(3)$ to include the two ($\uparrow$ and $\downarrow$) spin components; it has $SU(3)_{flavour} \times SU(2)_{spin}$ as a subgroup. Quarks transform as **6** which can be decomposed under $SU(3) \times SU(2)$ as $(\mathbf{3}, \frac{1}{2})$, while antiquarks are $\bar{\mathbf{6}} = (\bar{\mathbf{3}}, \frac{1}{2})$. Now the mesons can be decomposed as

$$\mathbf{6} \otimes \bar{\mathbf{6}} = \underbrace{(\mathbf{8}, 0)}_{\text{spin-0 octet}} \oplus \underbrace{(\mathbf{8}, 1)}_{\text{spin-1 octet}} \oplus \underbrace{(\mathbf{1}, 0)}_{\text{spin-0 singlet}} \oplus \underbrace{(\mathbf{1}, 1)}_{\text{spin-1 singlet}} \quad (7.200)$$

In terms of tensors, $\mathbf{6} \otimes \bar{\mathbf{6}}$ can be written as

$$M_{f\sigma\bar{f}\sigma'} \quad (7.201)$$

¹⁷This is related to the fact that the baryon octet wave function is also fully symmetric in the spatial part, in the combined flavour and spin part, and fully antisymmetric in the colour indices.

where $f\sigma$ is the quark multiplet index in $\mathbf{6}$, labelled by flavour $f = 1, 2, 3$ and spin $\sigma = \uparrow, \downarrow$, and similarly $\bar{f}\sigma'$ stands for the antiquark $\bar{\mathbf{6}}$. The number of states is given by

$$36 = 8 \cdot 1 + 8 \cdot 3 + 1 \cdot 1 + 1 \cdot 3 \quad (7.202)$$

The low-lying hadrons can be classified by the 56-dimensional three-index symmetric tensor of $SU(6)$

$$\mathbf{56} = (\mathbf{6} \otimes \mathbf{6} \otimes \mathbf{6})_{\text{symmetric}} = \underbrace{\left(10, \frac{3}{2}\right)}_{\text{baryon decuplet}} \oplus \underbrace{\left(8, \frac{1}{2}\right)}_{\text{baryon octet}} \quad (7.203)$$

The three-index totally symmetrical tensor $B_{(f_1\sigma_1, f_2\sigma_2, f_3\sigma_3)}$ (with the bracket denoting symmetrisation over the $f\sigma$ composite indices) of $SU(6)$ has

$$\binom{6+3-1}{3} = 56 \quad (7.204)$$

independent elements. The decuplet component is the one which is separately symmetric in the flavour indices ($\mathbf{10}$) and in the three spins (spin-3/2) and can be written as

$$\Phi_{(f_1 f_2 f_3)} \Sigma_{(\sigma_1 \sigma_2 \sigma_3)} \quad (7.205)$$

giving $10 \cdot 4$ states, while the rest can be ordered into a single baryon $SU(3)$ octet $\mathbf{8}$ with spin-1/2, giving $8 \cdot 2$ states. The index structure of the octet part also follows from this description. Here we only remark that this decomposition uniquely determines the flavour and spin wave functions of both the decuplet and octet baryons.

The $SU(6)$ symmetry breaking pattern corresponds to an operator which is a singlet under $SU(2)_{\text{spin}}$ (spin symmetry corresponds to spatial rotations, which are unbroken), and transforms as the eighth component of an octet under $SU(3)_{\text{flavour}}$. This pattern of symmetry breaking leads to the Gürsey-Radicati mass formula (7.175). For a more detailed discussion of $SU(6)$ cf. Sections 8.11-13 of W. Greiner and B. Müller: *Quantum Mechanics – Symmetries*.

7.3.7 Experimental evidence for quarks

Since the quark model is so successful in describing the hadrons, the natural question is whether there is any evidence for the existence of quarks. Despite the fact that no free quark has ever been observed, the answer is yes! The idea is essentially the same as the Rutherford experiment that led to the discovery of the atomic nucleus where atomic structure was probed with α particles. In the experiments performed in 1968 at the Stanford Linear Accelerator Center (SLAC), electrons were used to probe protons and neutrons in atomic nuclei; later similar experiments were performed with muons and neutrinos.

Let us consider first elastic scattering of an electron on a proton:

$$e^- + p \longrightarrow e^- + p \quad (7.206)$$

with four-momenta p_1 and p_2 for the incoming, and p'_1 and p'_2 for the outgoing electron/proton, respectively. We can consider the collision in the laboratory system where the initial proton is at

rest. Relativistic dispersion relations together with conservation of energy and momentum imposes the following constraints:

$$\begin{aligned} p_1^2 &= m^2 & p_2 &= (M, 0, 0, 0) \\ p_1 + p_2 &= p'_1 + p'_2 \\ p_1'^2 &= m^2 & p_2'^2 &= M^2 \end{aligned} \quad (7.207)$$

where m/M are the electron/proton masses. The outgoing momenta have eight components, and there are six relations they satisfy. The free variables are the angles characterising the direction of the outgoing electron, but the scattering can only depend on the deflection angle, i.e. the angle between $\vec{p}_1$ and $\vec{p}'_1$. It is convenient to introduce the 4-momentum transferred from the electron to the proton

$$q = p_1 - p'_1 \quad (7.208)$$

and define the two Lorentz-invariant parameters

$$Q^2 = -q^2 \quad \nu = \frac{p_2 \cdot q}{M} \quad (7.209)$$

In the lab system

$$\nu = \frac{M(E_1 - E'_1)}{M} = E_1 - E'_1 \quad (7.210)$$

is just the energy transferred to the proton, which is always positive so $\nu > 0$. It is also easy to compute that in the lab system

$$q^2 = (p'_2 - p_2)^2 = (E'_2 - M)^2 - \vec{p}_2'^2 = (E'_2 - M)^2 - E_2'^2 + M^2 = 2M(M - E'_2) < 0 \quad (7.211)$$

so Q^2 is always positive, too. Let us introduce the Bjorken variable

$$x = \frac{Q^2}{2M\nu} \quad (7.212)$$

Exercise 7.12. Show that for an elastic scattering $x = 1$.

Solution. Using energy-momentum conservation

$$p_1 - p'_1 + p_2 = p'_2 \quad (7.213)$$

and taking the Lorentz square of this equation gives

$$q^2 + p_2^2 + 2q \cdot p_2 = p_2'^2 \quad (7.214)$$

i.e

$$-Q^2 + M^2 + 2M\nu = M^2 \quad (7.215)$$

□

So an elastic scattering is just parametrised by the value of Q^2 , from which the deflection angle can be expressed in any system of reference with some effort.

Exercise 7.13. Find the expression of the deflection angle in the laboratory system using Q^2 and the initial electron energy E_1 .

Solution. Taking the square of q

$$\begin{aligned} -Q^2 &= (p_1 - p'_1)^2 = 2m^2 - 2p_1 \cdot p'_1 = 2m^2 - 2E_1 E'_1 + 2|\vec{p}_1||\vec{p}'_1| \cos \theta \\ &= 2m^2 - 2E_1(E_1 - \nu) + 2\sqrt{E_1^2 - m^2}\sqrt{(E_1 - \nu)^2 - m^2} \cos \theta \end{aligned} \quad (7.216)$$

from which $\cos \theta$ can be expressed with E_1 and Q^2 by eliminating ν using $x = 1$. $\square$

As a result, the differential cross section of elastic scattering is only a function of Q^2

$$\frac{d\sigma}{d\Omega} = f(Q^2) \quad (7.217)$$

For scattering on a point-like proton the theoretical result is

$$\frac{d\sigma}{d\Omega} = \frac{\alpha^2}{4p_1^2 \sin^4 \frac{\theta}{2}} \left(\cos^2 \frac{\theta}{2} + \frac{Q^2}{2M^2} \sin^2 \frac{\theta}{2} \right) \quad (7.218)$$

where the first term is the same as in Rutherford's formula and the second one accounts for proton recoil. However, the proton is an extended particle with a "shape", which is accounted for by form factors $\mathcal{F}_1$ and $\mathcal{F}_2$ which can only be functions of the single Lorentz invariant Q^2 , giving the expression

$$\frac{d\sigma}{d\Omega} = \frac{\alpha^2}{4E_1^2 \sin^4 \frac{\theta}{2}} \frac{E'_1}{E_1} \left\{ \left(\mathcal{F}_1^2 + \frac{Q^2}{4M^2} \mathcal{F}_2^2 \right) \cos^2 \frac{\theta}{2} + \frac{Q^2}{2M^2} (\mathcal{F}_1 + \mathcal{F}_2)^2 \sin^2 \frac{\theta}{2} \right\} \quad (7.219)$$

and satisfying

$$\mathcal{F}_1(0) = 1 \quad \mathcal{F}_2(0) = \kappa_p = 1.79 \quad (7.220)$$

where the first relation corresponds to the unit charge of the proton, while the second gives its anomalous magnetic moment.

In an inelastic scattering the end state contains other hadrons as well

$$e + p \longrightarrow e + X \quad (7.221)$$

where X denotes a state composed of hadrons. Since it must have baryon number 1, the invariant mass

$$M_X^2 = p_X^2 \quad p_X = \sum_{i \in X} p_i \quad (7.222)$$

must be larger than that of a proton which is the lightest baryon:

$$M_X > M \quad (7.223)$$

Exercise 7.14. Show that for an inelastic scattering $x < 1$.

Solution. Using energy-momentum conservation

$$p_1 - p'_1 + p_2 = p_X \quad (7.224)$$

and taking the Lorentz square of this equation gives

$$q^2 + p_2^2 + 2q \cdot p_2 = p_X^2 = M_X^2 \quad (7.225)$$

i.e

$$-Q^2 + M^2 + 2M\nu = M_X^2 \quad (7.226)$$

leading to

$$Q^2 < 2M\nu \quad (7.227)$$

□

The scattering cross-section, and therefore the form factors become dependent on x which takes values between 0 and 1, so

$$\frac{d\sigma}{d\Omega} = \frac{\alpha^2}{4E_1^2 \sin^4 \frac{\theta}{2}} \left\{ W_2(\nu, Q^2) \cos^2 \frac{\theta}{2} + 2W_1(\nu, Q^2) \sin^2 \frac{\theta}{2} \right\} \quad (7.228)$$

In the measurement performed at SLAC, it was observed that for high enough Q^2 (typically above a few GeV^2 , called the *deep inelastic regime*) the functions

$$F_1 = MW_1(\nu, Q^2) \quad F_2 = \nu W_2(\nu, Q^2) \quad (7.229)$$

only depend on x , which is called the *Bjorken scaling*.

Due to Bjorken scaling, the cross-section in the deep inelastic regime can be rewritten as

$$\frac{d\sigma}{d\Omega} = \frac{\alpha^2}{4E_1^2 \sin^4 \frac{\theta}{2}} \left\{ \frac{1}{\nu} F_1(x) \cos^2 \frac{\theta}{2} + \frac{2}{M} F_2(x) \sin^2 \frac{\theta}{2} \right\} \quad (7.230)$$

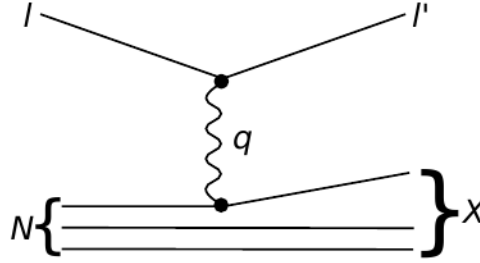
which can be interpreted as follows. *The electron scatters off spin-1/2 constituents* of the proton with a four-momentum k and an invariant mass $m_p^2 = k^2$ for which

$$\frac{Q^2}{2m_p\nu} = 1 \text{ i.e. } 2m_p\nu + q^2 = 0 \Rightarrow x = \frac{m_p}{M} \quad (7.231)$$

so x gives the fraction of the proton's 4-momentum carried by the given constituent¹⁸. The Q^2 independence implies that these constituents are point-like (have no form factor); they are called partons and further analysis identified them with constituent quarks¹⁹. In the deep inelastic regime, the scattering of the electron off a hadron can be illustrated as

¹⁸More precisely this is valid in the so-called infinite-momentum frame, where the proton is boosted to essentially the speed of light so that its relativistic factor $\gamma \gg 1$; in such a frame, the parton on which the electron scatters carries energy-momentum $k^\mu = xP^\mu$ where P^μ is the full energy-momentum of the proton. The whole picture of the proton composed of free partons is valid in the infinite-momentum frame, which is very different from the laboratory frame in which the proton is at rest. Fortunately, the scattering cross-section when expressed in terms of invariant variables such as ν and x is itself invariant, so conclusions drawn from the free parton picture are valid in all frames. Also note that the invariant mass m_p^2 has nothing to do with the physical quark mass; in an infinite momentum frame all masses are eventually negligible.

¹⁹In fact, measurements showed that the scaling is violated by a logarithmic dependence on Q^2 , which was later explained by interactions between the partons described by quantum chromodynamics. It was also found that around half of the momentum of the proton is carried by gluons, so they must also be included among the partons, despite not participating in the hard scattering itself.



In the above picture, time flows from left to right and the nucleon is drawn as a compound of three so-called *valence quarks* (uud for the proton and udd for the neutron), although the actual parton on which the initial scattering of the lepton happens may be a so-called *sea parton* (cf. Subsection 7.3.8). The initial energy-momentum transfer is q and corresponds to a hard scattering meaning that $Q^2 = -q^2$ is large with respect to typical strong interactions scales; the type of interaction (electromagnetic/weak) symbolised by the wavy line depends on the nature of the leptonic probe l . The hadronic system X is created after the initial hard scattering in a process called *hadronisation* that involves soft (low momentum transfer) physics; cf. also Subsection 12.4.

For their pioneering investigations at SLAC concerning deep inelastic scattering of electrons on protons and bound neutrons which confirmed and extended the quark model of hadrons, physicists J.I. Friedman, H.W. Kendall and R.E. Taylor received the Nobel Prize in Physics in 1990.

Later experiments carried out with muons confirmed the SLAC results, and scattering with neutrinos measured the structure of the proton and the neutron via weak interactions; today, we have a quite precise picture of the structure of these hadrons as seen by deep inelastic scattering probes. Indeed, precision measurements of the so-called structure function provide vital input data to modelling proton-proton collisions performed at the Large Hadron Collider.

7.3.8 Quark masses

The masses of the quarks cannot be determined from deep inelastic scattering as they turn out to be much smaller than the energies involved. In addition, due to absence of free quarks, *the quark masses depend on the precise definition used*. In the sense of mass parameters occurring in the QCD Lagrangian discussed later in Section 12.3 they can be determined by analysing the breakdown of flavour symmetries. The result for the strange quark mass comes from the size of $SU(3)$ breaking and the present best value is

$$m_s = 95 \pm 5 \text{ MeV} \quad (7.232)$$

The u and d quark masses are much smaller, and their determination involves the proton-neutron mass difference, where one must also separate the contribution of the electromagnetic interaction. The present values are

$$\begin{aligned} m_u &= 2.3 \pm 0.7 \pm 0.5 \text{ MeV} \\ m_d &= 4.8 \pm 0.5 \pm 0.3 \text{ MeV} \end{aligned} \quad (7.233)$$

where the first uncertainty is statistical, while the second is systematic.

It is clear that the above quark masses *cannot account for the hadron masses*, since they are too small. Therefore inside the hadrons the light quarks u, d, s are *highly relativistic objects* which have energies of order of hundreds of MeV. As a result, *hadrons cannot be described by ordinary quantum mechanics; they are very dynamic objects in which quarks, anti-quarks and gluons continuously*

created and annihilated. When considering the proton as uud , the two u s and the d are called valence quarks; they account for all quantum numbers of the proton, but *the inside of the particle is filled with a very dynamic relativistic quantum “mess”, consisting of quark-antiquark pairs and gluons, called sea partons as they originate from the quantum vacuum “sea”.* This is in contrast with atomic physics, where atomic structure can be described well by non-relativistic quantum mechanics. As discussed in Subsection 3.1.1, the simpler nature of atomic physics is due to the small value of the characteristic strength of electromagnetic interaction given by the fine structure constant $\alpha \approx 1/137$. In strong interactions, however, *the characteristic coupling strength is much larger and becomes very strong at low energies* (the energy scales corresponding to the bound states i.e. the hadrons); cf. Subsection 12.3.4.

On the other hand, the successful description of the magnetic moment of baryons (mentioned in Subsection 7.3.6) allows one to derive so-called *constituent masses* for these valence quarks, with the result²⁰

$$m_u \approx 338 \text{ MeV} \quad m_d \approx 322 \text{ MeV} \quad m_s \approx 510 \text{ MeV} \quad (7.234)$$

These can be interpreted as the full mass of the valence quarks acquired from the interacting medium inside the hadron. However, they also depend on the model they are derived from as a comparison with the values quoted later in Table 11.2 shows.

The different values of the quark masses obtained from symmetry breaking and from the constituent picture show that the strong interaction confining the quarks into the tight internal space of hadrons contributes a huge additional energy of several 100 MeV per quark to the mass of hadrons; this order of magnitude is consistent with the characteristic scale of quantum chromodynamics (QCD) discussed in Subsection 12.3.3. Indeed, the latest QCD calculations reproduce the spectrum of hadrons very well (cf. Fig 12.4), confirming the picture of hadrons as bound state of quarks and gluons.

²⁰For more details cf. Section 8.13 of W. Greiner and B. Müller: *Quantum Mechanics – Symmetries*.

8 Lagrangian field theory

8.1 Locality and fields

In a relativistic theory, there *can be no causal relation between events that are space-like separated*. To show that, let us take two space-time events A and B , with the four-vector

$$x^\mu = (c(t_A - t_B), \vec{x}_A - \vec{x}_B) \quad (8.1)$$

separating them space-like, i.e.

$$x^\mu x_\mu = c^2(t_A - t_B)^2 - (\vec{x}_A - \vec{x}_B)^2 < 0 \quad (8.2)$$

Note that therefore

$$|\vec{x}_A - \vec{x}_B| > c|t_A - t_B| \quad (8.3)$$

so that any causal connection between space-like separated events is equivalent to *superluminal signalling*. To simplify calculations, we take the x axis parallel to $\vec{x}_A - \vec{x}_B$ so that

$$x^\mu = (c\Delta t, \Delta x, 0, 0) \quad (8.4)$$

with $\Delta x > 0$. Let us suppose $\Delta t > 0$ as well, meaning that event A is later than B in reference system K .

In an inertial system K' moving with velocity v in the x direction:

$$\Delta t' = \frac{\Delta t - v\Delta x/c^2}{\sqrt{1 - v^2/c^2}} \quad \Delta x' = \frac{\Delta x - v\Delta t}{\sqrt{1 - v^2/c^2}} \quad (8.5)$$

Now $\Delta t' < 0$ is possible provided

$$v > c^2 / \left(\frac{\Delta x}{\Delta t} \right) \quad (8.6)$$

but this velocity is itself less than c since

$$\frac{\Delta x}{\Delta t} > c \quad (8.7)$$

So the *causal ordering of space-like events depends on the reference system*. Since physics is assumed to be relativistically invariant, *there cannot be physical influence between space-like separated events*. There even exists a system in which $\Delta t' = 0$ i.e. the two events are simultaneous.

This is called the

Principle of locality *Events having space-like separation cannot have causal connections.*

Note also that the same reasoning shows that the temporal order of time-like and light-like separated events is absolute. As a result, there cannot exist physical effects that propagate faster than light, i.e. *there is no superluminal signalling*.

This forces us to consider physical observable to be indexed by position and time, i.e. the *fundamental objects must be fields* $\Phi_\alpha(t, \vec{x})$ since it is necessary to keep a record of their causal relations. In quantum theory the observables are operators on a Hilbert space, therefore one is led to the introduction of *operator valued fields* $\hat{\Phi}_\alpha(t, \vec{x})$ to describe the observables in a relativistic quantum theory. However, two observables are only independent when the operators commute, so locality can be formulated as

Principle of locality (also called micro-causality) *Observable fields must commute at space-like separated space-time points:*

$$\left[\hat{\Phi}_\alpha(x), \hat{\Phi}_{\alpha'}(x') \right] = 0 \quad \text{whenever} \quad (x - x')^2 < 0 \quad (8.8)$$

This principle is the basis of *relativistic quantum field theory* which is the basic theoretical framework underlying elementary particle physics, but also plays an important role in the description of quantum many-body systems that occur in condensed matter and statistical mechanics. More generally, quantum field theory is the paradigm that forms the basis of the theory of quantum many-body systems, including non-relativistic ones.

Here and in Section [9](#) we review the necessary background in classical field theory; the basic concepts of quantum field theory are introduced in Section [10](#).

8.2 Action and Euler-Lagrange equations

The simplest way to specify field theory dynamics is via an action principle^{[21](#)} a properly formulated action ensures relativistic invariance and causality. Let us first recall the action for a mechanical system:

$$S = \int_{t_1}^{t_2} L(q_i(t), \dot{q}_i(t), t) dt \quad (8.9)$$

where L is the Lagrange function. We assume that the action only depends on the generalised coordinates q_i and their first derivatives $\dot{q}_i$, so demanding $\delta S = 0$ under a variation

$$\begin{aligned} q_i(t) &\rightarrow q_i(t) + \delta q_i(t) \\ \delta q_i(t_1) &= 0 \quad \delta q_i(t_2) = 0 \end{aligned} \quad (8.10)$$

leading to the Lagrange equation of motion

$$\frac{\partial L}{\partial q_i} - \frac{d}{dt} \frac{\partial L}{\partial \dot{q}_i} = 0 \quad (8.11)$$

which is second order in derivatives.

In field theory, locality dictates that the Lagrange function is integral of a *Lagrangian density* $\mathcal{L}$

$$L = \int d^3x \mathcal{L} \quad (8.12)$$

so the action is of the form

$$S = \int d^4x \mathcal{L} \quad (8.13)$$

²¹We note that there are other ways to formulate relativistic quantum field theories. These include: the so-called axiomatic field theory based on the Wightmann axioms, the algebraic quantum field theory based on the theory of C^* algebras, and other types of axiomatic approaches such as the one used in conformal field theory. In addition, it is also possible to formulate a relativistic theory of particle interaction based on the scattering matrix, the so-called bootstrap approach which can be carried to full success in the case of so-called integrable quantum field theories. However, all of these approaches incorporate insights from, and also (partially) build upon, the Lagrangian approach discussed here.

To obtain a Lorentz invariant dynamics, the action S and therefore also $\mathcal{L}$ must be Lorentz invariant so $\mathcal{L}$ must depend on the first spatial derivatives as well:

$$S = \int d^4x \mathcal{L}(\Phi_\alpha, \partial_\mu \Phi_\alpha) \quad (8.14)$$

where we also imposed that the system is translation invariant so $\mathcal{L}$ does not depend explicitly on x . The variation of the action can be evaluated as

$$\begin{aligned} \delta S &= \int d^4x \mathcal{L}(\Phi_\alpha + \delta\Phi_\alpha, \partial_\mu \Phi_\alpha + \partial_\mu \delta\Phi_\alpha) - \int d^4x \mathcal{L}(\Phi_\alpha, \partial_\mu \Phi_\alpha) \\ &= \int d^4x \left(\frac{\partial \mathcal{L}}{\partial \Phi_\alpha} \delta\Phi_\alpha + \frac{\partial \mathcal{L}}{\partial (\partial_\mu \Phi_\alpha)} \partial_\mu \delta\Phi_\alpha \right) \\ &= \int d^4x \left(\frac{\partial \mathcal{L}}{\partial \Phi_\alpha} - \partial_\mu \frac{\partial \mathcal{L}}{\partial (\partial_\mu \Phi_\alpha)} \right) \delta\Phi_\alpha + \int d^4x \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial (\partial_\mu \Phi_\alpha)} \delta\Phi_\alpha \right) \end{aligned} \quad (8.15)$$

Using Gauss's theorem, the second term can be cast into an integral over the boundary $\partial\Omega$ of the space time domain Ω in which the dynamics is considered

$$\int_\Omega d^4x \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial (\partial_\mu \Phi_\alpha)} \delta\Phi_\alpha \right) = \int_{\partial\Omega} d^3S_\mu \left(\frac{\partial \mathcal{L}}{\partial (\partial_\mu \Phi_\alpha)} \delta\Phi_\alpha \right) \quad (8.16)$$

where d^3S_μ is the normal vector oriented surface element of the boundary $\partial\Omega$. Now extending Ω to cover the whole Minkowski space-time and assuming

$$\begin{aligned} \delta\Phi_\alpha(t, \vec{x}) &\longrightarrow_{|t| \rightarrow \infty} 0 \\ \delta\Phi_\alpha(t, \vec{x}) &\longrightarrow_{|\vec{x}| \rightarrow \infty} 0 \end{aligned} \quad (8.17)$$

the surface term can be dropped, resulting in the *Euler-Lagrange equations of motion*

$$\frac{\partial \mathcal{L}}{\partial \Phi_\alpha} - \partial_\mu \frac{\partial \mathcal{L}}{\partial (\partial_\mu \Phi_\alpha)} = 0 \quad (8.18)$$

Exercise 8.1. Show that shifting the Lagrangian density by a total divergence

$$\mathcal{L} \rightarrow \mathcal{L} + \partial_\mu \mathcal{K}^\mu \quad (8.19)$$

leaves the equations of motion invariant.

Solution. The action is modified according to

$$\int_\Omega d^4x \mathcal{L} \rightarrow \int_\Omega d^4x \mathcal{L} + \int_{\partial\Omega} d^3S_\mu \mathcal{K}^\mu \quad (8.20)$$

The second term does not contribute to the variation of action since the field variations $\delta\Phi_\alpha$ vanish on $\partial\Omega$ when extending Ω to the whole Minkowski space. $\square$

8.2.1 The Klein-Gordon action for real and complex scalar fields

Let us consider the Lagrangian density

$$\mathcal{L} = \frac{1}{2} \partial_\mu \Phi \partial^\mu \Phi - \frac{m^2}{2} \Phi^2 \quad (8.21)$$

with a single real field Φ .

Exercise 8.2. Show that the above Lagrangian gives the Klein-Gordon equation as its equation of motion.

Solution. We can write

$$\mathcal{L} = \frac{1}{2} (\partial_0 \Phi)^2 - \frac{1}{2} (\vec{\nabla} \Phi)^2 - \frac{m^2}{2} \Phi^2 \quad (8.22)$$

so

$$\frac{\partial \mathcal{L}}{\partial(\partial_0 \Phi)} = \partial_0 \Phi = \partial^0 \Phi \quad \frac{\partial \mathcal{L}}{\partial(\partial_i \Phi)} = -\partial_i \Phi = \partial^i \Phi \quad (8.23)$$

i.e.

$$\frac{\partial \mathcal{L}}{\partial(\partial_\mu \Phi)} = \partial^\mu \Phi \quad (8.24)$$

We also have

$$\frac{\partial \mathcal{L}}{\partial \Phi} = -m^2 \Phi \quad (8.25)$$

so the equation of motion is

$$\frac{\partial \mathcal{L}}{\partial \Phi_\alpha} - \partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu \Phi_\alpha)} = -m^2 \Phi - \partial_\mu \partial^\mu \Phi = -(\square + m^2) \Phi = 0 \quad (8.26)$$

which is the Klein-Gordon equation in units $\hbar = c = 1$. □

One can also write the Lagrangian for a complex Klein-Gordon field Φ as

$$\mathcal{L} = \partial_\mu \Phi^* \partial^\mu \Phi - m^2 \Phi^* \Phi \quad (8.27)$$

There are two ways to see that this indeed gives the Klein-Gordon equation for Φ : the first is to write

$$\Phi = \frac{1}{\sqrt{2}} (\phi_1 + i\phi_2) \quad (8.28)$$

with $\phi_{1,2}$ real, which then gives

$$\mathcal{L} = \frac{1}{2} \partial_\mu \phi_1 \partial^\mu \phi_1 - \frac{m^2}{2} \phi_1^2 + \frac{1}{2} \partial_\mu \phi_2 \partial^\mu \phi_2 - \frac{m^2}{2} \phi_2^2 \quad (8.29)$$

So the real fields $\phi_{1,2}$ satisfy the Klein-Gordon equation and so does Φ .

The other way is to consider Φ and its conjugate Φ^* as independent variables instead of the real and the imaginary part of Φ and take the variation accordingly:

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial(\partial_\mu \Phi)} &= \partial^\mu \Phi^* & \frac{\partial \mathcal{L}}{\partial \Phi} &= -m^2 \Phi^* \\ \frac{\partial \mathcal{L}}{\partial(\partial_\mu \Phi^*)} &= \partial^\mu \Phi & \frac{\partial \mathcal{L}}{\partial \Phi^*} &= -m^2 \Phi \end{aligned} \quad (8.30)$$

so the equations of motion corresponding to the variations of Φ and Φ^* are

$$\begin{aligned}\delta\Phi &: -(\square + m^2)\Phi^* = 0 \\ \delta\Phi^* &: -(\square + m^2)\Phi = 0\end{aligned}\tag{8.31}$$

This is true in general: *for a complex field one can use the variational principle in terms of the field and its complex conjugate* instead of decomposing it into a real and an imaginary part.

8.2.2 The Dirac action

The Lagrangian which gives the Dirac equation is

$$\mathcal{L} = \bar{\Psi} (i\gamma^\mu \partial_\mu - m) \Psi\tag{8.32}$$

Since

$$\bar{\Psi} = \Psi^\dagger \gamma^0\tag{8.33}$$

variation with respect to Ψ^* can be substituted by variation with respect to $\bar{\Psi}$.

Exercise 8.3. Compute the Euler-Lagrange equations from the Dirac Lagrangian.

Solution. We can write

$$\frac{\partial \mathcal{L}}{\partial(\partial_\mu \bar{\Psi})} = 0 \quad \frac{\partial \mathcal{L}}{\partial \bar{\Psi}} = (i\gamma^\mu \partial_\mu - m) \Psi\tag{8.34}$$

so the equation of motion from the variation of $\bar{\Psi}$ is the Dirac equation itself

$$(i\gamma^\mu \partial_\mu - m) \Psi = 0\tag{8.35}$$

On the other hand

$$\frac{\partial \mathcal{L}}{\partial(\partial_\mu \Psi)} = i\bar{\Psi} \gamma^\mu \quad \frac{\partial \mathcal{L}}{\partial \Psi} = -m\bar{\Psi}\tag{8.36}$$

so the equation of motion from the variation of Ψ gives the conjugate Dirac equation

$$i\partial_\mu \bar{\Psi} \gamma^\mu + m\bar{\Psi} = \bar{\Psi} \left(i\gamma^\mu \overleftarrow{\partial}_\mu + m \right) = 0\tag{8.37}$$

again in units $\hbar = 1 = c$. □

8.3 Symmetries and Noether's theorem

An important fact in Lagrangian dynamics is that symmetries lead to conserved quantities, according to a theorem by Emmy Noether, a German Jewish mathematician who made outstanding contributions to abstract algebra and theoretical physics. The theorem is true both in mechanics and field theory; here we formulate and prove the field theoretic version.

Theorem. (*Noether's theorem for field theory*) *Let us suppose that there exists a one-parameter family of transformations $T(\epsilon)$ which forms a group*

$$T(\epsilon_1)T(\epsilon_2) = T(\epsilon_1 + \epsilon_2)\tag{8.38}$$

under which the fields transform as

$$\Phi_\alpha(x) \xrightarrow{T(\epsilon)} \Phi_\alpha(x) + \epsilon \Delta \Phi_\alpha(x) + O(\epsilon^2) \quad (8.39)$$

where we assumed that the transformation is infinitesimal i.e. can be expanded to first order in the parameter ϵ . Let us also assume the transformations preserve the action, i.e. the Lagrangian density is changed by a total derivative

$$\mathcal{L} \xrightarrow{T(\epsilon)} \mathcal{L} + \epsilon \partial_\mu \mathcal{K}^\mu + O(\epsilon^2) \quad (8.40)$$

In this case there exists a current with the explicit form

$$j^\mu = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \Phi_\alpha)} \Delta \Phi_\alpha - \mathcal{K}^\mu \quad (8.41)$$

which satisfies the continuity equation

$$\partial_\mu j^\mu = 0 \quad (8.42)$$

and its conservation follows from the equations of motion.

Proof. The proof of the above statement follows from examining

$$\mathcal{L}(\Phi_\alpha(x) + \epsilon \Delta \Phi_\alpha(x), \partial_\mu \Phi_\alpha(x) + \epsilon \partial_\mu \Delta \Phi_\alpha(x)) = \mathcal{L}(\Phi_\alpha(x), \partial_\mu \Phi_\alpha(x)) + \epsilon \partial_\mu \mathcal{K}^\mu + O(\epsilon^2) \quad (8.43)$$

which can be rewritten as

$$\frac{\partial \mathcal{L}}{\partial \Phi_\alpha} \Delta \Phi_\alpha + \frac{\partial \mathcal{L}}{\partial(\partial_\mu \Phi_\alpha)} \partial_\mu \Delta \Phi_\alpha = \partial_\mu \mathcal{K}^\mu \quad (8.44)$$

Using the Leibniz rule

$$\frac{\partial \mathcal{L}}{\partial(\partial_\mu \Phi_\alpha)} \partial_\mu \Delta \Phi_\alpha = \partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \Phi_\alpha)} \Delta \Phi_\alpha \right) - \partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu \Phi_\alpha)} \Delta \Phi_\alpha \quad (8.45)$$

Substituting this into the previous formula and using the Euler-Lagrange equations

$$\frac{\partial \mathcal{L}}{\partial \Phi_\alpha} - \partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu \Phi_\alpha)} = 0 \quad (8.46)$$

leads to

$$\partial_\mu \left(\frac{\partial \mathcal{L}}{\partial(\partial_\mu \Phi_\alpha)} \Delta \Phi_\alpha \right) = \partial_\mu \mathcal{K}^\mu \quad (8.47)$$

which is just

$$\partial_\mu j^\mu = 0 \quad \text{with} \quad j^\mu = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \Phi_\alpha)} \Delta \Phi_\alpha - \mathcal{K}^\mu \quad (8.48)$$

□

For the above theorem it follows that the quantity

$$Q(t) = \int d^3x j^0(t, \vec{x}) \quad (8.49)$$

is conserved:

$$\frac{dQ}{dt} = 0 \quad (8.50)$$

It is important to note that conservation of $Q(t)$ is a consequence of the dynamics corresponding to the equations of motions. Such conserved quantities are called *Noether charges*. We now proceed to demonstrate the power of this theorem by considering the most important conservation laws.

8.3.1 Energy-momentum tensor

Let us consider translation invariance in both space and time. The transformation $T_a(\epsilon)$ is

$$x^\mu \rightarrow x^\mu + \epsilon a^\mu \quad (8.51)$$

where a^μ is a constant 4-vector that gives the direction of the translation and ϵ is the transformation's parameter which is taken to be infinitesimal. Under translations all fields transform as

$$\Phi'_\alpha(x) = \Phi_\alpha(x - \epsilon a^\mu) = \Phi_\alpha(x) - \epsilon a^\mu \partial_\mu \Phi_\alpha(x) + O(\epsilon^2) \quad (8.52)$$

so

$$\Delta \Phi_\alpha = -a^\mu \partial_\mu \Phi_\alpha \quad (8.53)$$

In particular

$$\Delta \mathcal{L} = -a^\mu \partial_\mu \mathcal{L} \quad (8.54)$$

since $\mathcal{L}$ is a field itself; therefore

$$\mathcal{K}^\mu = -a^\mu \mathcal{L} \quad (8.55)$$

As a result

$$j^\mu = -\frac{\partial \mathcal{L}}{\partial(\partial_\mu \Phi_\alpha)} a^\nu \partial_\nu \Phi_\alpha + a^\mu \mathcal{L} = -a_\nu T^{\mu\nu} \quad (8.56)$$

where the quantity

$$T^{\mu\nu} = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \Phi_\alpha)} \partial^\nu \Phi_\alpha - \eta^{\mu\nu} \mathcal{L} \quad (8.57)$$

satisfies

$$\partial_\mu T^{\mu\nu} = 0 \quad (8.58)$$

What is the interpretation of the conserved quantities

$$P^\nu = \int d^3x T^{0\nu} \quad (8.59)$$

corresponding to translation invariance? These form a 4-vector with the temporal component

$$P^0 = \int d^3x T^{00} = \int d^3x \left(\frac{\partial \mathcal{L}}{\partial(\partial_0 \Phi_\alpha)} \partial_0 \Phi_\alpha - \mathcal{L} \right) \quad (8.60)$$

This resembles very much the expression

$$\frac{\partial L}{\partial \dot{q}_i} \dot{q}_i - L \quad (8.61)$$

which gives the total energy in Lagrangian mechanics. Indeed one can think of a field as a collection of canonical coordinates one for each space time points

$$q_{\vec{x},\alpha}(t) = \Phi_\alpha(t, \vec{x}) \quad (8.62)$$

The canonical momenta are then

$$p_{\vec{x},\alpha}(t) = \Pi_\alpha(t, \vec{x}) = \frac{\partial \mathcal{L}}{\partial(\partial_0 \Phi_\alpha)} \quad (8.63)$$

and the canonical energy expression can be written as

$$H = \sum_{\vec{x}, \alpha} p_{\vec{x}, \alpha} \dot{q}_{\vec{x}, \alpha} - L = \left(\sum_{\alpha} \int d^3x \frac{\partial \mathcal{L}}{\partial(\partial_0 \Phi_{\alpha})} \partial_0 \Phi_{\alpha} - \int d^3x \mathcal{L} \right) \equiv P_0 \quad (8.64)$$

The expression

$$\mathcal{H} = \left(\frac{\partial \mathcal{L}}{\partial(\partial_0 \Phi_{\alpha})} \partial_0 \Phi_{\alpha} - \mathcal{L} \right) \quad (8.65)$$

is called the *Hamiltonian density*. From special relativity we then know that the other components

$$P_i = \int d^3x \Pi_{\alpha} \partial_i \Phi_{\alpha} \quad (8.66)$$

form the *three-momentum*.

Due to the continuity equations, the components

$$T^{i0} = \frac{\partial \mathcal{L}}{\partial(\partial_i \Phi_{\alpha})} \partial^0 \Phi_{\alpha} \quad (8.67)$$

form the *energy current density 3-vector*²² while the components T^{ij} give the *current density of the three momentum components* $j = 1, 2, 3$ when considered as 3-vectors in i .

In general, *the energy-momentum tensor is not uniquely defined*, since $\mathcal{L}$ can be changed by a complete 4 divergence without altering the dynamics of the system. The most general possible redefinition has the form

$$T^{\mu\nu} \rightarrow T^{\mu\nu} + \partial_{\lambda} B^{\lambda\mu\nu} \quad (8.68)$$

where B is a rank three tensor that is antisymmetric in its first two indices

$$B^{\lambda\mu\nu} = -B^{\mu\lambda\nu} \quad (8.69)$$

This keeps the physical content of the energy-momentum intact. First of all, the new tensor is still conserved

$$\partial_{\mu} (T^{\mu\nu} + \partial_{\lambda} B^{\lambda\mu\nu}) = \partial_{\mu} T^{\mu\nu} + \partial_{\mu} \partial_{\lambda} B^{\lambda\mu\nu} = \partial_{\mu} T^{\mu\nu} = 0 \quad (8.70)$$

where we used that the second partial derivative is symmetric under interchanging the order of derivatives.

Exercise 8.4. Show that the redefined energy-momentum tensor gives the same conserved energy and momentum.

Solution. We can write

$$\begin{aligned} P^{\nu} &= \int d^3x T^{0\nu} = \int d^3x (T^{0\nu} + \partial_{\lambda} B^{\lambda 0\nu}) \\ &= \int d^3x (T^{0\nu} + \partial_i B^{i 0\nu}) \end{aligned} \quad (8.71)$$

since $B^{00\nu} = 0$. Applying Gauss's theorem

$$\int_V d^3x \partial_i B^{i 0\nu} = \oint_{\partial V} d^2f_i B^{i 0\nu} \quad (8.72)$$

²²In the case of electrodynamics, this is the Poynting vector.

and assuming that all fields satisfy the asymptotic property

$$\Phi_\alpha(t, \vec{x}) \xrightarrow{|\vec{x}| \rightarrow \infty} 0 \quad (8.73)$$

and also that the densities such as B formed from the fields decay sufficiently fast, in the infinite volume limit the surface integral term is 0. Therefore

$$P^\nu = P^\nu \quad (8.74)$$

(For a finite system, the correct assumption is that the system is closed so that all fluxes through the boundary vanish.) $\square$

As a result, the above redefinition leaves the conserved quantities invariant, which is in fact the physical content of the energy-momentum tensor. This shows that the energy-momentum tensor is not defined uniquely. In some cases, one prefers a different energy-momentum than the canonical one. For example in general relativity the Einstein equation reads

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R = \kappa T_{\mu\nu} \quad (8.75)$$

Now the geometric tensors on the left hand side (the metric tensor $g_{\mu\nu}$ and the Ricci tensor $R_{\mu\nu}$) are symmetric, so the energy-momentum tensor must also be chosen symmetric. This is not guaranteed by the canonical expression

$$T^{\mu\nu} = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \Phi_\alpha)} \partial^\nu \Phi_\alpha - \eta^{\mu\nu} \mathcal{L} \quad (8.76)$$

for an arbitrary choice of the Lagrangian density $\mathcal{L}$. The above freedom can be used to introduce a symmetric $T^{\mu\nu}$ that can serve as a source for the gravitational field described as the curvature of space-time; the appropriate choice for the Einstein equation is the so-called Belinfante-Rosenfeld energy-momentum tensor.

Exercise 8.5. Compute the energy-momentum tensor for a scalar field governed by the Lagrangian

$$\mathcal{L} = \frac{1}{2} \partial_\mu \Phi \partial^\mu \Phi - V(\Phi) \quad (8.77)$$

Solution.

$$\frac{\partial \mathcal{L}}{\partial(\partial_\mu \Phi)} = \partial^\mu \Phi \quad (8.78)$$

so

$$T^{\mu\nu} = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \Phi)} \partial^\nu \Phi - \eta^{\mu\nu} \mathcal{L} = \partial^\mu \Phi \partial^\nu \Phi - \eta^{\mu\nu} \left(\frac{1}{2} \partial_\mu \Phi \partial^\mu \Phi - V(\Phi) \right) \quad (8.79)$$

$\square$

Exercise 8.6. Compute the energy-momentum tensor for a Dirac field governed by the Lagrangian

$$\mathcal{L} = \bar{\Psi} (i \not{\partial} - m) \Psi \quad (8.80)$$

Solution.

$$\frac{\partial \mathcal{L}}{\partial(\partial_\mu \Psi)} = i\bar{\Psi}\gamma^\mu \quad \frac{\partial \mathcal{L}}{\partial(\partial_\mu \bar{\Psi})} = 0 \quad (8.81)$$

so

$$T^{\mu\nu} = \frac{\partial \mathcal{L}}{\partial(\partial_\mu \Psi)} \partial^\nu \Psi - \eta^{\mu\nu} \mathcal{L} = i\bar{\Psi}\gamma^\mu \partial^\nu \Psi - \eta^{\mu\nu} (\bar{\Psi} (i\not{\partial} - m) \Psi) \quad (8.82)$$

and using the equations of motion one can eliminate the second term to give

$$T^{\mu\nu} = i\bar{\Psi}\gamma^\mu \partial^\nu \Psi \quad (8.83)$$

□

The above result for the Dirac field is not the most convenient one. The resulting energy-momentum tensor is not symmetric and not even real! The reason for the latter is that $\mathcal{L}$ is not real to start with, as shown by the following calculation

$$\begin{aligned} \mathcal{L}^* &= [\Psi^\dagger \gamma^0 (i\gamma^\mu \partial_\mu - m) \Psi]^* = \Psi^\dagger (-i\gamma^{\mu\dagger} \overleftarrow{\partial}_\mu - m) \gamma^0 \Psi \\ &= \Psi^\dagger (-i\gamma^0 \gamma^\mu \gamma^0 \overleftarrow{\partial}_\mu - m) \gamma^0 \Psi = \Psi^\dagger \gamma^0 (-i\gamma^\mu \overleftarrow{\partial}_\mu - m) \Psi \\ &\quad \bar{\Psi} (-i\not{\partial} - m) \Psi \end{aligned} \quad (8.84)$$

However, the difference $\mathcal{L} - \mathcal{L}^*$ is a total derivative:

$$\mathcal{L} - \mathcal{L}^* = i (\bar{\Psi} \not{\partial} \Psi + \bar{\Psi} \overleftarrow{\not{\partial}} \Psi) = \partial_\mu (i\bar{\Psi} \gamma^\mu \Psi) \quad (8.85)$$

so one can introduce an equivalent real Lagrangian by

$$\mathcal{L}' = \frac{1}{2} (\mathcal{L} + \mathcal{L}^*) = \frac{i}{2} (\bar{\Psi} \not{\partial} \Psi - \bar{\Psi} \overleftarrow{\not{\partial}} \Psi) - m\bar{\Psi} \Psi \quad (8.86)$$

Exercise 8.7. Compute the energy-momentum tensor for a Dirac field governed by the Lagrangian $\mathcal{L}'$.

Solution.

$$\frac{\partial \mathcal{L}'}{\partial(\partial_\mu \Psi)} = \frac{i}{2} \bar{\Psi} \gamma^\mu \quad \frac{\partial \mathcal{L}'}{\partial(\partial_\mu \bar{\Psi})} = -\frac{i}{2} \gamma^\mu \Psi \quad (8.87)$$

so

$$T'^{\mu\nu} = \frac{\partial \mathcal{L}'}{\partial(\partial_\mu \Psi)} \partial^\nu \Psi + \partial^\nu \bar{\Psi} \frac{\partial \mathcal{L}'}{\partial(\partial_\mu \bar{\Psi})} - \eta^{\mu\nu} \mathcal{L}' \quad (8.88)$$

Note the position of the $\bar{\Psi}$ factor in the second term which is dictated by the fact that it is a row spinor, while

$$\frac{\partial \mathcal{L}}{\partial(\partial_\mu \bar{\Psi})} \quad (8.89)$$

is a column spinor, so the summation over the spinor indices implies the order of the product when written in matrix notation. This leads to

$$T'^{\mu\nu} = \frac{i}{2} \bar{\Psi} (\gamma^\mu \partial^\nu - \gamma^\mu \overleftarrow{\partial}^\nu) \Psi - \eta^{\mu\nu} \mathcal{L}' \quad (8.90)$$

The second term can again be eliminated using the equations of motions

$$(i\not{\partial} - m)\Psi = 0 \quad \bar{\Psi}(i\overleftarrow{\not{\partial}} - m) = 0 \quad (8.91)$$

leaving

$$T'^{\mu\nu} = \frac{i}{2}\bar{\Psi}\left(\gamma^\mu\partial^\nu - \gamma^\mu\overleftarrow{\partial}^\nu\right)\Psi = \frac{i}{2}\bar{\Psi}\gamma^\mu\overleftrightarrow{\partial}^\nu\Psi \quad (8.92)$$

□

The above expression for the energy-momentum tensor is still not symmetric, but at least it is real.

8.3.2 Angular momentum conservation

In mechanics, *angular momentum conservation is a consequence of rotational invariance*. In relativistic field theory, rotations are part of the Lorentz group, so one must consider an *arbitrary Lorentz transformation*. Under a Lorentz transformation

$$x' = \Lambda x \quad (8.93)$$

the fields transforms as

$$\Phi'_\alpha(x) = S(\Lambda)_\alpha^\beta \Phi_\beta(\Lambda^{-1}x) \quad (8.94)$$

where $S(\Lambda)$ describes the Lorentz property of the quantity corresponding to the field itself. This is analogous to electrodynamics where a spatial rotation $\mathcal{O}$ also rotates the electric and magnetic fields which are 3-vectors:

$$\vec{E}'(\vec{x}) = \mathcal{O}\vec{E}(\mathcal{O}^{-1}\vec{x}) \quad \vec{B}'(\vec{x}) = \mathcal{O}\vec{B}(\mathcal{O}^{-1}\vec{x}) \quad (8.95)$$

For example if our field is a covariant tensor of rank r then it transforms as

$$\Phi'_{\mu_1 \dots \mu_r}(x) = \Lambda_{\mu_1}^{\nu_1} \dots \Lambda_{\mu_r}^{\nu_r} \Phi_{\nu_1 \dots \nu_r}(\Lambda^{-1}x) \quad (8.96)$$

For integer spin fields, the $S(\Lambda)$ form a representation of the Lorentz group

$$S(\Lambda_1)S(\Lambda_2) = S(\Lambda_1\Lambda_2) \quad (8.97)$$

In the case of the Dirac field S is the spinor transformation matrix introduced earlier

$$\Psi'_{\xi'}(x) = S(\Lambda)_{\xi'}^\xi \Psi_\xi(\Lambda^{-1}x) \quad (8.98)$$

where ξ, ξ' are the bispinor indices, and in this case

$$S(\Lambda_1)S(\Lambda_2) = \pm S(\Lambda_1\Lambda_2) \quad (8.99)$$

as discussed in Subsection [5.4.1](#)

To derive the Noether current corresponding to Lorentz transformation, consider an infinitesimal transformation of the form

$$\Lambda_\nu^\mu = \delta_\nu^\mu + \epsilon \omega_\nu^\mu \quad (8.100)$$

where

$$\omega_{\mu\nu} = -\omega_{\nu\mu} \quad (8.101)$$

and ϵ is the small parameter. Then

$$(\Lambda^{-1})^\mu_\nu = \delta^\mu_\nu - \epsilon \omega^\mu_\nu \Rightarrow (\Lambda^{-1}x)^\mu = x^\mu - \epsilon \omega^\mu_\nu x^\nu \quad (8.102)$$

and one can also write

$$S(\Lambda)_\alpha^\beta = \mathbf{1} - \frac{i}{2} \epsilon \omega_{\mu\nu} (\mathcal{J}^{\mu\nu})_\alpha^\beta + O(\omega^2) \quad (8.103)$$

where the $\mathcal{J}^{\mu\nu} = -\mathcal{J}^{\nu\mu}$ are the infinitesimal generators of the Lorentz transformation acting on the field. For a Klein-Gordon field $\mathcal{J}^{\mu\nu} = 0$, while for a Dirac field they are

$$\mathcal{J}^{\mu\nu} = \frac{1}{2} \sigma^{\mu\nu} \quad (8.104)$$

As a result

$$\begin{aligned} \delta\Phi_\alpha(x) &= -\epsilon \omega^\mu_\nu x^\nu \partial_\mu \Phi_\alpha(x) - \frac{i}{2} \epsilon \omega_{\mu\nu} (\mathcal{J}^{\mu\nu})_\alpha^\beta \Phi_\beta(x) \\ &= \epsilon \Delta_\omega \Phi_\alpha(x) \end{aligned} \quad (8.105)$$

where

$$\Delta_\omega \Phi_\alpha(x) = -\frac{1}{2} \omega_{\mu\nu} \left[(x^\nu \partial^\mu - x^\mu \partial^\nu) \Phi_\alpha(x) + i (\mathcal{J}^{\mu\nu})_\alpha^\beta \Phi_\beta(x) \right] \quad (8.106)$$

The Lagrangian density itself is a Lorentz scalar field, so it transforms according to

$$\begin{aligned} \delta\mathcal{L}(x) &= -\frac{1}{2} \epsilon \omega_{\mu\nu} (x^\nu \partial^\mu - x^\mu \partial^\nu) \mathcal{L}(x) = \epsilon \partial_\lambda \mathcal{K}_\omega^\lambda \\ \text{with} \quad \mathcal{K}_\omega^\lambda &= -\frac{1}{2} \omega_{\mu\nu} (x^\nu \eta^{\lambda\mu} - x^\mu \eta^{\lambda\nu}) \mathcal{L} \end{aligned} \quad (8.107)$$

Exercise 8.8. Compute the Noether current J_ω^λ according to the above transformations.

Solution.

$$\begin{aligned} J_\omega^\lambda &= \frac{\partial \mathcal{L}}{\partial(\partial_\lambda \Phi_\alpha)} \Delta_\omega \Phi_\alpha - \mathcal{K}_\omega^\lambda \\ &= -\frac{1}{2} \omega_{\mu\nu} \left[\frac{\partial \mathcal{L}}{\partial(\partial_\lambda \Phi_\alpha)} \left[(x^\nu \partial^\mu - x^\mu \partial^\nu) \Phi_\alpha(x) + i (\mathcal{J}^{\mu\nu})_\alpha^\beta \Phi_\beta(x) \right] - (x^\nu \eta^{\lambda\mu} - x^\mu \eta^{\lambda\nu}) \mathcal{L} \right] \\ &= -\frac{1}{2} \omega_{\mu\nu} \left[(x^\nu T^{\lambda\mu} - x^\mu T^{\lambda\nu}) + i \frac{\partial \mathcal{L}}{\partial(\partial_\lambda \Phi_\alpha)} (\mathcal{J}^{\mu\nu})_\alpha^\beta \Phi_\beta(x) \right] \end{aligned} \quad (8.108)$$

where $T^{\mu\nu}$ is the canonical energy momentum tensor. □

One can introduce the following rank-3 tensor

$$\begin{aligned} J^{\lambda,\mu\nu} &= (x^\nu T^{\lambda\mu} - x^\mu T^{\lambda\nu}) + i \frac{\partial \mathcal{L}}{\partial(\partial_\lambda \Phi_\alpha)} (\mathcal{J}^{\mu\nu})_\alpha^\beta \Phi_\beta(x) \\ &= -J^{\lambda,\nu\mu} \end{aligned} \quad (8.109)$$

which is interpreted as the covariant angular momentum 4-current. To see that, compute the corresponding conserved quantities

$$M^{\mu\nu} = \int d^3x J^{0,\mu\nu} \quad (8.110)$$

They consist of two parts. The contribution

$$\int d^3x (x^\nu T^{0\mu} - x^\mu T^{0\nu}) \quad (8.111)$$

is the *orbital angular momentum tensor*. Indeed it is clear that its spatial parts give the usual orbital angular momentum by the formula

$$\begin{aligned} L_1 &= \int d^3x (x^2 T^{03} - x^3 T^{02}) \\ L_2 &= \int d^3x (x^3 T^{01} - x^1 T^{03}) \\ L_3 &= \int d^3x (x^1 T^{02} - x^2 T^{01}) \end{aligned} \quad (8.112)$$

i.e.

$$\epsilon^{ijk} L_k = \int d^3x (x^i T^{0j} - x^j T^{0i}) \quad L_i = \epsilon_{ijk} \int d^3x x^j T^{0k} \quad (8.113)$$

since T^{0k} is the density of the k component of the 3 momentum. The second term

$$S^{\lambda\nu\mu} = i \frac{\partial \mathcal{L}}{\partial (\partial_\lambda \Phi_\alpha)} (\mathcal{J}^{\mu\nu})_\alpha^\beta \Phi_\beta(x) \quad (8.114)$$

is then obviously the *internal contribution to the angular momentum*, which is the quantity usually called *spin*.

Exercise 8.9. Compute the tensor $J^{\lambda,\mu\nu}$ for a free Dirac field starting from

$$\mathcal{L}' = \frac{i}{2} (\bar{\Psi} \not{\partial} \Psi - \bar{\Psi} \overleftarrow{\not{\partial}} \Psi) - m \bar{\Psi} \Psi \quad (8.115)$$

Solution. The internal Lorentz transformation of the spinors is

$$\Psi \rightarrow \left(\mathbf{1} - \frac{i}{4} \omega_{\mu\nu} \sigma^{\mu\nu} \right) \Psi \quad \bar{\Psi} \rightarrow \bar{\Psi} \left(\mathbf{1} + \frac{i}{4} \omega_{\mu\nu} \sigma^{\mu\nu} \right) \quad (8.116)$$

from which

$$\mathcal{J}^{\mu\nu} = \frac{1}{2} \sigma^{\mu\nu} \quad (8.117)$$

and

$$\begin{aligned} \Delta_\omega \Psi &= -\frac{1}{2} \omega_{\mu\nu} \left[(x^\nu \partial^\mu - x^\mu \partial^\nu) \Psi(x) + \frac{i}{2} \sigma^{\mu\nu} \Psi(x) \right] \\ \Delta_\omega \bar{\Psi} &= -\frac{1}{2} \omega_{\mu\nu} \left[(x^\nu \partial^\mu - x^\mu \partial^\nu) \bar{\Psi}(x) - \frac{i}{2} \bar{\Psi}(x) \sigma^{\mu\nu} \right] \end{aligned} \quad (8.118)$$

Using

$$\frac{\partial \mathcal{L}'}{\partial(\partial_\mu \Psi)} = \frac{i}{2} \bar{\Psi} \gamma^\mu \quad \frac{\partial \mathcal{L}'}{\partial(\partial_\mu \bar{\Psi})} = -\frac{i}{2} \gamma^\mu \Psi \quad (8.119)$$

we obtain

$$J^{\lambda, \mu\nu} = (x^\nu T^{\lambda\mu} - x^\mu T^{\lambda\nu}) + \frac{1}{4} \bar{\Psi} \{ \gamma^\lambda, \sigma^{\mu\nu} \} \Psi \quad (8.120)$$

□

Let us consider the spin angular momentum density for the Dirac field. The spin part of $J^{0,jk}$ is

$$\frac{1}{4} \Psi^\dagger \gamma^0 \{ \gamma^0, \sigma^{jk} \} \Psi \quad (8.121)$$

However, γ^0 commutes with

$$\sigma^{jk} = \frac{i}{2} [\gamma^j, \gamma^k] \quad (8.122)$$

and for both the Dirac and Weyl representations we have

$$\sigma^{jk} = \epsilon^{jkl} \Sigma_l \quad \Sigma_l = \begin{pmatrix} \sigma_l & \\ & \sigma_l \end{pmatrix} \quad (8.123)$$

with σ_l , $l = 1, 2, 3$ denoting the usual Pauli matrices. So we obtain

$$\frac{1}{2} \Psi^\dagger \vec{\Sigma} \Psi \quad (8.124)$$

which is indeed the result we expect.

8.3.3 The Belinfante-Rosenfeld energy momentum tensor

We know that

$$\begin{aligned} J^{\lambda, \mu\nu} &= (x^\nu T^{\lambda\mu} - x^\mu T^{\lambda\nu}) + S^{\lambda\mu\nu} \\ S^{\lambda\mu\nu} &= i \frac{\partial \mathcal{L}}{\partial(\partial_\lambda \Phi_\alpha)} (\mathcal{J}^{\mu\nu})_\alpha^\beta \Phi_\beta(x) \end{aligned} \quad (8.125)$$

satisfies

$$\partial_\lambda J^{\lambda, \mu\nu} = 0 \quad (8.126)$$

by virtue of Noether's theorem. More explicitly

$$T^{\nu\mu} - T^{\mu\nu} + \partial_\lambda S^{\lambda\mu\nu} = 0 \quad (8.127)$$

where we used the continuity equation for the energy-momentum tensor. Therefore the antisymmetric part of the stress-energy tensor is given by the divergence of the spin part $S^{\lambda\mu\nu}$.

The Belinfante-Rosenfeld energy momentum tensor is defined by

$$\begin{aligned} \Theta^{\mu\nu} &= T^{\mu\nu} + \partial_\lambda B^{\lambda\mu\nu} \\ B^{\lambda\mu\nu} &= \frac{1}{2} (-S^{\lambda\mu\nu} - S^{\mu\lambda\nu} + S^{\nu\mu\lambda}) \end{aligned} \quad (8.128)$$

Since

$$B^{\lambda\mu\nu} = -B^{\mu\lambda\nu} \quad (8.129)$$

$\Theta^{\mu\nu}$ is as good an energy-momentum tensor as $T^{\mu\nu}$. In addition, it is also symmetric

$$\begin{aligned} \Theta^{\mu\nu} - \Theta^{\nu\mu} &= T^{\mu\nu} - T^{\nu\mu} + \frac{1}{2}\partial_\lambda (-S^{\lambda\mu\nu} - S^{\mu\lambda\nu} + S^{\nu\mu\lambda} + S^{\lambda\nu\mu} + S^{\nu\lambda\mu} - S^{\mu\nu\lambda}) \\ &\Downarrow \text{ using } S^{\lambda\mu\nu} = -S^{\lambda\nu\mu} \\ &= T^{\mu\nu} - T^{\nu\mu} - \partial_\lambda S^{\lambda\mu\nu} = 0 \end{aligned} \quad (8.130)$$

As mentioned before, the Belinfante-Rosenfeld energy-momentum tensor can be shown to be the appropriate one to use in the Einstein equation.

For the Dirac field a somewhat lengthy computation gives

$$\Theta^{\mu\nu} = \frac{1}{2}(T'^{\mu\nu} + T'^{\nu\mu}) = \frac{i}{4}\bar{\Psi}\gamma^\mu \overleftrightarrow{\partial}^\nu \Psi + \frac{i}{4}\bar{\Psi}\gamma^\nu \overleftrightarrow{\partial}^\mu \Psi \quad (8.131)$$

8.3.4 Other conserved charges

An important class of Noether charges correspond to so-called *internal transformations* which do not include any space-time part, and under which $\mathcal{L}$ is invariant. An example is the $U(1)$ *symmetry*

$$\Phi(x) \rightarrow e^{-i\epsilon}\Phi(x) \quad (8.132)$$

in the complex Klein-Gordon model

$$\mathcal{L} = \partial_\mu \Phi^* \partial^\mu \Phi - m^2 \Phi^* \Phi \quad (8.133)$$

For an infinitesimal transformation

$$\delta\Phi = -i\epsilon\Phi \quad \delta\Phi^* = +i\epsilon\Phi^* \quad (8.134)$$

so

$$\Delta\Phi = -i\Phi \quad \Delta\Phi^* = +i\Phi^* \quad (8.135)$$

Now the Noether current is

$$\begin{aligned} j^\mu &= \frac{\partial\mathcal{L}}{\partial(\partial_\mu\Phi)}\Delta\Phi + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\Phi^*)}\Delta\Phi^* \\ &= i(\Phi^*\partial^\mu\Phi - \Phi\partial^\mu\Phi^*) \end{aligned} \quad (8.136)$$

Note that this is exactly the current found in Subsection 4.3.3

Exercise 8.10. Compute the $U(1)$ current for a free Dirac field

$$\mathcal{L}' = \frac{i}{2}\left(\bar{\Psi}\not{\partial}\Psi - \bar{\Psi}\overleftarrow{\not{\partial}}\Psi\right) - m\bar{\Psi}\Psi \quad (8.137)$$

Solution. The transformations are

$$\Psi(x) \rightarrow e^{-i\epsilon}\Psi(x) \quad (8.138)$$

and for an infinitesimal transformation we have

$$\delta\Psi = -i\epsilon\Psi \quad \delta\bar\Psi = +i\epsilon\bar\Psi \quad (8.139)$$

so

$$\Delta\Psi = -i\Psi \quad \Delta\bar\Psi = +i\bar\Psi \quad (8.140)$$

The current is

$$\begin{aligned} j^\mu &= \frac{\partial\mathcal{L}}{\partial(\partial_\mu\Psi)}\Delta\Psi + \Delta\bar\Psi\frac{\partial\mathcal{L}}{\partial(\partial_\mu\bar\Psi)} = \frac{i}{2}\bar\Psi\gamma^\mu(-i)\Psi + i\bar\Psi\left(-\frac{i}{2}\right)\gamma^\mu\Psi \\ &= \bar\Psi\gamma^\mu\Psi \end{aligned} \quad (8.141)$$

which is the same as in Section [5.5](#). □

Note that in the presence of Majorana mass terms the Dirac $U(1)$ current is not conserved, since they explicitly break the symmetry according to [\(5.249, 5.250\)](#).

More generally one can consider a non-Abelian Lie group G acting as a symmetry. Suppose that the action of a group element is represented by some unitary matrices $R(g)$:

$$\Phi_\alpha(x) \xrightarrow[g]{} R(g)_\alpha^\beta \Phi_\beta(x) \quad (8.142)$$

Let us also suppose that the transformations are parametrised by some ϵ^a with $a = 1, \dots, \dim G$, and the transformation corresponding to $\epsilon^a = 0$ is the identity. Then one can expand the representation as

$$R(g)_\alpha^\beta = \delta_\alpha^\beta - i \sum_a \epsilon^a (t^a)_\alpha^\beta + O(\epsilon^2) \quad (8.143)$$

where t^a are the generators in the given representation, which are Hermitian matrices and satisfy the Lie algebra

$$[t^a, t^b] = i f_c^{ab} t^c \quad (8.144)$$

where f_c^{ab} are structure constants of the Lie group. Then we have

$$\delta\Phi_\alpha = -i \sum_a \epsilon^a (t^a)_\alpha^\beta \Phi_\beta \quad (8.145)$$

so under an infinitesimal transformation by generator t^a

$$\Delta_a \Phi_\alpha = -i (t^a)_\alpha^\beta \Phi_\beta \quad (8.146)$$

Noether's theorem then gives the following conserved currents

$$j^{a,\mu} = -\frac{\partial\mathcal{L}}{\partial(\partial_\mu\Phi_\alpha)} i (t^a)_\alpha^\beta \Phi_\beta \quad (8.147)$$

Note that the charges can be written as

$$Q^a(t) = \int d^3x j^{a,0} = -i \int d^3x \Pi^\alpha(t, \vec{x}) (t^a)_\alpha^\beta \Phi_\beta(t, \vec{x}) \quad (8.148)$$

and they are conserved

$$\frac{dQ^a(t)}{dt} = 0 \quad (8.149)$$

Note *these charges can also be defined when the symmetry is only approximate*, as was the case for the isospin and the $SU(3)$ symmetries of the strong interactions. In that case

$$\frac{dQ^a(t)}{dt} \neq 0 \quad (8.150)$$

but as long as symmetry breaking is small, the system can be studied first by ignoring it and then taking the symmetry breaking into account as a correction. Even if the symmetry is broken, the transformation properties of the symmetry breaking Hamiltonian lead to non-trivial predictions concerning physical quantities. This is exactly the idea used for the quark model in Section [7.3](#) and is frequently applied in other contexts as well.

Exercise 8.11. Suppose we have a Lagrangian with N fermions

$$\mathcal{L} = \frac{i}{2} \sum_{\alpha=1}^N \left(\bar{\Psi}^\alpha \not{\partial} \Psi_\alpha - \bar{\Psi}^\alpha \overleftarrow{\not{\partial}} \Psi_\alpha \right) - \sum_{\alpha} m \bar{\Psi}^\alpha \Psi_\alpha \quad (8.151)$$

which is invariant under $SU(N)$ transformations

$$\Psi_\alpha \rightarrow U_\alpha^\beta \Psi_\beta \quad \bar{\Psi}^\alpha \rightarrow (U^\dagger)^\alpha_\beta \bar{\Psi}^\beta \quad (8.152)$$

Derive the Noether currents.

Solution. The $SU(N)$ transformations can be parametrised as

$$U = \exp(-i\epsilon^a t^a) \quad (8.153)$$

where $t^a = \frac{\lambda^a}{2}$ are $N \times N$ traceless Hermitian matrices; there are exactly $N^2 - 1$ independent such matrices conventionally normalised as

$$\text{Tr } \lambda^a \lambda^b = 2\delta^{ab} \quad (8.154)$$

Then

$$\Delta_a \Psi_\alpha = -i(t^a)_\alpha^\beta \Psi_\beta \quad \Delta_a \bar{\Psi}^\beta = i\bar{\Psi}^\alpha (t^a)_\alpha^\beta \quad (8.155)$$

resulting in the currents

$$\begin{aligned} j^\mu &= \frac{\partial \mathcal{L}}{\partial(\partial_\mu \Psi_\alpha)} \Delta_a \Psi_\alpha + \Delta_a \bar{\Psi}^\beta \frac{\partial \mathcal{L}}{\partial(\partial_\mu \bar{\Psi}^\beta)} = \frac{i}{2} \bar{\Psi}^\alpha \gamma^\mu (-i)(t^a)_\alpha^\beta \Psi_\beta + i\bar{\Psi}^\alpha (t^a)_\alpha^\beta \left(-\frac{i}{2}\right) \gamma^\mu \Psi_\beta \\ &= \bar{\Psi}^\alpha \gamma^\mu (t^a)_\alpha^\beta \Psi_\beta \end{aligned} \quad (8.156)$$

□

More generally, the Dirac Lagrangian can be written as

$$\mathcal{L} = \frac{i}{2} \sum_{\alpha=1}^N \left(\bar{\Psi}^\alpha \not{\partial} \Psi_\alpha - \bar{\Psi}^\alpha \overleftarrow{\not{\partial}} \Psi_\alpha \right) - \sum_{\alpha} \bar{\Psi}^\beta m_\beta^\alpha \Psi_\alpha \quad (8.157)$$

where m is now a mass matrix. One can find the transformations under which it is invariant

$$U^\dagger m U = m \quad (8.158)$$

Such unitary transformations form, in general, a subgroup G of the unitary group $U(N)$, with generators t^a , $a = 1, \dots, \dim G$. That leads to currents of the same form as above, but only for the generators that are eventually included in G .

Similar considerations can be made for scalar theories of the form

$$\mathcal{L} = \partial_\mu \Phi^{*\alpha} \partial^\mu \Phi_\alpha - \Phi^{*\alpha} (M^2)_\alpha^\beta \Phi_\beta \quad (8.159)$$

The construction of the Noether currents for this case is left as an exercise for the reader.

9 Electromagnetic field

9.1 Relativistic electrodynamics

To start with, we first recall the formalism of relativistic electromagnetism in vacuum and the description of point-like charges in an external electromagnetic field, switching back to the familiar SI units for the time being.

9.1.1 Electromagnetic field in relativistic form

Let us consider Maxwell's equation in vacuum

$$\begin{aligned}\vec{\nabla} \times \vec{B} &= \mu_0 \vec{j} + \frac{1}{c^2} \frac{\partial \vec{E}}{\partial t} \\ \vec{\nabla} \cdot \vec{E} &= \rho / \epsilon_0 \\ \vec{\nabla} \times \vec{E} &= -\frac{\partial \vec{B}}{\partial t} \\ \vec{\nabla} \cdot \vec{B} &= 0\end{aligned}\tag{9.1}$$

Note that as used in standard vector calculus, the three-dimensional nabla is the vector

$$\vec{\nabla} = \left(\frac{\partial}{\partial x^1}, \frac{\partial}{\partial x^2}, \frac{\partial}{\partial x^3} \right) = (\partial_1, \partial_2, \partial_3)\tag{9.2}$$

which looks counter-intuitive at first, since the canonical form of all three-vectors such as $\vec{x}$ and $\vec{p}$ corresponds to upper indices. It is important to keep this in mind, otherwise one could commit sign mistakes due to

$$\partial_i = -\partial^i\tag{9.3}$$

So the usual 3-vector operations take the following form:

$$\begin{aligned}\text{grad } \Phi = \vec{\nabla} \Phi &: (\text{grad } \Phi)^i = \partial_i \Phi \\ \text{div } \vec{E} = \vec{\nabla} \cdot \vec{E} &: \text{div } \vec{E} = \partial_i E^i \\ \text{rot } \vec{E} = \vec{\nabla} \times \vec{E} &: \left(\text{rot } \vec{E} \right)^i = \epsilon_{ijk} \partial_j E^k\end{aligned}\tag{9.4}$$

With these preliminaries, the last two Maxwell equations can be solved by introducing the potentials Φ and $\vec{A}$:

$$\vec{B} = \vec{\nabla} \times \vec{A} \quad , \quad \vec{E} = -\vec{\nabla} \Phi - \frac{\partial \vec{A}}{\partial t}\tag{9.5}$$

which can be joined to form the *four-potential*

$$\begin{aligned}A^\mu &= (A^0, \vec{A}) \\ \text{with } A^0 &= \frac{1}{c} \Phi\end{aligned}\tag{9.6}$$

(again, a three-vector like $\vec{A}$ has upper indices when used as the spatial component of a four-vector quantity). Let us introduce the *field strength tensor*

$$F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu\tag{9.7}$$

which is antisymmetric i.e. $F^{\mu\nu} = -F^{\nu\mu}$, with the independent components

$$\begin{aligned} F^{i0} &= \partial^i A^0 - \partial^0 A^i = -\partial_i A^0 - \partial_0 A^i = -\frac{1}{c} \frac{\partial \Phi}{\partial x^i} - \frac{1}{c} \frac{\partial A^i}{\partial t} \\ &= E^i / c \\ F^{ij} &= \partial^i A^j - \partial^j A^i = -\partial_i A^j + \partial_j A^i = -\frac{\partial A^j}{\partial x^i} + \frac{\partial A^i}{\partial x^j} \\ &= -\epsilon_{ijk} B^k \end{aligned} \quad (9.8)$$

Note the following properties

$$F_{i0} = -F^{i0} = F^{0i} \quad F_{ij} = F^{ij} = -F^{ji} \quad (9.9)$$

It is important to keep in mind that the four-potential is not entirely physical. The following transformation

$$A_\mu \rightarrow A_\mu + \partial_\mu \Lambda \quad (9.10)$$

leaves all the field strengths invariant. This is called a *gauge transformation*, and the freedom in choosing the four-potential is called the *gauge freedom*. Due to the gauge freedom only quantities invariant under the gauge transformations can have any physical meaning.

Note that $\vec{E}$ and $\vec{B}$ despite being three-vectors, are not part of a 4-vector. Therefore the position of their indices is more of a useful convention, since their Lorentz transformation properties are inherited from the $F^{\mu\nu}$ tensor²³

Introducing the *four-current density*

$$j^\mu = (\rho c, \vec{j}) \quad (9.11)$$

Maxwell's equations can be written in the following form

$$\begin{aligned} F^{\mu\nu} &= \partial^\mu A^\nu - \partial^\nu A^\mu \\ \partial_\mu F^{\mu\nu} &= \mu_0 j^\nu \end{aligned} \quad (9.12)$$

which implies that *Maxwell's equations are relativistically invariant* provided A^μ and j^μ transform as four-vectors (and, as a consequence of the first relation, $F^{\mu\nu}$ is a (2,0)-tensor).

Exercise 9.1. Show that

$$\partial_\mu F^{\mu\nu} = \mu_0 j^\nu \quad (9.13)$$

Solution. Computing explicitly the four-divergence of the field strength tensor

$$\partial_\mu F^{\mu\nu} = \frac{\partial F^{\mu\nu}}{\partial x^\mu} \quad (9.14)$$

as

$$\begin{aligned} \partial_\mu F^{\mu 0} &= \partial_i F^{i0} = \frac{1}{c} \vec{\nabla} \cdot \vec{E} \\ &= \mu_0 c \rho \\ \partial_\mu F^{\mu l} &= \partial_0 F^{0l} + \partial_i F^{il} = -\frac{1}{c^2} \frac{\partial E^l}{\partial t} - \epsilon_{ilk} \frac{\partial B^k}{\partial x^i} = -\frac{1}{c^2} \frac{\partial E^l}{\partial t} + \left(\vec{\nabla} \times \vec{B} \right)^l \\ &= \mu_0 j^l \end{aligned} \quad (9.15)$$

establishes our claim. □

²³Similar consideration holds for the generators J^j and K^j as parts of the $J^{\mu\nu}$ tensor

Exercise 9.2. Show that the other two Maxwell-equations can be written as the vanishing of the curl of the field tensor

$$\epsilon_{\lambda\mu\nu\rho}\partial^\mu F^{\nu\rho} = 0 \quad (9.16)$$

where $\epsilon_{\lambda\mu\nu\rho}$ is the four-dimensional Levi-Civita defined by full antisymmetry in its indices and the normalisation

$$\epsilon_{0123} = -\epsilon^{0123} = +1 \quad (9.17)$$

In analogy with 3 dimensions, this allows for the field tensor to be written using a vector-potential as $F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu$.

Solution. The $\lambda = 0$ case can be computed using

$$\epsilon_{0ijk} = \epsilon_{ijk} \quad (9.18)$$

so the equation is

$$0 = \epsilon_{0\mu\nu\rho}\partial^\mu F^{\nu\rho} = -\epsilon_{ijk}\epsilon_{jkl}\partial^i B^l \quad (9.19)$$

Using

$$\epsilon_{ijk}\epsilon_{jkl} = 2\delta_{il} \quad (9.20)$$

one gets

$$\vec{\nabla} \cdot \vec{B} = 0 \quad (9.21)$$

For $\lambda = i \neq 0$ we obtain three terms

$$\begin{aligned} 0 = \epsilon_{i\mu\nu\rho}\partial^\mu F^{\nu\rho} &= \epsilon_{i0jk}\partial^0 F^{jk} + \epsilon_{ij0k}\partial^j F^{0k} + \epsilon_{ijk0}\partial^j F^{k0} \\ &= -\epsilon_{ijk}\frac{1}{c}\partial_t(-\epsilon_{jkl}B^l) + \epsilon_{ijk}(-\partial_j)\left(-\frac{1}{c}E^k\right) - \epsilon_{ijk}(-\partial_j)\left(+\frac{1}{c}E^k\right) \\ &= \frac{2}{c}\left(\partial_t B^i + (\vec{\nabla} \times \vec{E})^i\right) \end{aligned} \quad (9.22)$$

which gives

$$\vec{\nabla} \times \vec{E} = -\frac{\partial \vec{B}}{\partial t} \quad (9.23)$$

□

9.1.2 Relativistic particle in an external electromagnetic field

The *equation of motion of the relativistic particle* can be written as

$$\frac{dp^\mu}{d\tau} = qF^{\mu\nu}u_\nu \quad (9.24)$$

where

$$\begin{aligned} u^\mu &= \frac{dx^\mu}{d\tau} = \gamma \left(c, \dot{\vec{x}}(t) \right) \\ \text{with } \gamma &= \frac{1}{\sqrt{1 - \dot{\vec{x}}(t)^2/c^2}} \end{aligned} \quad (9.25)$$

is the four-velocity. It is clear that the equation must have this form for the following reasons:

1. p^μ is a four-vector, while τ is the Lorentz-invariant proper time. So the derivative

$$\frac{dp^\mu}{d\tau} \quad (9.26)$$

is a four-vector, and it is the covariant extension of the force.

2. The electromagnetic force is proportional to the charge q , and is linear in the field strength and the velocity. The only four-vector quantity with these properties is

$$qF^{\mu\nu}u_\nu \quad (9.27)$$

So the only issue is whether we got all the factors right. This is the task of the following two exercises.

Exercise 9.3. Show that

$$\frac{d\vec{p}}{dt} = q(\vec{E} + \vec{v} \times \vec{B}) \quad (9.28)$$

Solution.

□

Writing out the spatial components

$$\frac{dp^i}{d\tau} = q(F^{i0}u_0 + F^{ij}u_j) = q(F^{i0}u^0 - qF^{ij}u^j) = q\left(\frac{1}{c}E^i\gamma c + \epsilon_{ijk}B^k\gamma v^j\right) \quad (9.29)$$

and using

$$d\tau = dt/\gamma \quad (9.30)$$

we obtain

$$\frac{dp^i}{dt} = q\left(E^i + (\vec{v} \times \vec{B})^i\right) \quad (9.31)$$

Exercise 9.4. Show that the temporal component is just the conservation of energy.

Solution. Computing

$$\frac{dp^0}{d\tau} = qF^{0i}u_i = -qF^{0i}u^i = -q\left(-\frac{1}{c}E^i\right)\gamma v^i \quad (9.32)$$

leads to

$$\frac{dp^0}{dt} = \frac{q}{c}\vec{E} \cdot \vec{v} \quad (9.33)$$

Recalling that cp^0 is the kinetic energy of the particle, the above equation is just the conservation of energy with the usual expression for the power exerted by the electromagnetic field. □

One can also write an *action principle for the relativistic particle in an external electromagnetic field* in the following way. In the non-relativistic limit in a static electric field, the Lagrangian is

$$L = \frac{1}{2}m\dot{\vec{x}}^2 - q\Phi \quad (9.34)$$

The proper relativistic generalisation is then

$$L = \sqrt{1 - \dot{\vec{x}}(t)^2/c^2} (-mc^2 - qu^\mu A_\mu(t, \vec{x}(t))) \quad (9.35)$$

so that the action

$$\begin{aligned} S &= \int dt L = \int d\tau (-mc^2 - qu^\mu A_\mu(t, \vec{x}(t))) \\ &= -mc^2 \int d\tau - q \int dx^\mu A_\mu(x) \end{aligned} \quad (9.36)$$

is Lorentz invariant. In the last line the second integral is taken along the world-line.

Exercise 9.5. Show that the above action gives the correct equations of motion.

Solution. Substituting the expressions of

$$u^\mu = \frac{dx^\mu}{d\tau} = \frac{1}{\sqrt{1 - \dot{\vec{x}}(t)^2/c^2}} (c, \dot{\vec{x}}(t)) \quad (9.37)$$

and

$$A^\mu = \left(\frac{1}{c} \Phi, \vec{A} \right) \quad (9.38)$$

we obtain

$$L = -mc^2 \sqrt{1 - \dot{\vec{x}}^2/c^2} - q\Phi(t, \vec{x}) + q\dot{x}^j A^j \quad (9.39)$$

For the canonical momentum we obtain

$$\pi_i = \frac{\partial L}{\partial \dot{x}^i} = \frac{m\dot{x}^i}{\sqrt{1 - \dot{\vec{x}}^2/c^2}} + qA^i \quad (9.40)$$

The Euler-Lagrange equation can be written as

$$\frac{d}{dt} \frac{\partial L}{\partial \dot{x}^i} = \frac{\partial L}{\partial x^i} \quad (9.41)$$

with

$$\begin{aligned} \frac{\partial L}{\partial x^i} &= -q\partial_i \Phi + q\dot{x}^j \partial_i A^j \\ \frac{d}{dt} \frac{\partial L}{\partial \dot{x}^i} &= \frac{d}{dt} \left(\frac{m\dot{x}^i}{\sqrt{1 - \dot{\vec{x}}^2/c^2}} \right) + q \frac{\partial A^i}{\partial t} + q\partial_j A^i \dot{x}^j \end{aligned} \quad (9.42)$$

or more explicitly

$$\begin{aligned} \frac{d}{dt} \left(\frac{m\dot{x}^i}{\sqrt{1 - \dot{\vec{x}}^2/c^2}} \right) &= -q \frac{\partial A^i}{\partial t} - q\partial_j A^i \dot{x}^j - q\partial_i \Phi + q\dot{x}^j \partial_i A^j \\ &= q \left(-\frac{\partial A^i}{\partial t} - \partial_i \Phi \right) + qv^j (\partial_i A^j - \partial_j A^i) \\ &\quad \text{using: } \partial_i A^j - \partial_j A^i = \epsilon_{ijk} B^k \text{ and } -\frac{\partial A^i}{\partial t} - \partial_i \Phi = E^i \\ &= q \left(\vec{E} + \vec{v} \times \vec{B} \right) \end{aligned} \quad (9.43)$$

□

Note that in an external electromagnetic field *the mechanical and the canonical momentum do not coincide*: indeed

$$\vec{p} = \vec{\pi} - q\vec{A} \quad (9.44)$$

$$\begin{aligned} H &= \pi_i \dot{x}^i - L \\ &= \frac{m\dot{\vec{x}}^2}{\sqrt{1 - \dot{\vec{x}}^2/c^2}} + mc^2 \sqrt{1 - \dot{\vec{x}}^2/c^2} + q\Phi \\ &= c\sqrt{(\vec{\pi} - q\vec{A})^2 + m^2c^2} + q\Phi \end{aligned} \quad (9.45)$$

This can be written in an even more appealing form using that $H = cp^0$ is the energy, so

$$p^0 - qA^0 = \sqrt{(\vec{\pi} - q\vec{A})^2 + m^2c^2} \quad (9.46)$$

which shows introducing an external electromagnetic field is equivalent to the rule called *minimal substitution*:

$$p^\mu \rightarrow p^\mu - qA^\mu \quad (9.47)$$

which was already used in Section [5.8](#)

Exercise 9.6. Show that the above action is gauge invariant due to the fact that a gauge transformation

$$A_\mu \rightarrow A_\mu + \partial_\mu \Lambda \quad (9.48)$$

changes the Lagrangian by a total derivative.

Solution. A gauge transformation of the form

$$A_\mu \rightarrow A_\mu + \partial_\mu \Lambda \quad (9.49)$$

leaves to the following change in the action

$$\Delta S = - \int d\tau q \frac{dx^\mu}{d\tau} \partial_\mu \Lambda = -q \int d\tau \frac{d\Lambda}{d\tau} = -q \int dt \frac{d\Lambda}{dt} \quad (9.50)$$

This is the integral of a total derivative, and the Lagrangian itself changes by a total derivative

$$\Delta L = -q \frac{d\Lambda}{dt} \quad (9.51)$$

and so a gauge transformation does not affect the equations of motions. This could also be seen from their dependence on $F^{\mu\nu}$ only. $\square$

9.2 Action and energy-momentum tensor

Here we switch to units with $\epsilon_0 = c = \hbar = 1$. Besides the usual conventions this also entails $\mu_0 = 1$, and is equivalent to using a particular unit for the electric charge in which the elementary charge is a dimensionless number

$$e = \sqrt{4\pi\alpha} \quad (9.52)$$

with α denoting the fine structure constant.

The *Maxwell equations can be derived from an action principle* with the Lagrangian:

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} - j^\mu A_\mu \quad (9.53)$$

Indeed let us compute the Euler-Lagrange equations corresponding to A_μ :

$$\mathcal{L} = -\frac{1}{4}(\partial_\mu A_\nu - \partial_\nu A_\mu)(\partial^\mu A^\nu - \partial^\nu A^\mu) - j^\mu A_\mu \quad (9.54)$$

We obtain

$$\frac{\partial \mathcal{L}}{\partial(\partial_\mu A_\nu)} = -F^{\mu\nu} \quad \frac{\partial \mathcal{L}}{\partial A_\nu} = -j^\nu \quad (9.55)$$

so the equations of motion are

$$\partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu A_\nu)} - \frac{\partial \mathcal{L}}{\partial A_\nu} = -\partial_\mu F^{\mu\nu} + j^\nu = 0 \quad (9.56)$$

which is identical to the equation derived previously.

The action for the free electromagnetic field is

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} \quad (9.57)$$

Exercise 9.7. Compute the canonical energy-momentum tensor of the free electromagnetic field.

Solution.

$$\begin{aligned} T^{\mu\nu} &= \frac{\partial \mathcal{L}}{\partial(\partial_\mu A_\lambda)} \partial^\nu A_\lambda - \eta^{\mu\nu} \mathcal{L} \\ &= -F^{\mu\lambda} \partial^\nu A_\lambda + \frac{1}{4} \eta^{\mu\nu} F_{\rho\lambda} F^{\rho\lambda} \end{aligned} \quad (9.58)$$

□

Note that the above energy-momentum tensor is neither symmetric nor invariant under the gauge transformations

$$A_\mu \rightarrow A_\mu + \partial_\mu \Lambda \quad (9.59)$$

This can be resolved by introducing the modification

$$\Theta^{\mu\nu} = T^{\mu\nu} + \partial_\lambda B^{\lambda\mu\nu} \quad (9.60)$$

with

$$B^{\lambda\mu\nu} = F^{\lambda\mu} A^\nu \quad (9.61)$$

which results in the *symmetric energy-momentum tensor*

$$\Theta^{\mu\nu} = -F^{\mu\lambda} F^\nu_\lambda + \frac{1}{4} \eta^{\mu\nu} F_{\rho\lambda} F^{\rho\lambda} \quad (9.62)$$

where we used the field equation

$$\partial_\lambda F^{\lambda\mu} = 0 \quad (9.63)$$

9.3 Gauge invariance and minimal coupling

As discussed in Section [5.8](#), electromagnetic interaction can be added to the Dirac equation

$$(i\gamma^\mu \partial_\mu - m) \Psi(x) = 0 \quad (9.64)$$

using the principle of minimal subtraction

$$p^\mu \rightarrow p^\mu - qA^\mu \quad (9.65)$$

which leads to

$$[\gamma^\mu (i\partial_\mu - qA_\mu(x)) - m] \Psi(x) = 0 \quad (9.66)$$

or, using Dirac's notation

$$(i\not{\partial} - q\not{A} - m) \Psi = 0 \quad (9.67)$$

Now note that the $U(1)$ symmetry of the Dirac field can be extended to a local (i.e. space-time dependent) transformation of the form

$$\Psi(x) \rightarrow e^{-iq\Lambda(x)} \Psi(x) \quad (9.68)$$

provided the vector potential transforms as

$$A_\mu(x) \rightarrow A_\mu(x) + \partial_\mu \Lambda(x) \quad (9.69)$$

Exercise 9.8. Show that the above transformation leaves the dynamics of the Dirac field invariant.

Solution.

$$\begin{aligned} & [\gamma^\mu (i\partial_\mu - q(A_\mu(x) + \partial_\mu \Lambda(x))) - m] e^{-iq\Lambda(x)} \Psi(x) \\ = & e^{-iq\Lambda(x)} [\gamma^\mu (i\partial_\mu + q\partial_\mu \Lambda(x) - q(A_\mu(x) + \partial_\mu \Lambda(x))) - m] \Psi(x) \\ = & e^{-iq\Lambda(x)} [\gamma^\mu (i\partial_\mu - qA_\mu(x)) - m] \Psi(x) \end{aligned} \quad (9.70)$$

□

Switching to the level of the action, it can be shown that the electromagnetic interaction can be introduced by demanding that the global (i.e. position independent) $U(1)$ symmetry is extended to a local one. Taking the free fermion Lagrangian

$$\mathcal{L} = \bar{\Psi}(i\not{\partial} - m)\Psi \quad (9.71)$$

one can study its property under

$$\Psi(x) \rightarrow e^{-iq\Lambda(x)}\Psi(x) \quad (9.72)$$

The transformation of $\mathcal{L}$ is

$$\mathcal{L} \rightarrow \bar{\Psi}(i\not{\partial} - m)\Psi + q\bar{\Psi}\gamma^\mu\Psi\partial_\mu\Lambda \quad (9.73)$$

Note that the change in the action is proportional to the Noether current

$$j^\mu = q\bar{\Psi}\gamma^\mu\Psi \quad (9.74)$$

Therefore introducing the electromagnetic field coupled to this current gives a gauge invariant Lagrangian

$$\begin{aligned} \mathcal{L}_{QED} &= \bar{\Psi}(i\not{\partial} - m)\Psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} - j^\mu A_\mu \\ &= \bar{\Psi}(i\not{\partial} - q\not{A} - m)\Psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} \end{aligned} \quad (9.75)$$

As the notation suggests, this is the Lagrangian that defines the simplest form of *quantum electrodynamics* (*QED*) that describes the dynamics of electrons and positrons coupled to the electromagnetic fields. Adding other particles (muon, tau, quarks) necessitates the coupling of the fields describing them to the electromagnetic field.

Exercise 9.9. Compute the (classical) equations of motions for quantum electrodynamics.

Solution. Variation with respect to the different fields leads to the equations

$$\begin{aligned} \delta\bar{\Psi} &: (i\not{\partial} - q\not{A} - m)\Psi = 0 \\ \delta\Psi &: \bar{\Psi}(-i\overleftarrow{\not{\partial}} - q\not{A} - m) = 0 \\ \delta A &: \partial_\mu F^{\mu\nu} = -j^\nu \end{aligned} \quad (9.76)$$

with

$$j^\nu = q\bar{\Psi}\gamma^\nu\Psi \quad (9.77)$$

□

The derivative operator

$$D_\mu = \partial_\mu + iqA_\mu \quad (9.78)$$

has the particular properties that it is *covariant under gauge transformations*

$$e^{+iq\Lambda(x)}(\partial_\mu + iq(A_\mu(x) + \partial_\mu\Lambda(x)))e^{-iq\Lambda(x)} = \partial_\mu - iq\partial_\mu\Lambda(x) + iq(A_\mu(x) + \partial_\mu\Lambda(x)) = \partial_\mu + iqA_\mu(x) \quad (9.79)$$

With this notation, the Lagrangian can be written as

$$\mathcal{L}_{QED} = \bar{\Psi}(i\not{D} - m)\Psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} \quad (9.80)$$

This observation leads to a simple way of coupling fields to electromagnetism. For example, for a scalar field with the Lagrangian

$$\mathcal{L} = \partial_\mu\Phi^*\partial^\mu\Phi - m^2\Phi^*\Phi \quad (9.81)$$

of charge q , the derivative can be substituted with the covariant one to obtain the Lagrangian of scalar quantum electrodynamics

$$\begin{aligned}\mathcal{L}_{SQED} &= -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + (D_\mu\Phi)^*(D^\mu\Phi) - m^2\Phi^*\Phi \\ &= -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \partial_\mu\Phi^*\partial^\mu\Phi - m^2\Phi^*\Phi - j^\mu A_\mu + q^2 A_\mu A^\mu \Phi^*\Phi\end{aligned}\quad (9.82)$$

where the Noether current is

$$j^\mu = iq(\Phi^*\partial^\mu\Phi - \Phi\partial^\mu\Phi^*) \quad (9.83)$$

In this case, besides the current-potential coupling there is also a term quadratic in the potential; its presence is necessary to ensure gauge invariance.

Exercise 9.10. Compute the (classical) equations of motions for scalar quantum electrodynamics and show that they are gauge invariant.

Solution. The derivatives of the Lagrangian can be computed as follows

$$\begin{aligned}\frac{\partial\mathcal{L}}{\partial\Phi} &= -m^2\Phi^* + q^2 A_\mu A^\mu \Phi^* + iqA_\mu\partial^\mu\Phi^* & \frac{\partial\mathcal{L}}{\partial(\partial_\mu\Phi)} &= \partial^\mu\Phi^* - iqA^\mu\Phi^* \\ \frac{\partial\mathcal{L}}{\partial\Phi^*} &= -m^2\Phi + q^2 A_\mu A^\mu \Phi - iqA_\mu\partial^\mu\Phi & \frac{\partial\mathcal{L}}{\partial(\partial_\mu\Phi^*)} &= \partial^\mu\Phi + iqA^\mu\Phi \\ \frac{\partial\mathcal{L}}{\partial A_\nu} &= -j^\nu + 2q^2 A^\nu\Phi^*\Phi & \frac{\partial\mathcal{L}}{\partial(\partial_\mu A_\nu)} &= -F^{\mu\nu}\end{aligned}\quad (9.84)$$

The equations of motions are

$$\begin{aligned}\partial_\mu\frac{\partial\mathcal{L}}{\partial(\partial_\mu\Phi)} - \frac{\partial\mathcal{L}}{\partial\Phi} &= [(D_\mu D^\mu - m^2)\Phi]^* = 0 \\ \partial_\mu\frac{\partial\mathcal{L}}{\partial(\partial_\mu\Phi^*)} - \frac{\partial\mathcal{L}}{\partial\Phi^*} &= [(D_\mu D^\mu - m^2)\Phi] = 0 \\ \partial_\mu\frac{\partial\mathcal{L}}{\partial(\partial_\mu A_\nu)} - \frac{\partial\mathcal{L}}{\partial A_\nu} &= -\partial_\mu F^{\mu\nu} + j^\nu - 2q^2 A^\nu\Phi^*\Phi = 0\end{aligned}\quad (9.85)$$

The first two equations are obviously gauge invariant. For the last one, the transformation of the current

$$j^\nu \rightarrow j^\nu + 2q^2\partial^\nu\Lambda\Phi^*\Phi \quad (9.86)$$

under the transformation

$$\Phi(x) \rightarrow e^{-iq\Lambda(x)}\Phi(x) \quad (9.87)$$

is cancelled by the transformation of the other term under

$$A^\nu(x) \rightarrow A^\nu(x) + \partial^\nu\Lambda(x) \quad (9.88)$$

□

10 Elements of quantum field theory

The quantum dynamics of fields is described by quantum field theory, which is a vast subject. Here we only review some very basic elements that give a good intuitive picture that is sufficient for the present course.

10.1 Canonical quantisation of free fields

A field is called free if its Euler-Lagrange equations are linear. This happens when the corresponding Lagrangian density contains terms that are at most quadratic in field variables. An example is the real Klein-Gordon field:

$$\mathcal{L} = \frac{1}{2} \partial_\mu \Phi \partial^\mu \Phi - \frac{m^2}{2} \Phi^2 \quad (10.1)$$

which satisfies

$$(\square + m^2) \Phi(x) = 0 \quad (10.2)$$

Since this is a linear partial differential equation, it can be decoupled into an infinite number of harmonic oscillators using a Fourier transformation

$$\Phi(t, \vec{x}) = \int \frac{d^3 k}{(2\pi)^3} \tilde{\Phi}_{\vec{k}}(t) e^{i\vec{k} \cdot \vec{x}} \quad (10.3)$$

where due to reality of the field

$$\tilde{\Phi}_{\vec{k}}^*(t) = \tilde{\Phi}_{-\vec{k}}(t) \quad (10.4)$$

Then the equations of motion are

$$\ddot{\tilde{\Phi}}_{\vec{k}}(t) + \omega_{\vec{k}}^2 \tilde{\Phi}_{\vec{k}}(t) = 0 \quad \omega_{\vec{k}} = \sqrt{\vec{k}^2 + m^2} \quad (10.5)$$

and the Lagrangian takes the form

$$L = \int d^3 x \mathcal{L} = \int \frac{d^3 k}{(2\pi)^3} \frac{1}{2} \left(\left| \dot{\tilde{\Phi}}_{\vec{k}}(t) \right|^2 - \omega_{\vec{k}}^2 \left| \tilde{\Phi}_{\vec{k}}(t) \right|^2 \right) \quad (10.6)$$

which is just a family of coupled harmonic oscillators.

One introducing the canonical momentum field

$$\Pi(x) = \frac{\partial \mathcal{L}}{\partial \dot{\Phi}(x)} = \dot{\Phi}(x) \quad (10.7)$$

its Fourier transform can be written as

$$\begin{aligned} \Pi(t, \vec{x}) &= \int \frac{d^3 k}{(2\pi)^3} \tilde{\Pi}_{\vec{k}}(t) e^{i\vec{k} \cdot \vec{x}} \\ \tilde{\Pi}_{\vec{k}}(t) &= \dot{\tilde{\Phi}}_{\vec{k}}(t) \end{aligned} \quad (10.8)$$

Writing

$$L = \int \frac{d^3 k}{(2\pi)^3} \frac{1}{2} \left(\dot{\tilde{\Phi}}_{\vec{k}}(t) \dot{\tilde{\Phi}}_{-\vec{k}}(t) - \omega_{\vec{k}}^2 \tilde{\Phi}_{\vec{k}}(t) \tilde{\Phi}_{-\vec{k}}(t) \right) \quad (10.9)$$

it is obvious that the momentum conjugate to $\tilde{\Phi}_{\vec{k}}$ is $\frac{1}{(2\pi)^3}\tilde{\Pi}_{-\vec{k}}$. So the Klein-Gordon field can be quantised by promoting the classical variables to operators and requiring the *canonical commutation relations*

$$\begin{aligned}\left[\tilde{\Phi}_{\vec{k}}(t), \tilde{\Phi}_{\vec{k}'}(t)\right] &= 0 = \left[\tilde{\Pi}_{\vec{k}}(t), \tilde{\Pi}_{\vec{k}'}(t)\right] \\ \left[\tilde{\Phi}_{\vec{k}}(t), \tilde{\Pi}_{\vec{k}'}(t)\right] &= i(2\pi)^3\delta^{(3)}(\vec{k} + \vec{k}')\end{aligned}\quad (10.10)$$

which can be written in position space as

$$\begin{aligned}\left[\Phi(t, \vec{x}), \Phi(t, \vec{x}')\right] &= 0 = \left[\Pi(t, \vec{x}), \Pi(t, \vec{x}')\right] \\ \left[\Phi(t, \vec{x}), \Pi(t, \vec{x}')\right] &= i\delta^{(3)}(\vec{x} - \vec{x}')\end{aligned}\quad (10.11)$$

Exercise 10.1. Prove (10.11).

Solution. We only show the computation for the non-zero commutator, the others can be evaluated in a similar way:

$$\begin{aligned}\left[\Phi(t, \vec{x}), \Pi(t, \vec{x}')\right] &= \int \frac{d^3k}{(2\pi)^3} e^{i\vec{k}\cdot\vec{x}} \int \frac{d^3k'}{(2\pi)^3} e^{i\vec{k}'\cdot\vec{x}'} \left[\tilde{\Phi}_{\vec{k}}(t), \tilde{\Pi}_{\vec{k}'}(t)\right] \\ &= \int \frac{d^3k}{(2\pi)^3} e^{i\vec{k}\cdot\vec{x}} \int \frac{d^3k'}{(2\pi)^3} e^{i\vec{k}'\cdot\vec{x}'} i(2\pi)^3\delta^{(3)}(\vec{k} + \vec{k}') \\ &= i \int \frac{d^3k}{(2\pi)^3} e^{i\vec{k}\cdot(\vec{x}-\vec{x}')} = i\delta^{(3)}(\vec{x} - \vec{x}')\end{aligned}\quad (10.12)$$

□

Using reality and the equations of motion one can express

$$\begin{aligned}\tilde{\Phi}_{\vec{k}}(t) &= (2\pi)^{3/2} \frac{1}{\sqrt{2\omega_{\vec{k}}}} \left(a(\vec{k}) e^{-i\omega_{\vec{k}}t} + a^\dagger(-\vec{k}) e^{i\omega_{\vec{k}}t} \right) \\ \tilde{\Pi}_{\vec{k}}(t) = \dot{\tilde{\Phi}}_{\vec{k}}(t) &= -i(2\pi)^{3/2} \sqrt{\frac{\omega_{\vec{k}}}{2}} \left(a(\vec{k}) e^{-i\omega_{\vec{k}}t} - a^\dagger(-\vec{k}) e^{i\omega_{\vec{k}}t} \right)\end{aligned}\quad (10.13)$$

and so

$$\begin{aligned}\left[a(\vec{k}), a(\vec{k}')\right] &= 0 = \left[a^\dagger(\vec{k}), a^\dagger(\vec{k}')\right] \\ \left[a(\vec{k}), a^\dagger(\vec{k}')\right] &= \delta^{(3)}(\vec{k} - \vec{k}')\end{aligned}\quad (10.14)$$

are just a set of harmonic oscillator step operators.

Exercise 10.2. Prove (10.14).

Solution. One can express

$$\begin{aligned} a(\vec{k}) &= \frac{1}{(2\pi)^{3/2}} \left(\sqrt{\frac{\omega_{\vec{k}}}{2}} \tilde{\Phi}_{\vec{k}}(t) + \frac{i}{\sqrt{2\omega_{\vec{k}}}} \tilde{\Pi}_{\vec{k}}(t) \right) e^{+i\omega_{\vec{k}}t} \\ a^\dagger(\vec{k}) &= \frac{1}{(2\pi)^{3/2}} \left(\sqrt{\frac{\omega_{\vec{k}}}{2}} \tilde{\Phi}_{-\vec{k}}(t) - \frac{i}{\sqrt{2\omega_{\vec{k}}}} \tilde{\Pi}_{-\vec{k}}(t) \right) e^{-i\omega_{\vec{k}}t} \end{aligned} \quad (10.15)$$

and use the relations (10.10). □

So the final expansion for a real Klein-Gordon field is

$$\Phi(t, \vec{x}) = \int \frac{d^3k}{(2\pi)^{3/2}} \frac{1}{\sqrt{2\omega_{\vec{k}}}} \left(a(\vec{k}) e^{-i\omega_{\vec{k}}t + i\vec{k} \cdot \vec{x}} + a^\dagger(\vec{k}) e^{i\omega_{\vec{k}}t - i\vec{k} \cdot \vec{x}} \right) \quad (10.16)$$

where in the second term we changed the integration variable $\vec{k}$ to $-\vec{k}$. The energy and momentum can be expressed as

$$\begin{aligned} H &= \int d^3k \frac{1}{2} \omega_{\vec{k}} \left(a^\dagger(\vec{k}) a(\vec{k}) + a(\vec{k}) a^\dagger(\vec{k}) \right) \\ P_i &= \int d^3k \frac{1}{2} k_i \left(a^\dagger(\vec{k}) a(\vec{k}) + a(\vec{k}) a^\dagger(\vec{k}) \right) \end{aligned} \quad (10.17)$$

and satisfy

$$\begin{aligned} [H, a^\dagger(\vec{q})] &= \omega_{\vec{q}} a^\dagger(\vec{q}) & [H, a(\vec{q})] &= -\omega_{\vec{q}} a(\vec{q}) \\ [P_i, a^\dagger(\vec{q})] &= q_i a^\dagger(\vec{q}) & [P_i, a(\vec{q})] &= -q_i a(\vec{q}) \end{aligned} \quad (10.18)$$

Exercise 10.3. Compute (10.17) and (10.18).

Solution. Substitute the Fourier expansion (10.16) into

$$\begin{aligned} H &= \int d^3x \frac{1}{2} \left(\Pi(x) \Pi(x) + \partial_i \Phi(x) \partial_i \Phi(x) + m^2 \Phi(x) \Phi(x) \right) \\ P_i &= \int d^3x \Pi(x) \partial_i \Phi(x) \end{aligned} \quad (10.19)$$

and compute the spatial integral. For example

$$\begin{aligned}
H &= \int \frac{d^3k}{(2\pi)^{3/2}} \int \frac{d^3k'}{(2\pi)^{3/2}} \frac{1}{2\sqrt{2\omega_{\vec{k}}2\omega_{\vec{k}'}}} \\
&\quad \left[\left(-\omega_{\vec{k}}\omega_{\vec{k}'} - \vec{k} \cdot \vec{k}' \right) \left(a(\vec{k})e^{-i\omega_{\vec{k}}t+i\vec{k}\cdot\vec{x}} - a^\dagger(\vec{k})e^{i\omega_{\vec{k}}t-i\vec{k}\cdot\vec{x}} \right) \left(a(\vec{k}')e^{-i\omega_{\vec{k}'}t+i\vec{k}'\cdot\vec{x}} - a^\dagger(\vec{k}')e^{i\omega_{\vec{k}'}t-i\vec{k}'\cdot\vec{x}} \right) \right. \\
&\quad \left. + m^2 \left(a(\vec{k})e^{-i\omega_{\vec{k}}t+i\vec{k}\cdot\vec{x}} + a^\dagger(\vec{k})e^{i\omega_{\vec{k}}t-i\vec{k}\cdot\vec{x}} \right) \left(a(\vec{k}')e^{-i\omega_{\vec{k}'}t+i\vec{k}'\cdot\vec{x}} + a^\dagger(\vec{k}')e^{i\omega_{\vec{k}'}t-i\vec{k}'\cdot\vec{x}} \right) \right] \\
&= \int \frac{d^3k}{(2\pi)^{3/2}} \int \frac{d^3k'}{(2\pi)^{3/2}} \frac{1}{2\sqrt{2\omega_{\vec{k}}2\omega_{\vec{k}'}}} \\
&\quad \left[\left(-\omega_{\vec{k}}\omega_{\vec{k}'} - \vec{k} \cdot \vec{k}' + m^2 \right) \left(a(\vec{k})a(\vec{k}')e^{-i\omega_{\vec{k}}t-i\omega_{\vec{k}'}t} + a^\dagger(\vec{k})a^\dagger(\vec{k}')e^{i\omega_{\vec{k}}t+i\omega_{\vec{k}'}t} \right) \delta^{(3)}(\vec{k} + \vec{k}') \right. \\
&\quad \left. - \left(-\omega_{\vec{k}}\omega_{\vec{k}'} - \vec{k} \cdot \vec{k}' - m^2 \right) \left(a(\vec{k})a^\dagger(\vec{k}')e^{-i\omega_{\vec{k}}t+i\omega_{\vec{k}'}t} + a^\dagger(\vec{k})a(\vec{k}')e^{i\omega_{\vec{k}}t-i\omega_{\vec{k}'}t} \right) \delta^{(3)}(\vec{k} - \vec{k}') \right] \\
&\quad \text{recalling that } \omega_{\vec{k}} = \sqrt{\vec{k}^2 + m^2} \\
&= \int d^3k \frac{1}{2}\omega_{\vec{k}} \left(a^\dagger(\vec{k})a(\vec{k}) + a(\vec{k})a^\dagger(\vec{k}) \right) \tag{10.20}
\end{aligned}$$

For the commutators, an example is

$$\begin{aligned}
[H, a^\dagger(\vec{q})] &= \int d^3k \frac{1}{2}\omega_{\vec{k}} \left(a^\dagger(\vec{k})[a(\vec{k}), a^\dagger(\vec{q})] + [a(\vec{k}), a^\dagger(\vec{q})]a^\dagger(\vec{k}) \right) \\
&= \int d^3k \omega_{\vec{k}} \delta^{(3)}(\vec{k} - \vec{q}) a^\dagger(\vec{k}) = \omega_{\vec{q}} a^\dagger(\vec{q}) \tag{10.21}
\end{aligned}$$

all the others can be computed similarly. $\square$

Note that one can rewrite

$$H = \int d^3k \omega_{\vec{k}} a^\dagger(\vec{k})a(\vec{k}) + \frac{1}{2} \int d^3k \omega_{\vec{k}} \delta^{(3)}(\vec{0}) \tag{10.22}$$

The last term looks strange, but in fact it has a quite simple meaning. Note that

$$\delta^{(3)}(\vec{k}) = \int \frac{d^3x}{(2\pi)^3} e^{i\vec{k}\cdot\vec{x}} \tag{10.23}$$

so

$$\delta^{(3)}(\vec{0}) = \frac{V}{(2\pi)^3} \tag{10.24}$$

where V is the volume of the system. Therefore

$$\frac{1}{2} \int d^3k \omega_{\vec{k}} \delta^{(3)}(\vec{0}) = V \int \frac{d^3k}{(2\pi)^3} \frac{1}{2} \omega_{\vec{k}} \tag{10.25}$$

which is *nothing else than the total ground state energy of the system*, with the integral

$$V \int \frac{d^3k}{(2\pi)^3} \quad (10.26)$$

resulting from the summation over the momenta, which are discrete in finite volume, replaced by integral over the density of states. This term *can eventually be dropped, as the Hamiltonian is only defined up to a constant*²⁴ so the Hamiltonian can be redefined as

$$H = \int d^3k \omega_{\vec{k}} a^\dagger(\vec{k}) a(\vec{k}) \quad (10.27)$$

This is essentially a rewriting of the Hamiltonian by placing all creation operators in front, which is called *normal ordering*. Note that canonical quantisation does not uniquely define the Hamiltonian, since operators do not commute, so in addition to the classical form of the Hamiltonian, at the quantum level the ordering must also be specified. Normal ordering is simply a recipe to define an unambiguous ordering at the quantum level. Note that for the momentum the ground state contribution vanishes (and the corresponding ordering ambiguity does not exist), since the integrand is odd under $\vec{k} \rightarrow -\vec{k}$.

So far we considered a real Klein-Gordon field. For a complex field, the only difference is that

$$\begin{aligned} \Phi(t, \vec{x}) &= \int \frac{d^3k}{(2\pi)^{3/2}} \frac{1}{\sqrt{2\omega_{\vec{k}}}} \left(a(\vec{k}) e^{-i\omega_{\vec{k}}t + i\vec{k}\cdot\vec{x}} + b^\dagger(\vec{k}) e^{i\omega_{\vec{k}}t - i\vec{k}\cdot\vec{x}} \right) \\ \Phi^\dagger(t, \vec{x}) &= \int \frac{d^3k}{(2\pi)^{3/2}} \frac{1}{\sqrt{2\omega_{\vec{k}}}} \left(b(\vec{k}) e^{-i\omega_{\vec{k}}t + i\vec{k}\cdot\vec{x}} + a^\dagger(\vec{k}) e^{i\omega_{\vec{k}}t - i\vec{k}\cdot\vec{x}} \right) \end{aligned} \quad (10.28)$$

and the conjugate momenta are

$$\Pi(t, \vec{x}) = \dot{\Phi}^\dagger(t, \vec{x}) \quad \Pi^\dagger(t, \vec{x}) = \dot{\Phi}(t, \vec{x}) \quad (10.29)$$

Exercise 10.4. Show that

$$\begin{aligned} [a(\vec{k}), a(\vec{k}')] &= 0 = [a^\dagger(\vec{k}), a^\dagger(\vec{k}')] \\ [a(\vec{k}), a^\dagger(\vec{k}')] &= \delta^{(3)}(\vec{k} - \vec{k}') \\ [b(\vec{k}), b(\vec{k}')] &= 0 = [b^\dagger(\vec{k}), b^\dagger(\vec{k}')] \\ [b(\vec{k}), b^\dagger(\vec{k}')] &= \delta^{(3)}(\vec{k} - \vec{k}') \end{aligned} \quad (10.30)$$

while mixed commutators between a and b modes vanish.

Solution. Express the mode operators using from inverse Fourier transforms of Φ , $\Phi^\dagger$, Π and $\Pi^\dagger$ and use the canonical commutation relations

$$\begin{aligned} [\Phi(t, \vec{x}), \Pi(t, \vec{x}')] &= i\delta^{(3)}(\vec{x} - \vec{x}') \\ [\Phi^\dagger(t, \vec{x}), \Pi^\dagger(t, \vec{x}')] &= i\delta^{(3)}(\vec{x} - \vec{x}') \end{aligned} \quad (10.31)$$

(plus all the other combinations such as $[\Phi, \Phi]$, $[\Phi, \Pi^\dagger]$, $[\Phi, \Phi^\dagger]$ etc. that give zero) to compute the commutators between the modes. $\square$

²⁴The only exception is when coupling the system to gravity, since Einstein's equation couples the space-time metric to the full energy-momentum tensor. In that case, a constant ground state energy term in the Hamiltonian would correspond to vacuum energy density, which is the so-called cosmological constant. However, even in that case the remarks on the ambiguity due to ordering still applies, meaning that quantum field theory alone cannot predict the value of cosmological constant.

The Hamiltonian is

$$\begin{aligned}
H &= \int d^3x \left(\Pi(x) \dot{\Phi}(x) + \Pi^\dagger(x) \dot{\Phi}^\dagger(x) - \mathcal{L} \right) \\
&= \int d^3x \left(\Pi^\dagger(t, \vec{x}) \Pi(t, \vec{x}) + \vec{\nabla} \Phi^\dagger(t, \vec{x}) \cdot \vec{\nabla} \Phi(t, \vec{x}) + m^2 \Phi^\dagger(t, \vec{x}) \Phi(t, \vec{x}) \right) \\
&= \int d^3k \frac{\omega_{\vec{k}}}{2} \left(a^\dagger(\vec{k}) a(\vec{k}) + a(\vec{k}) a^\dagger(\vec{k}) + b^\dagger(\vec{k}) b(\vec{k}) + b(\vec{k}) b^\dagger(\vec{k}) \right)
\end{aligned} \tag{10.32}$$

or, after normal ordering

$$H = \int d^3k \omega_{\vec{k}} \left(a^\dagger(\vec{k}) a(\vec{k}) + b^\dagger(\vec{k}) b(\vec{k}) \right) \tag{10.33}$$

Similarly

$$\begin{aligned}
P_i &= \int d^3x \left(\Pi \partial_i \Phi + \Pi^\dagger \partial_i \Phi^\dagger \right) \\
&= \int d^3k k_i \left(a^\dagger(\vec{k}) a(\vec{k}) + b^\dagger(\vec{k}) b(\vec{k}) \right)
\end{aligned} \tag{10.34}$$

Exercise 10.5. Compute (10.33) and (10.34).

Solution. Substitute the Fourier expansion (10.32) into

$$H = \int d^3x \left(\Pi^\dagger(t, \vec{x}) \Pi(t, \vec{x}) + \vec{\nabla} \Phi^\dagger(t, \vec{x}) \cdot \vec{\nabla} \Phi(t, \vec{x}) + m^2 \Phi^\dagger(t, \vec{x}) \Phi(t, \vec{x}) \right) \tag{10.35}$$

and

$$P_i = \int d^3k k_i \left(a^\dagger(\vec{k}) a(\vec{k}) + b^\dagger(\vec{k}) b(\vec{k}) \right) \tag{10.36}$$

and compute the spatial integral as in Exercise 10.3 □

10.2 Free particles

The ground state of a quantum field system is called the *vacuum*. For free fields, it is the state annihilated by all the downward step operators:

$$a(\vec{k})|0\rangle = 0 \tag{10.37}$$

The excited states are

$$|\vec{k}_1, \dots, \vec{k}_n\rangle = a^\dagger(\vec{k}_1) \dots a^\dagger(\vec{k}_n)|0\rangle \tag{10.38}$$

These states span a Hilbert space called the *Fock space* and are eigenstates of energy and momentum:

$$\begin{aligned}
H|\vec{k}_1, \dots, \vec{k}_n\rangle &= \sum_{i=1}^n \omega_{\vec{k}_i} |\vec{k}_1, \dots, \vec{k}_n\rangle \\
\vec{P}|\vec{k}_1, \dots, \vec{k}_n\rangle &= \sum_{i=1}^n \vec{k}_i |\vec{k}_1, \dots, \vec{k}_n\rangle
\end{aligned} \tag{10.39}$$

Exercise 10.6. Derive (10.39).

Solution. For example, using that

$$H|0\rangle = 0 \quad (10.40)$$

one obtains

$$\begin{aligned} H|\vec{k}_1, \dots, \vec{k}_n\rangle &= [H, a^\dagger(\vec{k}_1) \dots a^\dagger(\vec{k}_n)]|0\rangle = [H, a^\dagger(\vec{k}_1)] \dots a^\dagger(\vec{k}_n)|0\rangle + \dots + a^\dagger(\vec{k}_1) \dots [H, a^\dagger(\vec{k}_n)]|0\rangle \\ &= \omega_{\vec{k}_1} a^\dagger(\vec{k}_1) \dots a^\dagger(\vec{k}_n)|0\rangle + \dots + a^\dagger(\vec{k}_1) \dots \omega_{\vec{k}_n} a^\dagger(\vec{k}_n)|0\rangle = \sum_{i=1}^n \omega_{\vec{k}_i} |\vec{k}_1, \dots, \vec{k}_n\rangle \end{aligned} \quad (10.41)$$

Similar computation gives the momentum eigenvalue. $\square$

Exercise 10.7. Assuming $\langle 0|0\rangle = 1$, compute the scalar product of the Fock states (10.38).

Solution. First of all, the scalar product vanishes when the number of particles in the two states are different:

$$\langle \vec{q}_1, \dots, \vec{q}_n | \vec{k}_1, \dots, \vec{k}_m \rangle = 0 \quad n \neq m \quad (10.42)$$

due to

$$a(\vec{k})|0\rangle = 0 \quad \text{or, by conjugation} \quad \langle 0|a^\dagger(\vec{q}) = 0 \quad (10.43)$$

For $n = m = 1$ we get

$$\langle q|k\rangle = \langle 0|a(q)a^\dagger(k)|0\rangle = \langle 0|[a(q), a^\dagger(k)]|0\rangle = \delta^{(3)}(\vec{k} - \vec{q}) \quad (10.44)$$

For $n = m = 2$

$$\begin{aligned} \langle q_1, q_2 | k_1, k_2 \rangle &= \langle 0|a(q_1)a(q_2)a^\dagger(k_1)a^\dagger(k_2)|0\rangle \\ &= \langle 0|[a(q_1)a(q_2), a^\dagger(k_1)a^\dagger(k_2)]|0\rangle \\ &= \delta^{(3)}(\vec{k}_1 - \vec{q}_1)\delta^{(3)}(\vec{k}_2 - \vec{q}_2) + \delta^{(3)}(\vec{k}_1 - \vec{q}_2)\delta^{(3)}(\vec{k}_2 - \vec{q}_1) \end{aligned} \quad (10.45)$$

For the general case

$$\langle q_1, \dots, q_n | k_1, \dots, k_n \rangle = \sum_{\pi \in \text{Perm}(n)} \delta^{(3)}(\vec{k}_1 - \vec{q}_{\pi(1)}) \dots \delta^{(3)}(\vec{k}_n - \vec{q}_{\pi(n)}) \quad (10.46)$$

$\square$

where the sum runs over all permutations π of the numbers $1, \dots, n$.

The physical interpretation of the state is $|\vec{k}_1, \dots, \vec{k}_n\rangle$ that it describes *a situation with n particles of definite momenta* (and therefore energies). Note that these states are *not normalisable* since the particles can be found uniformly over all space (plane waves); however, just as in ordinary quantum mechanics, one can form *wave packets* to obtain normalisable states. The operators $a^\dagger(\vec{k})$ and $a(\vec{k})$ create and destroy a particle, and are called *creation/annihilation operators*.

For the complex Klein-Gordon field there are two sets of particle states: the a -particles and the b -particles. Note that a complex field has a conserved $U(1)$ charge which can be computed

$$\begin{aligned} Q &= \int d^3x i (\Phi^\dagger \partial_0 \Phi - \Phi \partial_0 \Phi^\dagger) \\ &= \int d^3k \left(a^\dagger(\vec{k}) a(\vec{k}) - b^\dagger(\vec{k}) b(\vec{k}) \right) \quad (\text{after normal ordering}) \end{aligned} \quad (10.47)$$

so a -particles have charge $+1$, while b -particles have charge -1 :

$$\begin{aligned} [Q, a^\dagger(\vec{k})] &= +a^\dagger(\vec{k}) & [Q, a(\vec{k})] &= -a(\vec{k}) \\ [Q, b^\dagger(\vec{k})] &= -b^\dagger(\vec{k}) & [Q, b(\vec{k})] &= +b(\vec{k}) \end{aligned} \quad (10.48)$$

Since charge conjugation acts as

$$(C\Phi)(x) = \xi \Phi^\dagger(x) \quad (10.49)$$

it takes a -particles into b -particles and vice versa; *they are anti-particles of each other*.

In a more complicated case, such as e.g. Dirac or electromagnetic fields the formalism is a bit more involved, but the main outline is the same. Every (physical²⁵) Fourier mode of the field corresponds to a given particle state that can be empty or filled in with some number of particles.

For fermionic fields, antisymmetry of the wave function requires that the creation/annihilation operators must anticommute, so the field must be quantised with anticommutators taking the place of the commutators in (10.11)–(10.14). This enforces the *Pauli exclusion principle*: there can only be at most one particle corresponding to any given mode of a fermionic field. In all cases, real fields describe particles that are their own anti-particles, while complex fields describe particles together with their anti-particles:

- the electromagnetic field is real, so a photon is its own anti-particle and has no non-trivial conserved charges (apart from space-time related conserved quantities such as energy-momentum and angular momentum four-tensor);
- a Dirac field describes a spin $1/2$ fermion and its anti-particle;
- a Majorana field (see Subsection 5.10) describes a spin $1/2$ fermion which is its own anti-particle, and there exists a choice of basis (Majorana representation) in which the field is real. Majorana fermions can carry no non-trivial conserved charges since real fields can have no $U(1)$ transformation.

Remark 10.8. For later use we remark that similarly to the complex Klein-Gordon field, the Dirac field Ψ contains particle annihilation and anti-particle creation operators, while the conjugate field $\Psi^\dagger$ (or equivalently $\bar{\Psi} = \Psi^\dagger \gamma^0$) contains particle creation and anti-particle annihilation operators. For the electromagnetic field A_μ contains both the photon creation and annihilation operators, similarly to a real Klein-Gordon field.

Note that the above formalism solves the issue of positive definite conserved probabilities, since the probability is defined by the scalar product on the Fock space which is positive definite; the

²⁵For the electromagnetic field, only the modes with transverse polarisation are physical due to gauge invariance.

$U(1)$ Noether currents are related to conserved charges rather than probability densities. *Quantum field theory also gets rid of the problem of the Dirac sea since all states have energies strictly above the Fock vacuum*, and this solution works both for bosonic and fermionic systems without recourse to the Pauli principle to fill in the negative energy modes.

For any consistent relativistic quantum field theory, locality, relativistic invariance and Hermiticity of the Hamiltonian implies the following two theorems:

Theorem 10.9. (*CPT*) *The full dynamics is invariant under the combined application of charge conjugation, parity and time reversal.*

Theorem 10.10. (*Spin-statistics*) *Particles with integer spin satisfy commutation relations and therefore obey Bose-Einstein statistics (bosons), while particles with half-integer spin satisfy anti-commutation relations and therefore obey Fermi-Dirac statistics (fermions). In terms of more familiar quantum mechanics, the state (wave function) is symmetric under the exchange of two bosons, while it changes sign under the exchange of two fermions (leading to the Pauli exclusion principle).*

The behaviour under the exchange of a boson and fermion is a matter of conventions; the usual choice is that bosons and fermions commute, leading to the state being invariant under the exchange of a boson with a fermion.

Quantum field theory thus solves the issue of combining quantum theory and special relativity into a consistent theory; however, this comes at the price that a quantum field always describes a many-particle system and no consistent closed one-particle formulation exists in general.²⁶

10.3 Dyson series and Feynman diagrams

Turning now to interactions, note that by definition *an interacting field theory has non-linear equations of motion*, which corresponds to a Lagrangian containing terms higher than quadratic order in the fields. For example, consider the Lagrangian

$$\mathcal{L} = \frac{1}{2}\partial_\mu\Phi\partial^\mu\Phi - \frac{1}{2}m^2\Phi^2 - \frac{1}{4!}\lambda\Phi^4 \quad (10.50)$$

for a real scalar field Φ , which leads to the equation of motion

$$(\square + m^2)\Phi = -\frac{\lambda}{6}\Phi^3 \quad (10.51)$$

The Hamiltonian corresponding to this model is

$$H = \underbrace{\int d^3x \frac{1}{2} \left(\Pi^2 + (\vec{\nabla}\Phi)^2 + m^2\Phi^2 \right)}_{H_0} + \lambda \underbrace{\int d^3x \frac{1}{4!} \Phi^4}_V \quad (10.52)$$

²⁶The only known exception (apart from free fields) is the class of so-called integrable quantum field theories which conserve all particle numbers and so their Hilbert space can be split into subspaces with definite number of particles which are consistent relativistic interacting few-body systems. Such theories only exist in $1+1$ space-time dimensions.

where H_0 is the quadratic (free) part and V contains interactions, while λ is a so-called coupling constant. The most straightforward way to treat the dynamics of a quantum field theory is to use *time-dependent perturbation theory* and expand in the coupling constant. The formalism is very general and applies to all Lagrangian field theories; the free part H_0 is quantised in the canonical way, and the theory is treated in the interaction (Dirac) picture, where operators evolve by the free part H_0 and states evolve by the interacting part λV . As a result, the Hamiltonian can be rewritten in terms of the free field creation and annihilation operators by substituting the expansion (10.16) for Φ , leading to a model of infinitely many non-linearly coupled oscillators. From there on, one can proceed using standard quantum mechanical perturbation theory; however, there exists a much more efficient treatment using a so-called *Dyson series*.

In the interaction picture, the interacting part (as well as all physical operators) is itself time-dependent

$$V(t) = e^{iH_0 t} V(0) e^{-iH_0 t} \quad (10.53)$$

and the state of the system evolves according to

$$i \frac{\partial}{\partial t} |\Psi(t)\rangle = \lambda V(t) |\Psi(t)\rangle \quad (10.54)$$

which can be solved by *successive iterations* given the initial state $|\Psi(t_0)\rangle$ at time t_0 . Rewriting the Schrödinger equation in an integral form

$$|\Psi(t)\rangle = \Psi(t_0) - i\lambda \int_{t_0}^t d\tau V(\tau) |\Psi(\tau)\rangle \quad (10.55)$$

successive iteration starting from $|\Psi(t_0)\rangle$ gives the series

$$\begin{aligned} |\Psi(t)\rangle = & \left(1 - i\lambda \int_{t_0}^t d\tau V(\tau) + (-i\lambda)^2 \int_{t_0}^t d\tau_1 \int_{t_0}^{\tau_1} d\tau_2 V(\tau_1) V(\tau_2) \right. \\ & \left. + (-i\lambda)^3 \int_{t_0}^t d\tau_1 \int_{t_0}^{\tau_1} d\tau_2 \int_{t_0}^{\tau_2} d\tau_3 V(\tau_1) V(\tau_2) V(\tau_3) + \dots \right) |\Psi(t_0)\rangle \end{aligned} \quad (10.56)$$

Introducing the *time-ordered product*

$$T[V(t_1) \dots V(t_n)] = V(t_{\pi(1)}) \dots V(t_{\pi(n)}) \quad (10.57)$$

where the permutation π is a reordering of the times according to

$$t_{\pi(1)} \geq \dots \geq t_{\pi(n)} \quad (10.58)$$

one can rewrite

$$\begin{aligned} |\Psi(t)\rangle &= \left\{ 1 + \sum_{n=1}^{\infty} \frac{(-i\lambda)^n}{n!} \int_{t_0}^t d\tau_1 \dots \int_{t_0}^t d\tau_n T[V(\tau_1) \dots V(\tau_n)] \right\} |\Psi(t_0)\rangle \\ &= T \exp \left\{ -i\lambda \int_{t_0}^t d\tau V(\tau) \right\} |\Psi(t_0)\rangle \end{aligned} \quad (10.59)$$

where the last expression is just a *short-hand for the series expansion* using its structural similarity to an exponential expansion (it would exactly be an exponential in the absence of time-ordering).

Introducing a convenient basis on the Hilbert space, corresponding to the eigenbasis of the free Hamiltonian H_0 :

$$H_0|n\rangle = E_n|n\rangle \quad (10.60)$$

which are supposed to be orthonormal, and expanding the evolving state in this basis

$$|\Psi(t)\rangle = \sum_n c_n(t)|n\rangle \quad (10.61)$$

the time evolution can be written as

$$c_n(t) = \sum_m U_{nm}(t, t_0) c_m(t_0) \quad (10.62)$$

with

$$U_{nm}(t, t_0) = \langle n | T \exp \left\{ -i\lambda \int_{t_0}^t d\tau V(\tau) \right\} | m \rangle \quad (10.63)$$

These are the transition amplitudes for processes starting from a free state $|m\rangle$ and ending with a free state $|n\rangle$, induced by the interaction.

For example, in a scattering process one assumes that the incoming state in the asymptotic past and the outgoing state in the asymptotic future consists of very far separated particles, so that they are essentially free. The mathematical object describing a scattering is the so-called *scattering or S matrix*

$$S_{nm} = U_{nm}(\infty, -\infty) \quad (10.64)$$

from which the probabilities of scattering processes can be calculated. Instead of going into the details we show that the Dyson series allows a very appealing interpretation. Namely, we can consider inserting full sets of states as follows

$$U_{nm}(t, t_0) = \delta_{nm} - i\lambda \int_{t_0}^t d\tau \langle n | V(\tau) | m \rangle + (-i\lambda)^2 \int_{t_0}^t d\tau_1 \int_{t_0}^{\tau_1} d\tau_2 \sum_r \langle n | V(\tau_1) | r \rangle \langle r | V(\tau_2) | m \rangle + \dots \quad (10.65)$$

What is the meaning of these terms? The term, first order in λ describes that the initial state $|m\rangle$ exists up to time τ when it just happens to transition into the final state $|n\rangle$; since τ can be any instant between t_0 and t_1 , it has to be integrated over. The second term is a two-step process, which happens via an intermediate state $|r\rangle$: $|m\rangle$ transitions to $|r\rangle$ at time τ_2 and later at τ_1 $|r\rangle$ transitions into the final state $|n\rangle$. It is necessary now to integrate over the times $\tau_{1,2}$ and also to sum over the possible intermediate states $|r\rangle$. Now we know that V is itself the spatial integral of an interaction density $\mathcal{V}(x)$, so in fact we have

$$\begin{aligned} U_{nm}(t, t_0) = & \delta_{nm} - i\lambda \int_{t_0}^t d\tau \int d^3x \langle n | \mathcal{V}(\tau, \vec{x}) | m \rangle \\ & + (-i\lambda)^2 \int_{t_0}^t d\tau_1 \int d^3x_1 \int_{t_0}^{\tau_1} d\tau_2 \int d^3x_2 \sum_r \langle n | \mathcal{V}(\tau_1, \vec{x}_1) | r \rangle \langle r | \mathcal{V}(\tau_2, \vec{x}_2) | m \rangle \\ & + \dots \end{aligned} \quad (10.66)$$

so in fact the integrals are over space-time points (events) where/when the interactions eventually happen. By our choice of basis, the states $|n\rangle$, $|m\rangle$, $|r\rangle$ etc. are simply free particles propagating in space-time, so one can draw a diagram corresponding to the history embodied by each term

in this expansion, which is called a *Feynman diagram*. A *Feynman diagram* is in fact a way to encode a particular integral term contributing to the Dyson expansion. The possible histories are constrained by selection rules stating which are the states for which the interaction matrix elements do not vanish; these selection rules are in fact surprising simple to read out from the Lagrangian. The only computational task left is to perform the state summations and the integrals, for which powerful techniques exist. Instead of going into technical details, we give a few examples of particle processes and corresponding Feynman diagrams.

10.3.1 Interacting real scalar field

Let us now consider the theory

$$\mathcal{L}_{\text{scalar}} = \frac{1}{2} \partial_\mu \Phi \partial^\mu \Phi - \frac{1}{2} m^2 \Phi^2 - \frac{1}{4!} \lambda \Phi^4 \quad (10.67)$$

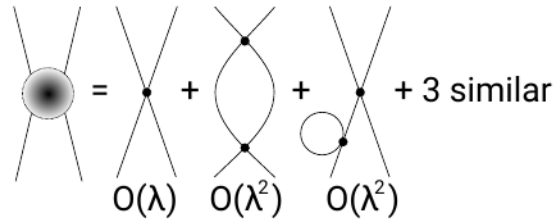
The interaction is given by the operator

$$\mathcal{V} = \frac{1}{4!} \lambda \Phi^4 \quad (10.68)$$

Using the expansion (10.16), the interaction contains terms consisting of 4 annihilation operators $aaaa$, terms of form $a^\dagger aaa$, $a^\dagger a^\dagger aa$, $a^\dagger a^\dagger a^\dagger a$ and $a^\dagger a^\dagger a^\dagger a^\dagger$. Considering a matrix element $\langle n | \mathcal{V} | m \rangle$, these correspond to the following processing happening between $|m\rangle$ and $|n\rangle$ at some space-time point:

- $aaaa$: 4 particles annihilating from $|m\rangle$, rest continuing to $|n\rangle$;
- $a^\dagger aaa$: 3 particles in $|m\rangle$ fusing into one particle in $|n\rangle$;
- $a^\dagger a^\dagger aa$: 2 particles in $|m\rangle$ scattering into two particles in $|n\rangle$;
- $a^\dagger a^\dagger a^\dagger a$: one particle in $|m\rangle$ splitting into three particles in $|n\rangle$; and
- $a^\dagger a^\dagger a^\dagger a^\dagger$: four particles created and propagating into $|n\rangle$.

Since at any step, energy (and also momentum) must be conserved, the $aaaa$ and $a^\dagger a^\dagger a^\dagger a^\dagger$ terms can never contribute, as energy conservation cannot be ensured. In any case, the histories that must be drawn are those of free bosonic particles propagating and interacting at four-legged vertices. For example, up to second order in λ two particles scattering into two particles have the following Feynman diagrams where time runs from bottom to top:



10.3.2 Quantum electrodynamics

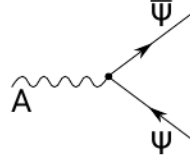
Let us now consider the simplest version of quantum electrodynamics

$$\mathcal{L}_{QED} = \bar{\Psi}(i\not{\partial} - q\not{A} - m)\Psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} \quad (10.69)$$

where A^μ is the electromagnetic field, while Ψ is a Dirac field describing electrons and positrons. The interaction vertex is

$$\mathcal{V}(x) = -e\bar{\Psi}(x)\gamma^\mu\Psi(x)A_\mu(x) \quad (10.70)$$

where we put $q = -e$, with $e = 1.6 \dots 10^{-19}\text{C}$ denoting the elementary charge. Now the field A is real and so contains both creation and annihilation operators for photons, so that at a given vertex one can either have a photon annihilated or created. Since Ψ is complex, it contains an electron annihilation and a positron creation operator, while $\bar{\Psi} = \Psi^\dagger\gamma^0$ contains an electron creation and a positron annihilation operator. The vertex is usually drawn in the following way:

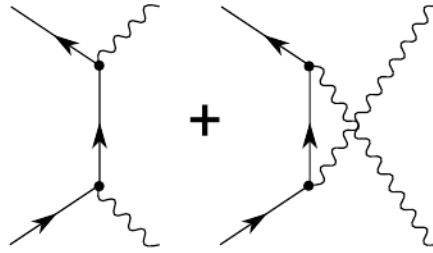


The *two fermionic legs are distinguished by arrows*, which also help properly drawing the diagram since *they can only be joined preserving orientation to ensure a non-zero value* via the proper ordering of creation and annihilation operators. According to remark [10.8](#) in Subsection [10.2](#), the leg Ψ corresponds either to creating a positron or destroying an electron, while the leg $\bar{\Psi}$ can do just the opposite; the leg A can either create or destroy a photon.

For example, a Compton scattering

$$e^- + \gamma \rightarrow e^- + \gamma \quad (10.71)$$

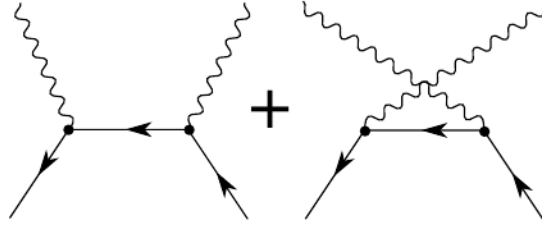
is given to $O(e^2)$ by the following two diagrams:



The amplitude of electron-positron annihilation

$$e^- + e^+ \rightarrow \gamma + \gamma \quad (10.72)$$

to $O(e^2)$ is given by



where the line with orientation running backwards in time corresponds to a positron, which can easily be ascertained by looking at the mode contents of Ψ and $\bar{\Psi}$.

11 Weak interactions revisited

We already covered the basics of weak interactions in Section 6 however, there is much more to learn in order to understand our current picture of them.

11.1 Lepton number conservation

Consider the following weak interaction processes:

$$\begin{aligned} n &\longrightarrow p + e^- + \nu \\ p + e^- &\longrightarrow n + \nu \\ \mu^\pm &\longrightarrow e^\pm + \nu + \nu \\ \pi^\pm &\longrightarrow \mu^\pm + \nu \end{aligned}$$

All of these involve one or more neutrinos, albeit all known conservation laws (energy, momentum, electric charge, baryon number) would allow for them to proceed without them (they do break some other conservation laws, such as isospin, though). However, their neutrinoless versions are never observed and it is natural to suppose that there is some conserved charge that forbids them. This is called *leptonic number* L and by convention e^- and μ^- have $L = +1$. In that case, one must assign $L = 1$ to the neutrino, which allows to designate the particle appearing in the neutron β -decay as an antineutrino, leading to

$$\begin{aligned} n &\longrightarrow p + e^- + \bar{\nu} \\ p + e^- &\longrightarrow n + \nu \\ \mu^\pm &\longrightarrow e^\pm + \bar{\nu} + \nu \\ \pi^+ &\longrightarrow \mu^+ + \nu \\ \pi^- &\longrightarrow \mu^- + \bar{\nu} \end{aligned}$$

However, it is still a question why we do not observe the process

$$\mu^\pm \longrightarrow e^\pm + \gamma$$

but only a decay with two neutrinos involved? This leads to the idea that the two neutrinos are necessary because they are not the same, and there is a separate *electronic and muonic number* to be conserved. The neutrino related to the electron is called the *electron neutrino* and denoted as ν_e , while the one corresponding to the muon is the *muon neutrino* and is denoted as ν_μ . Now we have the complete version of these processes

$$\begin{aligned} n &\longrightarrow p + e^- + \bar{\nu}_e \\ p + e^- &\longrightarrow n + \nu_e \\ \mu^- &\longrightarrow e^- + \bar{\nu}_e + \nu_\mu \\ \mu^+ &\longrightarrow e^+ + \bar{\nu}_\mu + \nu_e \\ \pi^+ &\longrightarrow \mu^+ + \nu_\mu \\ \pi^- &\longrightarrow \mu^- + \bar{\nu}_\mu \end{aligned}$$

This leads to the following assignment of lepton numbers:

	L_e	L_μ	$L = L_e + L_\mu$
e^-	1	0	1
ν_e	1	0	1
μ^-	0	1	1
ν_μ	0	1	1

with a separate conservation for the electronic lepton number L_e and the muonic one L_μ .

11.1.1 The Cowan-Reines neutrino experiment

The existence of the neutrino was proposed by Pauli in 1930, but it took more than twenty years until direct evidence was found by C.L. Cowan and F. Reines in 1956. They used the reaction

$$\bar{\nu}_e + p \longrightarrow n + e^+ \quad (11.1)$$

The cross section for this reaction can be computed using the lifetime of the neutron. Denoting the weak interaction Hamiltonian by H_W , the decay amplitude of the neutron computed to first order in perturbation theory involves the matrix element

$$\langle p e^- \bar{\nu}_e | H_W | n \rangle \quad (11.2)$$

which can be estimated from the lifetime of the neutron. The scattering (11.1) involves the matrix element

$$\langle n e^+ | H_W | p \bar{\nu}_e \rangle \quad (11.3)$$

which is related to (11.2) by the so-called crossing symmetry, which is generally valid in quantum field theory. Crossing symmetry relates the amplitudes of two processes which differ from each other by transferring a particle of momentum $\vec{p}$ from the initial state to its antiparticle in the final state with momentum $-\vec{p}$; in the case at hand it states that

$$\begin{aligned} \langle p(\vec{p}_1) e^-(\vec{p}_2) \bar{\nu}_e(\vec{p}_3) | H_W | n(\vec{p}_4) \rangle &= \langle p(\vec{p}_1) \bar{\nu}_e(\vec{p}_3) | H_W | n(\vec{p}_4) e^+(-\vec{p}_2) \rangle \\ &= \langle n(\vec{p}_4) e^+(-\vec{p}_2) | H_W | p(\vec{p}_1) \bar{\nu}_e(\vec{p}_3) \rangle^* \end{aligned} \quad (11.4)$$

Detailed calculation predicts that the cross section of reaction (11.1) is

$$\sigma \approx 6 \cdot 10^{-44} \text{cm}^2 \quad (11.5)$$

This is an extremely small value; for example, typical nuclear cross sections are of the order of 1 barn = 10^{-24}cm^2 .

Exercise 11.1. Compute the typical length travelled by an electron antineutrino without interaction in lead from the cross section (11.5).

Solution. Lead has atomic number 82, a density of 11.34 g/cm³ and an atomic weight of 207.2 g/mol, leading to a proton density

$$n = \frac{82 \cdot 11.34}{207.2} \cdot 6 \cdot 10^{23} \frac{1}{\text{cm}^3} = 2.7 \cdot 10^{24} \frac{1}{\text{cm}^3}$$

so the typical free path is

$$\lambda_\nu = \frac{1}{\sigma n} = 6.2 \cdot 10^{18} \text{cm} = 6.5 \text{ light years}$$

This result shows that neutrinos can pass through ordinary materials with astonishing ease! $\square$

To have a sufficiently strong signal, one needs a very potent neutrino source to compensate for the small interaction probability. Fortunately, the plutonium breeding reactors used for the development and building of nuclear weapons provided just such a source. In a fission process the reaction

$$n + {}^{235}\text{U} \longrightarrow \text{neutron rich elements} + 2.5 n + 200 \text{ MeV} \quad (11.6)$$

is followed by β -decays generating about 6 $\bar{\nu}$ per fission, which have an average energy of 3 MeV each. The neutrino intensity is given by

$$N_{\nu} = 6 \times \frac{P_{\text{reactor}}}{200 \text{ MeV}} \quad (11.7)$$

and using

$$1 \text{ MW} = \frac{1}{1.6 \cdot 10^{-19}} \frac{\text{MeV}}{\text{s}} \quad (11.8)$$

every MW of reactor power corresponds to a flux of about $1.9 \cdot 10^{17} \bar{\nu}/\text{s}$. A preliminary experiment built at the Hanford Site already achieved a tentative identification of neutrinos in 1953. The follow-up experiment was built 11 meters from reactor P at the Savannah River Site, where the neutrino flux density was of the order of $10^{13}/\text{cm}^2/\text{s}$.

Exercise 11.2. What is the neutrino flux density of a Paks reactor (1375 MW gross power) $d = 11$ meters away, assuming a spherical distribution of neutrino flux?

Solution. The full neutrino intensity is

$$N_{\nu} = 2.6 \cdot 10^{20} \frac{1}{\text{s}} \quad (11.9)$$

so the flux is

$$\Phi_{\nu} = \frac{N_{\nu}}{4\pi d^2} = 1.7 \cdot 10^{13} \frac{1}{\text{s} \cdot \text{cm}^2} \quad (11.10)$$

□

Since the reaction cross section is so small, much care had to be taken to exclude the cosmic ray background and so the experiment was built in shielded place 12 meters underground. The detector was made of two water tanks, with a total volume of 200 litres, and the idea of the detection was that the positrons created in the process would annihilate to two γ photons which could be subsequently detected by scintillation counters.

Exercise 11.3. Estimate the number of events expected with 200 litres of water as a detector and a neutrino flux density of $\Phi_{\nu} = 10^{13}/\text{cm}^2/\text{s}$.

Solution. With just a single proton as a target, the expected event rate would be

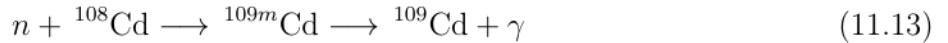
$$\sigma \Phi_{\nu} = 6 \cdot 10^{-31} \frac{1}{\text{s}} \quad (11.11)$$

However, 200 kg of water contains

$$\frac{200 \text{ kg}}{0.018 \text{ kg/mol}} \cdot 6 \cdot 10^{23} \frac{1}{\text{mol}} \cdot 10 = 6.7 \cdot 10^{28} \text{ protons} \quad (11.12)$$

leading to an event rate of around 145 neutrino events per hour. Note that this is a crude estimate, and the actual number of detected events is expected to be substantially lower due to several factors affecting detection efficiency. $\square$

However, the experiment was not conclusive enough, so another detection method was added by dissolving 40 kg of cadmium salt (CdCl_2) in the water, which is a highly efficient neutron absorber via the process



The delay between the γ rays from the positron and from the neutron absorption was computed to be within $5 \mu\text{s}$; in this way, signals from stray positrons could be excluded with coincidence measurements since their simultaneous double γ -rays were not followed by the neutron absorption signal γ rays within five microseconds. Months of data collection resulted in a detection rate of three neutrinos per hour. To demonstrate that these were indeed neutrino events from the reactor, they shut down the reactor and the signal went away. From the detection rate, they computed a measured cross section of

$$\sigma_{exp} = 6.3 \cdot 10^{-44} \text{ cm}^2 \quad (11.14)$$

which was in a very good agreement with expectations. Their results were published in Science in the issue of 20th July 1956. For his work on neutrino physics, F. Reines shared the Nobel Prize in 1995 with M.L. Perl, who discovered the τ lepton (C. Cowan died in 1974).

It is interesting to note that the antineutrinos from nuclear reactors never trigger the reaction



which is energetically favourable compared to (11.1); it can only proceed when the incoming particle is ν_e , which is not generated in the reactor. In fact, this was the observation that led to the idea of lepton number conservation proposed in 1952 independently by E. Konopinski and H. Mahmoud (US), J. Zeldovich (USSR) and G. Marx (Hungary).

11.1.2 Discovering the muon neutrino

As noted above, the two neutrinos produced in muon decay must be of two different types, otherwise there is no explanation for the absence of the decay $\mu^\pm \longrightarrow e^\pm + \gamma$. To establish this conclusively, a new experimental technique was used to create a neutrino beam, building on ideas proposed independently by B. Pontecorvo (Dubna, USSR) and M. Schwartz (Columbia, USA). M. Schwartz joined forces with L. Lederman and J. Steinberger to build a neutrino beam of sufficiently high intensity. They used the 15 GeV proton beam produced by the Alternating Gradient Synchrotron (AGS) at Brookhaven which was directed into a beryllium target, producing large numbers of pions (and other particles). The pions decay as

$$\pi^\pm \longrightarrow \mu^\pm + \nu \quad (11.16)$$

Since the original protons' large momentum is conserved, the end result is a beam consisting of pions, muons and neutrinos, some of the original protons and other particles produced in the

collision, moving in the direction of the original proton beam. To select the neutrinos, they applied a shielding wall of steel 13.5 meters thick, constructed of armour plates of a recently decommissioned battleship (USS *Missouri*). It was estimated that after this, only about 1 in every 10^{24} pions survived, with similar attenuation rates for all other particles except the neutrinos (some muons passed through as well, due to their high penetration ability in matter).

To detect the neutrinos, they built a 10-ton spark chamber consisting of ninety one-inch thick aluminium plates, with the gaps between the plates filled with gas. The neutrinos streaming through the chamber were expected to occasionally produce a positron or a muon by hitting a proton in an aluminium nucleus, and the resulting charged particle would create a spark track due to the high voltages applied to the plates. After several months of beam time and using $3.48 \cdot 10^{17}$ original beam protons, 113 candidate events were found; of these 51 events were identified as coming from the neutrino beam, the rest originating in various other processes (low-energy muons directly from the AGS beam, background neutrons and cosmic rays).

The tracks originating from these events were all due to muons, shown by the fact that they passed directly through the Al plates, which is very different from the behaviour of electrons. So the neutrinos making up the beam only created muons by the processes

$$\begin{aligned}\nu_\mu + n &\longrightarrow \mu^- + p \\ \bar{\nu}_\mu + p &\longrightarrow \mu^+ + n\end{aligned}\tag{11.17}$$

and therefore were of a different species than the electron neutrinos generated in beta decays. For the discovery of the muon neutrino, Schwartz, Lederman and Steinberger were awarded the Nobel Prize in 1988.

11.2 Charged current interactions

The first theory proposed to describe β -decay was Fermi's "tentativo". It describes a *point-like interaction between the neutron, proton, electron and neutrino described by the appropriate Dirac fields*. It is customary to use a notation in which the Dirac field Ψ_n of the neutron is denoted simply by n , and similarly for the other particles. In modern notations Fermi's interaction can be written as²⁷

$$\mathcal{L}_F = G'_F (\bar{p} \gamma^\mu n) (\bar{e} \gamma_\mu \nu_e) + \text{h.c.}\tag{11.18}$$

with h.c. denoting the Hermitian conjugate term, and G'_F is a coupling constant. This corresponds to a point-like interaction between the particles, and describes processes such as

$$\begin{aligned}n &\longrightarrow p + e^- + \bar{\nu}_e \\ p &\longrightarrow n + e^+ + \nu_e \\ \nu_e + n &\longrightarrow p + e^- \\ \bar{\nu}_e + p &\longrightarrow n + e^+\end{aligned}\tag{11.19}$$

The dimension of G_F can be computed as follows. The Dirac action

$$\int d^4x \bar{\Psi} \not{\partial} \Psi\tag{11.20}$$

²⁷To follow the arguments of this Section, it is important to recall the properties of the Dirac matrices derived in subsections 5.3, 5.5 and 5.7. To understand the transitions corresponding to the different interaction terms, consult also Remark 10.8 in Subsection 10.2

is dimensionless. Since x has dimension eV^{-1} and therefore ∂ has dimension of eV , we obtain

$$[\Psi] = \text{eV}^{3/2} \quad (11.21)$$

Since

$$\int d^4x \mathcal{L}_F \quad (11.22)$$

is a term in the action, it is similarly dimensionless, leading to

$$[G_F] = \text{eV}^{-2} \quad (11.23)$$

Fermi's Lagrangian does not contain parity violation: both the left and the right handed neutrinos participate with equal weight. Writing

$$\begin{aligned} \nu_e &= \nu_{eL} + \nu_{eR} \\ \nu_{eL} &= \frac{1 - \gamma_5}{2} \nu_e \\ \nu_{eR} &= \frac{1 + \gamma_5}{2} \nu_e \end{aligned} \quad (11.24)$$

and similarly for the electron, one gets

$$\bar{e} \gamma_\mu \nu_e = \bar{e}_L \gamma_\mu \nu_{eL} + \bar{e}_R \gamma_\mu \nu_{eR} \quad (11.25)$$

Exercise 11.4. Show that from

$$\Psi_{R,L} = \frac{1 \pm \gamma_5}{2} \Psi \quad (11.26)$$

it follows that

$$\bar{\Psi}_{R,L} = \bar{\Psi} \frac{1 \mp \gamma_5}{2} \quad (11.27)$$

Solution. E.g. starting from

$$\Psi_R = \frac{1 + \gamma_5}{2} \Psi \quad (11.28)$$

we get

$$\bar{\Psi}_R = \Psi_R^\dagger \gamma_0 = \Psi^\dagger \frac{1 + \gamma_5}{2} \gamma_0 = \Psi^\dagger \gamma_0 \frac{1 - \gamma_5}{2} = \bar{\Psi} \frac{1 - \gamma_5}{2} \quad (11.29)$$

Similarly for Ψ_L . □

The left-handedness of the neutrino can be enforced by modifying the leptonic term to

$$\bar{e} \gamma_\mu \nu_e \longrightarrow \bar{e} \gamma_\mu \nu_{eL} = \frac{1}{2} \bar{e} \gamma_\mu (1 - \gamma_5) \nu_e \quad (11.30)$$

which also enforces that only the left-handed electron participates, since

$$\gamma_\mu (1 - \gamma_5) = (1 + \gamma_5) \gamma_\mu \quad (11.31)$$

So we are lead to the Lagrangian

$$\mathcal{L}_F = \frac{G_F}{\sqrt{2}} (\bar{p} \gamma^\mu (1 - \alpha \gamma_5) n) (\bar{e} \gamma_\mu (1 - \gamma_5) \nu_e) + \text{h.c.} \quad (11.32)$$

where the coupling constant was slightly redefined, and also allow a γ_5 term for the hadronic terms, although its value is not fixed directly by any principle. For a free neutron the measured value is $\alpha = 1.25$, although its effective value is changed by the nuclear environment, reflecting the fact that the nucleon is a composite object. For the coupling constant, fitting its value from the neutron lifetime gives

$$G_F \approx 10^{-5} \frac{1}{m_p^2} \quad (11.33)$$

where m_p is the proton mass.

The decay of the muon

$$\mu^- \rightarrow e^- + \nu_\mu + \bar{\nu}_e \quad (11.34)$$

can be similarly described by a four-fermion interaction

$$\frac{G_\mu}{\sqrt{2}} (\bar{\nu}_\mu \gamma^\alpha (1 - \gamma_5) \mu) (\bar{e} \gamma_\alpha (1 - \gamma_5) \nu_e) + \text{h.c.} \quad (11.35)$$

where due to the fact that the muon is just like a heavier version of the electron, the form of the interaction term was chosen to match that of the electron. Comparing the prediction of this Lagrangian (with some electromagnetic corrections included) with the experimentally observed muon decay time leads to the value

$$G_\mu = 1.1663787(6) \cdot 10^{-5} \frac{1}{\text{GeV}^2} \quad (11.36)$$

Using $m_p \approx 938 \text{ MeV}$ it is easy to check that G_F and G_μ are eventually approximately equal. Indeed, this was the observation that eventually led to the idea that *these processes were essentially different manifestations of a universal weak interactions*, which takes the form

$$\mathcal{L}_{CC} = \frac{G_F}{\sqrt{2}} (J_\alpha^{C\dagger} J^{C\alpha} + \text{h.c.}) \quad (11.37)$$

where

$$J_\alpha^C = \bar{e} \gamma_\alpha (1 - \gamma_5) \nu_e + \bar{\mu} \gamma_\alpha (1 - \gamma_5) \nu_\mu + \text{hadronic part} \quad (11.38)$$

is a so-called *charged current* since it describes processes where the electric charge is changed by a unit: e.g. e^- turns into a ν_e while p turns into n , or μ^- turns into ν_μ while a pair of e^- and $\bar{\nu}_e$ are created. In contrast, the electromagnetic current which enters the interaction with the electromagnetic field (cf. Subsection [10.3.2](#))

$$J_\alpha^{EM} = \bar{e} \gamma_\alpha e + \bar{\mu} \gamma_\alpha \mu + \text{hadronic part} \quad (11.39)$$

is *neutral due to the force being mediated by a particle (the photon) which carries no charge*.

Exercise 11.5. Show that it follows that

$$(\bar{\mu} \gamma_\alpha (1 - \gamma_5) \nu_\mu)^\dagger = \bar{\nu}_\mu \gamma_\alpha (1 - \gamma_5) \mu \quad (11.40)$$

so the term [\(11.35\)](#) describing the muon decay really occurs in [\(11.37\)](#).

Solution.

$$\begin{aligned}
(\bar{\mu}\gamma_\alpha(1-\gamma_5)\nu_\mu)^\dagger &= (\mu^\dagger\gamma_0\gamma_\alpha(1-\gamma_5)\nu_\mu)^\dagger = \nu_\mu^\dagger(1-\gamma_5)\gamma_\alpha^\dagger\gamma_0\mu \\
&= \nu_\mu^\dagger(1-\gamma_5)\gamma_0\gamma_\alpha\mu = \nu_\mu^\dagger\gamma_0(1+\gamma_5)\gamma_\alpha\mu \\
&= \bar{\nu}_\mu\gamma_\alpha(1-\gamma_5)\mu
\end{aligned} \tag{11.41}$$

□

Since the nucleon is a composite object, we expect that the current can be written in a simpler form in terms of elementary quarks:

$$J_\alpha^C = \bar{e}\gamma_\alpha(1-\gamma_5)\nu_e + \bar{\mu}\gamma_\alpha(1-\gamma_5)\nu_\mu + \bar{d}\gamma_\alpha(1-\gamma_5)u \tag{11.42}$$

This current corresponds to $\Delta Q = -1$ while its conjugate contains $\Delta Q = +1$ terms; all terms in the interaction (11.37) therefore preserve the conservation of electric charge. Using this current, (11.37) contains e.g. the term

$$\frac{G_F}{\sqrt{2}} (\bar{e}\gamma^\alpha(1-\gamma_5)\nu_e) (\bar{u}\gamma_\alpha(1-\gamma_5)d) \tag{11.43}$$

responsible for the process

$$d \longrightarrow u + e^- + \bar{\nu}_e \tag{11.44}$$

which is the quark level process behind the neutron β -decay, or the conjugate term

$$\frac{G_F}{\sqrt{2}} (\bar{\nu}_e\gamma^\alpha(1-\gamma_5)e) (\bar{d}\gamma_\alpha(1-\gamma_5)u) \tag{11.45}$$

resulting in e.g.

$$u + \bar{\nu}_e \longrightarrow d + e^+ \tag{11.46}$$

which is the quark level process behind the reaction

$$\bar{\nu}_e + p \longrightarrow n + e^+ \tag{11.47}$$

used in the Cowan-Reines reactor neutrino detection experiment.

Exercise 11.6. Write down the interaction terms that are responsible for the processes

$$\begin{aligned}
\bar{\nu}_\mu + p &\longrightarrow n + \mu^+ \\
\nu_\mu + n &\longrightarrow p + \mu^-
\end{aligned} \tag{11.48}$$

Solution. The quark level process for the first one is

$$u + \bar{\nu}_\mu \longrightarrow d + \mu^+ \tag{11.49}$$

and the interaction term behind that is the muonic analogue of (11.45):

$$\frac{G_F}{\sqrt{2}} (\bar{\nu}_\mu\gamma^\alpha(1-\gamma_5)\mu) (\bar{d}\gamma_\alpha(1-\gamma_5)u) \tag{11.50}$$

while the second process corresponds to

$$d + \nu_\mu \longrightarrow u + \mu^- \quad (11.51)$$

and is generated by the Hermitian conjugate of (11.50):

$$\frac{G_F}{\sqrt{2}} (\bar{\mu}\gamma^\alpha(1 - \gamma_5)\nu_\mu) (\bar{u}\gamma_\alpha(1 - \gamma_5)d) \quad (11.52)$$

□

11.3 FCNC, the GIM mechanism and the discovery of charm

11.3.1 The Cabibbo angle

The current (11.42) does not contain the strange quark and therefore *cannot explain the weak decays of strange particles*. These require transitions between u and s , i.e., an inclusion of a term of the form

$$\bar{s}\gamma_\alpha(1 - \gamma_5)u \quad (11.53)$$

The proper way to do that was found out by N. Cabibbo in 1963. He started from the following observations:

- *Universality of weak interactions: transitions $u \longleftrightarrow d$, $e \longleftrightarrow \nu_e$ and $\mu \longleftrightarrow \nu_\mu$ all had similar amplitudes.*
- *Transitions with change in strangeness $\Delta S = 1$ had amplitudes equal to about 25% of those with $\Delta S = 0$.*

Adding considerations based on Gell-Mann's eightfold way, Cabibbo proposed *a mixing between $\Delta S = 0$ and $\Delta S = 1$ charged currents*, while keeping the universality of the weak coupling. It is easier to describe his idea in terms of quarks: it is equivalent to the statement that in the charged current (11.42), d must be replaced with

$$d_\theta = d \cos \theta + s \sin \theta \quad (11.54)$$

leading to the following expression:

$$\begin{aligned} J_\alpha^C &= \bar{e}\gamma_\alpha(1 - \gamma_5)\nu_e + \bar{\mu}\gamma_\alpha(1 - \gamma_5)\nu_\mu + \bar{d}_\theta\gamma_\alpha(1 - \gamma_5)u \\ &= \bar{e}\gamma_\alpha(1 - \gamma_5)\nu_e + \bar{\mu}\gamma_\alpha(1 - \gamma_5)\nu_\mu + \cos \theta \bar{d}\gamma_\alpha(1 - \gamma_5)u + \sin \theta \bar{s}\gamma_\alpha(1 - \gamma_5)u \end{aligned} \quad (11.55)$$

Note that the *Cabibbo rotation* ((11.54)) is an $SU(3)$ transformation in flavour space, which keeps universality of the weak coupling. Its content is that the mass eigenstates of quarks u, d, s differ from the weak interaction eigenstates u, d_θ, s_θ by a flavour rotation. Another way of viewing this transformation is that both the strong and weak interaction break $SU(3)$ flavour symmetry, but *there is a mismatch in the symmetry breaking direction in the three-dimensional flavour space* which is given by the Cabibbo rotation: neglecting isospin breaking, the strong interaction breaks $SU(3)$ with $m_s \gg m_u = m_d$, i.e. the direction of s , while according to (11.54), the combination

$$s_\theta = -d \sin \theta + s \cos \theta \quad (11.56)$$

does not couple to weak interactions at all, while u and d_θ couple with the universal strength G_F , so the weak interactions breaks flavour symmetry in the s_θ direction.

The value of the Cabibbo angle θ can be obtained from the ratio between $\Delta S = 0$ and $\Delta S = 1$ transition amplitudes:

$$\tan \theta \approx \frac{1}{4} \Rightarrow \theta \approx 14^\circ \quad (11.57)$$

and so

$$\cos \theta \approx 0.97 \quad \sin \theta \approx 0.24 \quad (11.58)$$

Therefore the term (11.35) responsible for muon decay is unchanged, but (11.43) is replaced by

$$\frac{G_F}{\sqrt{2}} \cos \theta (\bar{e} \gamma^\alpha (1 - \gamma_5) \nu_e) (\bar{u} \gamma_\alpha (1 - \gamma_5) d) \quad (11.59)$$

so the Fermi constant G_β measured from β -decay is expected to be slightly smaller than the value G_μ measured from the muon lifetime

$$G_\beta = G_F \cos \theta \approx 0.97 G_F < G_\mu = G_F \quad (11.60)$$

Indeed, measurements of G_β from the β -decay of ^{14}O were consistent with this prediction and so the charged current modified by the Cabibbo rotation (11.55) seemed consistent with available observations.

11.3.2 FCNC and the discovery of charm

The mixing between the d and s quarks, however, introduced a new problem: *flavour changing neutral currents* (FCNC). For example, such a decay would be

$$K_L^0 \rightarrow \mu^+ + \mu^- \quad (11.61)$$

This decay mode is in fact known and has a branching ratio $(6.84 \pm 0.11) \cdot 10^{-9}$, so it happens about seven times in every billion decays. Since K_L^0 is a combination of $K^0 = \bar{s}d$ and $\bar{K}^0 = s\bar{d}$, such a decay would, for example, follow from an interaction like

$$\frac{G_F}{\sqrt{2}} (\bar{\mu} \gamma^\alpha (1 - \gamma_5) \mu) (\bar{d} \gamma_\alpha (1 - \gamma_5) s) + \text{h.c.} \quad (11.62)$$

The currents in the product are neutral, since the outgoing lepton/quark has the same charge as the incoming one, but the quark current changes flavour, hence the name. As discussed later in Subsection 13.3.1, weak neutral current interactions do exist, but they preserve flavour. At this stage, however, the weak Lagrangian 11.37 only involves charged currents, so at first order of perturbation theory there are no neutral current interactions.

However, *flavour changing neutral currents occur at second order of perturbation theory*. For example, the decay 11.61 gets a contribution from including the interaction terms

$$\begin{aligned} & \frac{G_F}{\sqrt{2}} \cos \theta (\bar{\mu} \gamma^\alpha (1 - \gamma_5) \nu_\mu) (\bar{u} \gamma_\alpha (1 - \gamma_5) d) \\ & \frac{G_F}{\sqrt{2}} \sin \theta (\bar{\nu}_\mu \gamma^\alpha (1 - \gamma_5) \mu) (\bar{s} \gamma_\alpha (1 - \gamma_5) u) \end{aligned} \quad (11.63)$$

The first one allows, among others, the process

$$d \longrightarrow u + \mu^- + \bar{\nu}_\mu \quad (11.64)$$

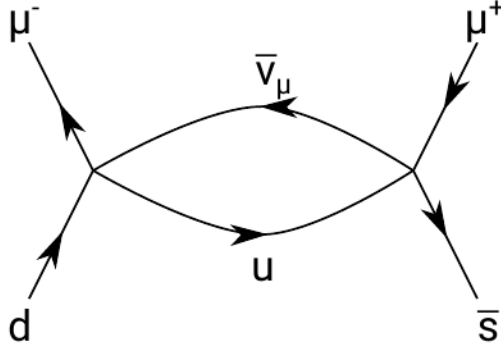
while the second one allows for

$$\bar{s} + u + \bar{\nu}_\mu \longrightarrow \mu^+ \quad (11.65)$$

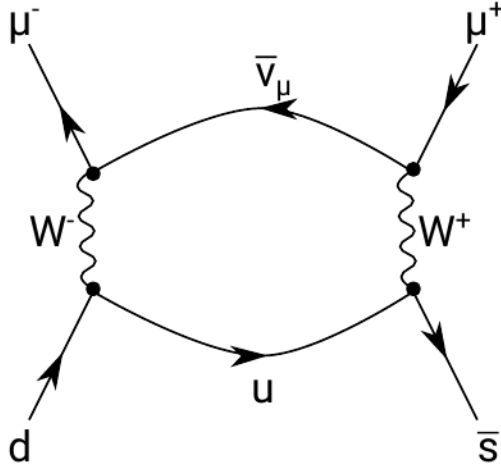
which can be combined as

$$\bar{s} + d \longrightarrow \mu^+ + \mu^- \quad (11.66)$$

This corresponds to the diagram (with time flowing from bottom to top)



The contribution of such diagrams cannot be estimated properly in the four-fermion theory, as they turn out to depend on the cut-off (cf. the discussion of loop divergences in Subsection [12.3.2](#)). A proper evaluation is possible once it is taken into account that the weak interactions are mediated by the $W^\pm$ bosons, which changes the diagram into



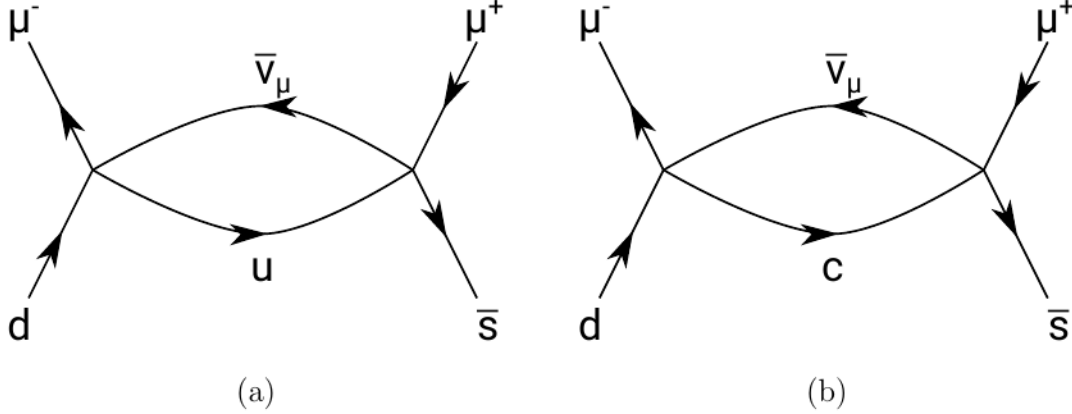
The contribution from these second order diagrams turns out to give a branching ratio of order 10^{-4} , which is too high.

In 1970, Glashow, Iliopoulos and Maiani proposed a mechanism (known as the *GIM mechanism*) to explain the observed low branching ratio of FCNC decays. By that time, people had already noticed the asymmetry resulting from d having a partner quark u , while s missing one. The GIM

mechanism is based on the assumption that the combination $s_\theta = -d \sin \theta + s \cos \theta$ enters in the weak interactions via a coupling to a *new quark* c , and the charge current looks like

$$\begin{aligned} J_\alpha^C &= \bar{e} \gamma_\alpha (1 - \gamma_5) \nu_e + \bar{\mu} \gamma_\alpha (1 - \gamma_5) \nu_\mu + \bar{d} \gamma_\alpha (1 - \gamma_5) u + \bar{s} \gamma_\alpha (1 - \gamma_5) c \\ &= \bar{e} \gamma_\alpha (1 - \gamma_5) \nu_e + \bar{\mu} \gamma_\alpha (1 - \gamma_5) \nu_\mu + \cos \theta \bar{d} \gamma_\alpha (1 - \gamma_5) u + \sin \theta \bar{s} \gamma_\alpha (1 - \gamma_5) u \\ &\quad + \cos \theta \bar{s} \gamma_\alpha (1 - \gamma_5) c - \sin \theta \bar{d} \gamma_\alpha (1 - \gamma_5) c \end{aligned} \quad (11.67)$$

Now there are two diagrams

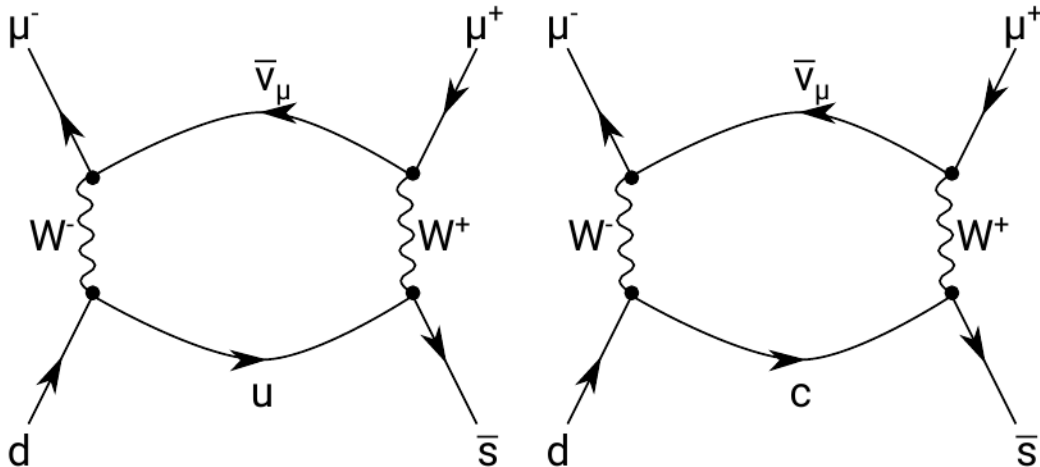


Neglecting any difference between the quark masses m_u and m_c , the only difference between these diagrams is between the interaction strengths involved in the vertices:

$$\begin{aligned} (a) &: \frac{G_F}{\sqrt{2}} \cos \theta \frac{G_F}{\sqrt{2}} \sin \theta \\ (b) &: \frac{G_F}{\sqrt{2}} (-\sin \theta) \frac{G_F}{\sqrt{2}} \cos \theta \end{aligned} \quad (11.68)$$

so the two contributions cancel each other! This is the GIM mechanism, which for degenerate quark masses leads to the complete absence of FCNC processes. It is essentially a *destructive quantum interference*: the transition from the state K_L^0 to $\mu^+ \mu^-$ can happen along two paths, but the corresponding amplitudes cancel, analogously to the explanation of dark stripes in a two-slit experiment.

In reality, the two diagrams are not exactly the same due to the difference between the u and c quark masses. A proper calculation of the amplitude, involving the intermediate vector bosons



leads to a *suppression factor given by the quark mass difference* as

$$\frac{m_c^2 - m_u^2}{M_W^2} \quad (11.69)$$

Given that M_W was expected to be of the order of 100 GeV (cf. Section 13.1), the observed value for the branching could be explained if m_c was of the order of GeV.

The predicted new quark was named *charm* by Glashow because its introduction solved some problems with the gauge theory proposed for weak interactions (discussed in Section 13), making it “work like a charm”. By the summer of 1974 the main expected properties of a $c\bar{c}$ meson were clarified. Besides having a mass of about a few GeV, its lifetime was expected to be much longer than naive estimates from other strongly decaying particles suggested. This is due to the OZI rule, discussed in Subsection 12.3.3, which states that strong decays which can only proceed via annihilation of heavy quarks are suppressed. In the case of the $c\bar{c}$ meson, this suppression was expected to give a very narrow resonance width of smaller than MeV magnitude, compared to a naive estimate (not taking into account the OZI rule) of the order of 100 MeV. A narrow resonance is very hard to find while scanning an energy range; in addition, the prediction was mainly ignored and the experimental work leading to the discovery was mainly motivated by exploring new energy regimes.

The discovery of the fourth quark was announced on the 11th November, 1974, ushering in what became known as the November Revolution due to the subsequent rapid changes in particle physics. Two groups, one at the Stanford Linear Accelerator Center (SLAC), headed by B. Richter, and another one at the Brookhaven National Laboratory (BNL) headed by S. Ting, found a new particle of rest mass of about 3.1 GeV and a mean lifetime of about $7 \cdot 10^{-21}$ s (about a thousand times longer than expected without the OZI rules). They discovered that they had eventually found the same particle, and announced it simultaneously on the same day. For their ground-breaking discovery, Richter and Ting shared the Nobel Prize in Physics in 1976.

Richter at first named the particle SP after the SPEAR accelerator used at SLAC, but following an objection by his co-workers, he chose the Greek letter ψ , which contains the name SP , but in reverse order. Ting named it J , which is a letter adjacent to K , already used for strange mesons, but is also the first letter of Ting’s eldest daughter’s name, Jeanne. Because of the simultaneous discovery, the particle was finally named the J/ψ . Due to the large mass of the c quark, a non-relativistic potential approximation works very well, and the J/ψ is essentially a charmonium bound state of quantum wave function $1S$, similarly to the ground state of the positronium, which is an electron-positron bound state²⁸. Richter’s group also found some other charmonium states such as η_c and ψ' . A few of these particles are summarised in Table 11.1.

Note the relatively large width of the states $\psi(3770)$ and $\psi(4040)$. The reason is that they are heavy enough to decay into charmed D mesons, with the end states being either $D^0 + \bar{D}^0$ or $D^+ + D^-$. Such decays go through creation of a light quark ($u\bar{u}$ or $d\bar{d}$) pair, and are not suppressed by the OZI rule. The D mesons have the quark content

$$\begin{aligned} D^+ &= c\bar{d} & D^- &= \bar{c}d \\ D^0 &= c\bar{u} & \bar{D}^0 &= \bar{c}u \end{aligned} \quad (11.70)$$

The mass of the charged ones is 1869 MeV, with 1865 MeV for the neutral ones.

There is a large number of other hadrons containing the charmed quark; for their properties the interested reader is referred to the latest Review of Particle Physics by the Particle Data Group.

²⁸In fact, a detailed analysis finds an almost perfect correlation between the spectra of charmonium and positronium (cf. D. Griffiths: *Introduction to Elementary Particles*, Section 5.4).

Name	State	Spin	Mass	Decay width
The ψ family				
J/ψ	$1S$	1	3096.916 ± 0.011 MeV	92.9 ± 2.8 keV
ψ'	$2S$	1	3686.109 ± 0.012 MeV	299 ± 8 keV
$\psi(3770)$	$3D$	1	3773.15 ± 0.012 MeV	27.2 ± 1.0 MeV
$\psi(4040)$	$3S$	1	4039 ± 1 MeV	80 ± 10 MeV
Some other states				
η_c	$1S$	0	2983.6 ± 0.7 MeV	32.2 ± 0.9 MeV
χ_{c0}	$1P$	0	3414.75 ± 0.31 MeV	10.5 ± 0.6 MeV
χ_{c1}	$1P$	1	3510.66 ± 0.07 MeV	0.84 ± 0.04 MeV

Table 11.1: A few members of the charmonium family

11.3.3 The third generation of leptons and quarks

After the discovery of charm, everything seemed perfect: the known fermions could be organised into two generations of leptons and quarks:

First generation		Second generation	
Leptons	Quarks	Leptons	Quarks
ν_e	u	ν_μ	c
e	d	μ	s

All of this made sense: under weak interactions, the leptons come in doublets, and transitions can happen between a charged lepton and its neutrino in the same doublet. Similarly for quarks, where the weak doublets are formed by u and d_θ , and by c and s_θ . The charge difference in each of these doublets is exactly the same. In addition, there is the Cabibbo mixing caused by difference between quark mass and weak eigenstates. *For the leptons, no such mixing was expected, since the neutrinos were thought to have zero mass*, so the only way to define which one was which was their transitions to charged leptons; there was no other basis which would have a different orientation. Indeed, all these symmetries were incorporated in the gauge theories of strong and weak interactions, discussed in Sections [12](#) and [13](#), respectively. A slight discomfort was expressed on the basis that considering the building blocks of everyday matter around us, the second generation had no reason to exist; the Nobel laureate Isidor Isaac Rabi famously asked “Who ordered that?” in reference to the muon.

However, the November Revolution started by the discovery of the charm quark brought new surprises. In a series of experiments between 1974 and 1977, M.L. Perl and his colleagues at SLAC SPEAR electron-positron collider ring (the same machine used by B. Richter’s group) found a new charged lepton, which was named τ after the starting letter of the Greek words triton (meaning “third” in English). The original discovery consisted of 64 events of the form

$$e^+ + e^- \longrightarrow e^\pm + \mu^\mp + \text{at least two undetected particles} \quad (11.71)$$

for which they had no explanation from known processes. The need for at least two undetected particles was found from energy-momentum conservation in the events. The proposed explanation was the creation of a new lepton pair

$$e^+ + e^- \longrightarrow \tau^+ + \tau^- \longrightarrow e^\pm + \mu^\pm + 4\nu \quad (11.72)$$

Adding this lepton with its own neutrino, the charged current (11.55) becomes

$$\begin{aligned}
J_\alpha^C = & \bar{e}\gamma_\alpha(1 - \gamma_5)\nu_e + \bar{\mu}\gamma_\alpha(1 - \gamma_5)\nu_\mu + \bar{\tau}\gamma_\alpha(1 - \gamma_5)\nu_\tau \\
& + \cos\theta\bar{d}\gamma_\alpha(1 - \gamma_5)u + \sin\theta\bar{s}\gamma_\alpha(1 - \gamma_5)u \\
& + \cos\theta\bar{s}\gamma_\alpha(1 - \gamma_5)c - \sin\theta\bar{d}\gamma_\alpha(1 - \gamma_5)c
\end{aligned} \tag{11.73}$$

and by looking at the terms in the weak Lagrangian (11.37) one obtains the following possible decays

$$\begin{aligned}
\tau^- & \longrightarrow e^- + \bar{\nu}_e + \nu_\tau \\
\tau^- & \longrightarrow \mu^- + \bar{\nu}_\mu + \nu_\tau \\
\tau^- & \longrightarrow d + \bar{u} + \nu_\tau
\end{aligned} \tag{11.74}$$

(with the charge conjugate processes for τ^+). The last one is a hadronic route, and eventually about 2/3 of τ s decay into various combinations of pions and a tau neutrino. The leptonic decays explaining the original observation happen about 1/3 of the time, and the branching ratio for the electronic and muonic routes is essentially the same, which is the consequence of *leptonic universality*²⁹

The reason why it took so long to identify the particle conclusively is that the energy necessary for a $\tau^+\tau^-$ pair production is similar to the threshold for D meson production. Work performed at SPEAR and at DESY at Hamburg, established that the τ is a spin-1/2 particle with mass $m = 1776.82 \pm 0.16$ MeV and a mean lifetime of $2.906(10) \cdot 10^{-13}$ s. For the discovery of τ , Perl shared the 1995 Nobel Prize in Physics with Reines, who was awarded his share the first direct detection of the neutrino.

Once the τ was found, it was expected to fit in a *third generation* of leptons and quarks. Indeed, a *new quark* b with a mass of about 4.5 GeV was discovered at the Fermilab by the E288 experiment lead by L.M. Lederman, in experiments similar to those that lead to J/ψ . This time they found a new generation of long-lived Υ (upsilon) mesons also called “bottomonium”, composed of a bottom and an anti-bottom quark. Similarly to the charm quark, the bottom quark is also a constituent of numerous hadrons: B mesons containing an u or d besides b , B_s mesons built from bottom and strange, and B_c composed of a bottom and a charm quark. Bottom baryons have also been found; an example is Λ_b^0 which is similar to Λ^0 but the s quark is replaced by the b .

The other quark of the third generation, the *top* was found by the CDF and D0 at Fermilab with a mass of about 173 GeV, which is close to the atomic mass of tungsten, making the top quark the heaviest fundamental particle known at present. *The lifetime of the top quark is so short that it decays before forming any hadrons, so there are no top mesons or baryons.*

Finally, the direct detection of the *tau neutrino* was achieved by the DONUT (Direct Observation of NU Tau) experiment run at the Tevatron proton accelerator where they were produced via decay of charmed mesons. The announcement in July 2000 was based on only 4 events; however this was far in excess of the expected background (< 0.2 events).

The presently known fundamental fermions are summarised in Table 11.2

²⁹Leptonic universality means that the weak couplings of all leptons are the same, so the electronic and muonic channels only differ because of $m_\mu \neq m_e$. However, this difference is very small compared to the mass of the τ , so the two branching ratio only differ slightly from each other.

Generation	Leptons				Quarks				
	Name	Mass	Charge	Spin	Name	Mass	Constituent mass	Charge	Spin
1st	ν_e	$< 2.2 \text{ eV}$	0	1/2	u	2.3 MeV	336 MeV	+2/3	1/2
	e	0.511 MeV	-1	1/2	d	4.8 MeV	340 MeV	-1/3	1/2
2nd	ν_μ	$< 0.17 \text{ MeV}$	0	1/2	c	1.275 GeV	1550 MeV	+2/3	1/2
	μ	105.7 MeV	-1	1/2	s	95 MeV	486 MeV	-1/3	1/2
3rd	ν_τ	$< 155 \text{ MeV}$	0	1/2	t	173.07 GeV	-	+2/3	1/2
	τ	1.711 GeV	-1	1/2	b	4.18 GeV	4730 MeV	-1/3	1/2

Table 11.2: The three generations of fundamental fermions. The u, d, s quark masses are current quark masses, while c and b masses are given in the so-called $\overline{MS}$ scheme. Their masses as hadronic constituents are different and their approximate values are also given (note that these values are depend on the phenomenological model used for hadron masses). The top mass is the value obtained by direct measurements, which is only possible in this case, since the top quark does not hadronise and therefore does not have a constituent mass either.

11.4 Flavour mixing, CP violation and neutrino oscillations

11.4.1 The Cabibbo-Kobayashi-Maskawa matrix

With three generations, the charged current takes the form

$$J_\alpha^C = \bar{e}\gamma_\alpha(1-\gamma_5)\nu_e + \bar{\mu}\gamma_\alpha(1-\gamma_5)\nu_\mu + \bar{\tau}\gamma_\alpha(1-\gamma_5)\nu_\tau + \bar{d}'\gamma_\alpha(1-\gamma_5)u + \bar{s}'\gamma_\alpha(1-\gamma_5)c + \bar{b}'\gamma_\alpha(1-\gamma_5)t \quad (11.75)$$

where

$$\begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} = V_{CKM} \begin{pmatrix} d \\ s \\ b \end{pmatrix} \quad (11.76)$$

with

$$V_{CKM} = \begin{pmatrix} V_{ud} & V_{us} & V_{ub} \\ V_{cd} & V_{cs} & V_{cb} \\ V_{td} & V_{ts} & V_{tb} \end{pmatrix} \quad (11.77)$$

denoting the 3×3 unitary matrix introduced by M. Kobayashi and T. Maskawa to extending the Cabibbo mixing to the case of three quark generations. Note that one could as well transfer the mixing on the u, c, t quarks, the important thing is the relative transformation of the up-type and down-type quarks with respect to each other. The unitarity of U_{KM} ensures that the GIM mechanism extends to the case with three generations.

Now a 3×3 unitary matrix has 9 real parameters which can be seen from the fact that it can be written as

$$V = e^{-iT} \quad (11.78)$$

with T Hermitian and a 3×3 real Hermitian matrix has exactly 9 real parameters. Real unitary matrices are just orthogonal matrices (rotations and reflections), and three-dimensional rotations have three real parameters, e.g. the Euler angles. So U depends on 3 real angles plus 6 complex phases. However, the phases of the 6 quark fields can be arbitrarily redefined; this makes 5 possible

phase redefinition, since an overall flavour independent phase redefinition cancels from the current J_α^C . So the physical parameters of V_{CKM} are three real angles plus a complex phase.

Exercise 11.7. Count the number of mixing parameters for two generations.

Solution. The mixing matrix is a 2×2 unitary matrix, which has 4 real parameters. Two-dimensional rotations are parametrised by a single angle, so there are three complex phases. However, with 4 quark fields, there are exactly $4 - 1 = 3$ non-trivial phase redefinitions, therefore all the phases can be eliminated. The remaining single real angle is the one introduced by Cabibbo. $\square$

The standard parametrisation of the Cabibbo-Kobayashi-Maskawa flavour mixing matrix is

$$V_{CKM} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_{23} & s_{23} \\ 0 & -s_{23} & c_{23} \end{pmatrix} \begin{pmatrix} c_{13} & 0 & s_{13}e^{-i\delta} \\ 0 & 1 & 0 \\ -s_{13}e^{i\delta} & 0 & c_{13} \end{pmatrix} \begin{pmatrix} c_{12} & s_{12} & 0 \\ -s_{12} & c_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (11.79)$$

given in terms of three Euler angles θ_{12} , θ_{13} , θ_{23} and a phase δ , with $c_{ij} = \cos \theta_{ij}$ and $s_{ij} = \sin \theta_{ij}$.

Kobayashi and Maskawa noted that *the existence of the non-trivial phase δ leads to CP violation in weak interactions*. One can understand this simply by writing down the weak Lagrangian (11.37):

$$\mathcal{L}_W = \frac{G_F}{\sqrt{2}} (J_\alpha^{C\dagger} J^{C\alpha} + \text{h.c.}) \quad (11.80)$$

and noting that for any process coming from the first term, its time reversal comes from the Hermitian conjugate one and vice versa. If the coupling constants are complex, these two terms are different; this is exactly the case if V_{CKM} has a physical complex phase (i.e. one that cannot be eliminated by field redefinitions). In fact, things are a little more complicated: in simple cases only the absolute value squared of the couplings appears; to find time reversal T violation it is necessary to consider processes which can proceed via two paths which involve the phase difference encoded by δ . In that case, analogously to our discussion of the GIM mechanism (cf. Subsection 11.3.2), there is an interference between the two contributions, which depends on the sign of δ , making the probability for the process and its time reversal different. However, all local relativistic field theories are CPT invariant, so *violating time reversal is the same as CP violation*.

The present values for the CKM parameters are

$$\begin{aligned} \theta_{12} &= 13.04 \pm 0.05^\circ & \theta_{13} &= 0.201 \pm 0.011^\circ & \theta_{23} &= 2.38 \pm 0.06^\circ \\ \delta &= 1.20 \pm 0.08 \text{ radians} \end{aligned} \quad (11.81)$$

The magnitudes of the CKM matrix elements are experimentally determined as

$$\begin{pmatrix} |V_{ud}| & |V_{us}| & |V_{ub}| \\ |V_{cd}| & |V_{cs}| & |V_{cb}| \\ |V_{td}| & |V_{ts}| & |V_{tb}| \end{pmatrix} = \begin{pmatrix} 0.97427 \pm 0.00015 & 0.22534 \pm 0.00065 & 0.00351^{+0.00015}_{-0.00014} \\ 0.22520 \pm 0.00065 & 0.97344 \pm 0.00016 & 0.0412^{+0.0011}_{-0.0005} \\ 0.00867^{+0.00029}_{-0.00031} & 0.0404^{+0.0011}_{-0.0005} & 0.999146^{+0.000021}_{-0.000046} \end{pmatrix} \quad (11.82)$$

Unitarity of the V matrix

$$V^\dagger V = 1 \quad (11.83)$$

ensures two key observed properties of weak interactions:

1. *Weak universality*: the weak coupling to all lepton doublets is known to be the same. For quarks, the total weak coupling strength must be summed up over the mixing channels. The relations

$$\sum_j |V_{ij}|^2 = \sum_j |V_{i'j}|^2 \text{ for } i \neq i' \quad (11.84)$$

ensures that all quarks couple with the same total strength to the weak interactions. This sum is normalised to 1 so the total strength of weak interaction of the quarks to agree with that of the leptons.

2. *Suppression of FCNC*: the relations

$$\sum_k V_{ik} V_{jk}^* = 0 \text{ for } i \neq j \quad (11.85)$$

result in the cancellation of flavour changing neutral current processes in the limit of equal quark masses, extending the GIM mechanism to three generations. When taking into account mass differences, it leads to a suppression mechanism that is an extension of the one discussed in Subsection [11.3](#)

Furthermore, the conclusion from Exercise [11.7](#) is that CP violation from weak interactions is only possible with at least three generations of quarks. On this basis, Kobayashi and Maskawa *predicted the existence of a third generations of quarks* in 1973, for which they were jointly awarded half of 2008 Nobel Prize in Physics (the other half going to Y. Nambu for his work on spontaneous symmetry breaking in subatomic physics). Recalling the discussion in Section [6.4](#), this leads to a tentative answer to Rabi's question: not only the muon, but the whole *existence of three generations is necessary to explain the observed asymmetry between matter and antimatter from flavour mixing in weak interactions*. As discussed in Section [6.4](#), supposing that the Universe started with no primordial matter-antimatter asymmetry, CP violation is also necessary to explain baryogenesis and the existence of enough matter left after annihilation with antimatter to create the galaxies, stars and ultimately life.

11.4.2 CP violation in the neutral kaon system

CP violation was first observed in the neutral kaon system. Neglecting weak interactions, strong interactions conserve strangeness so the eigenstates of strong interactions are

$$K^0 = d\bar{s} \quad \bar{K}^0 = \bar{d}s \quad (11.86)$$

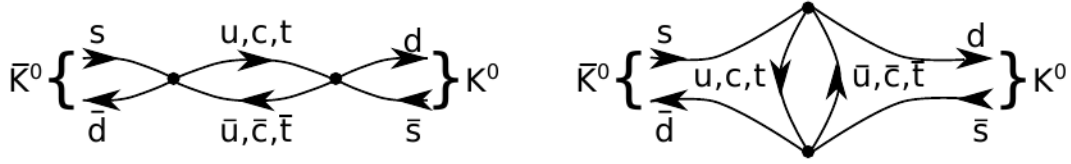
and these states are CP conjugates of each other

$$|K^0\rangle = \mathcal{CP}|\bar{K}^0\rangle \quad |\bar{K}^0\rangle = \mathcal{CP}|K^0\rangle \quad (11.87)$$

In rest frame, the mesons form a two-state system with the Hamiltonian

$$H = \begin{pmatrix} M_{11} & 0 \\ 0 & M_{22} \end{pmatrix} \quad (11.88)$$

with CPT symmetry implying that $M_{11} = M_{22}$. Weak interactions lead to a mixing between these states via diagrams such as



This is a special case of *neutral particle oscillations*; it also occurs in other systems such as the mesons $B^0 - \bar{B}^0$. Neglecting the decay of the particles for the moment, the Hamiltonian is now of the form

$$H = \begin{pmatrix} M & \Delta \\ \Delta^* & M \end{pmatrix} \quad (11.89)$$

Let us first neglect CP violation. Since CP symmetry flips 1 with 2, it implies the equality of the 12 and 21 matrix elements, leading to

$$H = \begin{pmatrix} M & \Delta \\ \Delta & M \end{pmatrix} \quad (11.90)$$

with Δ real. The eigenstates are then given by

$$|K_1\rangle = \frac{1}{\sqrt{2}} (|K^0\rangle + |\bar{K}^0\rangle) \quad |K_2\rangle = \frac{1}{\sqrt{2}} (|K^0\rangle - |\bar{K}^0\rangle) \quad (11.91)$$

which are CP -even/odd with mass $M \pm \Delta$, respectively.

Exercise 11.8. Supposing the system starts out in the state $|K^0\rangle$, compute the probability of finding it in state $|K^0\rangle$ or $|\bar{K}^0\rangle$ after time t .

Solution. Writing the state of the system as

$$|\Psi(0)\rangle = |K^0\rangle = \frac{1}{\sqrt{2}}|K_1\rangle + \frac{1}{\sqrt{2}}|K_2\rangle \quad (11.92)$$

it evolves as

$$|\Psi(t)\rangle = \frac{1}{\sqrt{2}}e^{-i(M+\Delta)t}|K_1\rangle + \frac{1}{\sqrt{2}}e^{-i(M-\Delta)t}|K_2\rangle = e^{-iMt} (\cos(\Delta t)|K^0\rangle + \sin(\Delta t)|\bar{K}^0\rangle) \quad (11.93)$$

The required probabilities are

$$P(K^0) = \cos^2 \Delta t = \frac{1 + \cos 2\Delta t}{2} \quad P(\bar{K}^0) = \sin^2 \Delta t = \frac{1 - \cos 2\Delta t}{2} \quad (11.94)$$

so the system oscillates between the two states with a frequency corresponding to the energy difference. This is the phenomenon of *Rabi oscillations*, which is well-known from quantum mechanics. $\square$

Let us now consider CP violation in the neutral kaon system. In this case the formalism needs to be extended, since it is necessary to account for the finite life-time of the kaons. For a single particle this could be accomplished via introducing a complex mass $M - i\Gamma/2$, which leads to the time evolution of the wave function amplitude

$$\psi(t) = \psi(0)e^{-iMt - \Gamma t/2} \quad (11.95)$$

i.e. the probability of finding the particle at time t is

$$|\psi(t)|^2 = |\psi(0)|^2 e^{-\Gamma t} \quad (11.96)$$

showing that Γ gives the decay rate, with average lifetime $\tau = 1/\Gamma$. For a two-state system the same idea results in the effective non-Hermitian³⁰ Hamiltonian

$$H = \begin{pmatrix} M_{11} & M_{12} \\ M_{12}^* & M_{22} \end{pmatrix} - \frac{i}{2} \begin{pmatrix} \Gamma_{11} & \Gamma_{12} \\ \Gamma_{12}^* & \Gamma_{22} \end{pmatrix} \quad (11.97)$$

which is just a way of writing a general complex matrix in terms of a Hermitian and anti-Hermitian parts. CPT symmetry implies $M_{11} = M_{22}$ and $\Gamma_{11} = \Gamma_{22}$ (mass of particle and antiparticle are equal to each other).

Again, neglecting CP violation leads to $M_{12} = M_{12}^*$ and $\Gamma_{12} = \Gamma_{12}^*$ so the eigenstates are the CP -even/odd states K_1 and K_2 , with different masses and decay rates, which explains the existence of K_L^0 and K_S^0 .

The true K_L^0 and K_S^0 , however, are different from K_1 and K_2 due to CP violation:

Exercise 11.9. Writing the eigenvalues of H as

$$\begin{aligned} M_L &= M_{11} - \frac{i}{2}\Gamma_{11} + \left(\Delta M - \frac{i}{2}\Delta\Gamma \right) \\ M_S &= M_{11} - \frac{i}{2}\Gamma_{11} - \left(\Delta M - \frac{i}{2}\Delta\Gamma \right) \end{aligned} \quad (11.98)$$

compute ΔM and $\Delta\Gamma$. Show that the eigenvectors are of the form

$$\begin{aligned} |K_L^0\rangle &= p|K\rangle - q|\bar{K}\rangle \\ |K_S^0\rangle &= p|K\rangle + q|\bar{K}\rangle \end{aligned} \quad (11.99)$$

with $|p|^2 + |q|^2 = 1$ and compute p/q .

Remark 1: for the notation L/S to match the long/short-lived kaon, one should choose the labels of the eigenvectors for which p/q is close to 1 (as opposed to being close to -1).

Remark 2: note that the eigenvectors are no more orthogonal, which is due to the non-Hermiticity of H (i.e. particle decay).

Solution.

$$\begin{aligned} \Delta m^2 - \frac{\Delta\Gamma^2}{4} &= 4|M_{12}|^2 - |\Gamma_{12}|^2 \\ \Delta m \Delta\Gamma &= 4 \operatorname{Re} (M_{12}\Gamma_{12}^*) \\ \left(\frac{p}{q} \right)^2 &= \frac{M_{12}^* - i\Gamma_{12}^*/2}{M_{12} - i\Gamma_{12}/2} \end{aligned} \quad (11.100)$$

□

³⁰The violation of Hermiticity corresponds to the fact that probability is not conserved in the two-state subspace, since the system decays into other physical states.

can be rewritten as

$$\begin{aligned} |K_L^0\rangle &= \frac{1}{\sqrt{1+|\epsilon|^2}} (|K_2\rangle + \epsilon|K_1\rangle) \\ |K_S^0\rangle &= \frac{1}{\sqrt{1+|\epsilon|^2}} (|K_1\rangle + \epsilon|K_2\rangle) \end{aligned} \quad (11.101)$$

with

$$\frac{p}{q} = \frac{1+\epsilon}{1-\epsilon} \quad (11.102)$$

So due to the mixing, K_L^0 is not purely the CP -odd K_2 , but contains a small admixture of the CP -even state K_1 . The state K_2 decays to 3 pions, while K_1 decays to two pions; here we assume that any direct CP violation effect (see below) in these decays can be neglected. Since the decay of K_L^0 to two pions only happens rarely, ϵ is expected to be small. The measured value is

$$|\epsilon| = (2.228 \pm 0.011) \cdot 10^{-3} \quad (11.103)$$

This is a case of indirect CP violation, when the CP violating mixing is observed via the (otherwise CP preserving) decays.

As discussed in Subsection [6.3.2](#), there is also evidence for direct CP violation in the decays

$$K_L^0 \longrightarrow \begin{cases} \pi^+ + e^- + \bar{\nu}_e \\ \pi^- + e^+ + \nu_e \end{cases} \quad (11.104)$$

Direct CP violation is when the decay rate of a particle state $|P\rangle$ to some final products $|f\rangle$ is not the same as the rate from the CP conjugate state $|\bar{P}\rangle = \mathcal{CP}|P\rangle$ to the CP conjugate final products $|\bar{f}\rangle = \mathcal{CP}|f\rangle$, i.e. the magnitudes of the corresponding amplitudes are different:

$$|\mathcal{A}(P \rightarrow f)| \neq |\mathcal{A}(\bar{P} \rightarrow \bar{f})| \quad (11.105)$$

11.4.3 Neutrino oscillations

The assumption that neutrinos have exactly zero mass excludes a mixing matrix for leptons, since it can be viewed as a *neutrino mixing matrix*. But with no other way to define neutrino species than weak interactions, there is no way of introducing a mixing between them.

However, experiments in the late 1960s challenged the picture of massless neutrinos. The Sun, being a thermonuclear reactor, is a source of electron neutrinos, with the net fusion process being

$$4p + 2e^- \longrightarrow {}^4_2\text{He} + 2\nu_e + 26.73 \text{ MeV} \quad (11.106)$$

The flux of these neutrinos was first measured by R. Davis's and J.N. Bahcall's Homestake Experiment, which found only about a third of the number expected on the basis of the standard solar model (which Bahcall himself helped to develop). Later experiments repeated this finding, which became known as the *solar neutrino problem*. Early attempts to explain the discrepancy focused on modifications of models of the Sun; however, later advances in helioseismology and neutrino measurement techniques excluded all such modifications.

The resolution came with the proposal of *neutrino masses*, which then allowed for *neutrino flavour mixings*, resulting in neutrinos changing their species in flight. This phenomenon is known

as *neutrino oscillation*; the original idea came from B. Pontecorvo in 1957 based on analogy with neutral kaon mixing.

The first evidence for neutrino oscillations came in 1999 from the Super-Kamiokande collaboration headed by T. Kajita, which observed that for muon neutrinos that are produced in the upper atmosphere by cosmic rays, fewer come from the direction through the Earth than from directly above the detector. This was interpreted as a conversion of muon neutrinos into tau neutrinos (the latter were not detected directly). Direct evidence for the solar neutrino oscillations was produced by A. McDonald's team at the Sudbury Neutrino Observatory in Canada in 2001. They detected all types of neutrinos from the Sun, and were able to distinguish electron neutrinos from the other two (muon and tau neutrinos were not distinguished). The total neutrino flux agreed well with the prediction of solar models, but only about 35% of them were electron neutrinos, demonstrating that *neutrino oscillation solves the solar neutrino problem*. For the confirmation of neutrino oscillation and showing that it resolves the solar neutrino problem, T. Kajita and A. McDonald shared the 2015 Nobel Prize in Physics. Presently, the details of neutrino oscillation are actively investigated, using also reactor sources and neutrino beams from accelerators.

For simplicity, let us illustrate neutrino oscillations with two species. Let the two flavours be ν_e and ν_μ , which are related to mass eigenstates ν_1 and ν_2 with masses m_1 and m_2 by the analogue of the Cabibbo mixing:

$$\begin{aligned} |\nu_e\rangle &= \cos\theta|\nu_1\rangle + \sin\theta|\nu_2\rangle \\ |\nu_\mu\rangle &= -\sin\theta|\nu_1\rangle + \cos\theta|\nu_2\rangle \end{aligned} \quad (11.107)$$

The quantum state of the neutrino can be written as

$$\begin{aligned} |\Psi\rangle &= \psi_1(t)|\nu_1\rangle + \psi_2(t)|\nu_2\rangle \\ &= \psi_e(t)|\nu_e\rangle + \psi_\mu(t)|\nu_\mu\rangle \end{aligned} \quad (11.108)$$

with

$$\begin{aligned} \psi_1(t) &= \psi_e(t)\cos\theta - \psi_\mu(t)\sin\theta \\ \psi_2(t) &= \psi_e(t)\sin\theta + \psi_\mu(t)\cos\theta \end{aligned} \quad (11.109)$$

Now supposing that the neutrino has a given momentum $\vec{k}$, states $|\nu_{1,2}\rangle$ have definite energies $E_j = \sqrt{k^2 + m_j^2}$, so their time evolution is given by

$$\psi_j(t) = \psi_j(0)e^{-iE_j t} \quad (11.110)$$

Exercise 11.10. Supposing the system starts out in the state $|\nu_e\rangle$, compute the probability of finding it in state $|\nu_\mu\rangle$ after time t .

Solution. Writing the initial state of the system as

$$\psi_1(0) = \cos\theta \quad \psi_2(0) = \sin\theta \quad (11.111)$$

it evolves as

$$|\Psi(t)\rangle = e^{-iE_1 t} \cos\theta|\nu_1\rangle + e^{-iE_2 t} \sin\theta|\nu_2\rangle \quad (11.112)$$

The probability amplitude is given by are

$$\langle \nu_\mu | \Psi(t) \rangle = -e^{-iE_1 t} \sin \theta \cos \theta + e^{-iE_2 t} \sin \theta \cos \theta \quad (11.113)$$

so

$$P(\nu_e \rightarrow \nu_\mu) = \left[\sin 2\theta \sin \left(\frac{E_1 - E_2}{2} t \right) \right]^2 \quad (11.114)$$

So ν_e converts to ν_μ and then back again, with an amplitude determined by the mixing angle θ , while the frequency is given by the energy difference. $\square$

Exercise 11.11. Assuming that the particles have very small mass (i.e. can be treated as ultra-relativistic) find the period T of oscillation.

Solution. The energies can be written as

$$E_j = \sqrt{k^2 + m_j^2} = k + \frac{m_j^2}{2k} + \dots \quad (11.115)$$

so

$$E_1 - E_2 \approx \frac{m_1^2 - m_2^2}{2k} \approx \frac{m_1^2 - m_2^2}{2E} \quad (11.116)$$

where E is just the energy neglecting the small difference between E_1 and E_2 . So the probability can be written as

$$P(\nu_e \rightarrow \nu_\mu) = \left[\sin 2\theta \sin \left(\frac{m_1^2 - m_2^2}{4E} t \right) \right]^2 \quad (11.117)$$

i.e. the period is

$$T = \frac{4\pi E}{m_1^2 - m_2^2} \quad (11.118)$$

$\square$

In particular, after travelling the *oscillation length*

$$L = \frac{T}{2} \quad (11.119)$$

or, restoring $\hbar$ and c

$$L = \frac{2\pi\hbar E}{(m_2^2 - m_1^2)c^3} \quad (11.120)$$

the conversion probability hits the maximum value $\sin^2(2\theta)$ and at $2L$ it is back to the original value. The distance travelled by neutrinos in an experiment is called the *baseline* and is an important parameter since its relation to the oscillation lengths characterizing the different neutrino conversions determines the parameters to which the experiment is sensitive.

With three flavours, neutrino mixing is described by the *Pontecorvo–Maki–Nakagawa–Sakata (PMNS) matrix*, which is analogous to the CKM quark flavour mixing matrix. We denote the

PMNS matrix by U_{PMNS} ; it can be parametrised the same way as V_{CKM} in [11.79](#). The present values of the parameters are

$$\begin{aligned}\theta_{12}[^{\circ}] &= 33.36^{+0.81}_{-0.78} & \theta_{23}[^{\circ}] &= 40.0^{+2.1}_{-1.5} \text{ or } 50.4^{+1.3}_{-1.3} \\ \theta_{13}[^{\circ}] &= 8.66^{+0.44}_{-0.46} & \delta[^{\circ}] &= 300^{+66}_{-138}\end{aligned}\tag{11.121}$$

and the matrix elements themselves as determined from experiments are

$$\begin{pmatrix} U_{e1} & U_{e2} & U_{e3} \\ U_{\mu 1} & U_{\mu 2} & U_{\mu 3} \\ U_{\tau 1} & U_{\tau 2} & U_{\tau 3} \end{pmatrix} = \begin{pmatrix} 0.82 \pm 0.01 & 0.54 \pm 0.02 & 0.15 \pm 0.03 \\ 0.35 \pm 0.06 & 0.70 \pm 0.06 & 0.62 \pm 0.06 \\ 0.44 \pm 0.06 & 0.45 \pm 0.06 & 0.77 \pm 0.06 \end{pmatrix}\tag{11.122}$$

Indices e , μ and ν refer to the weak eigenstates, while 1, 2 and 3 label the mass eigenstates. The ambiguity in θ_{23} arises from the fact that at present the neutrino masses are not known. The only thing that can be determined are the mass differences

$$\begin{aligned}|m_1^2 - m_2^2| &= (8 \pm 0.6) \cdot 10^{-5} \text{ eV}^2 \\ |m_2^2 - m_3^2| &= (2.4 \pm 0.6) \cdot 10^{-3} \text{ eV}^2\end{aligned}\tag{11.123}$$

Even the order of the masses is not known. By convention m_1 is taken to be the smaller of m_1 and m_2 , but then m_3 could either be much above (which is called *normal hierarchy*) or much below the two others (*inverted hierarchy*). It turns out that the angle θ_{23} is the one that is strongly sensitive to the nature of the mass hierarchy. Note also that the best fit value for the CP violating angle is consistent with zero, since 0 and 360 degrees are the same.

We note that transitions between neutrino species violate the separate conservation of lepton numbers L_e , L_μ and L_τ , leaving only $L = L_e + L_\mu + L_\tau$. This leads to processes like

$$\mu^\pm \longrightarrow e^\pm + \gamma\tag{11.124}$$

which has not been observed yet (cf. [Section 11.1](#)).

Another interesting possibility is that neutrino could be a Majorana particle, which would imply there is simply no conserved lepton number. In that case, ν_e and $\bar{\nu}_e$ are really just left/right handed components of the same Majorana fermion. The reason why they are different is that weak interactions couple the left-handed helicity neutrino to the (left-handed) electron, while the right-handed one couples to the (right-handed) positron. This makes possible a curious process called *neutrinoless double β decay*. Usual double β -decay has already been observed in several isotopes such as e.g. ^{48}Ca , when two β -decays happen just simultaneously. Neutrinoless double β -decay would turn 2 neutrons directly into 2 protons and 2 electrons without any neutrinos which would annihilate inside the diagram, which is possible if the neutrino is its own antiparticle. Presently, numerous experiments are searching for such a process.

The unitarity of U_{PMNS} is also an interesting issue: if it is violated, that would mean the existence of a so-called *sterile neutrino*. It is called sterile as it cannot participate in the weak interactions (and as a neutrino, neither in the electromagnetic nor the strong one), otherwise it would contradict the observed Z boson decay width. Cosmological constraints using the measured primordial abundances of nuclei ^4He and ^2H limit the number of light particle species according to

$$N_\nu = 3.14^{+0.70}_{-0.65} \text{ at } 68\% \text{ c.l.}\tag{11.125}$$

while limits from structure formation result in

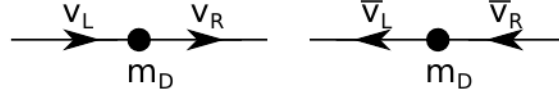
$$N_\nu = 4.34^{+0.88}_{-0.86} \text{ at } 68\% \text{ c.l.}\tag{11.126}$$

but there is still room for a sterile neutrino.

Concerning neutrino mass, not much is known except for the two mass differences and upper experimental limits. Cosmology also limits the masses of the neutrinos, with the sum of the masses expected to be at most about a quarter of an eV. Lepton number conservation only allows for the usual $m_i \bar{\nu}_i \nu_i$ Dirac mass terms in the Lagrangian; in terms of left/right handed components, these can be written as

$$m_D (\bar{\nu}_L \nu_R + \bar{\nu}_R \nu_L) \quad (11.127)$$

Written as an interaction vertex, this leads to processes such as



so a Dirac mass term couples the two helicities, as discussed in Section 5.7. However, if lepton number conservation is broken, there could also be so-called *Majorana mass terms*, with the form according to Eq. (5.243)

$$\frac{1}{2} m_L (\bar{\nu}_L^c \nu_L + \bar{\nu}_L \nu_L^c) + \frac{1}{2} m_R (\bar{\nu}_R^c \nu_R + \bar{\nu}_R \nu_R^c) \quad (11.128)$$

with c denoting charge conjugation as defined in Section 5.9. There are now two possibilities:

1. The neutrino may be a Majorana fermion identical to its anti-particle, in which case the right-handed field ν_R does not exist, and the only mass term is

$$\frac{1}{2} m_L (\bar{\nu}_L^c \nu_L + \bar{\nu}_L \nu_L^c)$$

Under a phase transformation $\nu \rightarrow e^{-i\alpha} \nu$ the two term acquires factors $e^{\mp 2\alpha}$, i.e. they violate lepton number conservation by ± 2 .

2. The neutrino may be a full Dirac fermion complete with the right-handed field ν_R , in which case one can introduce the self-conjugate fields

$$\begin{aligned} \chi &= \nu_L + \nu_L^c & \chi^c &= \chi \\ \omega &= \nu_R + \nu_R^c & \omega^c &= \omega \end{aligned}$$

Then the mass terms can be written as

$$\begin{aligned} \mathcal{L}_{\text{mass}} &= D \bar{\nu}_L \nu_R + A \bar{\nu}_L^c \nu_L + B \bar{\nu}_R^c \nu_R + \text{h.c.} \\ &= (\bar{\chi} \quad \bar{\omega}) \begin{pmatrix} A & D/2 \\ D/2 & B \end{pmatrix} \begin{pmatrix} \chi \\ \omega \end{pmatrix} \end{aligned} \quad (11.129)$$

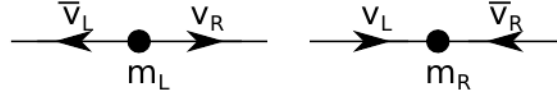
and can be diagonalised with the eigenvalues

$$m_{1,2} = \frac{1}{2} \left(A + B \pm \sqrt{(A - B)^2 + D^2} \right)$$

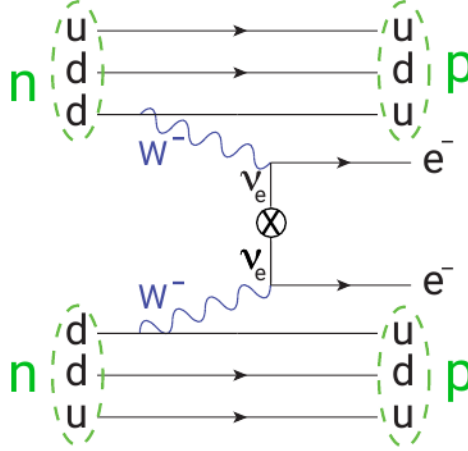
and eigenstates

$$\begin{aligned} \eta_1 &= \chi \cos \theta - \omega \sin \theta \\ \eta_2 &= \chi \sin \theta + \omega \cos \theta \quad \tan 2\theta = \frac{D}{A - B} \end{aligned}$$

Majorana mass terms couple a left-handed neutrino to a right-handed anti-neutrino and vice versa, leading to processes



which violate lepton number conservation and allow e.g. neutrinoless double β -decay, where two neutrons turn simultaneously into two protons and two electrons without the emission of neutrinos³¹ shown in the Feynman diagram



where the vertex induced by the neutrino mass is denoted by the crossed circle in the middle. Lepton number violation could, in turn, help to explain the cosmological matter-antimatter asymmetry via a process called *leptogenesis*, which would generate lepton asymmetry in the early Universe, leading to the observed matter-dominated Universe in later stages.

Additionally, when the neutrinos are Majorana fields (option 1), phase transformations of the neutrino fields are not a symmetry, which results in two more physical phases in the PMNS matrix, bringing the total number of parameters up to 3 angles and 3 CP violating phases. The see-saw model (option 2) results in even more freedom: the PMNS matrix is not even unitary as it is just a submatrix of a larger mixing matrix which involves the heavy neutrinos.

The exact nature of neutrino mixing is still unknown and is under active investigation. So neutrino physics is quite an open subject, and Nature can have more surprises in store for us!

³¹Ordinary double beta decay, where two neutrons decay simultaneously according to $2n \rightarrow 2p^+ + 2e^- + 2\bar{\nu}_e$ has been observed in about a dozen nuclei and has a very long half-life of the order of $10^{19} - 10^{21}$ years.

12 Non-Abelian gauge fields and quantum chromodynamics

Modern theories of particle interactions are built upon the concept of *gauge invariance*, which means local (i.e. space-time dependent) symmetry transformations. We have already met an example: the electromagnetic field and its interactions are specified by a local $U(1)$ invariance. Here we extend the notion of gauge invariance to more general symmetry groups, and sketch the building blocks of the Standard Model of elementary particle physics.

12.1 Non-Abelian gauge invariance

12.1.1 Global non-Abelian symmetry groups

Let us consider a Lagrangian of N free fermions:

$$\mathcal{L}_{\text{matter}} = \bar{\Psi}_k i \not{\partial} \Psi_k - \bar{\Psi}_k M_{kl} \Psi_l \quad (12.1)$$

The first term is invariant under any unitary transformation

$$\Psi'_k = U_{kl} \Psi_l \quad (12.2)$$

since

$$\bar{\Psi}'_k = \bar{\Psi}_m U_{km}^* = \bar{\Psi}_m U_{mk}^\dagger = \bar{\Psi}_m U_{mk}^{-1} \quad (12.3)$$

The mass term is invariant under all transformations that leave the mass matrix M invariant:

$$U^{-1} M U = M \quad \Leftrightarrow \quad M U = U M \quad (12.4)$$

Such transformations obviously form a group G . The matrices U give a *representation* R of the abstract group, so to any element γ of G there corresponds a symmetry transformation $R(\gamma)$ such that

$$\begin{aligned} \Psi'_k &= R(\gamma)_{kl} \Psi_l \\ \text{and} \quad M R(\gamma) &= R(\gamma) M \end{aligned} \quad (12.5)$$

We assume G is compact, and that it is a continuous i.e. Lie group, although this is not strictly necessary: gauge fields corresponding to discrete groups can be defined as well. A Lie group is usually given as a matrix group, i.e. one specifies a fundamental or defining representation; examples are

$$\begin{aligned} \text{orthogonal group} &: O(N) = \{O \in \mathbb{R}^{N \times N} \mid O^T O = \mathbf{1}_{N \times N}\} \\ \text{special orthogonal group} &: SO(N) = \{O \in \mathbb{R}^{N \times N} \mid O^T O = \mathbf{1}_{N \times N}, \det O = 1\} \\ \text{unitary group} &: U(N) = \{U \in \mathbb{C}^{N \times N} \mid U^\dagger U = \mathbf{1}_{N \times N}\} \\ \text{special unitary group} &: SU(N) = \{U \in \mathbb{C}^{N \times N} \mid U^\dagger U = \mathbf{1}_{N \times N}, \det U = 1\} \\ \text{unit complex numbers} &: U(1) = \{\xi \in \mathbb{C} \mid \xi^* \xi = 1\} \end{aligned} \quad (12.6)$$

Here we do not dwell on details of the theory of Lie groups and their classification; for our purposes it is enough to keep the above examples in mind. We only note that

$$U(N) \simeq SU(N) \times U(1) \quad (12.7)$$

The representation R is called a unitary representation of G ³²; well-known examples are the $2j+1$ -dimensional spin- j representations of $SU(2)$ which enter the quantum mechanical theory of angular momentum, or the representations of $SU(3)$ that play an important role in the quark model.

For a Lie group one can introduce continuous parameters ϵ^a where $a = 1, \dots, \dim G$. The origin of these coordinates is assumed to be the identity of the group, i.e.

$$\gamma(\epsilon^a = 0) = e \quad (12.8)$$

Introducing *infinitesimal transformations* by

$$R(\gamma(\epsilon))_{kl} = \delta_{kl} - i\epsilon^a (t_a)_{kl} + O(\epsilon^2) \quad (12.9)$$

the Hermitian generators t_a can be interpreted as giving the representation R for the Lie algebra of G :

$$t_a = R(T_a) \quad (12.10)$$

where the T_a are a basis of the Lie-algebra satisfying

$$[T_a, T_b] = if^{abc}T_c \quad (12.11)$$

where f^{abc} are the *structure constants* of the Lie algebra. For the defining representation $\mathbf{N}$ of $SU(N)$, the generators $R_{\mathbf{N}}(T_a)$ can be chosen as $N^2 - 1$ traceless Hermitian matrices satisfying

$$\text{Tr } R_{\mathbf{N}}(T_a)R_{\mathbf{N}}(T_b) = \frac{1}{2}\delta_{ab} \quad (12.12)$$

For $SU(2)$ the standard choice is

$$R_2(T_a) = \frac{\sigma_a}{2} \quad (12.13)$$

with σ^a the Pauli matrices and

$$f^{abc} = \epsilon^{abc} \quad (12.14)$$

while for $SU(3)$

$$R_3(T_a) = \frac{\lambda_a}{2} \quad (12.15)$$

where λ_a are the Gell-Mann matrices (7.78). Infinitesimal group elements can be expressed as

$$\gamma(\epsilon) = e - i\epsilon^a T_a + O(\epsilon^2) \quad (12.16)$$

12.1.2 Covariant derivative and the gauge fields

Now let's try to make the symmetry *local*

$$\Psi'(x) = R(\gamma(x))\Psi(x) \quad (12.17)$$

However, the Lagrangian $\mathcal{L}$ is not invariant any more

$$\mathcal{L}_{\text{matter}}(\Psi') - \mathcal{L}_{\text{matter}}(\Psi) = i\bar{\Psi}(x)\gamma^\mu \left(R(\gamma(x))^{-1}\partial_\mu R(\gamma(x)) \right) \Psi(x) \quad (12.18)$$

³²It is a mathematical theorem that all representations of compact Lie groups are equivalent (by similarity transformation) to unitary ones.

The reason for that is the presence of the derivative:

$$\partial_\mu \Psi(x) = \lim_{a \rightarrow 0} \frac{\Psi(x + a\hat{\mu}) - \Psi(x)}{a} \quad (12.19)$$

where $\hat{\mu}$ is the Minkowski basis vector along the direction μ .

However, the derivative ∂_μ is not covariant since it involves the difference of the field at two distinct space-time points where they are transformed differently by the position dependent symmetry. This suggests the introduction of a group valued field $\mathcal{U}(y, x)$ which transforms as

$$\mathcal{U}(y, x) \rightarrow \gamma(y)\mathcal{U}(y, x)\gamma(x)^{-1} \quad (12.20)$$

which can compensate for the difference. This plays the role of *parallel transport* so that $R(\mathcal{U}(y, x))\Psi(x)$ transforms the same way as $\Psi(y)$, so a *covariant derivative* can be defined as

$$D_\mu \Psi(x) = \lim_{a \rightarrow 0} \frac{\Psi(x + a\hat{\mu}) - R(\mathcal{U}(x + a\hat{\mu}, x))\Psi(x)}{a} \quad (12.21)$$

Strictly speaking $\mathcal{U}$ depends on the curve connecting the two points and can be more properly denoted as

$$\mathcal{U}(\mathcal{C}_{x,y}) \quad (12.22)$$

where $\mathcal{C}_{x,y}$ is a curve connecting the two end points, while its transformation still only depends on the end points:

$$\mathcal{U}(\mathcal{C}_{x,y}) \rightarrow \gamma(y)\mathcal{U}(\mathcal{C}_{x,y})\gamma(x)^{-1} \quad (12.23)$$

Obviously $\mathcal{U}(\mathcal{C}_{x,x}) = e$. Therefore, when x and y are close and $\mathcal{C}_{x,y}$ is short, it can be expanded as

$$\mathcal{U}(\mathcal{C}_{x,x+a\hat{\mu}}) = e - iaA_\mu^a(x)T_a + O(a^2) \quad (12.24)$$

where $A_\mu^a(x)$ are called the *gauge fields*.

Exercise 12.1. Show that

$$D_\mu \Psi(x) = \partial_\mu \Psi(x) + iA_\mu^a(x)t_a \Psi(x) \quad (12.25)$$

which is clearly analogous to the electromagnetic covariant derivative (9.78).

Solution. Writing (12.24) in representation R yields

$$R(\mathcal{U}(\mathcal{C}_{x,x+a\hat{\mu}}))_{kl} = \delta_{kl} - iaA_\mu^a(x)(t_a)_{kl} + O(a^2) \quad (12.26)$$

so

$$\begin{aligned} D_\mu \Psi(x)_k &= \lim_{a \rightarrow 0} \frac{\Psi(x + a\hat{\mu})_k - R(\mathcal{U}(x + a\hat{\mu}, x))_{kl} \Psi(x)_l}{a} \\ &= \lim_{a \rightarrow 0} \frac{\Psi(x + a\hat{\mu})_k - (\delta_{kl} - iaA_\mu^a(x)(t_a)_{kl} + O(a^2)) \Psi(x)_l}{a} \\ &= \lim_{a \rightarrow 0} \frac{\Psi(x + a\hat{\mu})_k - \Psi(x)_k}{a} + iA_\mu^a(x)(t_a)_{kl} \Psi(x)_l + O(a) \\ &= \partial_\mu \Psi_k(x) + iA_\mu^a(x)(t_a)_{kl} \Psi_l(x) \end{aligned} \quad (12.27)$$

□

Put differently, D_μ is a matrix differential operator with components

$$(D_\mu)_{kl} = \delta_{kl}\partial_\mu + iA_\mu^a(x)(t_a)_{kl} \quad (12.28)$$

One can also introduce the *Lie algebra valued gauge field*

$$A_\mu(x) = A_\mu^a(x)T_a \quad (12.29)$$

and the abstract operator

$$D_\mu = \partial_\mu + iA_\mu \quad (12.30)$$

in which case the concrete matrix operator (12.28) can be considered as the representation of the abstract operator under R :

$$R(D_\mu) = \partial_\mu + iR(A_\mu) = \partial_\mu + iA_\mu^a(x)t_a \quad (12.31)$$

From (12.23) one can derive the transformation of the gauge fields with the result

$$A'_\mu(x) = \gamma(x)A_\mu(x)\gamma(x)^{-1} + i(\partial_\mu\gamma(x))\gamma(x)^{-1} \quad (12.32)$$

or, for an infinitesimal transformation

$$\gamma(x) = e - i\epsilon^a(x)T_a \quad (12.33)$$

one obtains

$$\delta A_\mu^a(x) = \partial_\mu\epsilon^a(x) + f^{bca}\epsilon^b(x)A_\mu^c(x) \quad (12.34)$$

Exercise 12.2. Compute (12.32) and (12.34) using (12.23) and (12.24).

Solution. Starting from

$$\mathcal{U}(C_{x,x+a\hat{\mu}}) = e - iaA_\mu^a(x)T_a + O(a^2) \quad (12.35)$$

and writing (12.23) in the form

$$\begin{aligned} e - iaA'_\mu(x) &= \gamma(x + a\hat{\mu})(e - iaA_\mu(x))\gamma^{-1}(x) \\ &= (\gamma(x) + a\partial_\mu\gamma(x))(e - iaA_\mu(x))\gamma^{-1}(x) + O(a^2) \\ &= e - ia[\gamma(x)A_\mu(x)\gamma(x)^{-1} + i(\partial_\mu\gamma(x))\gamma(x)^{-1}] + O(a^2) \end{aligned} \quad (12.36)$$

gives

$$A'_\mu(x) = \gamma(x)A_\mu(x)\gamma(x)^{-1} + i(\partial_\mu\gamma(x))\gamma(x)^{-1} \quad (12.37)$$

(recall that all terms of order a^2 and higher must be neglected). Substituting the infinitesimal transformation (12.33) into this relation:

$$\begin{aligned} (A_\mu(x) + \delta A_\mu(x)) &= (e - i\epsilon^b(x)T_b)A_\mu(x)(e + i\epsilon^b(x)T_b) + i(\partial_\mu(e - i\epsilon^b(x)T_b))(e + i\epsilon^b(x)T_b) \\ &= A_\mu(x) - i\epsilon^b(x)[T_b, A_\mu(x)] + \partial_\mu\epsilon^b(x)T_b + O(\epsilon^2) \end{aligned} \quad (12.38)$$

leads to

$$\delta A_\mu(x) = \partial_\mu\epsilon^a(x)T_a - i\epsilon^b(x)[T_b, A_\mu(x)] \quad (12.39)$$

In components

$$\begin{aligned} \delta A_\mu^a(x)T_a &= \partial_\mu\epsilon^a(x)T_a - i\epsilon^b(x)[T_b, A_\mu^c(x)T_c] \\ &= \partial_\mu\epsilon^a(x)T_a + \epsilon^b(x)A_\mu^c(x)f^{bca}T_a \end{aligned} \quad (12.40)$$

which is just (12.34). $\square$

Note that the covariant derivative transforms as³³

$$\begin{aligned} D'_\mu &= \partial_\mu + iA'_\mu = \partial_\mu + i\gamma(x)A_\mu(x)\gamma(x)^{-1} - (\partial_\mu\gamma(x))\gamma(x)^{-1} = \gamma(x)(\partial_\mu + iA_\mu)\gamma(x)^{-1} \\ &= \gamma(x)D_\mu\gamma(x)^{-1} \end{aligned} \quad (12.41)$$

which is analogous to the homogeneous transformation (9.79) found in the electromagnetic case. We also remark that (12.39) shows that *the gauge fields transform in the adjoint representation apart from the additional gradient term.*

Now the gauge invariant fermion Lagrangian density can be obtained by *replacing the derivatives with the covariant ones*:

$$\begin{aligned} \mathcal{L}_{\text{inv}} &= \bar{\Psi}_k i(\not{D}\Psi)_k - \bar{\Psi}_k M_{kl}\Psi_l \\ &= \bar{\Psi}_k i\not{\partial}\Psi_k - \bar{\Psi}_k M_{kl}\Psi_l - \bar{\Psi}_k \gamma^\mu (t_a)_{kl}\Psi_l A_\mu^a \end{aligned} \quad (12.42)$$

Exercise 12.3. Show that ((12.42)) is indeed gauge invariant.

Solution. Let us write

$$\begin{aligned} \mathcal{L}_{\text{inv}} &= \mathcal{L}_{\text{matter}}(\Psi) + \mathcal{L}_{\text{int}}(\Psi, A_\mu) \\ \text{with } \mathcal{L}_{\text{matter}}(\Psi) &= \bar{\Psi}_k i\not{\partial}\Psi_k - \bar{\Psi}_k M_{kl}\Psi_l \\ \mathcal{L}_{\text{int}}(\Psi, A_\mu) &= -\bar{\Psi}_k \gamma^\mu t_{kl}^a \Psi_l A_\mu^a = -\bar{\Psi} \gamma^\mu R(A_\mu) \Psi \end{aligned} \quad (12.43)$$

From (12.18), under a gauge transformation

$$\mathcal{L}_{\text{matter}}(\Psi') - \mathcal{L}_{\text{matter}}(\Psi) = i\bar{\Psi}(x)\gamma^\mu (R(\gamma(x))^{-1}\partial_\mu R(\gamma(x))) \Psi(x) \quad (12.44)$$

while from (12.32)

$$\begin{aligned} \mathcal{L}_{\text{int}}(\Psi', A'_\mu) - \mathcal{L}_{\text{int}}(\Psi, A_\mu) &= -\bar{\Psi}(x)\gamma^\mu R(\gamma(x))^{-1} \left(R(\gamma(x))R(A_\mu(x))R(\gamma(x))^{-1} \right. \\ &\quad \left. + i\partial_\mu R(\gamma(x))R(\gamma(x))^{-1} \right) R(\gamma(x))\Psi(x) + \bar{\Psi}(x)\gamma^\mu R(A_\mu(x))\Psi(x) \\ &= -i\bar{\Psi}(x)\gamma^\mu (R(\gamma(x))^{-1}\partial_\mu R(\gamma(x))) \Psi(x) \end{aligned} \quad (12.45)$$

so the transformation of the gauge interaction compensates for the kinetic term. $\square$

Note that the interaction with the gauge fields is of the form $j^{a\mu}A_\mu^a$, where

$$j^{a\mu} = \bar{\Psi}_k \gamma^\mu (t_a)_{kl}\Psi_l \quad (12.46)$$

are *currents for the gauge charges* labelled by $a = 1, \dots, \dim G$. This is a straightforward generalisation of the electromagnetic interaction to charges corresponding to the non-Abelian symmetry group. In contrast to electromagnetism, which has only a single type of charge, a non-Abelian theory with a gauge group G has as many charges and corresponding gauge fields as the dimension of the gauge group.

³³Here we used the identity $\partial_\mu(\gamma(x)^{-1}) = -\gamma(x)^{-1}(\partial_\mu\gamma(x))\gamma(x)^{-1}$.

Exercise 12.4. Show that $j^{a\mu} = \bar{\Psi}_k \gamma^\mu t_{kl}^a \Psi_l$ are indeed the Noether currents corresponding to the global transformations under G .

Solution. Under an infinitesimal global transformation

$$\begin{aligned}\delta\Psi_k &= -i\epsilon^a(t_a)_{kl}\Psi_l \\ \delta\bar{\Psi}_k &= i\epsilon^a(t_a)_{kl}\bar{\Psi}_l\end{aligned}\tag{12.47}$$

so according to Section 8.3 under the a th transformation one has

$$\Delta_a\Psi_k = -i(t_a)_{kl}\Psi_l\tag{12.48}$$

and so we obtain

$$j^{a\mu} = \frac{\partial\mathcal{L}_{\text{matter}}}{\partial(\partial_\mu\bar{\Psi}_k)}\Delta_a\Psi_k = i\bar{\Psi}_k\gamma^\mu(-i)(t_a)_{kl}\Psi_l = \bar{\Psi}_k\gamma^\mu(t_a)_{kl}\Psi_l\tag{12.49}$$

for the corresponding conserved current. $\square$

It is also possible to couple scalar fields to the gauge fields. Let us consider the scalar action with N' complex scalars

$$\mathcal{L}_{\text{scalar}} = (\partial_\mu\Phi_k)^*\partial^\mu\Phi_k - \Phi_k^*M_{kl}^2\Phi_l\tag{12.50}$$

invariant under an N' -dimensional unitary representation S of the group G :

$$S(g)M^2S(g)^{-1} = M^2\tag{12.51}$$

with generators

$$S(T_a)_{kl} = (s_a)_{kl}\tag{12.52}$$

The gauge invariant action can then be written as

$$\mathcal{L}_{\text{inv}} = (D_\mu\Phi)_k^*(D^\mu\Phi)_k - \Phi_k^*M_{kl}^2\Phi_l\tag{12.53}$$

with

$$(D_\mu\Phi)_k = \partial_\mu\Phi_k + iA_\mu^a(s_a)_{kl}\Phi_l\tag{12.54}$$

12.1.3 The gauge field action

In order to construct an action for the dynamics of the gauge field, it is necessary to introduce the field strength. In electrodynamics, the field strength is

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu\tag{12.55}$$

which is gauge invariant. It turns out that there is no similar gauge invariant for non-Abelian fields. However, guidance is provided by the geometric interpretation as parallel transport. Namely, parallel transport depends on the path; for infinitesimal movements, this is reflected in the fact that covariant derivatives in different directions do not commute, so taking first an infinitesimal

shift in direction ν and then in direction μ does not give the same result as in the opposite order. Indeed, it is easy to compute

$$\begin{aligned}[D_\mu, D_\nu] &= [\partial_\mu + iA_\mu^a T_a, \partial_\nu + iA_\nu^b T_b] \\ &= i(\partial_\mu A_\nu^c - \partial_\nu A_\mu^c) T_c - A_\mu^a A_\nu^b [T_a, T_b] \\ &= iT_c(\partial_\mu A_\nu^c - \partial_\nu A_\mu^c - f^{abc} A_\mu^a A_\nu^b)\end{aligned}\tag{12.56}$$

The *non-Abelian gauge field strength* can be defined as

$$F_{\mu\nu}^c = \partial_\mu A_\nu^c - \partial_\nu A_\mu^c - f^{abc} A_\mu^a A_\nu^b\tag{12.57}$$

or in a matrix form

$$F_{\mu\nu} = F_{\mu\nu}^c T^c\tag{12.58}$$

Note that there is exactly one field strength tensor per gauge field. Due to the transformation rule (12.41) of the covariant derivative, *the field tensor transforms homogeneously*

$$F'_{\mu\nu}(x) = \gamma(x) F_{\mu\nu} \gamma(x)^{-1}\tag{12.59}$$

so

$$\text{Tr} F_{\mu\nu}(x) F^{\mu\nu}(x)\tag{12.60}$$

is invariant. Let's assume that the gauge group can be decomposed into factors³⁴

$$G = G_1 \times G_2 \times \dots G_n\tag{12.61}$$

where each G_i is either a simple (i.e. non factorizable) non-Abelian group or a $U(1)$ factor. Then we can consider each factor separately. For a non-Abelian simple factor, one can always choose generators such that³⁵

$$\text{Tr} R_d(T_a) R_d(T_b) = \frac{1}{2} \delta_{ab}\tag{12.62}$$

where R_d is the defining representation of the group ($d = \mathbf{N}$ for $SU(N)$) so

$$\text{Tr} F_{\mu\nu}(x) F^{\mu\nu}(x) = \frac{1}{2} F_{\mu\nu}^a(x) F^{a\mu\nu}(x)\tag{12.63}$$

So it is possible to write down the *non-Abelian analogue of the Maxwell action*

$$\mathcal{L}_{\text{gauge}} = -\frac{1}{4g^2} F_{\mu\nu}^a(x) F^{a\mu\nu}(x)\tag{12.64}$$

where we introduced an arbitrary overall normalisation g . As a result, a general non-Abelian gauge theory Lagrangian can be written as

$$\begin{aligned}\mathcal{L}_{\text{inv}} &= \bar{\Psi}_k (i\not{\partial} \delta_{kl} - \mathcal{A}^a(t_a)_{kl}) \Psi_l + ((\partial^\mu \delta_{rt} + iA^{a\mu}(s_a)_{rt}) \Phi_t)^* ((\partial_\mu \delta_{ru} + iA_\mu^a(s_a)_{ru}) \Phi_u) \\ &\quad - \bar{\Psi}_k M_{kl}^{(f)} \Psi_l - \Phi_k^* M_{kl}^{2(s)} \Phi_l + \mathcal{L}_1(\Phi, \Psi) - \frac{1}{4g^2} F_{\mu\nu}^a(x) F^{a\mu\nu}(x)\end{aligned}\tag{12.65}$$

³⁴This decomposition is possible and unique for all compact Lie groups.

³⁵This is ensured by the existence of an essentially unique invariant scalar product (the so-called Killing form) on the Lie algebra.

where $\mathcal{L}_1$ is some gauge invariant interaction involving the scalar and fermion fields, composed of terms such as e.g.

$$\begin{aligned} \text{scalar self-interaction:} & \quad \lambda_{rstu} \Phi_r^* \Phi_s^* \Phi_t \Phi_u + \text{h.c.} \\ \text{Yukawa interaction:} & \quad y_{rkl} \Phi_r \bar{\Psi}_k \Psi_l + \text{h.c.} \end{aligned} \quad (12.66)$$

where

$$\begin{aligned} \lambda_{rstu} &= S(\gamma)_{rr'}^* S(\gamma)_{ss'}^* S(\gamma)_{tt'} S(\gamma)_{uu'} \lambda_{r's't'u'} \\ y_{rkl} &= S(\gamma)_{rr'} R(\gamma)_{kk'}^* R(\gamma)_{ll'} y_{r'k'l'} \end{aligned} \quad (12.67)$$

for any gauge group element γ , and h.c. stands for Hermitian conjugate.

12.1.4 Recovering the Abelian case

It is possible to rescale the fields

$$A_\mu \rightarrow g A_\mu \quad F_{\mu\nu} \rightarrow g F_{\mu\nu} \quad (12.68)$$

so that

$$\begin{aligned} \mathcal{L}_{\text{inv}} &= \bar{\Psi}_k (i \not{\partial} \delta_{kl} - g \mathcal{A}^a (t_a)_{kl}) \Psi_l + (\partial_\mu \Phi_r + i g A_\mu^a (s_a)_{rt} \Phi_t)^* (\partial_\mu \Phi_r + i g A_\mu^a (s_a)_{ru} \Phi_u) \\ &\quad - \bar{\Psi}_k M_{kl}^{(f)} \Psi_l - \Phi_k^* M_{kl}^{2(s)} \Phi_l + \mathcal{L}_1(\Phi, \Psi) - \frac{1}{4} F_{\mu\nu}^a(x) F^{a\mu\nu}(x) \end{aligned} \quad (12.69)$$

and

$$F_{\mu\nu}^c = \partial_\mu A_\nu^c - \partial_\nu A_\mu^c - g f^{abc} A_\mu^a A_\nu^b \quad (12.70)$$

In this case the gauge kinetic terms are canonically normalised

$$\mathcal{L}_{\text{gkin}} = -\frac{1}{4} (\partial_\mu A_\nu^c - \partial_\nu A_\mu^c) (\partial^\mu A^{c\nu} - \partial^\nu A^{c\mu}) \quad (12.71)$$

just as for the electromagnetic field. This form of the Lagrangian makes it clear that *for each factor group the interaction strength is determined by the corresponding coupling constant g , and the non-Abelian charges of any field are determined by its representation via the group generators: t for the fermions, s for the scalars, and the adjoint generators for the gauge fields.*

For the Abelian case $U(1)$ the structure constants vanish ($f^{abc} = 0$), so the gauge field is neutral as is well-known for the photon in electromagnetism. The coupling constant g can be chosen as the elementary charge e . All irreducible representations of the $U(1)$ group are one dimensional, and are parametrised by a single real number Q :

$$R(e^{-i\alpha}) = e^{-i\alpha Q} \quad (12.72)$$

Therefore the representations of the single generator T are

$$\begin{aligned} t &= R(T) = \begin{pmatrix} Q_1^{(\Psi)} & & & \\ & Q_2^{(\Psi)} & & \\ & & \dots & \\ & & & Q_N^{(\Psi)} \end{pmatrix} \\ s &= S(T) = \begin{pmatrix} Q_1^{(\Phi)} & & & \\ & Q_2^{(\Phi)} & & \\ & & \dots & \\ & & & Q_{N'}^{(\Phi)} \end{pmatrix} \end{aligned} \quad (12.73)$$

and the action is

$$\begin{aligned}\mathcal{L}_{\text{inv}} = & \sum_{k=1}^N \bar{\Psi}_k \left(i\not{\partial} - Q_k^{(\Psi)} e A \right) \Psi_k + \sum_{r=1}^{N'} \left(\partial^\mu \Phi_r^* - i Q_r^{(\Phi)} e A^\mu \Phi_r^* \right) \left(\partial_\mu \Phi_r + i Q_r^{(\Phi)} e A_\mu \Phi_r \right) \\ & - \bar{\Psi}_k M_{kl}^{(f)} \Psi_l - \Phi_k^* M_{kl}^{2(s)} \Phi_l + \mathcal{L}_1(\Phi, \Psi) - \frac{1}{4} F_{\mu\nu}(x) F^{\mu\nu}(x)\end{aligned}\quad (12.74)$$

So we conclude that the combinations $Q_k^{(\Psi)} e$ and $Q_r^{(\Phi)} e$ give the electric charges of the particles created by the fields Ψ_k and Φ_r (with a minus sign for anti-particles). Therefore *a crucial difference between Abelian and non-Abelian theories is that in the former the charges can take any arbitrary value, while in the latter the charges are all quantised in units of a single gauge coupling* and are determined by the representations under which the fields transform. Therefore *in a non-Abelian theory the gauge fields are themselves charged under the gauge interaction according to the adjoint representation*.

12.2 Discovery of quark colour

The quark model introduced in Section 7.3 describes a lot of features of the hadronic spectrum well. However, it also leads to some puzzles, the chief among them being an *apparent violation of Pauli's principle by the baryon decuplet*. As discussed before, the wave function of the decuplet is fully symmetric in the flavour indices and, since the spin of these baryons is 3/2, it is fully symmetric in the spin indices as well. In addition, these hadrons are in the ground state of the three-quark system, and simple quantum mechanics dictates that their wave-function is then symmetric in real space coordinates, too. However, due to Pauli principle the full wave function must be antisymmetric under exchanging any pair of the constituent quarks.

One may suggest an escape route by observing that flavour $SU(3)$ is only approximate, so maybe the symmetry in flavour indices is a weak link in the argument. However, for particles composed of three identical quarks such as $\Delta^{++} = uuu$, $\Delta^- = ddd$ and $\Omega^- = sss$ the wave function is trivially symmetric in flavour, so the contradiction is definitely valid at least for these.

The eventual way out was proposed by Murray Gell-Mann and Harald Fritzsch: it is to *introduce a new quantum number called colour*, which takes three values³⁶ and assume that the wave function is antisymmetric in the colour index. So the total baryonic wave function for the decuplet can be written as

$$\sum_{\sigma_i, f_i, c_i} \int d^3 p_1 \int d^3 p_2 \int d^3 p_3 \Psi(\vec{p}_1, \vec{p}_2, \vec{p}_3) \Sigma_{\sigma_1 \sigma_2 \sigma_3} \Phi_{f_1 f_2 f_3} \epsilon_{c_1 c_2 c_3} |\vec{p}_1, \sigma_1, f_1, c_1; \vec{p}_2, \sigma_2, f_2, c_2; \vec{p}_3, \sigma_3, f_3, c_3\rangle \quad (12.75)$$

where $\vec{p}_i, \sigma_i, f_i, c_i$ denote the momentum, spin- z component, flavour and colour of the i th quark, $i = 1, 2, 3$ with Ψ, Σ and Φ totally symmetric in their arguments/indices, and ϵ the usual 3-dimensional Levi-Civita tensor.

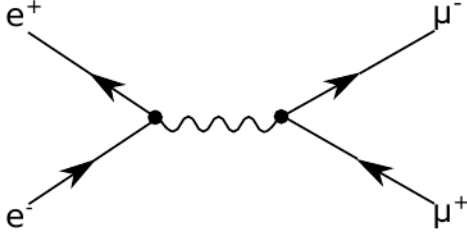
So far this is just a theoretical argument; however, soon after its introduction it became clearly supported by experimental results:

1. Electron-positron annihilation into hadrons

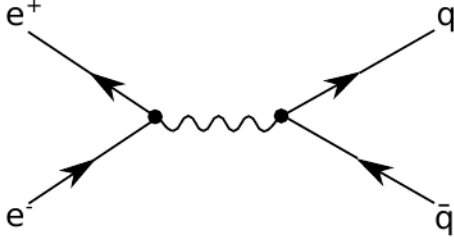
In e^+e^- colliders, electron-positron annihilation leads to the creation of particles, the dominant route being an electromagnetic exchange. For example, to leading order in α muon pair

³⁶In principle, the puzzle can be resolved with a quantum number taking more than three values, but as shown shortly, experimental results indicate the existence of exactly three independent colours.

creation is described by the Feynman diagram



while hadron creation proceeds dominantly via quark-antiquark pair creation



which is then followed by a process called hadronisation. Now the amplitudes corresponding to muon and quark pair creations differ both due to the difference in charge and in mass. However, provided the collision energy is much higher than the muon and quark masses, the mass dependence can be neglected, and the only difference is the charge at the second vertex which leads to the following relation between the amplitudes:

$$\begin{aligned}\mathcal{A}(e^+e^- \rightarrow u\bar{u}) &= -\frac{2}{3}\mathcal{A}(e^+e^- \rightarrow \mu^+\mu^-) \\ \mathcal{A}(e^+e^- \rightarrow d\bar{d}) &= \frac{1}{3}\mathcal{A}(e^+e^- \rightarrow \mu^+\mu^-) \\ \mathcal{A}(e^+e^- \rightarrow s\bar{s}) &= \frac{1}{3}\mathcal{A}(e^+e^- \rightarrow \mu^+\mu^-)\end{aligned}\quad (12.76)$$

Since the probabilities are obtained as the modulus squared of the amplitudes and are additive for independent channels, this leads to the following ratio of probabilities:

$$R = \frac{\sigma(e^+e^- \rightarrow \text{hadrons})}{\sigma(e^+e^- \rightarrow \mu^+\mu^-)} = \left(-\frac{2}{3}\right)^2 + \left(\frac{1}{3}\right)^2 + \left(\frac{1}{3}\right)^2 = \frac{2}{3} \quad (12.77)$$

as long as the collision energy is much higher than the strange quark mass (cc. 135 MeV) but lower than the threshold for the creation of a charm-anticharm pair (cc. 3 GeV). However, the observed value for this ratio is

$$R \approx 2.0 \quad (12.78)$$

in the range of energy between 1.5 and 3 GeV. This is explained by the presence of colour, since then there are three times as many independent end states for quark-antiquark pairs, which leads to

$$R = \frac{\sigma(e^+e^- \rightarrow \text{hadrons})}{\sigma(e^+e^- \rightarrow \mu^+\mu^-)} = 3 \left[\left(-\frac{2}{3}\right)^2 + \left(\frac{1}{3}\right)^2 + \left(\frac{1}{3}\right)^2 \right] = 2 \quad (12.79)$$

Above the charm threshold this changes to

$$R = \frac{\sigma(e^+e^- \rightarrow \text{hadrons})}{\sigma(e^+e^- \rightarrow \mu^+\mu^-)} = 3 \left[\left(-\frac{2}{3}\right)^2 + \left(\frac{1}{3}\right)^2 + \left(\frac{1}{3}\right)^2 + \left(-\frac{2}{3}\right)^2 \right] = \frac{10}{3} \quad (12.80)$$

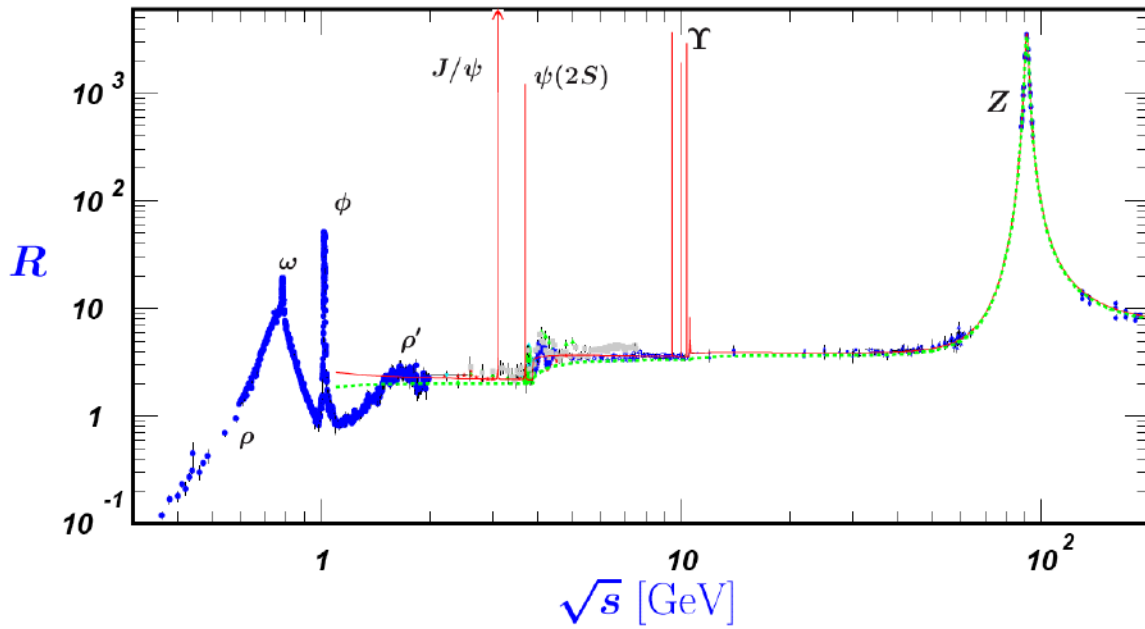


Figure 12.1: The ratio R as a function of the centre-of-mass energy of the electron-positron collision. The green curve is the value predicted by the naive quark model with three colours, including thresholds and the resonance corresponding to the Z boson that mediates neutral current weak interactions. The red curve indicates the prediction improved by taking into account quark interactions using third-order perturbative QCD corrections, including heavy quark-antiquark meson resonances such as the ψ ($c\bar{c}$) and the Υ ($b\bar{b}$) family. The lower the energies, the more the interaction contribution matters, and at energies below 1.5 GeV even perturbation theory fails: the quarks are strongly interacting, and the cross-section is dominated by contributions from quark-antiquark bound states, corresponding to the vector meson resonances indicated in the plot.

From: Review of Particle Physics (K.A. Olive et al. (Particle Data Group), Chinese Phys. C38 (2014) No 9.)

while above the bottom threshold the result is

$$R = \frac{\sigma(e^+e^- \rightarrow \text{hadrons})}{\sigma(e^+e^- \rightarrow \mu^+\mu^-)} = 3 \left[\left(-\frac{2}{3}\right)^2 + \left(\frac{1}{3}\right)^2 + \left(\frac{1}{3}\right)^2 + \left(-\frac{2}{3}\right)^2 + \left(\frac{1}{3}\right)^2 \right] = \frac{11}{3} \quad (12.81)$$

Again, this is in full agreement with experimental data as shown in Fig. 12.1. The crossover between the two values happens at the charm threshold where R depends on energy; this energy dependence disappears again for energies much higher than the charm threshold.

2. Decay rate of π^0

The decay rate of the π^0 meson can be calculated using the so-called Adler-Bell-Jackiw (ABJ) triangle anomaly and the partially conserved axial current (PCAC) hypothesis³⁷ it involves an integration over intermediate quark states. Again, the decay amplitude comes out a factor 3 smaller compared to the experimental result, which is exactly accounted for by the colour

³⁷See e.g. Section 6.2 of the book T.-P. Cheng and L.-F. Li: *Gauge theory of elementary particle physics*, Clarendon Press, Oxford, 1982.

degrees of freedom. The decay rate itself, being the square of the amplitude, is proportional to the number of colours squared:

$$\Gamma(\pi^0 \rightarrow 2\gamma) = N_c^2(Q_u^2 - Q_d^2)^2 \frac{\alpha^2 m_{\pi^0}^3}{64\pi^3 F_\pi^2} \approx 7.76 \text{ eV} \quad (12.82)$$

where $F_\pi = 92 \text{ MeV}$ is the pion decay constant entering the PCAC, $Q_u = 2/3$ and $Q_d = -1/3$. Using the experimentally measured lifetime $\tau = (8.4 \pm 0.6) \times 10^{-17} \text{ s}$, this must be compared to

$$\Gamma_{\text{exp}} = \hbar/\tau = 7.83 \pm 0.56 \text{ eV} \quad (12.83)$$

which agrees well with the theoretical prediction.

3. Furthermore, assuming that colour $SU(3)$ is a gauge symmetry allows the construction of *quantum chromodynamics*, which is supported by overwhelming evidence as the correct fundamental theory of strong interactions. This is the topic of the next Section.

12.3 Quantum chromodynamics: the gauge theory of strong interactions

Quantum chromodynamics (*QCD*) is a non-Abelian gauge theory, where the gauge group is the $SU(3)$ symmetry acting on the colour degree of freedom of quarks. Quarks transform in representation $\mathbf{3}$, while antiquarks in the conjugate representation $\bar{\mathbf{3}}$ of the group $SU(3)_c$. It is very important to distinguish this from the Gell-Mann $SU(3)_f$, which instead acts on flavours and is only an approximate global symmetry for energies below the charm threshold acting on flavours u, d, s . In contrast, $SU(3)_c$ is an exact and local (gauge) symmetry.

The full QCD Lagrangian can be written as

$$\mathcal{L}_{\text{QCD}} = -\frac{1}{4}F_{\mu\nu}^a F^{a\mu\nu} + \sum_q \bar{q}(i\not{D} - M_q)q \quad (12.84)$$

where q runs over the quarks u, d, s, c, b, t , their masses denoted by M_q . As just said, the quarks transform in representation $\mathbf{3}$; besides the Dirac bispinor index, the fields q have an additional colour index i , transforming under a 3×3 $SU(3)$ matrix U as

$$q'_i(x) = U_{ij}q_j(x) \quad (12.85)$$

and in the product $\bar{q}(i\not{D} - M_q)q$ a summation over colour indices is understood. The generators of $SU(3)_c$ in this representation are

$$t_a = \frac{1}{2}\lambda_a \quad a = 1, \dots, 8 \quad (12.86)$$

where λ_a are the Gell-Mann matrices introduced in Section [7.3](#). The covariant derivative is given by

$$(D_\mu q)_k = \partial_\mu q_k - ig_s A_\mu^a \frac{(\lambda_a)_{kl}}{2} q_l \quad (12.87)$$

where g_s is the strong coupling constant. The colour field strengths are given by

$$F_{\mu\nu}^c = \partial_\mu A_\nu^c - \partial_\nu A_\mu^c - f_3^{abc} A_\mu^a A_\nu^b \quad (12.88)$$

where f_3^{abc} are the structure constants of $SU(3)$:

$$\left[\frac{\lambda_a}{2}, \frac{\lambda_b}{2} \right] = i f_3^{abc} \frac{\lambda_c}{2} \quad (12.89)$$

We note that it is possible to add an additional gauge-invariant term to the Lagrangian

$$\mathcal{L}_\theta = \frac{\theta}{32\pi^2} \epsilon^{\mu\nu\lambda\rho} F_{\mu\nu}^a F_{\lambda\rho}^a \quad (12.90)$$

which is, *despite being a total derivative and so irrelevant for the classical equations of motion, does contribute at the quantum level*. The parameter θ is defined as an angle, i.e. up to a period of 2π . It leads to *strong CP violation*, for example, an electric dipole moment for the neutron. Current limits on such effects mean that

$$|\theta| < 10^{-9} \quad (12.91)$$

Since this term is allowed by all known symmetry principles, its absence (or smallness) is a challenge to explain and constitutes a major open question at present. A leading scenario involves the existence of a novel light particle called the axion, which has not been observed yet.

12.3.1 Confinement and asymptotic freedom

The theory of quantum chromodynamics must give an account of two fundamental observations:

1. *No free quarks have ever been observed* (their presence would be obvious from their fractional charge). This suggests a very strong interaction binding them together, but even that is not necessarily enough. The binding energy between quarks cannot be much more than a few hundred MeV without being in tension with the observed hadron masses. However, the quarks do not separate even in collisions which involve energy transfers that are orders of magnitude larger, even up to the TeV scale. This is called *quark confinement*.
2. When considering the evidence for quarks in Section 7.3.7, we saw that during the scattering of a high energy probe *in the kinematical regime of large momentum transfer, the quarks behave as almost free particles*. This so-called *asymptotic freedom* seems to contradict the presence of very strong forces responsible for confinement.

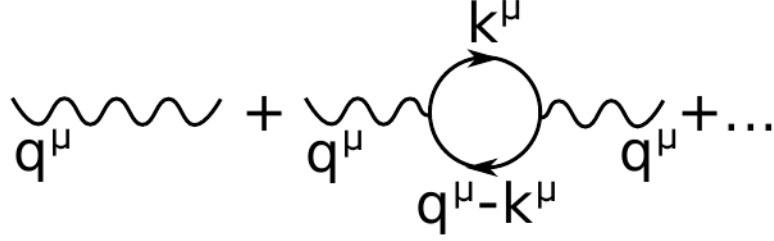
Below we give a quick sketch on how QCD solves both problems at the same time.

12.3.2 Running coupling, effective theories and the Theory of Everything

The explanation for *asymptotic freedom* is simpler as it only requires considering QCD where it is weakly interacting and so perturbation theory can be applied. However, the eventual calculation is still quite involved to below we give just a sketch of the results. The interested reader may consult Appendix B where a relatively elementary derivation is given.

Let us first consider quantum electrodynamics where the interaction is mediated by photons. However, once a photon is emitted by some charge and is propagating towards the other charge participating in the interaction, it can interact with the vacuum itself. The true photon propagator³⁸ is described as the sum of all Feynman diagrams with one incoming and one outgoing photon leg, and is given by

³⁸The propagator is nothing else but the amplitude describing the propagation of the field between space-time points. In classical electrodynamics it is also known as the Green's function, which can be retarded or advanced, depending on the boundary conditions. Quantum field theory uses neither retarded nor advanced, but a third version called the Feynman propagator.



where the first term is the naive photon propagator, the second term is the $O(e^2)$ i.e. $O(\alpha)$ correction and the omitted terms are all of at least fourth order in e (second order in α). The four-momentum q^μ is the one carried by the photon, while there is also a so-called loop momentum k to be integrated over when summing over the intermediate states.

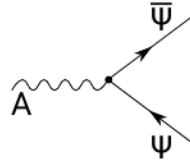
Now it turns out that the integral over k is divergent, so some care must be taken. However, we know that QED is not valid to arbitrarily high energies, and so a cut-off scale Λ can be introduced. Essentially, one only includes intermediate states below this energy scale. It turns out that the divergence is logarithmic, so the dependence on the actual value of Λ is mild. Even better, this dependence can be absorbed into a redefinition of the electric charge. So the idea is that the bare electron would have a charge e_0 , but the measured quantity is the so-called renormalised charge, which is defined to be the strength of interaction as observed in low energy processes, and is given by the elementary charge e . Eventually, the value of the bare charge can be computed in terms of the elementary charge and the cut-off, and the relation is given by

$$e_0^2 = e^2 \left(1 + \frac{\alpha}{3\pi} \log \frac{\Lambda^2}{m^2} + O(\alpha^2) \right) \quad (12.92)$$

where m is the physical electron mass. We mention here that the physical mass is similarly different from the bare mass of the electron by interaction corrections (electromagnetic self-energy):

$$\delta m = m \frac{3\alpha}{4\pi} \left(\log \frac{\Lambda^2}{m^2} + \frac{1}{2} \right) + O(\alpha^2) \quad (12.93)$$

and the vertex



receives corrections as well, which lead to a shift in the gyromagnetic factor of the electron, away from the value 2 predicted by the Dirac equation:

$$g = 2 \left(1 + \frac{\alpha}{2\pi} \right) + O(\alpha^2) = 2.00232 \dots \quad (12.94)$$

These formulae only take into account the corrections due to the electron/positron field and only to the lowest order; eventually, other particles and their interactions must be taken into account for more precise calculations, as well as higher orders in α .

The observation relevant to the problem of asymptotic freedom is that *the propagator correction depends on q , leading to an effective interaction strength that depends on the momentum transfer*

characterised by q^μ . In fact, it can solely depend on the only Lorentz invariant q^2 and an explicit calculation gives the differential equation

$$\mu \frac{d\alpha}{d\mu} = \frac{2\alpha^2}{3\pi} + O(\alpha^3) \quad (12.95)$$

where $\mu^2 = -q^2$ is the relevant energy scale in the process. The interested reader can find a relatively elementary explanation of this result in Subsection [B.3](#) of Appendix [B](#). Taking into account the existence of other fermion species this equation is modified to

$$\mu \frac{d\alpha}{d\mu} = z_f \frac{2\alpha^2}{3\pi} + O(\alpha^3) \quad (12.96)$$

with $z_f = \sum_i Q_i^2$ for the charged fermions of charge $Q_i e$ relevant at scale μ . At the minimum value $\mu = 2m_e \approx 1$ MeV the value is $z_f = 1$, increasing to $z_f = 4$ at 1 GeV and $z_f = 20/3$ at 100 GeV (note that quarks count with multiplicity 3 due to their colour). For energies above $2M_W$ (where $M_W \approx 83$ GeV is the mass of the $W^\pm$) the contribution of the W bosons must also be added.

For simplicity we consider the solution of [\(12.95\)](#) which is

$$\alpha(\mu) = \frac{\alpha(\mu_0)}{1 - \frac{2\alpha(\mu_0)}{3\pi} \log \frac{\mu}{\mu_0}} \quad (12.97)$$

The usual choice for the fine structure constant is that measured from atomic physics, which corresponds to the threshold scale $\mu_0 = 2m_e \approx 1$ MeV and has the value

$$\alpha(\mu_0)^{-1} = 137.035999139(31) \quad (12.98)$$

It is clear that the effective fine structure constant increases with energy. The energy dependence of interaction strength such as in [\(12.97\)](#) is known as the *running coupling*. The measured value at the Z resonance is

$$\alpha(M_Z)^{-1} \approx 129 \quad (12.99)$$

This agrees well with the prediction computed from the more complete equation [\(12.96\)](#).

Why does the coupling increase with the energy scale? Recall that high energy is the same as short distance, so the physical picture is that the electron has a larger charge when viewed from closer. The value of the elementary charge is the charge of the electron from a very large distance compared to the Compton wave-length

$$\frac{\hbar}{m_e c} = 2.4263102389(16) \cdot 10^{-12} \text{ meters} \quad (12.100)$$

Exercise 12.5. A somewhat inaccurate, but useful analogy is provided by the Debye screening in plasma. When placing a charged object inside a plasma, its electric potential field is modified according to

$$V(r) = \frac{Q}{4\pi\epsilon_0 r} e^{-r/d} \quad (12.101)$$

where d is the so-called Debye screening length.

1. Compute the electric field $\vec{E}(\vec{r})$!

2. What is the effective charge $Q_{\text{eff}}(R)$ seen from distance R ?

Solution. The electric field strength is then

$$\vec{E}(\vec{r}) = -\text{grad}V(r) = -\frac{\partial V}{\partial r} \frac{\vec{r}}{r} = \frac{Q\vec{r}}{4\pi\epsilon_0 r^3} \left(1 + \frac{r}{d}\right) e^{-r/d} \quad (12.102)$$

The charge seen at distance R can be computed by Gauss's theorem

$$Q_{\text{eff}}(R) = \epsilon_0 \oint_R d\vec{f} \cdot \vec{E}(\vec{r}) = Q \left(1 + \frac{R}{d}\right) e^{-R/d} \quad (12.103)$$

Note that right at the place of the charge one sees $Q_{\text{eff}}(0) = Q$, which then monotonically decreases to 0 as $R \gg d$. $\square$

What is the medium polarised when the electron is in a vacuum? Surprisingly, it is the vacuum itself! Remember that the physical vacuum is not empty space: it is filled by the quantum fields whose elementary excitations correspond to the photon, electron/positron and other particles³⁹. Since these fields are quantum, they do fluctuate even in the vacuum, so the vacuum is effectively an active medium like the plasma in Debye screening. A charged object distorts these ever-present vacuum fluctuations: what happens is that the fluctuations become preferentially polarised in the presence of the charge, building up an effective background of opposite charge density around the object, which cancels some of the electric flux. When approaching the object closer one measures a larger flux, and therefore a larger charge, since part of the polarisation cloud is left further out. The result is the charge renormalisation (12.92) which includes the effect of fluctuating electron-positron modes up to the cut-off Λ . These fluctuations are represented by the electron-positron loop in the perturbative expansion. Alternatively, their leading order effect can be computed by borrowing methods from condensed matter theory of magnetic materials, as discussed in Appendix B. Vacuum fluctuations have many consequences, such as the running coupling, the anomalous magnetic moment of the electron (and the muon), the Casimir effect, the Lamb shift in atomic spectra, and the yet unobserved phenomena of vacuum birefringence and photon-photon scattering.

Note that the running coupling (12.97) blows up at the energy

$$\Lambda_0 = \mu_0 e^{\frac{3\pi}{2\alpha(\mu_0)}} \approx 10^{280} \text{ MeV} \quad (12.104)$$

which is astronomically high, a phenomenon noted by L. Landau and called a *Landau singularity*. Adding the other known charged particles brings this down to lower values, but it still remains much higher than the Planck scale 10^{19} GeV. Eventually, at this scale perturbation theory breaks down, as also manifest from (12.92) (12.93), so *to investigate the presence of this singularity one needs to go beyond perturbation theory*. However, this is not necessary in practice since pure QED ceases to be valid at energies around 100 GeV due to electroweak unification, and it is generally expected that standard quantum field theory ceases to be valid due to quantum gravitational effects around the Planck scale anyway.

³⁹More precisely, in quantum field theory the vacuum is the ground state of the quantum fields that make up matter at the fundamental level. Since these fields have a non-linear dynamics due to interactions, the vacuum is one of the most complicated dynamical objects that exist in Nature, which seems surprising in light of the everyday notion of the vacuum being just empty space.

QED is an example of a theory with a coupling that increases with energy. It turns out that this is a very general phenomenon; the chief exception is just the case of non-Abelian gauge theories which manifest asymptotic freedom. For all other quantum field theories⁴⁰ such as those with scalar self-interactions and/or fermion-scalar Yukawa interactions in addition to Abelian gauge fields, the couplings increase with energy and even taking into account non-perturbative effects they are expected to blow up at some Landau singularity Λ_0 . The proper interpretation of this phenomenon is that these theories only work for energies below Λ_0 so the ultraviolet cut-off Λ cannot be increased any further. *Such theories can only be used as low-energy effective descriptions, and for high energies they must be replaced by some more complete theory.* However, as shown by the example of asymptotically free theories, there can exist models that can in principle be valid to arbitrary high energies. *The hypothetical model which provides a description of physical phenomena to arbitrary high energies is called the Theory of Everything (TOE); such a theory is called ultraviolet complete meaning that it does not have an upper energy limit for its validity.* The TOE must include quantum gravity and due to problems of reconciling gravity with ordinary quantum theory, it is generally expected to be beyond the reach of quantum field theory (although there is a recent attempt to avoid such problems, the so-called asymptotic safety scenario).

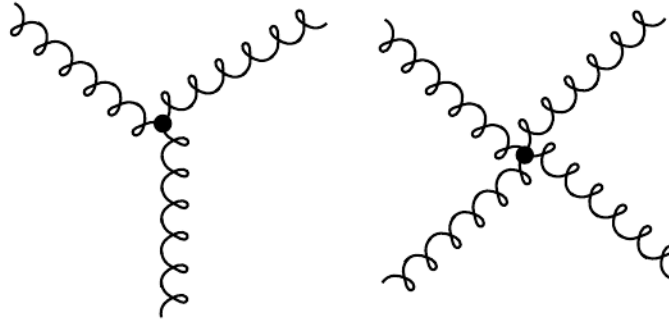
In contrast, a theory with a Landau singularity is called *trivial*, since the Landau singularity is generally of the form

$$\Lambda_0 \propto e^{C/g^2} \quad (12.105)$$

where g is the coupling constant ($= e$ for QED), so the only way to get a UV complete theory is to set the coupling to zero, i.e. a trivial (non-interacting) theory. The value of Λ_0 is called the triviality bound; in general it must be found from non-perturbative arguments, since it lies at (infinitely) strong coupling.

12.3.3 Asymptotic freedom

For a non-Abelian gauge theory there are terms in the Lagrangian involving the product of three or four gauge fields, which lead to interaction vertices of the form



The fermionic contribution to the running coupling has the same effect as in QED. However, the gluon contribution arising from the above vertices has the opposite sign⁴¹ and for an $SU(N_c)$ gauge theory with N_f quarks in fundamental representation the running coupling is given by

$$\mu \frac{d\alpha_s}{d\mu} = - \left(\frac{11}{3} N_c - \frac{2}{3} N_f \right) \frac{\alpha_s^2}{2\pi} + O(\alpha_s^3) \quad (12.106)$$

⁴⁰This is valid in 4 dimensions or higher; in two space-time dimensions there is a number of theories manifesting asymptotic freedom without non-Abelian gauge fields.

⁴¹Unfortunately, in contrast to the fermion contribution which was also present in QED, there is no good intuitive way to understand why the sign is opposite for the gluons; this just happens to be the result of the field theoretic evaluation of the necessary Feynman diagrams.

For the origin of the antiscreening contribution the interested reader is referred to Appendix [B.4](#). For QCD with $N_c = 3$ the combination

$$\frac{11}{3}N_c - \frac{2}{3}N_f = 11 - \frac{2}{3}N_f \quad (12.107)$$

is positive, which leads to *a coupling decreasing with energy* as long as $N_f \leq 16$; since there are only six flavours at presently accessible energies, *the QCD coupling decreases with energy. This phenomenon is called asymptotic freedom.*

The solution of the differential equation [\(12.97\)](#) is given by

$$\alpha_s(\mu) = \frac{1}{b_0 \log \frac{\mu}{\Lambda_{\text{QCD}}}} \quad b_0 = \frac{1}{2\pi} \left(\frac{11}{3}N_c - \frac{2}{3}N_f \right) \quad (12.108)$$

as long as $b_0 > 0$. The scale Λ_{QCD} parametrises the initial conditions; the solution is only valid for $\mu \gg \Lambda_{\text{QCD}}$ where the coupling is small enough. Moving closer to Λ_{QCD} from above higher order corrections must be taken into account; these are dependent on further conventions, leading to a scheme dependence in the scale Λ_{QCD} . The singularity of the coupling at the scale Λ_{QCD} is not physical; there is no Landau singularity due to the fact that for low energies the physical behaviour is governed by a new non-perturbative phenomenon, namely colour confinement. The value of Λ_{QCD} depends on the scheme used to compute higher corrections, but is in the range 200...300 MeV. For processes involving momentum transfers far above this scale ($\gtrsim 1$ GeV) QCD can be treated perturbatively.

The theoretically predicted running coupling can be compared to the measured value of the QCD coupling as function of the energy scale, obtained from various processes involving strong interactions. The experimental situation is summarised in Fig. [\(12.2\)](#) and (together with a vast number of other data) confirms QCD as the correct theory of strong interactions, at least in the high-energy (perturbative) regime.

A few important experimental facts explained by the running coupling:

1. Deep inelastic scattering

As discussed in Section [7.3.7](#) the inclusive cross section of the deep inelastic scattering depends on two kinematical variables x and Q^2 . In particular, for high values of Q^2 the structure functions depend mostly on x only, which corresponds to the presence of essentially independent i.e. free point-like partons on which the scattering happens, which were identified with the quarks. Asymptotic freedom explains their why the interactions between partons can be neglected for processes at high momentum transfers Q^2 .

A more detailed analysis shows that the structure functions show a slow (logarithmic) dependence on Q^2 , which can be explained by perturbative corrections from the QCD interaction.

2. Flavour symmetry breaking

Exact $SU(3)$ flavour symmetry would hold in the case when quark masses are equal: $m_u = m_d = m_s$. Breaking of flavour symmetry depends on how quark mass differences compare to the QCD scale Λ_{QCD} . It is clear that isospin is a quite good symmetry, since it is broken partly by the $m_u - m_d$ mass difference, which is of the order of a few MeV, and by electromagnetic interactions, which are much weaker than the strong interactions on typical hadronic energy scales. $SU(3)$ breaking is much stronger, as the strange mass m_s is smaller than Λ_{QCD} by only a factor of about 2...3.

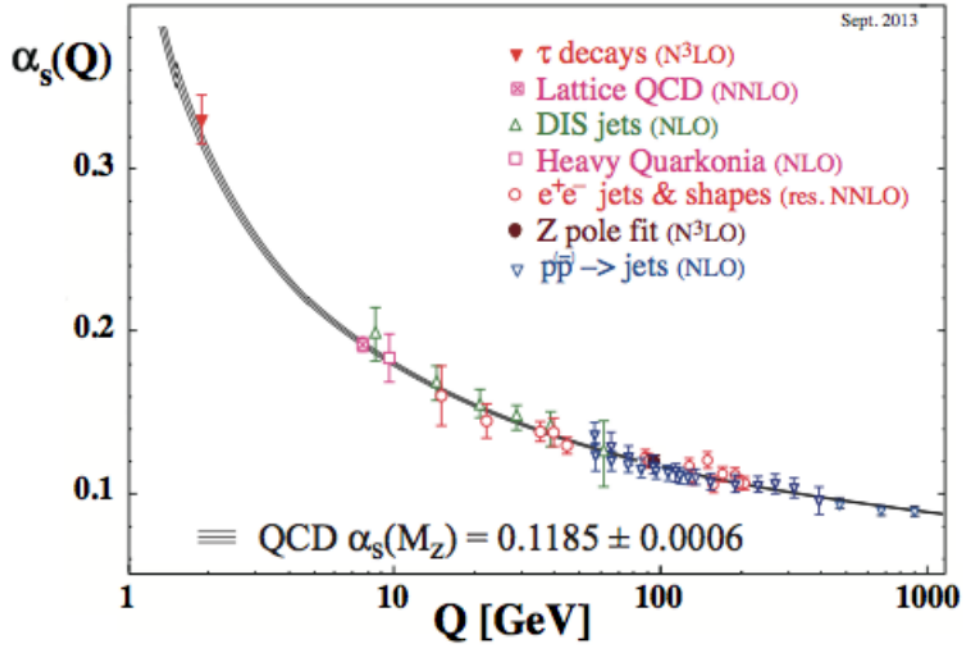


Figure 12.2: The QCD coupling as the function of energy scale: comparison of theory and experiments. NLO means 2nd, NNLO means 3rd while $N^3\text{LO}$ means 4th order theoretical calculations involved in the comparison.

From: Review of Particle Physics (K.A. Olive et al. (Particle Data Group), Chinese Phys. C38 (2014) No 9.)

Since the masses of the other quarks c, b, t are much larger than Λ_{QCD} , extended flavour symmetries such as $SU(4)$ perform much worse than $SU(3)$, and so are not very useful. In fact, due to their large mass in many cases the heavy quarks c and b can be treated in a non-relativistic approximation when describing hadrons containing them, in contrast to the light flavours u, d, s .

3. The Okubo-Zweig-Iizuka (OZI) rule

During the studies of the quark model it was observed that certain strong decays were suppressed according to an interesting rule. To understand the observation, the first step is to note that decay rates can in many cases be estimated from other similar processes. Recall that the strong interaction is supposed to be flavour symmetric, so decays proceeding in a similar way with just some flavours relabelled would have the same transition matrix element. However, the decay rate also involves summing over the final states; this leads to a dependence on the momentum space volume that is available for the decay products. This is the reason why among two otherwise similar decays, the one with a higher value for liberated energy proceeds faster. Cf. for example the decay rates of neutral kaons to two or three pions: neglecting the small correction from CP violation, the one decaying to two pions is the short-lived K_S^0 , while the one decaying to three pions is the long-lived K_L^0 .

However, it was observed that decays in which the process involved annihilation of heavy quarks proceeded much slower than estimated or were suppressed compared to other channels. An example is the vector meson ϕ which is an almost pure $s\bar{s}$ state (cf. the maximal mixing discussed in Section 7.3.6): the decay mode $\phi \rightarrow \pi^+ + \pi^0 + \pi^-$ is much suppressed compared to $\phi \rightarrow K^+ + K^-$, although phase space considerations suggest otherwise. However, it is clear that the suppressed decay eventually involves the annihilation of the s and $\bar{s}$, since the decay products are all non-strange particles. Typical diagrams for the two processes are shown in Fig. 12.3

To accommodate this observation, the following rule was introduced:

OZI (Okubo-Zweig-Iizuka) rule: *strong decay processes that can only proceed via annihilation of heavy quarks are suppressed with respect to those not involving such steps.*

In diagram 12.3 (a) there exists a cut axis at some time where the heavy quarks are absent, so they annihilated before. The reason why the leading process is drawn with three gluons instead of one or two is due to symmetry constraints. The colour singlet nature of the end state (cf. Section 12.3.4) excludes a single gluon intermediate states since that transforms as an octet $\mathbf{8}$. For two gluons there is a singlet channel according to the decomposition $\mathbf{8} \times \mathbf{8} = \mathbf{1} + \mathbf{8} + \mathbf{8} + \mathbf{10} + \mathbf{\overline{10}} + \mathbf{27}$, but the vector meson ϕ has eigenvalue -1 under charge conjugation $\mathcal{C}$, while a two-gluon state has $\mathcal{C} = +1$ (with each gluon having $\mathcal{C} = -1$), and strong interactions (as well as electromagnetism) preserve $\mathcal{C}$ symmetry. So the intermediate state must contain at least three gluons. The OZI rule itself is explained by the fact that *the three gluons created in this annihilation carry a relatively large q^2 among them*, equal to the square of the mass of ϕ (which is $m_\phi = 1020$ MeV), *so the corresponding strong couplings are suppressed by asymptotic freedom.*

In diagram 12.3 (b) no such cut axis exists, and *the gluons involved can be very soft* (their q^2 need only cover the creation of a light $u\bar{u}$ pair), so the corresponding couplings are much stronger. This explains why the left process dominates despite being at a disadvantage when it comes to available momentum space volume for the final products. We remark that it is possible to draw diagrams mediated by photons instead of gluons, but these contribute little since the electromagnetic interaction strength α is so much smaller than α_s .

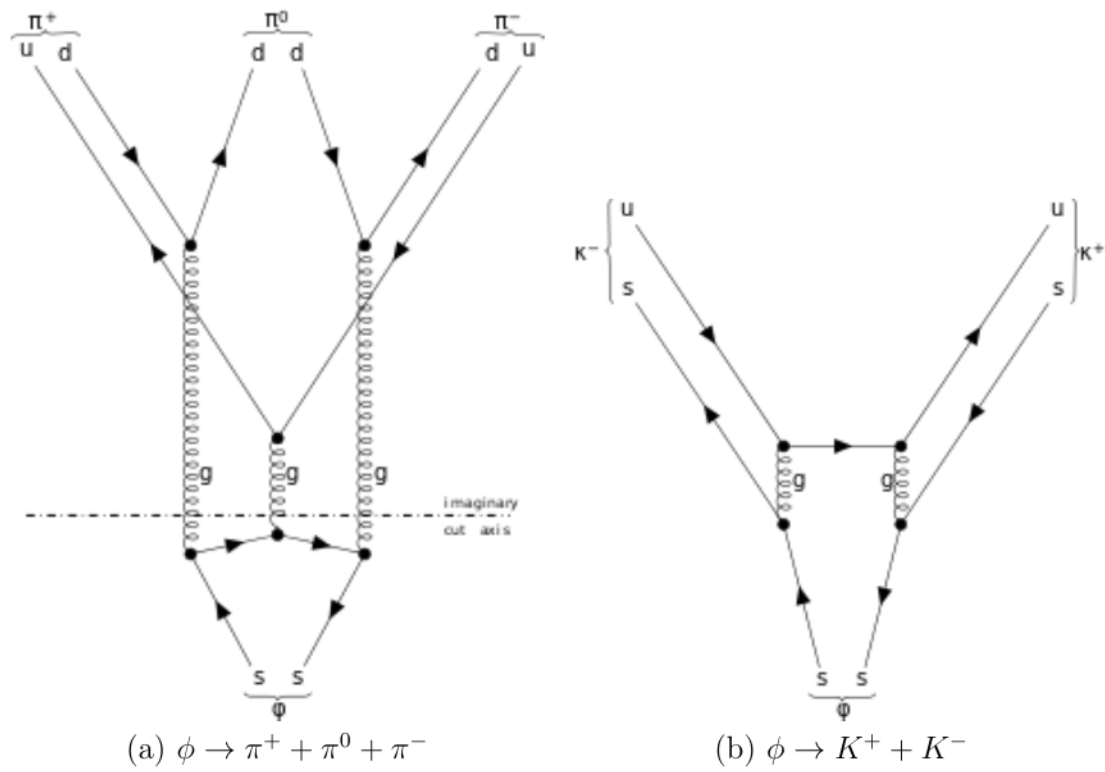


Figure 12.3: Typical Feynman diagrams contributing to ϕ decay.

The diagrams are from the book D. Griffiths: *Introduction to Elementary Particles*, Wiley-VCH, 2008, redrawn by Krishnavedala for the Wikipedia page https://en.wikipedia.org/wiki/OZI_rule.

For the case of the J/Ψ with quark content $c\bar{c}$, the effect is much more pronounced and can even be treated perturbatively, finding a suppression by a factor of order $\alpha_s(m_c)^3$, where $m_c \approx 1.5$ GeV is the charm quark mass. As a result, the lifetime of the J/Ψ is three orders of magnitude longer than expected otherwise, leading to a very narrow resonance at energy 3.1 GeV, of width 93 keV compared to the expected width of order of 100 MeV. Due the J/Ψ being too light compared to other charmed particles, its decay mainly proceeds via the analogue of the OZI suppressed diagram for $\phi \rightarrow \pi^+ + \pi^0 + \pi^-$, but with the strange quarks replaced by charm. The OZI suppression for this case is so strong that electromagnetic decays proceeding via a single intermediate photon eventually begin to compete with hadronic ones, and so the J/Ψ has significant branching ratios to leptonic end states like e^+e^- , $e^+e^-\gamma$ and $\mu^+\mu^-$.

In the case of the family of upsilon vector mesons ($\Upsilon = b\bar{b}$) a similar OZI suppression happens, and the lifetime of the three lowest lying states⁴² is again much longer than expected, with significant branching ratios to leptons.

12.3.4 Confinement

Asymptotic freedom implies that the QCD coupling constant decreases with increasing momentum transfer, which explains why quarks behave as (almost) free particles under deep inelastic scattering. This contrast with the experimental fact that no free quarks have ever been observed. So can QCD explain this phenomenon, called *quark confinement*?

Since the strong interaction gets stronger with decreasing energy, one may argue that this explains why the quarks are tightly bound into hadrons when the energy is low. Still, binding energies are finite, and by looking at the masses of the hadrons, are expected to be around a few 100 MeV per quark. So why don't we see quarks when we create collisions with much higher energies?

Confinement is clearly some strongly coupling phenomenon, so there is no hope to address it using perturbation theory. The leading framework to study non-perturbative physics of QCD directly from first principles (i.e. without involving additional assumptions) is lattice field theory, which is a numerical approach based on Monte Carlo methods borrowed from statistical physics. At the present state of art, the best simulations are able to reproduce the hadron spectrum quite well, as shown in Fig. 12.4. Albeit they do have their limitations, lattice QCD simulations demonstrate the QCD describes the physics of strong interactions not only at high energies, but also in the low-energy, non-perturbative regime.

However, can one understand confinement using some simple qualitative arguments, without resorting to not very transparent numerical simulations? This is indeed possible using the so-called hadronic (or QCD) string picture. Let us first consider interaction between electric charges. The field of a positive and a negative electric charge can be depicted as

⁴²For the state $\Upsilon(4S)$ there is enough energy to create a $d\bar{d}$ or $u\bar{u}$ pair and decay to a pair of B mesons; this provides a process which is not OZI suppressed and therefore proceeds much faster, so this state has a large width compared to the lowest three Υ mesons. The branching to these unsuppressed modes is more than 96%. Similar phenomenon happens for excited states of the J/Ψ charmonium system, where the first two states only have OZI suppressed decay modes, but from the third one there exist decays into charmed D mesons via creation of light quarks (cf. the discussion at the end of Subsection 11.3.2).

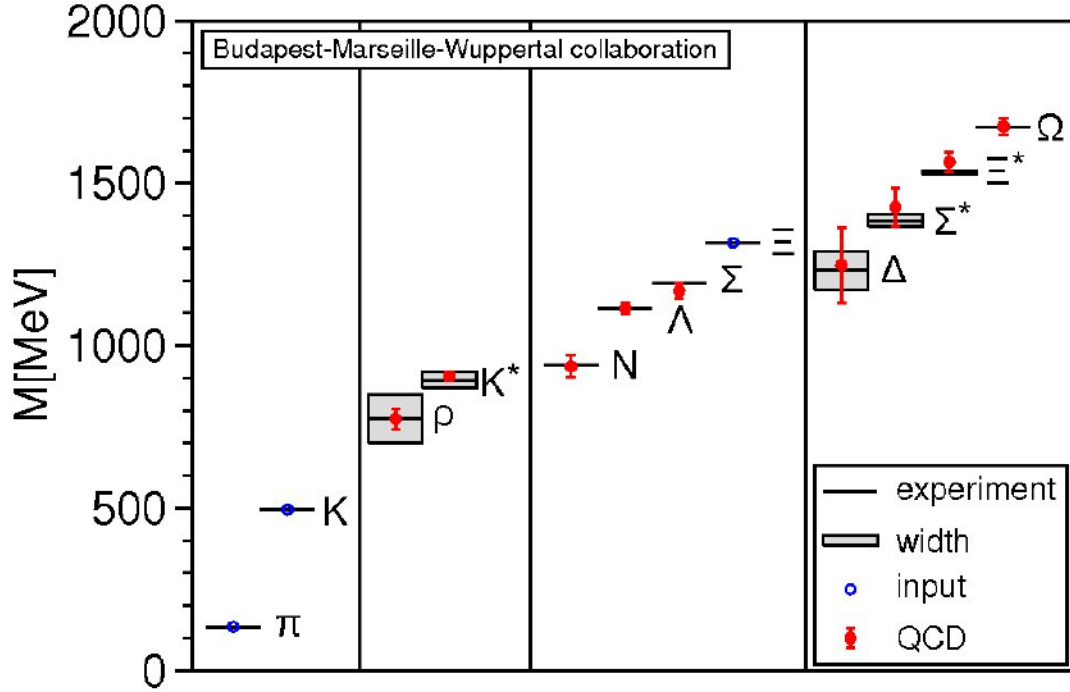
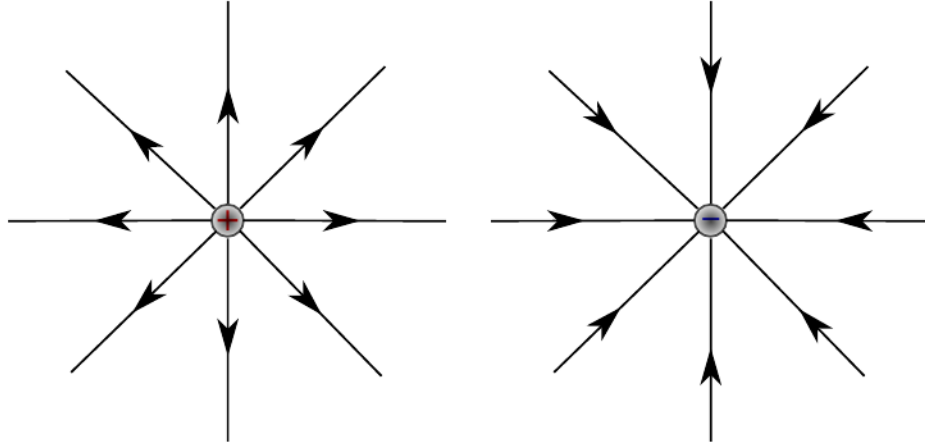
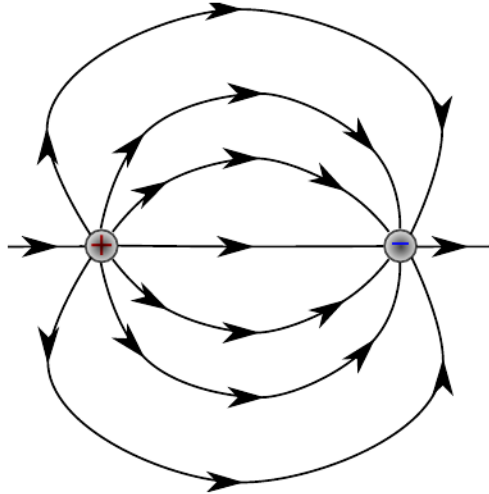


Figure 12.4: The hadronic spectrum from QCD lattice simulations, computed by the Budapest-Marseille-Wuppertal collaboration. The computation supposes three quarks with unbroken isospin symmetry and flavour $SU(3)$ breaking ($m_s > m_u = m_d$), which means there are two mass parameters plus the QCD scale. Therefore, three of the masses are fixed from experiments as inputs, and the rest are determined from the simulation.

From: S. Dürer et al: Ab initio determination of light hadron masses, *Science* **322** (2008) 1224.

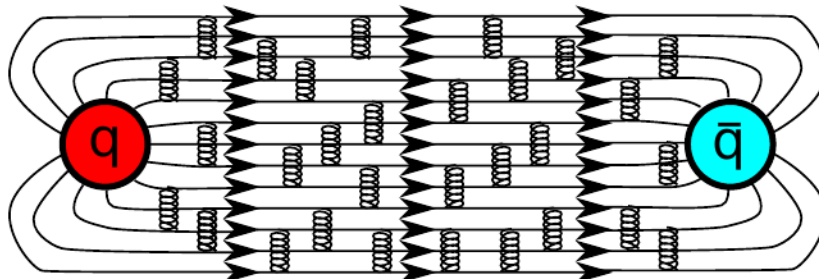


Note that since the total electric flux is constant the field strength varies as $1/R^2$, where R is the distance between the charges. When both charges are present, due to the superposition principle their fields can be added, resulting in the field



The force on any of the charges, which is given by the charge multiplied by the field strength created by the other, is also proportional to $1/R^2$, which is Coulomb's law.

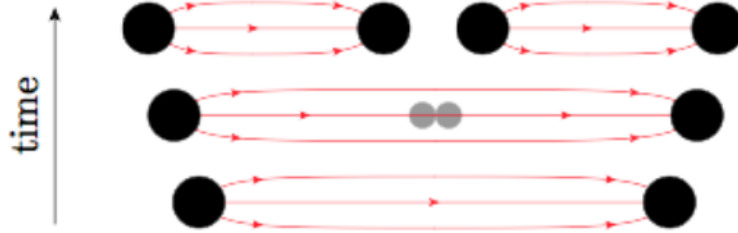
However, the QCD Lagrangian (12.84) has terms which are cubic and quartic in the gauge field, due to the quadratic term in the colour field strength (12.88). *This results in a field equation which is non-linear in the gauge fields A_μ^a , so the superposition principle is invalid.* A consequence is that *the field lines are interacting*, so field lines connecting two coloured objects do not correspond to a flux uniformly spreading with distance. The interaction between the field lines turns out to be attractive, leading to *colour flux tubes* as depicted for the case of a quark and an antiquark:



As a result, the *flux per unit area and therefore the force stays constant with distance*. This leads to a quark-antiquark potential that is linear for large distances. A phenomenological potential function is

$$V(R) = \frac{\alpha}{R} + \beta R \quad (12.109)$$

where the short distance behaviour is modelled by a Coulomb term, while the long-distance asymptotic behaviour is set by the linear one. These colour tubes behave as strings: stretching them increases their energy, linearly with their length. Since this is a quantum theory, when the energy stored in a string reaches the threshold necessary to create a quark-antiquark pair, it becomes unstable: the string snaps and a new pair is created. So just pulling apart a meson does not lead to freeing the quarks as illustrated below:



From: Review of Particle Physics (K.A. Olive et al. (Particle Data Group), Chinese Phys. C38 (2014) No 9.)

In fact, the above argument shows that *no coloured object would be stable*, since any colour source would act as the endpoint of a QCD string, and when trying to free it from other coloured objects, the string would snap until all excess energy is converted into hadrons. Therefore *all observable QCD states are colour singlets*. Since a quark transforms as $\mathbf{3}$, while an anti-quark as $\bar{\mathbf{3}}$, the simplest uncoloured objects can be realised as follows:

$$\begin{array}{lll} \text{mesons:} & q\bar{q} & \mathbf{3} \times \bar{\mathbf{3}} = \mathbf{1} + \mathbf{8} \\ \text{baryons} & qqq & \mathbf{3} \times \mathbf{3} \times \mathbf{3} = \mathbf{1} + \mathbf{8} + \mathbf{8} + \mathbf{10} \\ \text{antibaryons:} & \bar{q}\bar{q}\bar{q} & \bar{\mathbf{3}} \times \bar{\mathbf{3}} \times \bar{\mathbf{3}} = \mathbf{1} + \mathbf{8} + \mathbf{8} + \mathbf{10} \end{array} \quad (12.110)$$

where the group composition rules show that for such combinations it is possible to find colour singlet states.

Exercise 12.6. Which of the following states are possible: qq , $qq\bar{q}\bar{q}$, $qqqq\bar{q}$?

Solution. It is only necessary to find the decomposition of these states into irreducible colour components:

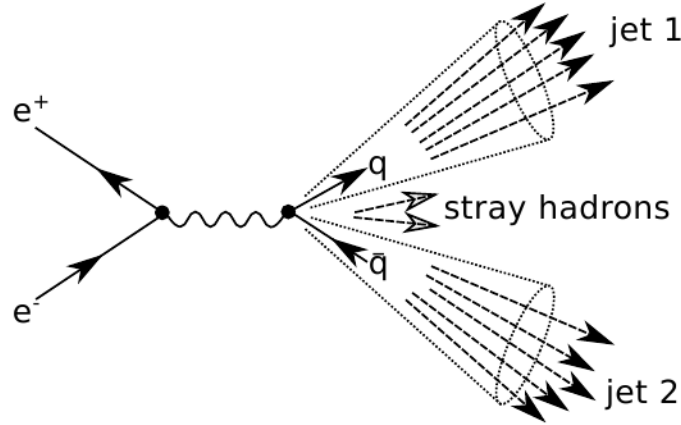
$$\begin{aligned} qq &: \mathbf{3} \times \mathbf{3} = \bar{\mathbf{3}} + \mathbf{6} \\ qq\bar{q}\bar{q} &: \mathbf{3} \times \bar{\mathbf{3}} \times \mathbf{3} \times \bar{\mathbf{3}} = (\mathbf{1} + \mathbf{8}) \times (\mathbf{1} + \mathbf{8}) \\ &= \mathbf{1} + \mathbf{8} + \mathbf{8} + \mathbf{1} + \mathbf{8} + \mathbf{8} + \mathbf{10} + \bar{\mathbf{10}} + \mathbf{27} \end{aligned} \quad (12.111)$$

so qq is not possible, while $qq\bar{q}\bar{q}$ is and would be a so-called tetraquark state. It is also easy to see that $qqqq\bar{q}$ is possible, since it is $(\mathbf{1} + \mathbf{8} + \mathbf{8} + \mathbf{10}) \times (\mathbf{1} + \mathbf{8})$ contains at least three singlets from

the 1×1 and the two 8×8 products; this would be a pentaquark. In fact, some more advanced group theory shows that the sufficient and necessary condition for a state composed of N_q quarks and $N_{\bar{q}}$ anti-quarks to contain at least one colour singlet component is that $N_q - N_{\bar{q}}$ divisible by three. $\square$

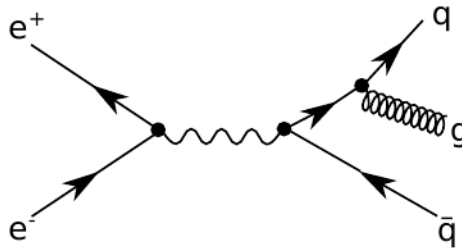
12.4 Jet physics and evidence for gluons

To test QCD, it is also desirable to have evidence for gluons. This leads us to consider the physics of so-called jets which are typically produced in high-energy collisions with hadronic end states. Let us consider the collision $e^+e^- \rightarrow \text{hadrons}$. We already know from Section 12.2 that at leading order the observed data are well-explained by the creation of a quark-antiquark pair. After the creation of the quark pair which involves high momentum transfer and is therefore called a hard process, two hadron showers produced by so-called hadronisation involving soft physics, that can be modelled e.g. using the string picture⁴³ Because of the large momentum of the initial quarks, the generated showers are confined to some narrow cones forming so-called *jets*:



In addition, due to the break-up of strings connecting the two lead quarks some “stray” hadrons are also produced in the space between the two jets as depicted above.

However, due to the QCD interaction it is possible to create *more than two jets*. For example, a hard process leading to a $q\bar{q}g$ intermediate state is



In this process, if the gluon is hard enough and is emitted at a large enough angle, it leads to a formation of a distinct third jet, with a probability that is suppressed with respect to the two-jet

⁴³In fact, detailed predictions for hadronic end states are derived using Monte-Carlo event generators, which use either string or cluster based models for the formation of hadrons.

process by α_s . Soft and almost collinear gluon emissions form part of the quark jet; the distinction must be defined by a so-called jet algorithm, which gives a prescription of how to group particles into jets. A gluonic jet can be distinguished from other processes (e.g. involving a high-energy photon) by the existence of stray hadrons between it and the quark jets. *Processes with three and more jets have indeed been observed, with the characteristics matching those predicted by QCD.*

Many characteristics of jets are determined by the initial stages of evolution which happen at energy scales above 1 GeV, which makes perturbative calculations of these details possible. In particular, the strong coupling measured from jet physics matches measurements from other process very well as shown in Fig. [12.2](#). The angular distribution of the jets in three jet events depends on gluon spin, and can be used to confirm that the gluon has spin 1, as expected for a vector boson. It is also possible to distinguish jets according to whether the initial particle was a light quark (u, d, s), heavy quark (c, b) or a gluon, with a quite good efficiency.

13 Gauge theory of weak interactions

13.1 Intermediate vector bosons

13.1.1 Problems with the four-fermion interaction

So far, we described weak interactions using the four-fermion interaction (11.37):

$$\mathcal{L}_{CC} = \frac{G_F}{\sqrt{2}} (J_\alpha^{C\dagger} J^{C\alpha} + \text{h.c.}) \quad (13.1)$$

involving the charged current

$$J_\alpha^C = \sum_{l=e,\mu,\tau} \bar{\nu}_l \gamma_\alpha (1 - \gamma_5) l + \sum_{i=1}^3 \bar{u}_i \gamma_\alpha (1 - \gamma_5) d'_i \quad (13.2)$$

Here we flipped the current with respect to (11.55) which amounts to Hermitian conjugation, and also introduced a notation for the three generations of quarks as follows

$$\begin{aligned} u_1 &= u & u_2 &= c & u_3 &= t \\ d_1 &= d & d_2 &= s & d_3 &= b \end{aligned} \quad (13.3)$$

and the flavour mixings

$$\begin{aligned} d'_i &= \sum_{j=1}^3 (V_{CKM})_{ij} d_j \\ i &= 1, 2, 3 \\ \nu_l &= \sum_{j=1}^3 (U_{PMNS})_{lj} \nu_j \\ l &= e, \mu, \tau \end{aligned} \quad (13.4)$$

with $\nu_{1,2,3}$ denoting the neutrino mass eigenstates, while ν_l denoting the neutrino flavours selected according to the lepton $l = e, \mu, \tau$ they are coupled with. The quark fields d_j are the ones entering the QCD Lagrangian (12.84), while d'_j are the weak interaction eigenstates, i.e. the combinations coupled to the u -type quarks u, c, t by weak interactions.

While this theory successfully describes weak decays of hadrons and leptons, and accommodates the observed flavour mixing, parity and CP violation, it has serious problems which prevent it from giving a final and complete description of weak interactions. All of that is related to the dimension of G_F ; recall from Section 11.2 that

$$G_F \approx 10^{-5} \frac{1}{m_p^2} \quad (13.5)$$

so its dimension is

$$[G_F] = \text{eV}^{-2} \quad (13.6)$$

Consider now high energy scattering processes with two incoming particles of 4-momenta p_1 and p_2 ; the characteristic energy scale is given by $s = (p_1 + p_2)^2$, which is just the square of total energy measured in the center-of-mass frame. By dimensional analysis, the effective interaction parameter is given by

$$G_F s \quad (13.7)$$

This results in two, eventually closely related problems:

1. Violation of unitarity

The scattering matrix introduced in (10.64) is unitary by definition, which is related to *conservation of probability*. Namely, unitarity implies that for a fixed initial state i summing over all final states f gives unity:

$$\sum_f |S_{fi}|^2 = 1 \quad (13.8)$$

This imposes upper bounds on the size of matrix elements, and ultimately on the scattering cross sections. In general, cross sections at high energy must decrease as s^{-1} , while simple dimensional analysis using $[\sigma] = \text{eV}^{-2}$ results in

$$\sigma = G_F^2 s \quad (13.9)$$

since the lowest order interaction amplitude is proportional to G_F , while the transitional probabilities and so the cross sections are proportional to the square of the amplitudes. For example, direct computation of the $J = 1$ partial wave cross section for the reaction

$$\nu_e + e^- \longrightarrow \nu_\mu + \mu^- \quad (13.10)$$

gives a result behaving as (13.9), while unitarity bounds the amplitude as

$$\sigma_{J=1} < s^{-1} \quad (13.11)$$

So for energies above

$$\sqrt{s} \sim G_F^{-1/2} \approx 300 \text{ GeV} \quad (13.12)$$

unitarity is violated. One could hope that this would be rectified by higher-order perturbative corrections, but there is a serious problem with the perturbation expansion in the four-fermion model.

2. Cut-off dependence

As discussed in Subsection 12.3.2, *higher order perturbative corrections in quantum field theories are generally only defined by introducing a high energy cut-off Λ* . This is due to the presence of sums over intermediate states in the perturbative expansion (10.66), which are usually divergent without truncation. The cut-off dependence can be absorbed in redefining the parameters of the theory, such as couplings and masses (cf. (12.92) and (12.93)). This process is called *renormalisation*, and its physical content is that far below the cut-off scale, *physics can be parametrised in terms of effective low-energy parameters which are to be determined from experiments*. For quantum electrodynamics and quantum chromodynamics, this results in a finite set of parameters: the fine structure constant, the QCD scale Λ_{QCD} and the fermion masses. After their experimental determination, other observational results can be predicted in terms of these parameters.

However, for the four-fermion interaction every new order in G_F brings an additional power of Λ^2 , so a new power of cut-off dependence appears. This means that at every new order, new parameters must be introduced to absorb the cut-off dependence. This spoils the predictivity of the theory since the more precisely one computes, the higher order of perturbation theory and consequently the larger number of experimentally determined parameters are needed. The net result is that *the theory ceases to give a meaningful description of physics at around the scale (13.12)*.

From the above discussion it is also clear that these problems are not present for theories with coupling constants that are either dimensionless or have dimensions of a positive power of energy, since for a coupling constant g with dimension $[g] = \text{eV}^d$ one expects

$$\sigma \sim \frac{g^2}{s^{1+d}} \quad (13.13)$$

or, in terms of the cut-off dependence, every new order introduces a factor Λ^{-d} . Therefore as long as $d \geq 0$ the number of parameters introduced by the cut-off dependence does not grow. Such theories are called *renormalisable*, while the four-fermion model is an example of a *non-renormalisable* theory.

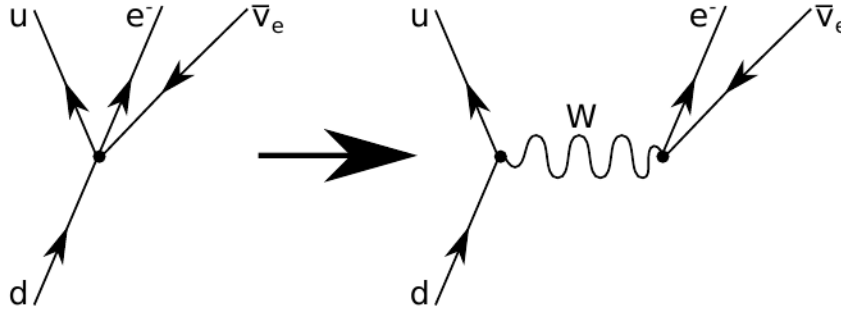
The preceding discussion is *only valid in perturbation theory*; non-perturbatively there could be additional problems such as the Landau singularity discussed in Subsection [12.3.2](#) which leads to an upper energy bound on the field theory (cf. the notion of *triviality*).

13.1.2 The charged weak intermediate vector bosons

To remedy the situation, one can apply the lessons learnt from QED and QCD, and *introduce a particle to mediate weak interactions*. Since charged currents involve charge transfer, the intermediate particle must be charged, and since the current is a vector, its field must also carry a Lorentz index. So we modify the weak interaction Lagrangian ([13.1](#)) to

$$\mathcal{L}_{IVB} = g_W (J_\mu^C W^\mu + \text{h.c.}) \quad (13.14)$$

This corresponds to resolving processes as in the following example corresponding to neutron β -decay



so the conversion of the d quark into a u quark is accompanied by the emission of a W^- excitation which then decays into an electron and an anti electron neutrino. This also demonstrates that in contrast to the photon, W must be charged. Unlike the photon, it must also have a non-zero mass since weak interactions have a very short range. Indeed, since transitions happen at the quark level, their range must be smaller than the proton/neutron radius which is of order GeV^{-1} as determined from their elastic scattering form factors, so M_W must be larger than GeV . The Lagrangian for the W field can be written down in analogy to the photon case, however it is necessary to account for charge (i.e. the field must be complex) and for its mass:

$$\mathcal{L}_W = -\frac{1}{2}(\partial_\mu W_\nu^\dagger - \partial_\nu W_\mu^\dagger)(\partial^\mu W^\nu - \partial^\nu W^\mu) + M_W^2 W_\mu^\dagger W^\mu \quad (13.15)$$

Neglecting interactions, the field equations for W are

$$\square W^\mu - \partial^\mu (\partial_\lambda W^\lambda) + M_W^2 W^\mu = 0 \quad (13.16)$$

Exercise 13.1. Derive the field equations (13.16).

Solution. Using the Euler-Lagrange formalism with respect to $W_\mu^\dagger$

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial W_\mu^\dagger} &= M_W^2 W^\mu \\ \frac{\partial \mathcal{L}}{\partial (\partial_\nu W_\mu^\dagger)} &= \partial^\mu W^\nu - \partial^\nu W^\mu \end{aligned} \quad (13.17)$$

so

$$\frac{\partial \mathcal{L}}{\partial W_\mu^\dagger} - \partial_\nu \frac{\partial \mathcal{L}}{\partial (\partial_\nu W_\mu^\dagger)} = M_W^2 W^\mu - \partial_\nu (\partial^\mu W^\nu - \partial^\nu W^\mu) \quad (13.18)$$

which is just (13.16). Varying the action with respect to W_μ gives the complex conjugate of the above equation of motion. $\square$

Computing the divergence of (13.16) leads to

$$\partial_\mu W^\mu = 0 \quad (13.19)$$

and so

$$(\square + M_W^2) W^\mu = 0 \quad (13.20)$$

As usual, this equation can be solved in terms of plane waves

$$W_\mu = \int \frac{d^3 k}{(2\pi)^{3/2}} \frac{1}{\sqrt{2E_k}} \sum_{\alpha=1}^3 \left(a_\alpha(\vec{k}) \epsilon_\mu^{(\alpha)}(\vec{k}) e^{-ik \cdot x} + b_\alpha^\dagger(\vec{k}) \epsilon_\mu^{(\alpha)}(\vec{k})^* e^{ik \cdot x} \right) \quad (13.21)$$

where the four-vector k^μ is $(E_k, \vec{k})$ with $E_k = \sqrt{\vec{k}^2 + M_W^2}$, $\epsilon_\mu^{(\alpha)}(k)$ are three independent vectors satisfying

$$k^\mu \epsilon_\mu^{(\alpha)}(k) = 0 \quad (13.22)$$

and $a_\alpha(\vec{k})$, $b_\alpha^\dagger(\vec{k})$ are complex coefficients, which after quantisation become the annihilation operators of particles and the creation operators of anti-particles, in analogy to the complex Klein-Gordon field discussed in Subsections 10.1 and 10.2. A convenient basis for the three independent solution corresponds to

$$\begin{aligned} \text{longitudinal polarisation} &: \epsilon^{(1)}(k) = \mathcal{N}_0(\epsilon_0, \vec{k}) \\ \text{transversal polarisations} &: \epsilon^{(2)}(k) = (0, \vec{e}_2) \\ &\epsilon^{(3)}(k) = (0, \vec{e}_3) \end{aligned} \quad (13.23)$$

where the parameters are determined by (13.22) and by orthonormality

$$\epsilon_\mu^{(\alpha)}(k) \epsilon^{(\beta)\mu}(k)^* = -\delta^{\alpha\beta} \quad (13.24)$$

(the Lorentz norms can only be normalised to -1 since k^μ is time-like, and as result of (13.22), the polarisation vectors are space-like). The result of non-zero mass is therefore the existence of physical longitudinal polarisation; zero mass vector bosons such as the photon can only have transversal polarisations⁴⁴. The transverse vectors are simply orthogonal to $\vec{k}$

$$\vec{\epsilon}_i \vec{k} = 0 \quad i = 2, 3$$

and orthonormal in the usual sense

$$\vec{\epsilon}_i \vec{\epsilon}_j^* = \delta_{ij} \quad i, j = 2, 3$$

Exercise 13.2. Compute ϵ_0 and $\mathcal{N}_0$.

Solution. ϵ_0 can be obtained from

$$k \cdot \epsilon^{(1)}(k) = \mathcal{N}_0 \left(\epsilon_0 E_k - \vec{k}^2 \right) = 0 \quad (13.25)$$

so

$$\epsilon_0 = \frac{\vec{k}^2}{E_k} \quad (13.26)$$

while $\mathcal{N}_0$ can be obtained from

$$\epsilon^{(1)}(k) \cdot \epsilon^{(1)}(k)^* = |\mathcal{N}_0|^2 \left(\frac{k^4}{E_k^2} - \vec{k}^2 \right) = -|\mathcal{N}_0|^2 \frac{\vec{k}^2 M_W^2}{\vec{k}^2 + M_W^2} = -1 \quad (13.27)$$

leading to

$$|\mathcal{N}_0|^2 = \frac{\vec{k}^2 + M_W^2}{\vec{k}^2 M_W^2} \quad (13.28)$$

(the phase of $\mathcal{N}_0$ is physically irrelevant and can be chosen arbitrarily). $\square$

Canonical quantisation leads to

$$\begin{aligned} \left[a_\alpha(\vec{k}), a_{\alpha'}^\dagger(\vec{k}') \right] &= \delta_{\alpha\alpha'} \delta^{(3)}(\vec{k} - \vec{k}') \\ \left[b_\alpha(\vec{k}), b_{\alpha'}^\dagger(\vec{k}') \right] &= \delta_{\alpha\alpha'} \delta^{(3)}(\vec{k} - \vec{k}') \end{aligned} \quad (13.29)$$

with all other commutators vanishing, and computing the interactions using perturbation theory leads to the relation

$$\frac{G_F}{\sqrt{2}} = \frac{g_W^2}{M_W^2} \quad (13.30)$$

Supposing that the interaction strength is not very much different from 1, this leads to M_W being order of 100 GeV.

⁴⁴We already know this from classical electromagnetism; a short explanation is that the longitudinal mode of a zero mass vector particle turns out to be a pure gauge component, and is thus not physical.

13.2 Spontaneous symmetry breaking

Introduction of intermediate vector bosons resolves the immediate problems with the fermion-fermion scattering. However, processes involving longitudinal components of the $W^\pm$ bosons are still problematic because the longitudinal polarisation vector (computed in Exercise 13.2) has components that grow linearly with momentum and bring back the same problems as before.

Vector bosons corresponding to a non-Abelian gauge symmetry would have no such problems (since longitudinal polarisation is missing), but gauge symmetry forbids the introduction of mass terms since they are not gauge invariant. It seems we have arrived at an impasse: there seems to be no way to have a theory of weak interactions that is valid for energy scales over a few hundred GeV. Fortunately, this problem can be solved using the so-called Higgs mechanism. To understand its origin it is necessary to learn about spontaneous symmetry breaking first.

13.2.1 Spontaneous symmetry breaking and Goldstone's theorem

Let us start with the simple example of a complex scalar field Φ with a $U(1)$ invariant self-interaction, governed by the Lagrangian density

$$\mathcal{L}_\Phi = \partial_\nu \Phi^* \partial^\nu \Phi - V(\Phi^* \Phi) \quad (13.31)$$

The Hamiltonian of such a scalar field is given by

$$H = \int d^3x \left(\partial_t \Phi \partial_t \Phi^* + \vec{\nabla} \Phi \vec{\nabla} \Phi^* + V(\Phi^* \Phi) \right) \quad (13.32)$$

The classical vacuum corresponds to a configuration for which H is minimal, which implies

$$\partial_\mu \Phi = 0 \text{ and } V = \min. \quad (13.33)$$

Let us choose

$$V = -\mu^2 |\Phi|^2 + \frac{\lambda}{4} |\Phi|^4 \quad (13.34)$$

with $\lambda > 0$

which is minimal when

$$|\Phi| = \Phi_0 = \sqrt{\frac{2\mu^2}{\lambda}} \quad (13.35)$$

So the classical vacua are degenerate; the independent solutions are

$$\begin{aligned} \Phi &= \frac{v}{\sqrt{2}} e^{i\vartheta} \\ v &= \frac{2\mu}{\sqrt{\lambda}} \end{aligned} \quad (13.36)$$

and are related by the $U(1)$ symmetry

$$\Phi \longrightarrow e^{-i\alpha} \Phi \quad (13.37)$$

When quantising the system, a choice of the vacuum must be made first; all possible choices are equivalent by $U(1)$ symmetry, so we choose to expand the field around $\Phi = \Phi_0$ (i.e. $\vartheta = 0$):

$$\Phi = \frac{1}{\sqrt{2}} (v + \rho(x)) e^{i\Theta(x)/v} \quad (13.38)$$

where ρ describes the radial, while Θ corresponds to the angular degrees of freedom in Φ space. This can be interpreted as a non-zero expectation value of the scalar field in the vacuum:

$$\langle \Phi(x) \rangle = \frac{v}{\sqrt{2}} \quad (13.39)$$

its independence of x reflects the translational invariance of the vacuum state.

Exercise 13.3. Express the Lagrangian (13.31) in terms of ρ and Θ .

Solution. One can compute

$$\begin{aligned} \partial_\mu \Phi &= e^{i\Theta(x)/v} \frac{1}{\sqrt{2}} \left(\partial_\mu \rho + \frac{i}{v} (v + \rho(x)) \partial_\mu \Theta \right) \\ V &= -\frac{\mu^4}{\lambda} + \mu^2 \rho^2 + \frac{\sqrt{\mu^2 \lambda}}{2} \rho^3 + \frac{\lambda}{16} \rho^4 \end{aligned} \quad (13.40)$$

and substituting into (13.31) gives

$$\begin{aligned} \mathcal{L}_\Phi &= \frac{1}{2} \left(\partial_\mu \rho \partial^\mu \rho + \left(1 + \frac{2\rho(x)}{v} + \frac{\rho^2}{v^2} \right) \partial_\mu \Theta \partial^\mu \Theta \right) \\ &\quad + \frac{\mu^4}{\lambda} - \mu^2 \rho^2 - \frac{\sqrt{\mu^2 \lambda}}{2} \rho^3 - \frac{\lambda}{16} \rho^4 \end{aligned} \quad (13.41)$$

We note that the term $-\mu^4/\lambda$ corresponds to a shift in the vacuum energy density, and can be dropped. $\square$

Separating the quadratic, i.e. free field part

$$\mathcal{L}_{\text{free}} = \frac{1}{2} \partial_\mu \rho \partial^\mu \rho + \frac{1}{2} \partial_\mu \Theta \partial^\mu \Theta - \mu^2 \rho^2 \quad (13.42)$$

shows that ρ describes excitations with mass $\sqrt{2}\mu$, while Θ is a real scalar field with zero mass. Furthermore, the interactions only depend on the derivative of Θ , so the Lagrangian is invariant under the shift $\Theta(x) \rightarrow \Theta(x) - \alpha$, which is just the original $U(1)$ invariance. This makes it manifest that the Lagrangian is independent of the choice of the vacuum, i.e. the angle ϑ .

The zero mass excitations are called *Goldstone bosons*. Despite the classical nature of the computation above, *their existence is eventually a general conclusion valid also in the quantum theory*, due to the

Theorem 13.4. (*Goldstone's theorem*) *When the ground state of a field theory is only invariant under a proper subgroup H of a continuous symmetry group G , there exist zero mass scalar boson excitations, in one-to-one correspondence with the generators which correspond to broken symmetries (i.e. those not leaving the vacuum invariant). Due to the correspondence, their number is exactly $\dim G - \dim H$.*

Without discussing the detailed proof here, we note that the validity of the theorem is clear from the following heuristic reasoning. Acting on the vacuum with a broken generator produces a physical excitation due to the non-invariance of the vacuum, which has zero energy since the

transformation is a symmetry of the Hamiltonian. Such an excitation, due to being global (i.e. position independent) has wave number $\vec{k} = 0$, so it satisfies

$$E^2 - \vec{k}^2 = 0 \quad (13.43)$$

and therefore has zero mass. In addition, due to rotational invariance of the vacuum, the excitation has spin zero, i.e. it corresponds to a massless scalar boson particle (cf. our discussion between the relation of zero spin to rotational invariance in the rest frame in Subsection 4.3.1).

A well-known example of spontaneous symmetry breaking occurs in ferromagnetic materials, where the existence of non-zero equilibrium magnetisation below the Curie point breaks rotational symmetry. The spontaneous magnetisation is called an *order parameter*; in the above example this role is played by the scalar field Φ according to (13.39).

13.2.2 Chiral symmetry breaking in strong interactions

Another example, this time in particle physics, is related to the scalar meson octet composed of the pions, the kaons and the η meson (cf. Subsection 7.3.3). Keeping only three flavours and in the limit of zero quark masses $m_u = m_d = m_s$, the strong interaction Lagrangian is invariant under separate $SU(3) \times U(1)$ transformations of the left and right handed quark fields:

$$\begin{aligned} SU(3)_R : \quad q_{iR} &\longrightarrow U_{ij}^{(R)} q_{jR} \\ U(1)_R : \quad q_{iR} &\longrightarrow e^{-i\alpha_R} q_{iR} \\ SU(3)_L : \quad q_{iL} &\longrightarrow U_{ij}^{(L)} q_{jL} \\ U(1)_L : \quad q_{iL} &\longrightarrow e^{-i\alpha_L} q_{iL} \end{aligned} \quad (13.44)$$

called a *chiral symmetry* (i, j are flavour indices 1, 2, 3 corresponding to u, d, s). The combination $SU(3)_R \times SU(3)_L$ is assumed to be spontaneously broken by strong interactions to the usual flavour $SU(3)$ which consists of the “diagonal” transformations $U^{(R)} = U^{(L)}$. This breaks eight of the sixteen $SU(3)_R \times SU(3)_L$ symmetries, which implies there should be eight Goldstone bosons. These are exactly the members of the scalar meson octet; their odd parity corresponds to the chiral nature of the transformations broken, which are the ones with $U^{(R)} = U^{(L)-1}$. The reason why the octet masses are not exactly zero is that non-zero quark masses explicitly break the chiral part of $SU(3)_R \times SU(3)_L$, so the symmetry is not exact to begin with. However, this pattern of *chiral symmetry breaking* explains why the masses of the scalar mesons, especially the pions are so small compared to other hadron masses (remember that $m_s \gg m_{u,d}$, so explicit symmetry breaking of strange particles such as the pions is much larger than for non-strange ones). Due to their non-zero mass and relation to chiral symmetry breaking, the scalar meson octet particles are called *pseudo-Goldstone bosons*; chiral symmetry breaking is a part of standard strong interaction phenomenology.

Spontaneous breaking of chiral symmetry corresponds to the following vacuum expectation value (order parameter)

$$\langle \bar{q}_i q_j \rangle = \delta_{ij} \Lambda_{\text{chiral}}^3 \quad (13.45)$$

where the quantity Λ_{chiral} has dimensions of energy. Such an expectation value is Lorentz invariant, but it is not invariant under $SU(3)_R \times SU(3)_L$ since

$$\bar{q}_i q_j = \bar{q}_{Ri} q_{Lj} + \bar{q}_{Li} q_{Rj} \quad (13.46)$$

so a non-zero expectation value like (13.45) is only left invariant by diagonal transformations $U^{(R)} = U^{(L)}$. Non-perturbative Monte-Carlo studies show that chiral symmetry breaking indeed happens in quantum chromodynamics.

Switching on non-zero quark masses leads to *explicit chiral symmetry breaking* and gives masses for the pseudo-Goldstone bosons that are linear in the symmetry breaking parameter given by the quark masses (cf. Exercise 13.6 below); the small value of the scalar masses is explained by the quark masses being small compared to the strong interaction energy scale, which holds very well for u and d , and much less for s .

The diagonal $U(1)$ symmetry $\alpha_R = \alpha_L$ is just baryon number; the so-called axial $U(1)$ corresponding to $\alpha_R = -\alpha_L$ is broken by quantum effects (not spontaneously, so it does not correspond to any vacuum degeneracy), which in fact explains why the $SU(3)$ singlet η' is so much heavier than the octet particle η , but we do not enter into details here.

Exercise 13.5. A toy model of chiral symmetry breaking is given by the *linear sigma model*

$$\mathcal{L}_\sigma = \frac{1}{2} \partial_\mu \vec{\pi} \partial^\mu \vec{\pi} + \frac{1}{2} \partial_\mu \sigma \partial^\mu \sigma - \frac{\lambda}{8} (\vec{\pi}^2 + \sigma^2 - f_\pi^2)^2 \quad (13.47)$$

where $\vec{\pi} = (\pi_1, \pi_2, \pi_3)$ are three scalar fields that model the pions (note that their indices do not correspond to space components!), while σ is a scalar field that does not correspond to any known particle. This model is eventually invariant under $SO(4)$ rotations of the four-component vector of fields

$$\begin{pmatrix} \sigma \\ \pi_1 \\ \pi_2 \\ \pi_3 \end{pmatrix} \longrightarrow \begin{pmatrix} & & & \\ & \mathcal{O} & & \\ & & & \\ & & & \end{pmatrix} \begin{pmatrix} \sigma \\ \pi_1 \\ \pi_2 \\ \pi_3 \end{pmatrix} \quad (13.48)$$

with $\mathcal{O}^T \mathcal{O} = \mathbf{1}_{4 \times 4}$. The full symmetry group can be decomposed as $SO(4) \simeq SO(3)_1 \times SO(3)_2$. $SO(3)_1$ is just rotations acting on $\vec{\pi}$, leaving σ unchanged.

Show that the remaining symmetry transformations are given in infinitesimal form as

$$\begin{aligned} \vec{\pi} &\rightarrow \vec{\pi} + \vec{\omega} \sigma \\ \sigma &\rightarrow \sigma - \vec{\omega} \vec{\pi} \end{aligned} \quad (13.49)$$

with $\vec{\omega}$ being the vector of three infinitesimal parameters of $SO(3)_2$. Compute the Noether currents corresponding to this symmetry.

Solution. One can write an infinitesimal 4-dimensional rotation matrix as

$$\mathcal{O}_{ab} = \delta_{ab} + \omega_{ab} \quad a, b = 0, 1, 2, 3 \quad (13.50)$$

with 0 corresponding to the direction σ , and $\mathcal{O} \mathcal{O}^T = \mathbf{1}$ implies that $\omega_{ab} = -\omega_{ba}$. $SO(3)_1$ is generated by transformations for which $\omega_{0a} = 0$, while $SO(3)_2$ is generated by matrices satisfying

$$\begin{aligned} \omega_{ij} &= 0 \\ \omega_{i0} &= -\omega_{0i} = \omega_i \end{aligned}$$

where $i, j = 1, 2, 3$ and ω_i makes up the vector $\vec{\omega}$ in (13.49). One can then show $\sigma^2 + \vec{\pi}^2$ is an invariant under these transformations

$$(\sigma - \vec{\omega} \vec{\pi})^2 + (\vec{\pi} + \vec{\omega} \sigma)^2 = \sigma^2 + \vec{\pi}^2 + O(\omega^2) \quad (13.51)$$

Similar computation shows the invariance of the kinetic term $\frac{1}{2}\partial_\mu\vec{\pi}\partial^\mu\vec{\pi} + \frac{1}{2}\partial_\mu\sigma\partial^\mu\sigma$, so (13.47) is indeed invariant.

To compute the Noether current, we notice that, using the notations of Section 8.3, the field variations corresponding to ω_k , $k = 1, 2, 3$ are

$$\Delta_k\pi_j = \delta_{kj}\sigma \quad \Delta_k\sigma = -\pi_k$$

so the three currents are

$$\begin{aligned} J_k^\mu &= \frac{\partial\mathcal{L}}{\partial(\partial_\mu\pi_j)}\Delta_k\pi_j + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\sigma)}\Delta_k\sigma \\ &= \partial^\mu\pi_j\delta_{kj}\sigma + \partial^\mu\sigma(-\pi_k) \\ &= \sigma\partial^\mu\pi_k - \pi_k\partial^\mu\sigma \end{aligned}$$

□

Exercise 13.6. Consider the effect of spontaneous and explicit symmetry breaking in the linear sigma model (13.47).

1. Taking the vacuum configuration as $\sigma = f_\pi$, show that expanding around the vacuum one obtains three massless and one massive scalar fields! This breaks the $SO(3)_2$ symmetry, but leaves $SO(3)_1$ unaffected.
2. Adding a small explicit $SO(3)_2$ breaking term

$$\mathcal{L}'_\sigma = \mathcal{L}_\sigma + \epsilon\sigma \tag{13.52}$$

show that it results in a mass term for the pseudo-Goldstone boson fields $\vec{\pi}$!

Solution.

1. Just substitute

$$\sigma = \sigma' + f_\pi \tag{13.53}$$

and expand the Lagrangian (13.47) in powers of the fields.

2. The potential to be minimised is now

$$\frac{\lambda}{8}(\vec{\pi}^2 + \sigma^2 - f_\pi^2)^2 - \epsilon\sigma \tag{13.54}$$

and the minimum is given by $\vec{\pi}_0 = 0$ and the solution of

$$\frac{\lambda}{2}\sigma(\sigma^2 - f_\pi^2) = \epsilon \tag{13.55}$$

To first order in ϵ the minimum is shifted as

$$\sigma_0 = f_\pi - \frac{\epsilon}{\lambda f_\pi^2} \tag{13.56}$$

Substituting $\sigma = \sigma' + \sigma_0$ results in a non-zero mass for $\vec{\pi}$, which is given by

$$m_\pi^2 = \frac{\epsilon}{f_\pi} \tag{13.57}$$

We note that the mass of σ is also shifted by a small amount

$$m_\sigma^2 = \lambda f_\pi^2 + \frac{3\epsilon}{2f_\pi} \tag{13.58}$$

□

It is an interesting observation that a simple extension of the Noether theorem can be used to compute what happens to the current under the explicit symmetry breaking. Repeating the derivation of Section 8.3 with a Lagrangian density transforming as

$$\mathcal{L} \rightarrow \mathcal{L} + \omega_k \Delta_k \mathcal{L} \quad (13.59)$$

modifies (8.47) to

$$\partial_\mu \left(\frac{\partial \mathcal{L}}{\partial (\partial_\mu \Phi_\alpha)} \Delta_k \Phi_\alpha \right) = \Delta_k \mathcal{L} \quad (13.60)$$

In our case applying the transformation (13.49) to the Lagrangian (13.52) gives

$$\Delta_k \mathcal{L}'_\sigma = \epsilon \Delta_k \sigma = -\epsilon \pi_k \quad (13.61)$$

so we obtain

$$\partial_\mu J_k^\mu = -\epsilon \pi_k \quad (13.62)$$

so the divergence of the current is just the pseudo-Goldstone field multiplied by the magnitude of the explicit symmetry breaking.

In the case of the quark model, the part of the chiral symmetries which is broken is called *axial*, and the corresponding equation (13.62) is called the *partially conserved axial current (PCAC)* relation, which leads to relations between physical properties of pions and the parameters of the symmetry breaking, such as e.g. the following relation between the π - K mass splitting and current quark masses (i.e. quark masses defined from chiral symmetry breaking):

$$\frac{m_u + m_d}{2m_s} = \frac{m_\pi^2}{2m_K^2 - m_\pi^2}$$

and other interesting results⁴⁵. For a detailed discussion the reader is referred to Chapter 5 of the book T.-P. Cheng and L.-F. Li: *Gauge theory of elementary particle physics*.

13.2.3 The Higgs mechanism

The problem of massive gauge bosons is solved by a mechanism which occurs when one attempts to break spontaneously a gauge symmetry. As an example, let us consider the model (13.31), but this time coupled to a $U(1)$ gauge field:

$$\begin{aligned} \mathcal{L}_H &= -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} + (\partial_\nu - iqA) \Phi^* (\partial^\nu + iqA^\nu) \Phi - V(\Phi^* \Phi) \\ F_{\mu\nu} &= \partial_\mu A_\nu - \partial_\nu A_\mu \end{aligned} \quad (13.63)$$

We assume that in the absence of gauge fields V induces spontaneous symmetry breaking so the vacuum corresponds to

$$|\Phi_0| = \frac{v}{\sqrt{2}} \quad (13.64)$$

⁴⁵Examples for the latter are the Goldberger-Treiman relation and the Adler-Weisberger theorem, both of which which allow to express the nucleon axial coupling strength α in (11.32) in terms of the pion-nucleon interaction.

Substituting

$$\Phi = \frac{1}{\sqrt{2}} (v + \rho(x)) e^{i\Theta(x)/v} \quad (13.65)$$

and assuming again a potential of the form

$$V = -\mu^2 |\Phi|^2 + \frac{\lambda}{4} |\Phi|^4 \quad \text{with } \lambda > 0 \quad (13.66)$$

one can write the Lagrangian in the form

$$\begin{aligned} \mathcal{L}_H = & -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} + \frac{1}{2} \partial_\mu \rho \partial^\mu \rho + \frac{1}{2} (v + \rho)^2 \left(\frac{1}{v} \partial_\mu \Theta + q A_\mu \right) \left(\frac{1}{v} \partial^\mu \Theta + q A^\mu \right) \\ & + \frac{\mu^4}{\lambda} - \mu^2 \rho^2 - \frac{\sqrt{\mu^2 \lambda}}{2} \rho^3 - \frac{\lambda}{16} \rho^4 \end{aligned} \quad (13.67)$$

Exercise 13.7. Derive (13.67).

Solution. Using the results of Exercise 13.3 we can write

$$\begin{aligned} (\partial_\mu + i q A_\mu) \Phi &= e^{i\Theta/v} \frac{1}{\sqrt{2}} \left(\partial_\mu \rho + i (v + \rho) \left(\frac{1}{v} \partial_\mu \Theta + q A_\mu \right) \right) \\ V &= -\frac{\mu^4}{\lambda} + \mu^2 \rho^2 + \frac{\sqrt{\mu^2 \lambda}}{2} \rho^3 + \frac{\lambda}{16} \rho^4 \end{aligned} \quad (13.68)$$

and substituting into (13.63) gives (13.67). $\square$

The first important point is that *the symmetry is not really spontaneously broken*, as sometimes is mistakenly claimed. Note that in contrast to the case in Subsection 13.2, the symmetry transformation is now a gauge transformation, so although it connects vacua

$$\langle \Phi \rangle = \frac{v}{\sqrt{2}} e^{i\vartheta} \quad (13.69)$$

with different ϑ , these states are connected by a constant (i.e. global) gauge transformation, and are therefore physically equivalent. Put in another way, gauge symmetry is really an artifact of describing the model with redundant variables such as the gauge potential A_μ , so “breaking” it is not a genuine physical symmetry breaking.

However, something else does happen. Let us introduce a new field B_μ by writing

$$q B_\mu = \frac{1}{v} \partial_\mu \Theta + q A_\mu \quad (13.70)$$

i.e.

$$B_\mu = A_\mu + \frac{1}{qv} \partial_\mu \Theta \quad (13.71)$$

This corresponds to a gauge transformation under which

$$\Phi(x) \rightarrow \Phi(x) e^{-\frac{i}{v} \Theta(x)} = \frac{1}{\sqrt{2}} (v + \rho) \quad (13.72)$$

Rewriting the Lagrangian in this so-called *unitary gauge* and dropping the vacuum energy term results in

$$\begin{aligned}\mathcal{L}_H = & -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} - \frac{1}{2}\partial_\mu\rho\partial^\mu\rho + \frac{1}{2}(v+\rho)^2 q^2 B_\mu B^\mu \\ & -\mu^2\rho^2 - \frac{\sqrt{\mu^2\lambda}}{2}\rho^3 - \frac{\lambda}{16}\rho^4\end{aligned}\tag{13.73}$$

Looking at the quadratic terms in the Lagrangian

$$\mathcal{L}_H = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \frac{q^2 v^2}{2}B_\mu B^\mu + \frac{1}{2}\partial_\mu\rho\partial^\mu\rho - \mu^2\rho^2 + \text{interaction terms}$$

we see that it describes a vector boson of mass $M_B = qv$ and a scalar particle of mass $M_H = \sqrt{2}\mu$. Note that the massive vector boson B_μ is a composite made from the massless gauge boson and the would-be Goldstone boson field Θ , with the latter supplying the missing longitudinal polarisation degree of freedom as is manifest from (13.73) (note that in Fourier space ∂_μ corresponds to $-ik_\mu$ in terms of the wave number k_μ).

This is called the *Higgs mechanism* (in full: the *Englert-Brout-Higgs-Guralnik-Hagen-Kibble mechanism*), and it can be formulated as follows:

When a gauge symmetry is “spontaneously broken”, the gauge bosons corresponding to the broken generators become massive by absorbing the would-be Goldstone bosons as their longitudinal polarisation components. The massive scalar boson left over in the above example is called the Higgs boson; depending on details of the symmetry breaking there may be more than one Higgs bosons.

In a non-relativistic setting, this mechanism was proposed in 1962 by P. W. Anderson following work in the late 1950s on symmetry breaking in superconductivity. Its application in particle physics was first discussed in 1960 by Y. Nambu; however, it was not clear how to evade the Goldstone theorem in relativistic field theories that seems to imply that there should always be zero mass Goldstone bosons. Since no zero mass bosonic particles are known in Nature, this seemed a serious obstacle to application of the mechanism in particle physics. In fact, as we see from the above example the Goldstone modes are simply non-physical due to gauge invariance, and the physical degrees of freedom consist of a vector and a scalar boson, all of them massive.

Therefore the discovery of the Higgs mechanism in 1964 was far from trivial; it was achieved independently by three groups of authors:

- R. Brout and P. Englert;
- P. Higgs;
- G.S. Guralnik, C.R. Hagen and T.W.B. Kibble;

listed in temporal order of publication. Their papers eventually contain much more than the simple classical argument above, including explaining how the mechanism works at the level of quantum theory and how it evades the Goldstone theorem due to gauge invariance. However, these papers were largely ignored at the time because of the problem with high energy behaviour of models containing massive vector particles. It was only shown by G. 't Hooft and M. Veltman in 1971-72 that in theories with spontaneous symmetry breaking these problems are absent. Their work on the *perturbative renormalisation of non-Abelian gauge theories* was a crucial step leading to the acceptance of such theories as the basis of particle physics model building. In addition it also has

direct phenomenological importance for precision tests of models of particle physics; 't Hooft and Veltman were awarded the Nobel Prize in Physics in 1999.

In 1967, S. Weinberg and A. Salam showed that the Higgs mechanism can be used to break the electroweak symmetry of S. Glashow's *unification of weak and electromagnetic interactions*, which is discussed in the Section [13.3](#)

The electroweak unification combined with quantum chromodynamics makes up the *Standard Model* of particle physics, which comprises the state-of-the-art knowledge we have about the fundamental constituents of matter and their interactions (excluding gravity).

The famous prediction of the existence of the Higgs boson was made by P. Higgs, which is why in most places the mechanism is simply called the Higgs mechanism. After the discovery of the Higgs boson responsible for electroweak symmetry breaking (discussed in Subsection [\(13.3.2\)](#)) in 2012, P. Englert and P. Higgs shared the 2013 Nobel Prize in Physics^{[46](#)}.

13.3 Elements of the Glashow-Salam-Weinberg model

13.3.1 Gauge symmetry and fermion multiplets

We recall that under the weak interactions, left-handed fermions organise into doublets

$$\begin{pmatrix} \nu_e \\ e \end{pmatrix}_L \begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix}_L \begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix}_L \begin{pmatrix} u \\ d \end{pmatrix}_L \begin{pmatrix} c \\ s \end{pmatrix}_L \begin{pmatrix} t \\ b \end{pmatrix}_L \quad (13.74)$$

The charged current interaction induces transitions between members of these doublets. It is natural to assume that the $W^\pm$ vector bosons are gauge bosons that acquire mass via Higgs mechanism, which solves the problems with high-energy behaviour of scattering amplitudes. The simplest gauge group to accommodate the fermions is $SU(2)_L$, where the index L signifies that it only acts on left-handed fermions; the right-handed components do not couple to $W^\pm$ and must therefore be singlets.

$SU(2)_L$ has three generators which in the two-dimensional fundamental representation are

$$R_2(T_i) = \frac{\tau_i}{2} \quad i = 1, 2, 3 \quad (13.75)$$

where

$$\tau_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad , \quad \tau_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad , \quad \tau_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad (13.76)$$

The generators

$$\begin{aligned} T_+ &= T_1 + iT_2 \text{ with } R_2(T_+) = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \\ T_- &= T_1 - iT_2 \text{ with } R_2(T_-) = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \end{aligned} \quad (13.77)$$

correspond to transitions upward/downward within the doublets, and are mediated by the $W^\pm$ bosons. However, the generator

$$R_2(T_0) = \frac{\tau_3}{2} = \frac{1}{2} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \quad (13.78)$$

⁴⁶R. Brout died in 2011, and the rules of the Nobel Prize exclude splitting it among more than three recipients.

corresponds to a neutral current interaction. A proposal would be to take T_0 to be the electromagnetic interaction. This would mean that the gauge coupling is the elementary charge $g = e = \sqrt{4\pi\alpha} \approx 0.3$, which leads to $M_W \approx 100$ GeV from (13.30).

However, this is impossible for two reasons:

1. The gauge bosons of $SU(2)_L$ break parity invariance, while the electromagnetic interaction is known to preserve parity.
2. In any representation $\text{Tr } T_0 = 0$, but the sum of charges in each of the doublets (13.74) is non-zero.

One way to make T_0 correspond to the electric charge is to enlarge the fermions into triplets by involving new fermions, which must be heavy enough so that they escaped detection so far. Such a model was indeed proposed by H. Georgi and S. Glashow in 1972 but it quickly fell from favour after the discovery of neutral current weak interactions in 1973.

The other alternative is to enlarge the gauge group with a $U(1)$ factor to account for electromagnetism. The full gauge group is

$$SU(2)_L \times U(1)_Y \quad (13.79)$$

with the $U(1)$ generator $Y/2$ where Y is called *weak hypercharge*. As we see shortly, electric charge is a combination of T_3 and Y , and we choose the normalisation of Y so that

$$Q = T_3 + \frac{Y}{2} \quad (13.80)$$

in analogy with the relation (7.76), but with different symmetries involved: the T_i *weak isospin* and the Y *weak hypercharge*. Note also that the hadronic isospin and hypercharge are *global symmetries*, while their weak counterparts are *local (gauge) symmetries*.

In any case, relation (13.80) allows us to assign the following weak properties to the fermions:

- the left-handed fermion doublets have weak isospin $1/2$, with $T_3 = +1/2$ for their upper components and $T_3 = -1/2$ for their lower components. Using the known electric charges, left-handed lepton doublets have $Y = -1$, while quark doublets have $Y = 1/3$ (since Y corresponds to a $U(1)$ transformation, these are the same for both members of the doublets).
- right-handed fermions

$$u_R, d_R, c_R, s_R, t_R, b_R, e_R, \mu_R, \tau_R$$

are singlets under $SU(2)_L$, so they have $T = T_3 = 0$ and their weak hypercharge $Y = 2Q$, where Q is the electric charge (in units of e); i.e. $Y(l_R) = -2$, $Y(q_R) = 4/3$ for $q = u, c, t$ and $Y(q_R) = -2/3$ for $q = d, s, b$.

- If the neutrinos are Dirac particles, the right-handed neutrinos exist due to their non-zero mass. They have $T = T_3 = 0$ and $Y = 0$, and do not have electric or weak interactions (neither strong ones, by definition).

From the above assignment, the part of the Lagrangian involving the fermion fields and gauge bosons can be explicitly written down using the rules of gauge theory (cf. Section 12.1). We do not do that here; we only note that there are two coupling constants g and g' , corresponding to $SU(2)_L$ and $U(1)_Y$, respectively. Note that the assignment of weak quantum numbers ensures that left and right fields corresponding to the same flavour carry the same electric charge, therefore

the electromagnetic interaction preserves parity. The eventual electromagnetic field is a mixture between the neutral component of $SU(2)_L$ and $U(1)_Y$, as discussed later in [13.3.2](#). The pattern of symmetry breaking underlying the Higgs mechanism is

$$SU(2)_L \times U(1)_Y \longrightarrow U(1)_{EM} \quad (13.81)$$

where the electromagnetic subgroup is generated by Q given in [\(13.80\)](#).

However, there is another neutral current interaction corresponding to the other independent combination of T_3 and Y , mediated by a vector boson Z of zero electric charge, which has a mass M_Z that is expected to be similar in magnitude to M_W . This mediates a neutral current interaction, which preserves flavour since T_3 and Y are diagonal in flavour space. Since left-handed neutrinos transform non-trivially under $SU(2)_L$, they do couple to the weak neutral current interaction mediated by Z . As a result, the Glashow-Salam-Weinberg model predicts the existence of weak neutral current interactions, resulting processes such as

$$\begin{aligned} \nu_l + p &\longrightarrow \nu_l + p \\ \nu_l + n &\longrightarrow \nu_l + n \end{aligned} \quad (13.82)$$

(compare these to the charged current neutrino processes in [\(11.1\)](#) and [\(11.17\)](#), which always change flavour as well).

Neutral current interactions of neutrinos were discovered at CERN in 1973 by experiments using the Gargamelle bubble chamber, which then disproved the Georgi-Glashow model and confirmed the predictions of the $SU(2)_L \times U(1)_Y$ electroweak theory, for which Glashow, Salam and Weinberg were awarded the Nobel Prize in Physics in 1979.

13.3.2 Higgs mechanism and vector boson masses

The origin of vector boson masses is clear: it is connected to spontaneous symmetry breaking. The simplest way to implement is by introducing a complex scalar field doublet called the electroweak Higgs field, which has the form

$$\Phi = \begin{pmatrix} \phi_+ \\ \phi_0 \end{pmatrix} \quad (13.83)$$

with both components being complex scalar fields, and which transforms as the fundamental representation **2** under $SU(2)_L$. One member of this doublet should be electrically neutral to allow for a vacuum expectation value that leaves $U(1)_{EM}$ unbroken; choosing $Y(\Phi) = +1$ makes the lower component neutral according to [\(13.80\)](#). Then a scalar potential

$$V = -\mu^2 \Phi^\dagger \Phi + \lambda (\Phi^\dagger \Phi)^2 \quad (13.84)$$

leads to an expectation value with

$$|\Phi_0| = \frac{v}{\sqrt{2}} \quad v = \sqrt{\frac{\mu^2}{\lambda}} \quad (13.85)$$

for which one can choose the following orientation:

$$\langle \Phi \rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix} \quad (13.86)$$

In polar variables, the field can be written

$$\Phi(x) = \exp \left\{ -i \frac{\zeta_k(x) \tau_k}{v} \right\} \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + \eta(x) \end{pmatrix} \quad (13.87)$$

Let us consider now the pattern of symmetry breaking. The general $SU(2)_L \times U(1)_Y$ gauge transformation is

$$e^{-i\epsilon^a(x)T^a} e^{-i\epsilon(x)Y}$$

where T^a are the $SU(2)_L$ generators (weak isospin) while Y is the $U(1)_Y$ generator (weak hypercharge). Since the Higgs field is a doublet and has weak hypercharge -1 we must substitute

$$T^a = \frac{\sigma^a}{2} \quad Y = -1$$

Therefore the Higgs doublet is gauge transformed as

$$\begin{pmatrix} \Phi_1 \\ \Phi_2 \end{pmatrix} \rightarrow \exp \left\{ -i \begin{pmatrix} \frac{1}{2}\epsilon^3(x) + \epsilon(x) & \frac{1}{2}(\epsilon^1(x) - i\epsilon^2(x)) \\ \frac{1}{2}(\epsilon^1(x) + i\epsilon^2(x)) & -\frac{1}{2}\epsilon^3(x) + \epsilon(x) \end{pmatrix} \right\} \begin{pmatrix} \Phi_1 \\ \Phi_2 \end{pmatrix}$$

Now if the Higgs has a vacuum value, it can always be rotated into the form

$$\frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix}$$

where v is real. This is left invariant by the transformation with $\epsilon^{1,2}(x) = 0$ and $\epsilon(x) = \epsilon^3(x)/2$, which are

$$e^{-i\epsilon^3(x)(T^3+Y/2)}$$

So the one-parameter gauge symmetry subgroup generated by

$$Q = T^3 + \frac{Y}{2}$$

is unaffected by the Higgs mechanism. The gauge boson corresponding to this generator does not couple to the Higgs vacuum field, so it does not gain mass and a gauge field for the electromagnetic gauge group $U(1)_Q \equiv U(1)_{EM}$.

It is also clear that the physical Higgs boson is electrically neutral, since it corresponds to quanta of the excitations $\eta(x)$ of the Higgs vacuum field

$$\frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + \eta(x) \end{pmatrix}$$

which is invariant under $U(1)_Q$ gauge transformations and so does not couple to the electromagnetic field.

As for the $U(1)$ Higgs model discussed in Subsection 13.2.3, the dependence on the Goldstone modes $\zeta_k(x)$ can be eliminated by introducing the unitary gauge. The gauge covariant derivative is

$$D_\mu = \partial_\mu - igT_k W_{k\mu} - \frac{i}{2}g'Y B_\mu \quad (13.88)$$

with g denoting the $SU(2)_L$, while g' denoting the $U(1)_Y$ coupling. For the Higgs doublet it has the form

$$D_\mu \Phi = \left(\partial_\mu - \frac{i}{2}g\tau_k W_{k\mu} - \frac{i}{2}g'B_\mu \right) \Phi(x) \quad (13.89)$$

In the unitary gauge

$$D_\mu \Phi = \left(\partial_\mu - \frac{i}{2} g \tau_k W_{k\mu} - \frac{i}{2} g' B_\mu \right) \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + \eta(x) \end{pmatrix}$$

The gauge boson mass terms are the ones independent of $\eta(x)$ in

$$D_\mu \Phi^\dagger D^\mu \Phi$$

which are

$$\begin{aligned} \mathcal{L}_{gm} &= \frac{1}{2} \begin{pmatrix} 0 & v \end{pmatrix} \left(\frac{g}{2} \tau_k W_{k\mu} + \frac{g'}{2} B_\mu \right) \left(\frac{g}{2} \tau_k W_k^\mu + \frac{g'}{2} B^\mu \right) \begin{pmatrix} 0 \\ v \end{pmatrix} \\ &= \frac{v^2}{8} \{ g^2 (W_{1\mu} W_1^\mu + W_{2\mu} W_2^\mu) + (g W_{3\mu} - g' B_\mu) (g W_3^\mu - g' B^\mu) \} \end{aligned} \quad (13.90)$$

Exercise 13.8. The neutral gauge boson mass terms have the form

$$\frac{v^2}{8} \begin{pmatrix} W_{3\mu} & B_\mu \end{pmatrix} \begin{pmatrix} g^2 & -gg' \\ -gg' & g'^2 \end{pmatrix} \begin{pmatrix} W_3^\mu \\ B^\mu \end{pmatrix} \quad (13.91)$$

Find the eigenvalues, and also the eigenvectors in the form

$$\begin{aligned} Z_\mu &= W_{3\mu} \cos \theta_W - B_\mu \sin \theta_W \\ A_\mu &= W_{3\mu} \sin \theta_W + B_\mu \cos \theta_W \end{aligned} \quad (13.92)$$

Solution.

$$\begin{aligned} \tan \theta_W &= \frac{g'}{g} \\ M_Z &= v^2 \frac{g^2 + g'^2}{4} \\ M_A &= 0 \end{aligned} \quad (13.93)$$

□

The *gauge field* A_μ mediates the electromagnetic interaction and from (13.92) the interaction strength is given by

$$e = g \sin \theta_W = g' \cos \theta_W \quad (13.94)$$

which is equal to the elementary charge ($e = \sqrt{4\pi\alpha}$ in our units). The mass of the *weak charged bosons* $W^\pm$ described by the fields

$$W_\mu^\pm = \frac{1}{\sqrt{2}} (W_{1\mu} \pm i W_{2\mu}) \quad (13.95)$$

are given by

$$M_W = \frac{gv}{2} \quad (13.96)$$

while the mass of the *weak neutral vector boson* Z is

$$M_Z = \frac{v}{2} \sqrt{g^2 + g'^2} \quad (13.97)$$

Taking properly into account the normalisation factors of 2 in (13.89) and $\sqrt{2}$ in (13.95), (13.30) is changed to

$$\frac{G_F}{\sqrt{2}} = \frac{g^2}{8M_W^2} \quad (13.98)$$

and combining this relation with (13.96), the Higgs vacuum expectation value can be determined

$$v = 2^{-1/4} G_F^{-1/2} \approx 246 \text{ GeV} \quad (13.99)$$

The angle θ_W is called the *Weinberg angle*. There is the following relation between θ_W and the gauge boson masses:

$$\rho = \frac{M_W^2}{M_Z^2 \cos^2 \theta_W} = 1 \quad (13.100)$$

which can be shown to result from an additional (so-called *custodial*) $SU(2)'$ symmetry which is due to the doublet representation chosen for the Higgs field. This parameter receives corrections at higher orders of perturbation theory and its comparison to observations is a sensitive way to test for deviations from the Standard Model.

θ_W can be directly measured from parity violating effects e.g. in the Møller scattering $e^- + p \rightarrow e^- + p$, or the muon pair creation process $e^- + e^+ \rightarrow \mu^- + \mu^+$. These result from the fact that these processes can proceed via the exchange of a Z boson in addition to the usual electromagnetic contribution. The present measured value is

$$\sin^2 \theta_W = 0.23126(5) \quad (13.101)$$

at the scale⁴⁷ M_Z (in the $\overline{MS}$ scheme), i.e. θ_W is about 30° .

The weak vector bosons were discovered by the UA1 and the UA2 collaborations at CERN's Super Proton Synchrotron (SPS) in 1983. For their decisive contributions to the experimental discovery of the W and Z bosons, C. Rubbia and S. van der Meer were awarded the Nobel Prize in Physics in 1984.

Subsequent generations of experiments have greatly improved and extended the experimental tests of the Standard Model, determining parameters more precisely and excluding deviations with higher and higher certainty⁴⁸. The best values for the masses of Z and W at present are

$$\begin{aligned} M_W &= 80.385(15) \text{ GeV} \\ M_Z &= 91.1876(21) \text{ GeV} \end{aligned} \quad (13.102)$$

and their properties (such as decay width and branching ratios) are in very precise agreement with the predictions of the Standard Model (including higher order corrections).

⁴⁷Since coupling constants are energy dependent (cf. Subsection (12.3.3)), the scale and the renormalisation scheme must be given.

⁴⁸More precisely, models proposing observable physical effects beyond the Standard Model (in the jargon: *BSM physics*) have some parameters corresponding to the new effects, and part of the parameter space is excluded by experiments. As there has yet been no clear evidence for BSM physics, the excluded part of these parameter spaces has grown considerably for the majority of BSM models. In some cases the allowed parameter space has shrunk to zero, making the model fully excluded.

An interesting part of the story is that the number of weakly interacting light⁴⁹ neutrinos can be determined from studying the decay of the Z boson, with the present result

$$N_\nu = 2.984 \pm 0.008 \quad (13.103)$$

13.3.3 Fermion masses and mixing angles

There is a problem that we avoided so far. *The fermions are chiral: the left-handed and the right-handed fermions transform in different representations of the $SU(2)_L \times U(1)_Y$ gauge symmetry, and therefore mass terms are excluded!* For example, a mass term for the electron takes the form

$$m_e \bar{e} e = m_e (\bar{e}_L e_R + \bar{e}_R e_L) \quad (13.104)$$

(where e now denotes the electron Dirac field and m_e is the electron mass). However e_L is part of a doublet and e_R is a singlet under $SU(2)_L$, so (13.104) is not invariant under $SU(2)_L$. In addition, it also violates $U(1)_Y$ since the weak hypercharges of e_L and e_R are not equal, so under a $U(1)_Y$ transformation $\exp(-i\epsilon Y/2)$

$$\begin{aligned} e_L &\rightarrow \exp(-i\epsilon/2) e_L \\ e_R &\rightarrow \exp(-i\epsilon) e_R \end{aligned} \quad (13.105)$$

corresponding to the assignments determined in Subsection 13.3.1. This is also the reason while there is no flavour mixing involved in (13.74) since there is no physical significance of a mixing matrix as it can be eliminated by relabelling the flavours as long as the fermions are massless.

However, *the Higgs field solves this issue as well via the so-called Yukawa couplings.* These are the same sort of interactions that Yukawa proposed to describe the strong interactions and couple a spin-1/2 fermion field to a scalar boson field. Let's see how it works for the mass of the u and d quarks. Introducing the notation

$$Q = \begin{pmatrix} u \\ d \end{pmatrix}_L \quad (13.106)$$

for the doublet, the term

$$\bar{Q}_\alpha \Phi_\alpha d_R \quad (13.107)$$

is $SU(2)_L \times U(1)_Y$ invariant

$$\begin{aligned} SU(2)_L &: \bar{Q}_\alpha \Phi_\alpha d_R \rightarrow U_{\alpha\beta}^* \bar{Q}_\beta U_{\alpha\gamma} \Phi_\gamma d_R = \bar{Q}_\alpha \Phi_\alpha d_R \quad \text{by } U^\dagger U = \mathbf{1}_{2 \times 2} \\ U(1)_Y &: \bar{Q}_\alpha \Phi_\alpha d_R \rightarrow e^{i\alpha/3} \bar{Q}_\alpha e^{-i\alpha} \Phi_\alpha e^{i2\alpha/3} d_R = \bar{Q}_\alpha \Phi_\alpha d_R \end{aligned} \quad (13.108)$$

(note that $U(1)_Y$ transformations have the form $e^{-i\alpha Y}$). Adding a term to the Lagrangian

$$y \bar{Q}_\alpha \Phi_\alpha d_R + \text{h.c.} \quad (13.109)$$

produces a mass term and a coupling to Higgs field upon the substitution of the following expression (valid in the unitary gauge)

$$\Phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + \eta(x) \end{pmatrix} \quad (13.110)$$

as follows

$$y \bar{Q}_\alpha \Phi_\alpha d_R + \text{h.c.} \rightarrow \frac{yv}{\sqrt{2}} \bar{d}_L d_R + \frac{y}{\sqrt{2}} \eta \bar{d}_L d_R + \text{h.c.} \quad (13.111)$$

⁴⁹Here light means that $M_\nu < M_Z/2$, so the Z is kinematically allowed to decay into a neutrino pair.

so d acquires the mass

$$m_d = \frac{yv}{\sqrt{2}} \quad (13.112)$$

One can generate mass terms in the same way for the lower components of all doublets, including the leptons. For the upper components of the doublets, we recall from Exercise (7.7) that the conjugate doublet defined as

$$\tilde{\Phi} = \begin{pmatrix} \phi_0^* \\ -\phi_+^* \end{pmatrix}$$

transforms as the defining representation **2** under $SU(2)_L$:

$$SU(2)_L : \tilde{\Phi}_\alpha \rightarrow U_{\alpha\beta} \tilde{\Phi}_\beta$$

but due to the conjugation it has the opposite hypercharge $Y = -1$. As a result, the term

$$y \bar{Q}_\alpha \tilde{\Phi}_\alpha u_R + \text{h.c.} \quad (13.113)$$

is $SU(2)_L$ invariant (with the same reasoning as in (13.108)), while it is also $U(1)_Y$ invariant according to

$$U(1)_Y : \bar{Q}_\alpha \tilde{\Phi}_\alpha u_R \rightarrow e^{i\alpha/3} \bar{Q}_\alpha e^{i\alpha} \tilde{\Phi}_\alpha e^{-i4\alpha/3} u_R = \bar{Q}_\alpha \tilde{\Phi}_\alpha u_R \quad (13.114)$$

Upon substituting (13.110) leads to masses for the u quarks

$$y \bar{Q}_\alpha \tilde{\Phi}_\alpha u_R + \text{h.c.} \rightarrow \frac{yv}{\sqrt{2}} \bar{u}_L u_R + \frac{y}{\sqrt{2}} \eta \bar{u}_L u_R + \text{h.c.} \quad (13.115)$$

and similarly for the upper components of all other doublets.

Including three generations of quarks and leptons the general Yukawa terms can be written as

$$y_{ij}^{(l)} \bar{L}_{i\alpha} \Phi_\alpha l_{jR} + y_{ij}^{(\nu)} \bar{L}_{i\alpha} \tilde{\Phi}_\alpha \nu_{jR} + y_{ij}^{(d)} \bar{Q}_{i\alpha} \Phi_\alpha d_{jR} + y_{ij}^{(u)} \bar{Q}_{i\alpha} \tilde{\Phi}_\alpha u_{jR} \quad (13.116)$$

where the generation indices i, j run from 1 to 3, with the notation

$$Q_i = \begin{pmatrix} u_i \\ d_i \end{pmatrix} \quad L_i = \begin{pmatrix} \nu_i \\ l_i \end{pmatrix} \quad (13.117)$$

$$\begin{aligned} u_1 &= u & \nu_1 &= \nu_e \\ u_2 &= c & \nu_2 &= \nu_\mu \\ u_3 &= t & \nu_3 &= \nu_\tau \\ d_1 &= d & l_1 &= e \\ d_2 &= s & l_2 &= \mu \\ d_3 &= b & l_3 &= \tau \end{aligned}$$

(note that here the neutrino indices 1, 2, 3 label the flavour eigenstates e, μ, τ , in contrast to Subsection (11.4.3) where they labelled the mass eigenstates). This leads to the following mass matrices:

$$\begin{aligned} m_{ij}^{(u)} &= y_{ij}^{(u)} \frac{v}{\sqrt{2}} & m_{ij}^{(\nu)} &= y_{ij}^{(\nu)} \frac{v}{\sqrt{2}} \\ m_{ij}^{(d)} &= y_{ij}^{(d)} \frac{v}{\sqrt{2}} & m_{ij}^{(l)} &= y_{ij}^{(l)} \frac{v}{\sqrt{2}} \end{aligned} \quad (13.118)$$

Using a theorem of linear algebra, *any matrix can be diagonalised by bi-unitary transformations*, e.g.

$$S_L^{(u)\dagger} m^{(u)} S_R^{(u)} = \begin{pmatrix} m_u & 0 & 0 \\ 0 & m_c & 0 \\ 0 & 0 & m_t \end{pmatrix} \quad (13.119)$$

where $S_{L,R}^{(u)}$ are unitary matrices (determined up to phase redefinitions), and similarly for the other three cases $m^{(d)}$, $m^{(\nu)}$ and $m^{(l)}$. *Mass eigenstates are obtained by redefining the flavour basis states* as

$$\begin{aligned} u_{iL} &= \left(S_L^{(u)} \right)_{ij} u'_{jL} \\ u_{iR} &= \left(S_R^{(u)} \right)_{ij} u'_{jR} \end{aligned} \quad (13.120)$$

so the mass terms become diagonal

$$\bar{u}_{iL} m_{ij}^{(u)} u_{jR} \rightarrow \bar{u}'_{kL} \left(S_L^{(u)\dagger} \right)_{ki} m_{ij}^{(u)} \left(S_R^{(u)} \right)_{jl} u'_{lR} = m_u \bar{u}' u' + m_c \bar{c}' c' + m_t \bar{t}' t' \quad (13.121)$$

and the CKM flavour mixing matrix is obtained as the *relative transformation of the left-handed d quarks with respect to the left-handed u quarks*, i.e.

$$V_{CKM} = S_L^{(u)\dagger} S_L^{(d)} \quad (13.122)$$

while the neutrino mixing matrix is

$$U_{PMNS} = S_L^{(l)\dagger} S_L^{(\nu)} \quad (13.123)$$

Note that all these mass terms are Dirac type; Majorana mass terms (cf. [\(11.128\)](#)), if exist, cannot come from Yukawa couplings to the Standard Model Higgs field. Therefore if the existence of Majorana mass terms is confirmed by experiments (e.g. by observing neutrinoless double β -decay), it would be a signal of physics beyond the Standard Model.

We close by noting that the existence of Majorana masses opens the possibility of explaining the smallness of neutrino masses. Using the results in Subsection [11.4.3](#) that the neutrino mass matrix appering in [11.129](#) is of the form

$$\begin{pmatrix} A & D/2 \\ D/2 & B \end{pmatrix} = \begin{pmatrix} 0 & m \\ m & M \end{pmatrix} \quad (13.124)$$

where the zero entry comes from the fact that Majorana mass cannot be produced by the Standard Model Higgs, m is the Dirac mass which is of order of the fermion masses produced by the Higgs, while M is a Majorana mass term that comes from physics beyond the Standard Model. For example, if one considers the so-called Grand Unification Theories, the typical ratio is expected to be $m/M \sim 10^{-13}$. Then the two eigenvalues are

$$m_1 \sim -\frac{m^2}{M} \quad m_2 \sim M$$

and one naturally ends up with a very light Majorana species^{[50](#)}, which corresponds to the observed neutrino, and a 'superheavy' one which cannot be observed at presently available energies.

⁵⁰Note that the sign of the fermion mass in the Dirac equation carries no physical significance, and can be simply redefined.

13.3.4 The Higgs boson and present status of the Standard Model

Note that *the coupling of each particle to the Higgs boson is determined by the particle's mass*. This is true even for the Higgs itself; due to the relation (13.85) the Higgs mass $M_H = \sqrt{2}\mu$ is determined, since the vacuum expectation value v is known (cf. (13.99)). Since prior to the discovery of the Higgs all other relevant particle masses were known⁵¹ once the Higgs mass of $M_H = 125.7 \pm 0.3$ GeV was pinned down from the data obtained by the LHC experiments CMS and ATLAS, *all other properties of the particle such as lifetime and decay branching ratios were predicted by the Standard Model (SM)*. So measuring these it can be decided experimentally whether the discovered particle was really a proper candidate for the SM Higgs boson; up to now, all the data are consistent with this identification, and possible deviations are continually becoming more and more constrained by the data.

Deviations from SM predictions could result from the existence of other Higgs bosons which also contribute to particle masses, making it possible for the Higgs couplings to deviate from the SM values, or from the Higgs boson being a composite, rather than elementary object. Indeed the Higgs boson is a very unique particle, as all other spinless bosons (such as pions or kaons) are composite particles; however up to now there are no indications of compositeness for the Higgs boson.

Presently the Standard Model is well-established and consistent with all observations. It is not an elegant model, however, and there are several reasons why it is not considered to be the final word in particle physics:

1. *It has a large number of parameters.* The gauge group is a product $SU(3)_c \times SU(2)_L \times U(1)_Y$, resulting in 3 couplings (plus eventually the QCD θ angle, cf. (12.90)). There are the fermion masses: three charged lepton plus three neutrino⁵² plus six quark masses. The two flavour mixing matrices contribute three angles and one CP violating phase each.⁵³ In addition to these there are also the Higgs mass and vacuum expectation value. Taken altogether these are 26 parameters, which is not very elegant.⁵⁴ However, it must also be kept in mind that the model explains a vast amount of observed data, and typically for a given class of phenomena only a few of these parameters are actually relevant.
2. *There is no explanation for the fermion generations, nor for the parity and CP violations;* they are simply built into the structure of the model. There is a clear hierarchy between masses, which span many orders of magnitudes, as well as between family mixings, but the standard model does not provide any explanation for these.
3. *Similarly, it does not provide any explanation for the absence of strong CP violation* coming from the QCD θ angle.
4. It is also unclear where the *neutrino species mixing* comes from. In the case of the quarks, the mixing matrix eventually corresponds to a misalignment between the flavour basis of the strong and weak interactions. The neutrinos only participate in the weak interactions,

⁵¹Except for neutrinos, but due to their small mass their coupling to the Higgs is minuscule anyway.

⁵²Excluding Majorana masses.

⁵³In the presence of Majorana masses there would be further parameters in the neutrino mixing.

⁵⁴Some people, in fact, see a resemblance to the epicycles in the geocentric model (exemplified by the inclusion of neutrino oscillations at the price of adding seven new parameters, not counting Majorana masses and related additional mixing parameters), and wish for a Copernican revolution that would put all the phenomena in a new light!

so it is not clear where the misalignment described by their flavour mixing matrix comes from. Although their mixing can be included by introducing the necessary parameters in the standard model, its origin is as mysterious as that of the three generations itself.

5. The *Higgs sector seems ad hoc*, simply the minimal choice for symmetry breaking. The symmetry breaking mechanism is destabilised in the presence of other high-energy scales, and there seems to be at least one such scale given by the Planck units (3.16) and (3.18). This is called the *hierarchy problem*.
6. *It cannot accommodate the dark matter inferred from cosmology, or explain the dark energy* making up about 2/3 of the energy density of the Universe.
7. *It does not include gravity*: it is a notorious outstanding problem of theoretical physics to harmonise or unify general relativity with quantum theory in some way.
8. *The Higgs sector has a problem of triviality* (cf. Subsection 12.3.2).

The last two problems mean that *the Standard Model is theoretically incomplete as it stands*. However, it is not entirely clear what is the scale where new physics should manifest itself, nor it is clear what this beyond-the-standard-model (BSM) physics should be. In the decades since the '70s, a large number of such models were constructed, involving ideas such as grand unification, supersymmetry, extra dimensions, string theory etc. However, at present there is no observational evidence for any violation of the Standard Model apart from the cues coming from cosmology and astrophysics (existence of dark matter and dark energy), which are not specific enough to pin down any details yet. These issues are at the forefront of contemporary theoretical and experimental research in particle physics and related areas.

14 Accelerators

14.1 Linac, cyclotron, synchrotron

14.2 Cross section and luminosity

15 Detectors

15.1 Anatomy of high energy particle detectors

15.2 Principles of particle identification

A Tensor method for Lie group representations

A.1 Tensors for $GL(n, \mathbb{C})$

Consider an n -dimensional complex vector space V with an inner product

$$\begin{aligned} u, v \in V &\rightarrow (u, v) \\ (\alpha u_1 + \beta u_2, v) &= \alpha^*(u_1, v) + \beta^*(u_2, v) \\ (u, \alpha v_1 + \beta v_2) &= \alpha(u, v_1) + \beta(u, v_2) \end{aligned}$$

The dual of this vector space

$$V^\vee = \{f : V \rightarrow \mathbb{C} \mid f(\alpha u_1 + \beta u_2) = \alpha f(u_1) + \beta f(u_2)\}$$

is isomorphic with V via the mapping

$$u \in V \mapsto f_u \in V^\vee : f_u(v) = (u, v)$$

A basis of V is given by n linearly independent vectors $e_1, \dots, e_n$; any $u \in V$ can be uniquely written as

$$u = \sum_{i=1}^n u_i e_i$$

The dual basis for $V^\vee \equiv V$ is given by elements $f_{\bar{1}}, \dots, f_{\bar{n}}$ such that

$$(f_{\bar{i}}, e_j) = \delta_{\bar{i}j}$$

If the basis $e_1, \dots, e_n$ is orthonormal i.e.

$$(e_i, e_j) = \delta_{ij}$$

then it is self-dual. From now on we work in orthonormal basis, and then the scalar product of two vectors

$$u = u_i e_i \quad v = v_i e_i$$

can be written as

$$(u, v) = u_i^* v_i$$

Under an invertible linear map $V \rightarrow V$, the components of a vector transform as

$$u'_i = A_{ij} u_j$$

where the matrix A_i^j is invertible. The set of all such matrices is called the general linear group and denoted by $GL(n, \mathbb{C})$. The above transformation gives an n -dimensional representation of this group on V :

$$\begin{aligned} A \in GL(n, \mathbb{C}) &\mapsto D_d(A) : V \rightarrow V \\ D_d(A)(u_i e_i) &= u'_i e_i \end{aligned}$$

which is called the defining representation. One can also define a dual representation $D_d^\vee(A)$ on $V^\vee \equiv V$ by requiring that

$$(D_d^\vee(A)u, D_d(A)v) = (u, v)$$

Denoting the matrix $D_d^\vee(A)$ by $\bar{A}$, this means in components

$$\bar{A}_{ij}u_j A_{ik}v_k = u_i v_i$$

i.e.

$$\bar{A}_{ij}A_{ik} = \delta_{jk}$$

which in matrix form means that

$$\bar{A} = (A^T)^{-1}$$

which is called the contragredient of A , and the dual representation is also called the contragredient representation

The space of tensors of type (r, s) is defined as

$$\mathcal{T}^{(r,s)} = \underbrace{V \otimes V \otimes \dots V}_r \otimes \underbrace{V^\vee \otimes V^\vee \otimes \dots V^\vee}_s = V^{\otimes r} \otimes (V^\vee)^{\otimes s}$$

If we have a basis $e_1, \dots, e_n$ for V with a dual basis $f_1, \dots, f_n$ for $V^\vee$, then a general tensor $T \in \mathcal{T}^{(r,s)}$ can be expanded as

$$T = T_{i_1 \dots i_r \bar{j}_1 \dots \bar{j}_s} e_{i_1} \otimes \dots \otimes e_{i_r} \otimes f_{\bar{j}_1} \otimes \dots \otimes f_{\bar{j}_s}$$

The group $GL(n, \mathbb{C})$ is represented on the tensor product by the product representation $D^{(r,s)}$:

$$D^{(r,s)}(A) = \underbrace{D_d(A) \otimes \dots \otimes D_d(A)}_r \otimes \underbrace{D_d^\vee(A) \otimes \dots \otimes D_d^\vee(A)}_s$$

which can be written in tensor components as

$$T'_{i'_1 \dots i'_r \bar{j}'_1 \dots \bar{j}'_s} = A_{i'_1 i_1} \dots A_{i'_r i_r} \bar{A}_{\bar{j}'_1 \bar{j}_1} \dots \bar{A}_{\bar{j}'_s \bar{j}_s} T_{i_1 \dots i_r \bar{j}_1 \dots \bar{j}_s}$$

A.2 Interesting subgroups of $GL(n, \mathbb{C})$

Special linear group

$$SL(n, \mathbb{C}) = \{A \in GL(n, \mathbb{C}) | \det A = 1\}$$

The condition on the determinant can also be written as

$$\begin{aligned} \epsilon_{i_1 \dots i_n} A_{i_1 j_1} \dots A_{i_n j_n} &= \epsilon_{j_1 \dots j_n} \\ \epsilon_{j_1 \dots j_n} A_{i_1 j_1} \dots A_{i_n j_n} &= \epsilon_{i_1 \dots i_n} \end{aligned}$$

Unitary group

$$U(n) = \{A \in GL(n, \mathbb{C}) | AA^\dagger = \mathbb{I}\}$$

i.e.

$$A_{ij}A_{ik}^* = \delta_{jk}$$

Special unitary group

$$SU(n) = \{A \in U(n) | \det A = 1\}$$

A.3 Relation to the permutation group

The permutation group S_r of order r acts on the space $\mathcal{T}^{(r,0)}$ of tensors in the following way. An element of S_r is a bijective map $\pi : \{1, \dots, r\} \rightarrow \{1, \dots, r\}$, which simply acts on tensors by permuting their indices:

$$T \rightarrow T^\pi$$

$$(T^\pi)_{i_1 \dots i_r} = T_{i_{\pi(1)} \dots i_{\pi(r)}}$$

This defines a representation $\mathcal{P}_r$ of S_r , where the permutation $\pi \in S_r$ is represented by the linear map

$$\mathcal{P}_r(\pi) : \mathcal{T}^{(r,0)} \rightarrow \mathcal{T}^{(r,0)}$$

$$T \rightarrow T^\pi$$

It is important to realise that the action of the permutation group commutes with the representation of $GL(n, \mathbb{C})$ i.e.

$$D^{(r,0)}(A) \mathcal{P}_r(\pi) = \mathcal{P}_r(\pi) D^{(r,0)}(A)$$

which is simply the fact that if we permute the indices of a tensor and then apply the transformation by $A \in GL(n, \mathbb{C})$ then we get the same result if first apply the transformation A and permute the indices afterwards. This is a consequence of the fact that all indices are transformed by the same matrix A .

The representation $\mathcal{P}_r$ of the permutation group S_r can be decomposed into several irreducible components, and due to the above observation they are also invariant subspaces of the representation $R^{(r,0)}(A)$. In fact, these subspaces are also irreducible:

Theorem A.1. *Any subspace of $\mathcal{T}^{(r,0)}$ which forms an irreducible subrepresentation under the action of the representation $\mathcal{P}_r$ is also irreducible under the action of $D^{(r,0)}$.*

This fundamental theorem (whose proof is omitted here) means that one can generate new irreducible representations of $GL(n, \mathbb{C})$ by decomposing tensors of type $(r, 0)$ according to their symmetry under permutation of indices. For tensors of type (r, s) the situation is somewhat different as demonstrated below. Examples:

1. $(2, 0)$ tensors

Such tensors can be decomposed into a symmetric and antisymmetric part:

$$T_{ij} = T_{ij}^+ + T_{ij}^-$$

$$T_{ij}^+ = \frac{1}{2}(T_{ij} + T_{ji}) \quad T_{ij}^- = \frac{1}{2}(T_{ij} - T_{ji})$$

The representation formed by the symmetric component is an $n(n+1)/2$ -dimensional irreducible representation, while the antisymmetric one is $n(n-1)/2$ -dimensional.

For the case of $GL(2, \mathbb{C})$ the latter is a one-dimensional, but non-trivial representation, since T^- can be written in the form

$$T_{ij}^- = t \epsilon_{ij}$$

and the representation of $GL(2, \mathbb{C})$ acts as

$$t' \epsilon_{i'j'} = A_{i'i} A_{j'j} t \epsilon_{ij}$$

We can then read off

$$t' = (\det A)t$$

So on T^- the matrix $A \in GL(n, \mathbb{C})$ is represented by its determinant.

2. $(1, 1)$ tensors

In this case the two indices transform differently:

$$T'_{i\bar{j}} = A_{ii'} T_{i'j'} (A^{-1})_{j'j}$$

so there is no permutation symmetry to use for decomposition. Instead, this representation has two irreducible components. One of them corresponds to the traceless part

$$\tilde{T}_{i\bar{j}} = T_{i\bar{j}} - \frac{1}{n} (\text{Tr } T) \delta_{i\bar{j}}$$

which forms an $n^2 - 1$ -dimensional representation (since a traceless $n \times n$ matrix has $n^2 - 1$ independent components), while the trace

$$\text{Tr } T = T_{1\bar{1}} + T_{2\bar{2}} + \cdots + T_{n\bar{n}}$$

itself is invariant so it transforms as the trivial representation.

For a general (r, s) tensor, it is necessary to consider symmetry classes in both index types separately, and also to decompose using lower rank invariant tensors obtained by taking traces on index pairs formed by one fundamental and one contragredient fundamental index in all possible independent ways. Here we do not pursue this further, as it turns out there is a much simpler way to construct all irreducible representations for $SU(n)$, which is the case of most interest for us.

A.4 Tensors for $SU(2)$

For $SU(2)$ things are much simpler. First of all, the antisymmetric tensor T_{ij}^- now transforms trivially due to $\det A = 1$. In fact, there is no need to consider tensors which are antisymmetric in any of their indices because

- If the number of indices which are antisymmetric under exchange is larger than 2, the tensor is identically zero;
- If a tensor is antisymmetric under swapping any two of its indices

$$T_{i_1 \dots i_p k i_{p+1} \dots i_q l i_{q+1} \dots i_n} = -T_{i_1 \dots i_p l i_{p+1} \dots i_q k i_{q+1} \dots i_n}$$

then it can be written as

$$T_{i_1 \dots i_p k i_{p+1} \dots i_q l i_{q+1} \dots i_n} = T_{i_1 \dots i_n} \epsilon_{kl}$$

and since the $SU(2)$ transformation acts trivially in the indices k, l they can be simply eliminated and the rank of the tensor reduced by two.

An additional simplification is that the fundamental representation is equivalent to its contragredient: the condition for the unity of determinant

$$\epsilon_{i_1 i_2} A_{i_1 j_1} A_{i_2 j_2} = \epsilon_{j_1 j_2}$$

can be written in matrix form as

$$A^T \epsilon A = \epsilon$$

which can be rearranged as

$$(A^T)^{-1} = \epsilon A \epsilon^{-1} \quad (\text{A.1})$$

so the two representations are equivalent - they are related by a basis transformation defined by ϵ . Note that the rank 2 Levi-Civita symbol satisfies the following matrix properties:

$$\epsilon^T = \epsilon^{-1} = -\epsilon$$

Therefore for a construction of the representations of $SU(2)$ it is sufficient to consider only symmetric tensors. The two examples of the previous subsection can be specialised to the case $n = 2$ as follows:

1. $(2, 0)$ tensors

Such tensors can be decomposed into a symmetric and antisymmetric part:

$$\begin{aligned} T_{ij} &= T_{ij}^+ + T_{ij}^- \\ T_{ij}^+ &= \frac{1}{2} (T_{ij} + T_{ji}) \quad T_{ij}^- = \frac{1}{2} (T_{ij} - T_{ji}) \end{aligned}$$

The representation formed by the symmetric component is an 3-dimensional irreducible representation, while the second one is one-dimensional and can be written in the form

$$T_{ij}^- = t \epsilon_{ij}$$

The representation of $SU(2)$ acts as

$$t' \epsilon_{ij} = A_{ii'} A_{jj'} t \epsilon_{ij}$$

We can then read off

$$t' = (\det A) t = t$$

So the antisymmetric part forms the trivial representation of $SU(2)$.

2. $(1, 1)$ tensors

In this case the two indices transform differently:

$$T'_{i\bar{j}} = A_{ii'} T_{i'j'} (A^{-1})_{j'j}$$

so there is no permutation symmetry to use for decomposition. Instead, this representation has two irreducible components. One of them corresponds to the traceless part

$$\tilde{T}_{i\bar{j}} = T_{i\bar{j}} - \frac{1}{2} (T_{1\bar{1}} + T_{2\bar{2}}) \delta_{i\bar{j}}$$

which forms a three-dimensional representation (since a traceless 2×2 matrix has 3 independent components), while the trace

$$\text{Tr } T = T_{1\bar{1}} + T_{2\bar{2}}$$

itself is invariant so it transforms as the trivial representation.

A.4.1 Irreducible representations of $SU(2)$

Now let us consider the representations of $SU(2)$. According to the above properties and the main theorem, all irreducible tensors can be obtained as the fully symmetric component of $\mathcal{T}^{(r,0)}$. These can be written as

$$T_{i_1 \dots i_r}$$

where the tensor is symmetric under all permutations of its indices.

In the language of quantum mechanics, the fundamental representation corresponds to a spin- $1/2$ degree of freedom, with the basis vectors interpreted as spin-up and down along some chosen direction z :

$$e_1 = |\uparrow\rangle \quad e_2 = |\downarrow\rangle$$

with the spin projection being $S^z = \pm 1/2$.

Due to permutational symmetry, the independent components of the tensor $T_{i_1 \dots i_r}$ can be labelled by the number of indices taking the value 1 with the rest taking value 2. There are exactly $r + 1$ independent choices, which correspond to the following spin states

$$\begin{aligned} T_{11\dots 11} &: |\uparrow\uparrow \dots \uparrow\uparrow\rangle & S^z &= r/2 \\ T_{11\dots 12} &: |\uparrow\uparrow \dots \uparrow\downarrow\rangle & S^z &= r/2 - 1 \\ &\vdots \\ T_{12\dots 22} &: |\uparrow\downarrow \dots \downarrow\downarrow\rangle & S^z &= -r/2 + 1 \\ T_{22\dots 22} &: |\downarrow\downarrow \dots \downarrow\downarrow\rangle & S^z &= -r/2 \end{aligned}$$

Here we omitted symmetrisation in the positions of the spins which is implied by the symmetry of the tensor; e.g. the second state when properly written corresponds to

$$\frac{1}{\sqrt{r}} (|\downarrow\uparrow \dots \uparrow\uparrow\rangle + |\uparrow\downarrow \dots \uparrow\uparrow\rangle + \dots + |\uparrow\uparrow \dots \downarrow\uparrow\rangle + |\uparrow\uparrow \dots \uparrow\downarrow\rangle)$$

since

$$T_{21\dots 11} = T_{12\dots 11} = \dots = T_{11\dots 21} = T_{11\dots 12}$$

Defining $j = r/2$, this is nothing else but the states of a spin- j system, and gives the irreducible representation of $SU(2)$ corresponding to spin j , with

$$j = 0, \frac{1}{2}, 1, \dots$$

This means that one can construct all the irreducible representations of $SU(2)$ using only $\mathcal{T}^{(r,0)}$ tensors.

Note that this can easily be extended to construct representations of $U(2)$. The determinant of any unitary matrix $A \in U(2)$ is a phase

$$\det A = e^{2i\phi}$$

We see that any such matrix can be written as the product of a phase and an element of $\tilde{A} = SU(2)$ as follows:

$$A = e^{i\phi} \tilde{A}$$

so we have the isomorphism

$$U(2) \simeq U(1) \times SU(2)$$

Now any representation of $U(1)$ is one-dimensional; they can be parametrised by a real number $t \in \mathbb{R}$, so that a group element $e^{it\phi} \in U(1)$ can be represented by multiplication with a phase $e^{it\phi}$. Combining these representations with those of $SU(2)$ we get a general irreducible representation of $U(1)$ parametrised by (t, j) . It is described by a fully symmetric tensor

$$T_{i_1 \dots i_{2j}}$$

which transforms under the unitary $A \in U(2)$ decomposed in the form

$$A = e^{it\phi} \tilde{A} \quad \det \tilde{A} = 1$$

as

$$T'_{i'_1 \dots i'_{2j}} = e^{it\phi} \tilde{A}_{i'_1 i_1} \dots \tilde{A}_{i'_{2j} i_{2j}} T_{i_1 \dots i_{2j}}$$

A.4.2 Relation between $SU(2)$ and $SO(3)$

Let us now look at the 3-dimensional representation of $SU(2)$, which is given by a symmetric tensor with 2 indices:

$$T'_{ij} = A_{ii'} A_{jj'} T_{i'j'}$$

i.e.

$$T' = A T A^T$$

Using

$$A^T \epsilon A = \epsilon \Rightarrow A^T = \epsilon^{-1} A^{-1} \epsilon$$

we can write

$$T' \epsilon^{-1} = A (T \epsilon^{-1}) A^{-1}$$

so we can rearrange T into a new tensor $V = T \epsilon^{-1}$ which transforms as

$$V' = A V A^{-1}$$

In components

$$T = \begin{pmatrix} T_{11} & T_{12} \\ T_{12} & T_{22} \end{pmatrix} \Rightarrow V = \begin{pmatrix} T_{12} & -T_{11} \\ T_{22} & -T_{12} \end{pmatrix}$$

Note that $\text{Tr } V = 0$, therefore V can be written as

$$V = \sum_{k=1}^3 v_k \sigma_k \quad v_k \in \mathbb{C}$$

with $\det V = -(v_1^2 + v_2^2 + v_3^2)$. The action $V' = A V A^{-1}$ of $A \in SU(2)$ can be written as

$$\sum_{k=1}^3 v'_k \sigma_k = A \left(\sum_{k=1}^3 v_k \sigma_k \right) A^{-1} = \sum_{k=1}^3 v_k A \sigma_k A^{-1}$$

Since $A \sigma_k A^{-1}$ are Hermitian and traceless, there exists a real matrix $\mathcal{O}(A)_{lk}$ such that

$$A \sigma_k A^{-1} = \sum_{l=1}^3 \sigma_l \mathcal{O}(A)_{lk} \tag{A.2}$$

Now we have

$$\sum_{l=1}^3 v'_l \sigma_l = \sum_{l=1}^3 \sigma_l \sum_{k=1}^3 \mathcal{O}(A)_{lk} v_k$$

i.e.

$$v'_l = \sum_{k=1}^3 \mathcal{O}(A)_{lk} v_k$$

Due to

$$V' = AVA^{-1} \Rightarrow \det V' = \det V$$

we have

$$v_1'^2 + v_2'^2 + v_3'^2 = v_1^2 + v_2^2 + v_3^2$$

so the matrix $\mathcal{O}(A)$ is a 3×3 orthogonal matrix. With some more effort one can show its determinant is +1 and so it is an element of $SO(3)$. Therefore the 3-dimensional representation of $SU(2)$ is the same as the complexified version of the defining representation of $SO(3)$. It is complexified due to the vector $\vec{v} = (v_1, v_2, v_3)$ being complex; however its real and imaginary parts transform independently due to $\mathcal{O}(A)$ being real, so the representation can be restricted to real vectors $\vec{v}$.

This correspondence can be stated in the following way. The 3-dimensional group $SO(3)$ leaves the 3-dimensional scalar product

$$\vec{u} \cdot \vec{v} = u_i v_i = u_i v_j \delta_{ij}$$

The corresponding metric is given by δ_{ij} , so the Clifford algebra corresponding to $SO(3)$ is defined by the anti-commutation relations

$$\{\Gamma_i, \Gamma_j\} = 2\delta_{ij}\mathbb{I}$$

By Clifford's theorem, all irreducible matrix solutions to this relations are related by similarity transformation, so all of them are equivalent up to redefining the basis. This is solved by the Pauli matrices i.e.

$$\Gamma_i = \sigma_i$$

As a result, spinors corresponding to 3-dimensional rotation group form a 2-dimensional complex space. The spinor transformation corresponding to rotations can be obtained by embedding the 3-vectors into spinor space as

$$\hat{x} = \vec{x} \cdot \vec{\sigma} = \sum_{k=1}^3 x_k \sigma_k$$

Note that $\hat{x}$ is a Hermitian matrix. The spinor transformation $S(\mathcal{O})$ corresponding to a rotation $\mathcal{O} \in SO(3)$ is defined from the relation

$$\widehat{\mathcal{O}x} = S(\mathcal{O})\hat{x}S(\mathcal{O})^\dagger$$

Note that the right hand side is automatically hermitian as it should be. Orthogonality of $\mathcal{O}$ implies that $\det \widehat{\mathcal{O}x} = \det \hat{x}$, which can be guaranteed by $S(\mathcal{O})^\dagger = S(\mathcal{O})^{-1}$ i.e. $S(\mathcal{O}) \in SU(2)$. In components the above relation is

$$\sum_{k,l=1}^3 \mathcal{O}_{kl} x_l \sigma_k = \sum_{l=1}^3 S(\mathcal{O}) x_l \sigma_l S(\mathcal{O})^\dagger$$

i.e.

$$\sum_{k=1}^3 \mathcal{O}_{kl} \sigma_k = S(\mathcal{O}) \sigma_l S(\mathcal{O})^\dagger$$

which is the same correspondence as (A.2). Note that this relation always has two independent solutions: if $S(\mathcal{O})$ is a solution, then $-S(\mathcal{O})$ works as well. So the correspondence between $SU(2)$ and $SO(3)$ is 2-to-1, which means that $SU(2)$ is a double covering of $SO(3)$.

We can also obtain an explicit solution for $S(\mathcal{O})$. Let us first consider the following element of $SU(2)$:

$$A_3(\varphi) = e^{-i\varphi\sigma_3/2} = \begin{pmatrix} e^{-i\varphi/2} & 0 \\ 0 & e^{i\varphi/2} \end{pmatrix}$$

A simple computation gives

$$\begin{aligned} A_3(\varphi) \sigma_1 A_3(\varphi)^{-1} &= \sigma_1 \cos \varphi - \sigma_2 \sin \varphi \\ A_3(\varphi) \sigma_2 A_3(\varphi)^{-1} &= \sigma_1 \sin \varphi + \sigma_2 \cos \varphi \\ A_3(\varphi) \sigma_3 A_3(\varphi)^{-1} &= \sigma_3 \end{aligned}$$

so

$$A_3(\varphi) \sigma_k A_3(\varphi)^{-1} = \sum_{l=1}^3 \sigma_l \mathcal{O}_3(\varphi)_{lk}$$

where

$$\mathcal{O}_3(\varphi) = \begin{pmatrix} \cos \varphi & -\sin \varphi & 0 \\ \sin \varphi & \cos \varphi & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

which is the rotation around axis 3 with angle φ .

We can replace axis 3 by a general direction given by a unit vector $\vec{e}$ and denote by $\mathcal{O}(\varphi\vec{e})$ the rotation around $\vec{e}$ with angle φ . Then the corresponding spinor transformation is given by

$$S(\mathcal{O}(\varphi\vec{e})) = \pm e^{-i\varphi\vec{e}\cdot\vec{\sigma}/2}$$

Denoting one of the solutions

$$A_{\vec{e}}(\varphi) = e^{-i\varphi\vec{e}\cdot\vec{\sigma}/2}$$

it is easy to see that the other solution can be written as

$$A_{\vec{e}}(\varphi + 2\pi) = -A_{\vec{e}}(\varphi)$$

Therefore rotating real space by an angle 2π (which is identity in terms of real rotations $SO(3)$) corresponds to a sign change in spinors. In particular

$$A_{\vec{e}}(2\pi) = -A_{\vec{e}}(0) = -\mathbb{I}$$

This is the reflection of $SU(2)$ being a double covering of $SO(3)$.

A.4.3 The Lie algebra $su(2)$ and its representations

Any matrix $A \in SU(2)$ can be written in the form

$$A = e^{-iH}$$

where

$$AA^\dagger = \mathbb{I} : H = H^\dagger$$

$$\det A = 1 : \text{Tr } H = 0$$

The matrices H span the Lie algebra $su(2)$, and can be parametrised by 3 independent real parameters and can be written in the form

$$H = \sum_{k=1}^3 \epsilon_k \frac{\sigma_k}{2}$$

where

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

$$\sigma_k \sigma_l = \delta_{kl} \mathbb{I} + \epsilon_{klm} \sigma_m$$

Therefore the group elements can be written as

$$A = \exp \left(-i \sum_{k=1}^3 \epsilon_k t_k \right)$$

where $\epsilon_k \in \mathbb{R}$ are the parameters of the group, while

$$t_k = \frac{\sigma_k}{2}$$

are its generators, which satisfy the Lie algebra relations

$$[t_k, t_l] = i \epsilon_{klm} t_m$$

In quantum theory these are the commutation relations satisfied by the three component of angular momentum; therefore, angular momentum operators form a representation of $su(2)$.

Note that according to our previous result, the spinor transformation $A_{\vec{e}}(\varphi) = e^{-i\varphi \vec{e} \cdot \vec{\sigma}/2}$ corresponding to the rotation $\mathcal{O}(\varphi \vec{e})$ around the unit vector $\vec{e}$ with angle φ , is given by the parameter choice

$$\vec{\epsilon} = \varphi \vec{e}$$

Representations of the Lie group always correspond to representation of the Lie algebra. Take a representation of $SU(2)$ on some vector space V :

$$A \in SU(2) \mapsto D(A) : V \rightarrow V$$

We can then consider infinitesimal group elements, i.e. assuming that $\epsilon_k \in \mathbb{R}$ is small we can write

$$A = \exp \left(-i \sum_{k=1}^3 \epsilon_k t_k \right) = \mathbb{I} - i \sum_{k=1}^3 \epsilon_k t_k + O(\epsilon^2)$$

Expanding the representation matrices the same way

$$D(A) = \mathbb{I}_V - i \sum_{k=1}^3 \epsilon_k J_k + O(\epsilon^2)$$

we can identify

$$J_k = D(t_k)$$

as the representation of the Lie algebra induced by D . Note that applying the group representation property

$$D(A_1) D(A_2) = D(A_1 A_2)$$

to infinitesimal group elements implies that J_k satisfy the same relations as the t_k i.e.

$$[J_k, J_l] = i\epsilon_{klm} J_m$$

The irreducible representations of these commutation relations are all known and can be obtained by introducing the step operators

$$\begin{aligned} J_{\pm} &= J_1 \pm iJ_2 \\ [J_3, J_{\pm}] &= \pm J_{\pm} \\ [J_+, J_-] &= 2J_3 \end{aligned}$$

Then the irreducible representation corresponding to spin $j \in \frac{1}{2}\mathbb{N}$ has a basis spanned by the $2j+1$ vectors

$$|j, m\rangle \quad m = -j, -j+1, \dots, j$$

with the following action

$$\begin{aligned} J_3 |j, m\rangle &= m |j, m\rangle \\ J_+ |j, m\rangle &= \sqrt{(j-m)(j+m+1)} |j, m+1\rangle \\ J_- |j, m\rangle &= \sqrt{(j+m)(j-m+1)} |j, m-1\rangle \end{aligned} \tag{A.3}$$

A.4.4 $su(2)$ representations in the tensor product method

Let us assume we take n representations of the Lie group $SU(2)$

$$A \in SU(2) \mapsto D_k(A) : V_k \rightarrow V_k \quad k = 1, \dots, n$$

Then the group acts on the tensor product space $V = V_1 \otimes \dots \otimes V_n$ according to the representation

$$A \in SU(2) \mapsto D(A) = D_1(A) \otimes \dots \otimes D_n(A) : V_1 \otimes \dots \otimes V_n \rightarrow V_1 \otimes \dots \otimes V_n$$

and denote Now we can get the infinitesimal form the tensor product action by considering

$$\left(\mathbb{I}_{V_1} - i \sum_{k=1}^3 \epsilon_k D_1(t_k) \right) \otimes \dots \otimes \left(\mathbb{I}_{V_n} - i \sum_{k=1}^3 \epsilon_k D_n(t_k) \right) = \mathbb{I}_V - i \sum_{k=1}^3 \epsilon_k D(t_k) + O(\epsilon^2)$$

where $D(t_k) = D_1(t_k) \otimes \mathbb{I}_{V_2} \otimes \dots \otimes \mathbb{I}_{V_n} + \mathbb{I}_{V_1} \otimes D_2(t_k) \otimes \dots \otimes \mathbb{I}_{V_n} + \dots + \mathbb{I}_{V_1} \otimes \mathbb{I}_{V_2} \otimes \dots \otimes D_n(t_k)$

So the action of a Lie algebra element in the product representation is obtained by summing over terms in which it terms on individual factors separately.

Let us now use this knowledge to make a connection with the tensor product method. Consider the product of $2j$ copies of the spin-1/2 representation. The highest vector in this representation has $m = j$ and corresponds to a spin- j multiplet:

$$|j, j\rangle = \underbrace{|\uparrow\uparrow \dots \uparrow\rangle}_{2j \text{ spins up}}$$

In the tensor notation, this corresponds to the tensor component with all indices taking the value 1:

$$T_{11\dots 1}$$

Now we can act with T_- once. According to the above rule, it acts on all components separately, meaning that it turns one spin down in all possible ways:

$$T_- |j, j\rangle = |\downarrow\uparrow \dots \uparrow\rangle + |\uparrow\downarrow \dots \uparrow\rangle + \dots + |\uparrow\uparrow \dots \downarrow\rangle$$

where in each term we have exactly one spin down. In terms of the tensor notation, this corresponds to the tensor component with exactly one index taking the value 2, completely symmetrised:

$$\mathcal{T}_{21\dots 1} + \mathcal{T}_{12\dots 1} + \mathcal{T}_{11\dots 2}$$

Note that the properly normalised basis vector can be obtained by comparing this result to

$$T_- |j, j\rangle = \sqrt{2j} |j, j-1\rangle \quad (\text{A.4})$$

resulting from (A.3), so

$$|j, j-1\rangle = \frac{1}{\sqrt{2j}} (|\downarrow\uparrow \dots \uparrow\rangle + |\uparrow\downarrow \dots \uparrow\rangle + \dots + |\uparrow\uparrow \dots \downarrow\rangle)$$

Going one further, we can act with T_- once more. Again we must act on each spin separately in all terms, but the terms where we act on a spin already down give zero. Therefore we get all possible terms with any of the two spins down with all others up:

$$T_-^2 |j, j\rangle = 2 \sum_{1 \leq k < l \leq n} |\dots \downarrow_k \dots \downarrow_l \dots\rangle$$

The factor of 2 comes from the two possible orders of turning down the k th and the l th spin. This corresponds to the tensor component with exactly two indices taking the value 2, completely symmetrised. Using (A.3)

$$T_- |j, j-1\rangle = \sqrt{(2j-1)2} |j, j-2\rangle$$

which upon substituting (A.4) leads to

$$T_-^2 |j, j\rangle = \sqrt{2j(2j-1)2} |j, j-2\rangle$$

and we obtain

$$\begin{aligned} |j, j-2\rangle &= \sqrt{\frac{2}{2j(2j-1)}} \sum_{1 \leq k < l \leq n} |\dots \downarrow_k \dots \downarrow_l \dots\rangle \\ &= \frac{1}{\sqrt{\binom{2j}{2}}} \sum_{1 \leq k < l \leq n} |\dots \downarrow_k \dots \downarrow_l \dots\rangle \end{aligned}$$

Now the general pattern is

$$T_-^r |j, j\rangle = r! \sum_{1 \leq k_1 < \dots < k_r \leq n} |\dots \downarrow_{k_1} \dots \downarrow_{k_r} \dots\rangle$$

which corresponds to the symmetrisation of the $2j$ -index tensor components with exactly r taking the value 2. On the other hand

$$T_-^r |j, j\rangle = \sqrt{2j(2j-1) \dots (2j-r+1)} r! |j, j-r\rangle$$

so we obtain

$$\begin{aligned} |j, j-r\rangle &= \sqrt{\frac{r!}{2j(2j-1) \dots (2j-r+1)}} \sum_{1 \leq k_1 < \dots < k_r \leq n} |\dots \downarrow_{k_1} \dots \downarrow_{k_r} \dots\rangle \\ &= \frac{1}{\sqrt{\binom{2j}{r}}} \sum_{1 \leq k_1 < \dots < k_r \leq n} |\dots \downarrow_{k_1} \dots \downarrow_{k_r} \dots\rangle \end{aligned}$$

A.5 Tensor method for $SU(n)$

A.5.1 Reducing the problem to $\mathcal{T}^{(N,0)}$

First we show that the contragredient of the fundamental representation is contained in $\mathcal{T}^{(n-1,0)}$. We start by noting that the completely antisymmetric part of this tensor space has exactly

$$\binom{n}{n-1} = n$$

components, since all $n-1$ indices must be different and can be chosen from n possible values, with different orderings giving tensor elements that are related to each other by an appropriate sign (equal to the signature of the permutation which relates the two orderings).

Taking an element $A \in SU(n)$ we can write

$$\epsilon_{i_1 \dots i_n} A_{i_1 j_1} \dots A_{i_n j_n} = \epsilon_{j_1 \dots j_n}$$

and so

$$\begin{aligned} \epsilon_{i_1 \dots i_n} A_{i_1 j_1} \dots A_{i_{n-1} j_{n-1}} &= \epsilon_{j_1 \dots j_n} (A^{-1})_{j_n i_n} \\ &= A_{i_n j_n}^* \epsilon_{j_1 \dots j_n} \end{aligned}$$

Consider now the completely antisymmetric component of $\mathcal{T}^{(n-1,0)}$ which is described by a tensor

$$T_{[i_1 \dots i_{n-1}]}$$

where the brackets denote the antisymmetry in the indices. This has exactly n components $\bar{T}_k$ ($k = 1, \dots, n$) and can be written as

$$T_{[i_1 \dots i_{n-1}]} = \epsilon_{i_1 \dots i_{n-1} k} \bar{T}_k$$

The inverse of this relation is

$$\bar{T}_k = \frac{1}{(n-1)!} \epsilon_{i_1 \dots i_{n-1} k} T_{[i_1 \dots i_{n-1}]}$$

Now consider transforming the tensor with $A \in SU(n)$:

$$T'_{[i_1 \dots i_{n-1}]} = A_{i_1 j_1} \dots A_{i_{n-1} j_{n-1}} T_{[j_1 \dots j_{n-1}]}$$

Then we obtain

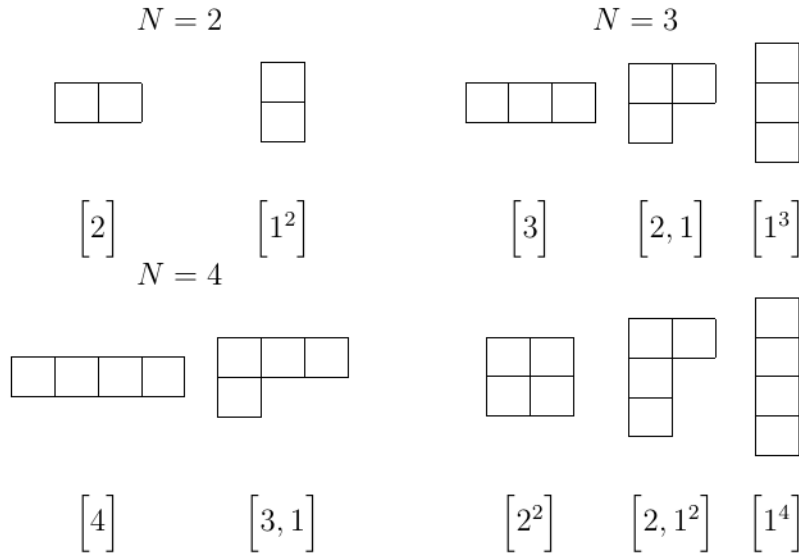
$$\begin{aligned} \bar{T}'_k &= \frac{1}{(n-1)!} \epsilon_{i_1 \dots i_{n-1} k} T'_{[i_1 \dots i_{n-1}]} \\ &= \frac{1}{(n-1)!} \epsilon_{i_1 \dots i_{n-1} k} A_{i_1 j_1} \dots A_{i_{n-1} j_{n-1}} T_{[j_1 \dots j_{n-1}]} \\ &= \frac{1}{(n-1)!} A_{kl}^* \epsilon_{j_1 \dots j_{n-1} l} T_{[j_1 \dots j_{n-1}]} \\ &= A_{kl}^* \bar{T}_l \end{aligned}$$

Therefore the tensor $\bar{T}_k$ belongs to the space $\mathcal{T}^{(0,1)}$: the contragredient of the fundamental representation is identical to the fully antisymmetric component of the product of $n-1$ fundamental representations. In order to construct all representations we can therefore restrict our attentions to tensors $\mathcal{T}^{(N,0)}$ which are products of the fundamental representation only.

A.5.2 Tensor symmetry classes and Young tableaux

Let us consider the action of $SU(n)$ on the tensor product of N copies of the fundamental representation, i.e. the space of tensors $\mathcal{T}^{(N,0)}$. This space carries an action of the permutation group S_N , which is generally reducible. As we have seen, decomposing this space into irreducible representations of $SU(n)$ is essentially equivalent to decomposing it into irreducible subspaces under the action of S_N .

Now we recall that irreducible representations of S_N can be labelled by Young tableaux, examples of which are shown below.



The Young tableaux of S_N have N boxes arranged into rows whose length monotonically decreases from top to bottom. When considering a tensor $T_{i_1 \dots i_N}$, the standard fillings of the tableaux dictated by the representation theory of the permutation group are obtained by writing the i_k into the boxes in such a way that k increases from top to bottom in each column and from left to right in each row. The corresponding symmetry class is given by anti-symmetrising over indices occurring in each column, and then symmetrising the tensor over the indices which occur in the same row. Let us consider the cases $N = 2$ and $N = 3$ to illustrate the decomposition.

1. $N = 2$ has two symmetry classes:

(a) Class $[2]$: the only filling of the Young tableaux is

$$\begin{array}{|c|c|} \hline i_1 & i_2 \\ \hline \end{array}$$

corresponding to the symmetrised tensor

$$T_{i_1 i_2} + T_{i_2 i_1}$$

The dimension of this representation is

$$\frac{n(n+1)}{2}$$

(b) Class $[1^2]$ has the filling

$$\begin{array}{|c|} \hline i_1 \\ \hline i_2 \\ \hline \end{array}$$

which gives the antisymmetric tensor

$$T_{i_1 i_2} - T_{i_2 i_1}$$

with dimension

$$\frac{n(n-1)}{2}$$

2. $N = 3$

(a) Class $[3]$ with only one standard filling

$$\begin{array}{|c|c|c|} \hline i_1 & i_2 & i_3 \\ \hline \end{array} : \text{fully symmetric tensor } T_{i_1 i_2 i_3} + T_{i_2 i_1 i_3} + T_{i_3 i_2 i_1} + T_{i_1 i_3 i_2} + T_{i_3 i_1 i_2} + T_{i_2 i_3 i_1}$$

with dimension

$$\frac{n(n+1)(n+2)}{6}$$

(b) Class $[2, 1]$ with two standard fillings

$$\text{i. } \begin{array}{|c|c|} \hline i_1 & i_2 \\ \hline i_3 & \end{array} : \text{mixed symmetry tensor } T_{i_1 i_2 i_3} - T_{i_3 i_2 i_1} + T_{i_2 i_1 i_3} - T_{i_3 i_1 i_2}$$

with dimension

$$\frac{n(n+1)(n-1)}{3}$$

ii. $\begin{array}{|c|c|} \hline i_1 & i_3 \\ \hline i_2 & \\ \hline \end{array}$: mixed symmetry tensor $T_{i_1 i_2 i_3} - T_{i_2 i_1 i_3} + T_{i_3 i_2 i_1} - T_{i_2 i_3 i_1}$

with dimension

$$\frac{n(n+1)(n-1)}{3}$$

(c) Class $[1^3]$ with only one standard filling

$\begin{array}{|c|} \hline i_1 \\ \hline i_2 \\ \hline i_3 \\ \hline \end{array}$: fully antisymmetric tensor $T_{i_1 i_2 i_3} - T_{i_2 i_1 i_3} - T_{i_3 i_2 i_1} - T_{i_1 i_3 i_2} + T_{i_3 i_1 i_2} + T_{i_2 i_3 i_1}$

with dimension

$$\frac{n(n-1)(n-2)}{6}$$

The total dimension works out as follows:

$$\frac{n(n+1)(n+2)}{6} + 2\frac{n(n+1)(n-1)}{3} + \frac{n(n-1)(n-2)}{6} = n^3$$

The contragredient fundamental representation of $SU(n)$ corresponds to the diagram $[1^{n-1}]$, i.e. $n-1$ boxes arranged in a single column. Note that it does not make sense to consider more than n rows since anti-symmetrisation over more than n indices would vanish. Columns made of exactly n boxes can be eliminated from the diagram since anti-symmetrising over n indices each running from 1 to n gives the n index Levi-Civita as a factor in the tensor which can be eliminated analogously to $SU(2)$, so to obtain all the inequivalent irreducible representation it is enough to consider diagrams with $n-1$ rows only.

For the case of $SU(2)$ the only surviving Young tableau are composed of one single row, giving the fully symmetric tensors as the irreducible representations of $SU(2)$.

A.5.3 The adjoint representation

An interesting irreducible representation is obtained by considering the product of the fundamental and contragredient fundamental representation, corresponding to a tensor $M_{i\bar{j}}$ which transforms as

$$M'_{i\bar{j}} = A_{ii'} A_{\bar{j}\bar{j}}^* M_{i'\bar{j}'}$$

which can also be written in the matrix form

$$M' = A M A^\dagger = A M A^{-1}$$

This can be decomposed in the traceless and the trace part, both of which are irreducible under the action of the group. The trace part is invariant so it is the fundamental representation, while the traceless part

$$\tilde{M}_{i\bar{j}} = M_{i\bar{j}} - \frac{1}{n} (\text{Tr } M) \delta_{i\bar{j}}$$

is an $n^2 - 1$ dimensional representation, which form a complex vector space. These are $n \times n$ complex matrices which can be decomposed into hermitian and anti-hermitian parts:

$$\tilde{M} = \frac{1}{2} (\tilde{M} + \tilde{M}^\dagger) + \frac{1}{2} (\tilde{M} - \tilde{M}^\dagger)$$

The transformation $M' = AMA^\dagger$ keeps this decomposition invariant, so the representation can be restricted to the Hermitian part, which is now a real vector space (only real linear combinations of hermitian matrices preserve Hermiticity). The Young tableau for this representation can be shown to be the diagram $[2, 1^{n-2}]$ (one row of length 2 and $n - 2$ rows of length 1):



Just like for $SU(2)$, any $A \in SU(n)$ can be written as

$$A = e^{-iH}$$

with H a hermitian traceless matrix which has $n^2 - 1$ independent real parameters. Therefore the Lie algebra $su(n)$ of the group $SU(n)$ is exactly the space of hermitian traceless matrices, and so the real version of the above $n^2 - 1$ dimensional representation gives an action of the group on its Lie algebra. This is called the adjoint representation, which is irreducible for $SU(n)$. Lie groups for which the adjoint representation is irreducible are called simple, so $SU(n)$ is a simple Lie group; another example of a simple Lie group is $SO(n)$ with $n \geq 3$ with the exception of $n = 4$ ⁵⁵.

There are exactly $n^2 - 1$ linearly independent traceless hermitian matrices λ_a which are normally chosen to be orthogonal under the matrix inner product defined by the trace operation:

$$\text{Tr } \lambda_a \lambda_b = 2\delta_{ab}$$

(the factor 2 is included for conformity with the $SU(2)$ case where the λ_a are the Pauli matrices). The general group element can then be written as

$$A = \exp \left(-i \sum_{a=1}^{n^2-1} \epsilon_a \frac{\lambda_a}{2} \right)$$

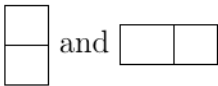
with $\epsilon_a \in \mathbb{R}$.

A.6 $SU(2)$ examples

Since there is only one irreducible representation for every dimension d , they can be denoted as $\mathbf{d}$, so we write $\mathbf{1}$ for spin-0, $\mathbf{2}$ for spin-1/2, $\mathbf{3}$ for spin-1, $\mathbf{4}$ for spin-3/2 etc. Now let us consider the irreducible components of products of the fundamental representation $\mathbf{2}$.

1. $\mathbf{2} \otimes \mathbf{2} = \mathbf{1} \oplus \mathbf{3}$

There are two Young tableaux



⁵⁵ $SO(4)$ has the same Lie algebra as the product group $SU(2) \times SU(2)$ and its 6-dimensional adjoint representation decomposes into two invariant subspaces of dimension 3.

The first one corresponds to the antisymmetric part, which has only a single non-trivial component corresponding to the spin state

$$\frac{1}{\sqrt{2}} (|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle)$$

forming a spin-0 (singlet), while the second one is the symmetric part with three independent components corresponding to the states

$$\begin{aligned} &|\uparrow\uparrow\rangle \\ &\frac{1}{\sqrt{2}} (|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle) \\ &|\downarrow\downarrow\rangle \end{aligned}$$

and describes a spin-1 (triplet).

2. $\mathbf{2} \otimes \mathbf{2} \otimes \mathbf{2} = \mathbf{2} \oplus \mathbf{2} \oplus \mathbf{4}$

There are now three Young tableaux:

$$\begin{array}{|c|} \hline \square \\ \hline \square \\ \hline \square \\ \hline \end{array}, \quad \begin{array}{|c|c|} \hline \square & \square \\ \hline \square & \\ \hline \end{array} \text{ and } \begin{array}{|c|c|c|} \hline \square & \square & \square \\ \hline \end{array}$$

- (a) The first one corresponds to the completely antisymmetric part, which vanishes since the indices only take two values.
- (b) The second one corresponds to two mixed tensors

$$\begin{array}{|c|c|} \hline i_1 & i_2 \\ \hline i_3 & \\ \hline \end{array} [2, 1]_1 : T_{i_1 i_2 i_3} - T_{i_3 i_2 i_1} + T_{i_2 i_1 i_3} - T_{i_3 i_1 i_2}$$

$$\begin{array}{|c|c|} \hline i_1 & i_3 \\ \hline i_2 & \\ \hline \end{array} [2, 1]_2 : T_{i_1 i_2 i_3} - T_{i_2 i_1 i_3} + T_{i_3 i_2 i_1} - T_{i_2 i_3 i_1}$$

$[2, 1]_1$ can only have non-zero elements when $i_1 \neq i_3$, so the two non-trivial components are

$$\begin{aligned} &|\uparrow\uparrow\downarrow\rangle - |\downarrow\uparrow\uparrow\rangle + |\uparrow\uparrow\downarrow\rangle - |\downarrow\uparrow\uparrow\rangle \\ &|\uparrow\downarrow\downarrow\rangle - |\downarrow\downarrow\uparrow\rangle + |\downarrow\uparrow\downarrow\rangle - |\downarrow\uparrow\downarrow\rangle \end{aligned}$$

which results in the spin-1/2 doublet

$$\begin{aligned} [2, 1]_1 : & \frac{1}{\sqrt{2}} (|\uparrow\uparrow\downarrow\rangle - |\downarrow\uparrow\uparrow\rangle) \\ & \frac{1}{\sqrt{2}} (|\uparrow\downarrow\downarrow\rangle - |\downarrow\downarrow\uparrow\rangle) \end{aligned} \tag{A.5}$$

How does the representation work? On the generators t_k it is given by

$$J_k = D_{\mathbf{2} \otimes \mathbf{2} \otimes \mathbf{2}}(t_k) = \frac{\sigma_k}{2} \otimes \mathbb{I} \otimes \mathbb{I} + \mathbb{I} \otimes \frac{\sigma_k}{2} \otimes \mathbb{I} + \mathbb{I} \otimes \mathbb{I} \otimes \frac{\sigma_k}{2}$$

We can then easily compute

$$J_3 (|\uparrow\uparrow\downarrow\rangle - |\downarrow\uparrow\uparrow\rangle) = +\frac{1}{2} (|\uparrow\uparrow\downarrow\rangle - |\downarrow\uparrow\uparrow\rangle)$$

$$J_3 (|\uparrow\downarrow\downarrow\rangle - |\downarrow\downarrow\uparrow\rangle) = -\frac{1}{2} (|\uparrow\downarrow\downarrow\rangle - |\downarrow\downarrow\uparrow\rangle)$$

The step operators act as

$$J_+ (|\uparrow\uparrow\downarrow\rangle - |\downarrow\uparrow\uparrow\rangle) = |\uparrow\uparrow\uparrow\rangle - |\uparrow\uparrow\uparrow\rangle = 0$$

$$J_- (|\uparrow\uparrow\downarrow\rangle - |\downarrow\uparrow\uparrow\rangle) = |\downarrow\uparrow\downarrow\rangle + |\uparrow\downarrow\downarrow\rangle - |\downarrow\downarrow\uparrow\rangle - |\downarrow\uparrow\downarrow\rangle = |\uparrow\downarrow\downarrow\rangle - |\downarrow\downarrow\uparrow\rangle$$

$$J_+ (|\uparrow\downarrow\downarrow\rangle - |\downarrow\downarrow\uparrow\rangle) = |\uparrow\uparrow\downarrow\rangle + |\uparrow\downarrow\uparrow\rangle - |\uparrow\downarrow\uparrow\rangle - |\downarrow\uparrow\uparrow\rangle = |\uparrow\uparrow\downarrow\rangle - |\downarrow\uparrow\uparrow\rangle$$

$$J_- (|\uparrow\downarrow\downarrow\rangle - |\downarrow\downarrow\uparrow\rangle) = |\downarrow\downarrow\downarrow\rangle - |\downarrow\downarrow\downarrow\rangle = 0$$

so the two states (A.5) form the up and down component of a spin-1/2 doublet under the action of the Lie algebra $su(2)$.

$[2, 1]_2$ can only have non-zero elements when $i_1 \neq i_2$, so the two non-trivial components are

$$|\uparrow\downarrow\uparrow\rangle - |\downarrow\uparrow\uparrow\rangle + |\uparrow\downarrow\uparrow\rangle - |\downarrow\uparrow\uparrow\rangle$$

$$|\uparrow\downarrow\downarrow\rangle - |\downarrow\uparrow\downarrow\rangle + |\downarrow\downarrow\uparrow\rangle - |\downarrow\downarrow\uparrow\rangle$$

which results in the spin-1/2 doublet

$$[2, 1]_2 : \frac{1}{\sqrt{2}} (|\uparrow\downarrow\uparrow\rangle - |\downarrow\uparrow\uparrow\rangle)$$

$$\frac{1}{\sqrt{2}} (|\uparrow\downarrow\downarrow\rangle - |\downarrow\uparrow\downarrow\rangle)$$

One can again establish that these two states form the up and down component of a spin-1/2 doublet.

It is interesting to observe that in principle one could form a spin-1/2 object by projecting two spins to a singlet, which can be performed in three different ways:

Singlet combination	(12)	(13)	(23)
Spin-up state	$\frac{1}{\sqrt{2}} (\uparrow\downarrow\uparrow\rangle - \downarrow\uparrow\uparrow\rangle)$	$\frac{1}{\sqrt{2}} (\uparrow\uparrow\downarrow\rangle - \downarrow\uparrow\uparrow\rangle)$	$\frac{1}{\sqrt{2}} (\uparrow\uparrow\downarrow\rangle - \uparrow\downarrow\uparrow\rangle)$
Spin-down state	$\frac{1}{\sqrt{2}} (\uparrow\downarrow\downarrow\rangle - \downarrow\uparrow\downarrow\rangle)$	$\frac{1}{\sqrt{2}} (\uparrow\downarrow\downarrow\rangle - \downarrow\downarrow\uparrow\rangle)$	$\frac{1}{\sqrt{2}} (\downarrow\uparrow\downarrow\rangle - \downarrow\downarrow\uparrow\rangle)$

However, these three doublets are not independent: the states in (12) can be obtained by taking the difference of the corresponding states in (13) and (23). The two independent doublets $[2, 1]_1$ and $[2, 1]_2$ can be identified as (13) and (12), respectively.

- (c) $\square\square\square$ corresponds to the completely symmetric component, which comprises four states forming a spin-3/2 multiplet

$$|3/2, +3/2\rangle = |\uparrow\uparrow\uparrow\rangle$$

$$|3/2, +1/2\rangle = \frac{1}{\sqrt{3}} (|\downarrow\uparrow\uparrow\rangle + |\uparrow\downarrow\uparrow\rangle + |\uparrow\uparrow\downarrow\rangle)$$

$$|3/2, -1/2\rangle = \frac{1}{\sqrt{3}} (|\uparrow\downarrow\downarrow\rangle + |\downarrow\uparrow\downarrow\rangle + |\downarrow\downarrow\uparrow\rangle)$$

$$|3/2, -3/2\rangle = |\downarrow\downarrow\downarrow\rangle$$

It can be established by direct calculation that the action of the generators J_k on these states agrees with the $j = 3/2$ case of (A.3).

A.7 $SU(3)$ examples

For $SU(3)$, a standard choice of the basis in the Lie algebra $su(3)$ is given by the eight Gell-Mann matrices

$$\begin{aligned}\lambda_1 &= \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} & \lambda_2 &= \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} & \lambda_3 &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \\ \lambda_4 &= \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} & \lambda_5 &= \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix} & \lambda_6 &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \\ \lambda_7 &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix} & \lambda_8 &= \frac{1}{\sqrt{3}} \begin{pmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 2 \end{pmatrix}\end{aligned}\tag{A.6}$$

An element $A \in SU(3)$ can be written as

$$A = \exp \left(-i \sum_{a=1}^8 \epsilon_a \frac{\lambda_a}{2} \right)$$

The action of the matrix A on a 3-dimensional complex vector space is the fundamental representation $\mathbf{3}$. The Lie algebra $su(3)$ has 8 generators t_a and

$$D_{\mathbf{3}}(t_a) = \frac{\lambda_a}{2}$$

Two of the generators can be simultaneously diagonalised, here they are taken to be $I_3 = t_3$ (third component of isospin) and $Y = -2t_8/\sqrt{3}$ (hypercharge), under which the three basis vectors of $\mathbf{3}$ have the following eigenvalues:

	I_3	Y
$ 1\rangle$	$+1/2$	$+1/3$
$ 2\rangle$	$-1/2$	$+1/3$
$ 3\rangle$	0	$-2/3$

The contragredient representation $\bar{\mathbf{3}}$ is given by

$$(A^T)^{-1} = A^* = \exp \left(i \sum_{a=1}^8 \epsilon_a \frac{\lambda_a^T}{2} \right)$$

so the generators are represented as

$$D_{\bar{\mathbf{3}}}(t_a) = -\frac{\lambda_a^T}{2}$$

The three basis vectors of $\bar{\mathbf{3}}$ have the following eigenvalues:

	I_3	Y
$ 1\rangle$	$-1/2$	$-1/3$
$ 2\rangle$	$+1/2$	$-1/3$
$ 3\rangle$	0	$+2/3$

A.7.1 $\mathbf{3} \otimes \mathbf{3} = \bar{\mathbf{3}} \oplus \mathbf{6}$

There are two Young tableaux

$$\begin{array}{|c|} \hline \square \\ \hline \square \\ \hline \end{array} \text{ and } \begin{array}{|c|c|} \hline \square & \square \\ \hline \end{array}$$

The first one corresponds to the antisymmetric part, which is spanned by the three states

	I_3	Y	$\bar{\mathbf{3}}$
$\frac{1}{\sqrt{2}}(23\rangle - 32\rangle)$	$-1/2$	$-1/3$	$ \bar{1}\rangle$
$\frac{1}{\sqrt{2}}(31\rangle - 13\rangle)$	$+1/2$	$-1/3$	$ \bar{2}\rangle$
$\frac{1}{\sqrt{2}}(12\rangle - 21\rangle)$	0	$+2/3$	$ \bar{3}\rangle$

which form a representation $\bar{\mathbf{3}}$ with the identification of basis vectors given in the table. The second one is the symmetric part corresponding to a representation denoted $\mathbf{6}$.

A.7.2 $\mathbf{3} \otimes \bar{\mathbf{3}} = \mathbf{1} \oplus \mathbf{8}$

In this case we can avoid using the Young tableaux, instead we can simply decompose into a traceless and a trace part. The trace part is the singlet $\mathbf{1}$, spanned by the state

$$\frac{1}{\sqrt{3}}(|1\bar{1}\rangle + |2\bar{2}\rangle + |3\bar{3}\rangle)$$

The traceless part is the adjoint representation $\mathbf{8}$ which is spanned by the states

	I_3	Y		I_3	Y		I_3	Y		I_3	Y
$ 12\rangle$	$+1$	0	$ 13\rangle$	$+1/2$	1	$ 32\rangle$	$+1/2$	-1	$\frac{1}{\sqrt{2}}(1\bar{1}\rangle - 2\bar{2}\rangle)$	0	0
$\frac{1}{\sqrt{2}}(1\bar{1}\rangle - 2\bar{2}\rangle)$	0	0	$ 23\rangle$	$-1/2$	1	$ 31\rangle$	$-1/2$	-1			
$ 21\rangle$	-1	0									

The way we split up the representation into these tables is not accidental and can be understood using isospin symmetry. The first three generators t_1 , t_2 and t_3 form an $su(2)$ subalgebra which is identified with the isospin. Under this $su(2)$ the three states in the first table form an isospin-1 triplet, the states in the second and third table form an isospin doublet, and the state in the fourth table is a isospin singlet.

How do we know this? Let us look at the generators of the isospin subgroup in representation $\mathbf{3}$:

$$D_{\mathbf{3}}(t_1) = \frac{1}{2} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad D_{\mathbf{3}}(t_2) = \frac{1}{2} \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad D_{\mathbf{3}}(t_3) = \frac{1}{2} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

So $\mathbf{3}$ splits into an isospin doublet formed by $|1\rangle$ and $|2\rangle$, while $|3\rangle$ is an isospin singlet. One can write the following correspondence with the usual interpretation of $su(2)$ as a spin:

$$|1\rangle \leftrightarrow |\uparrow\rangle \quad |2\rangle \leftrightarrow |\downarrow\rangle$$

Similarly $\bar{\mathbf{3}}$ splits into an ‘anti-doublet’ (contragredient doublet) formed by $|\bar{1}\rangle$ and $|\bar{2}\rangle$, while $|\bar{3}\rangle$ is an isospin singlet. From Section [A.4](#) we know that the anti-doublet is related to the doublet by

a basis redefinition with the 2-index Levi-Civita ϵ , cf. eqn. (A.1) so we have

$$|\bar{1}\rangle \leftrightarrow |\downarrow\rangle \quad |\bar{2}\rangle \leftrightarrow -|\uparrow\rangle$$

Now we see that the three states in the first table can be interpreted in $su(2)$ language as

$$\begin{aligned} |1\bar{2}\rangle &\leftrightarrow -|\uparrow\uparrow\rangle \\ \frac{1}{\sqrt{2}}(|1\bar{1}\rangle - |2\bar{2}\rangle) &\leftrightarrow \frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle) \\ |2\bar{1}\rangle &\leftrightarrow -|\downarrow\downarrow\rangle \end{aligned}$$

Apart from some phase redefinitions, these are indeed the three components of an $su(2)$ (isospin) triplet.

The states in the next two tables obviously form isospin doublets, since $|3\rangle$ is a singlet. Finally, the state in the last table can be written as

$$\frac{1}{\sqrt{2}}(|1\bar{1}\rangle - |2\bar{2}\rangle) \leftrightarrow \frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle)$$

which is indeed the correct expression for an $su(2)$ (isospin) singlet.

A.7.3 $3 \otimes 3 \otimes 3 = 1 \oplus 8 \oplus 8 \oplus 10$

$3 \otimes 3 \otimes 3$ is represented by a 3-index tensor $T_{i_1 i_2 i_3}$, transforming as

$$T'_{i_1 i_2 i_3} = A_{i_1 j_1} A_{i_2 j_2} A_{i_3 j_3} T_{j_1 j_2 j_3} \quad A \in SU(3)$$

There are now three Young tableaux:

$$\begin{array}{|c|} \hline \square \\ \hline \square \\ \hline \square \\ \hline \end{array}, \quad \begin{array}{|c|c|} \hline \square & \square \\ \hline \square & \square \\ \hline \end{array} \text{ and } \begin{array}{|c|c|c|} \hline \square & \square & \square \\ \hline \end{array}$$

The first one is $[1^3]$ which is the fully antisymmetric component which is the trivial one-dimensional representation since the fully antisymmetric component has only one component proportional to the three-index Levi-Civita tensor, which is invariant under $SU(3)$ due to $\det A = 1$ for all $A \in SU(3)$. The third one is the fully symmetric component **10**, which decomposes under isospin as follows:

Isospin 3/2 (quadruplet)	I_3	Y
$ 111\rangle$	$+3/2$	$+1$
$\frac{1}{\sqrt{3}}(112\rangle + 121\rangle + 211\rangle)$	$+1/2$	$+1$
$\frac{1}{\sqrt{3}}(122\rangle + 212\rangle + 221\rangle)$	$-1/2$	$+1$
$ 222\rangle$	$-3/2$	$+1$
Isospin 1/2 (doublet)	I_3	Y
$\frac{1}{\sqrt{3}}(133\rangle + 313\rangle + 331\rangle)$	$+1/2$	-1
$\frac{1}{\sqrt{3}}(233\rangle + 323\rangle + 332\rangle)$	$-1/2$	-1

Isospin 1 (triplet)	I_3	Y
$\frac{1}{\sqrt{3}}(113\rangle + 131\rangle + 311\rangle)$	$+1$	0
$\frac{1}{\sqrt{6}}(123\rangle + 213\rangle + 132\rangle + 231\rangle + 312\rangle + 321\rangle)$	0	0
$\frac{1}{\sqrt{3}}(223\rangle + 232\rangle + 322\rangle)$	-1	0

Isospin 0 (singlet)	I_3	Y
$ 333\rangle$	0	-2

This describes the decuplet formed by the spin-3/2 baryons.

There are two mixed tensors corresponding to the tableau $[2, 1]$. The first one is

$$[2, 1]_1 : T_{i_1 i_2 i_3} - T_{i_3 i_2 i_1} + T_{i_2 i_1 i_3} - T_{i_3 i_1 i_2}$$

Here $i_1 \neq i_3$ by antisymmetry. Let us list the possibilities:

1. An isospin doublet formed by

$$\begin{aligned} \frac{1}{\sqrt{2}} (|112\rangle - |211\rangle) : I_3 = +1/2, Y = 1 \\ \frac{1}{\sqrt{2}} (|122\rangle - |221\rangle) : I_3 = -1/2, Y = 1 \end{aligned}$$

2. An isospin triplet formed by

$$\begin{aligned} \frac{1}{\sqrt{2}} (|113\rangle - |311\rangle) : I_3 = +1, Y = 0 \\ \frac{1}{2} (|123\rangle + |213\rangle - |312\rangle - |321\rangle) : I_3 = 0, Y = 0 \\ \frac{1}{\sqrt{2}} (|223\rangle - |322\rangle) : I_3 = -1, Y = 0 \end{aligned}$$

3. An isospin singlet

$$\frac{1}{2} (|312\rangle - |321\rangle + |132\rangle - |231\rangle) : I_3 = 0, Y = 0$$

4. Another isospin doublet

$$\begin{aligned} \frac{1}{\sqrt{2}} (|332\rangle - |233\rangle) : I_3 = +1/2, Y = -1 \\ \frac{1}{\sqrt{2}} (|331\rangle - |133\rangle) : I_3 = -1/2, Y = -1 \end{aligned}$$

In the octet formed by the spin-1/2 baryons, the first doublet corresponds to the the proton/neutron isospin doublet (contains no strange quark), the triplet is the Σ hyperon isospin triplet, the isospin singlet is the Λ hyperon, and the last doublet is the Ξ hyperon isospin doublet.

In fact this is not yet the full story since $\mathbf{3} \otimes \mathbf{3} \otimes \mathbf{3} = \mathbf{1} \oplus \mathbf{8} \oplus \mathbf{8} \oplus \mathbf{10}$; however, we only see an octet and a decuplet of baryons in Nature. The reason is that the quarks also have spin which decomposes according to $\mathbf{2} \otimes \mathbf{2} \otimes \mathbf{2} = \mathbf{2} \oplus \mathbf{2} \oplus \mathbf{4}$, i.e. two spin-1/2 and one spin-3/2 components. For the decuplet, the wave function is entirely symmetric in both flavour and spin since it is the product of representation $\mathbf{10}$ for the $SU(3)$ flavour and the spin-3/2 representation $\mathbf{4}$ for the spin part.

It turns out the the general rule is that the combined spin/flavour wave function must be symmetric. The decompositions of the flavour and spin parts can be written as

$$\begin{aligned} \mathbf{2} \otimes \mathbf{2} \otimes \mathbf{2} &= \mathbf{2}_1 \oplus \mathbf{2}_2 \oplus \mathbf{4} \\ \mathbf{3} \otimes \mathbf{3} \otimes \mathbf{3} &= \mathbf{1} \oplus \mathbf{8}_1 \oplus \mathbf{8}_2 \oplus \mathbf{10} \end{aligned}$$

So there are four independent ways to compose states which transform as an octet of **8** and have spin 1/2, but none of them is symmetric under simultaneous permutations of spin and flavour indices. By explicit examination it can be seen that there is only one combination of these which is completely symmetric under permuting flavours and spins together: this gives the spin-1/2 baryon octet. In addition, we can see that there is no way to see to make the flavour singlet **1** symmetric by extending it with a spin component, since the complete antisymmetry of **1** requires the corresponding spin wave-function to be completely antisymmetric as well, but such a wave function vanishes as shown in Section (A.6).

A final remark: enlarging the combined spin-flavour symmetry to an $SU(6)$ with a fundamental representation **6** and its contragredient $\bar{\mathbf{6}}$ allows to fit the scalar and vector mesons into the representation $\mathbf{6} \otimes \bar{\mathbf{6}}$, while baryons can be accommodated in the fully symmetric part of $\mathbf{6} \otimes \mathbf{6} \otimes \mathbf{6}$ (and antibaryons in the fully symmetric part of $\bar{\mathbf{6}} \otimes \bar{\mathbf{6}} \otimes \bar{\mathbf{6}}$). The fact that the combined flavour-spin wave-function of baryons is fully symmetric under exchange of the particles is an essential part of the puzzle which necessitates the introduction of the colour $SU(3)$.

B Running coupling and asymptotic freedom

Here we consider the running coupling to leading order using an approach that elucidates the mechanism behind screening and asymptotic freedom adapting ideas from the description of magnetic materials⁵⁶

B.1 The running coupling

The interaction potential between two opposite charges at a distance r can be written as

$$V(r) = -\frac{\alpha}{r}. \quad (\text{B.1})$$

Taking into account vacuum polarisation the interaction strength depends on the energy/distance scale. Introducing the energy scale $\Lambda = 1/r$ corresponding to distance r one can write⁵⁷

$$\alpha(\Lambda) = \frac{e^2(\Lambda)}{4\pi}. \quad (\text{B.2})$$

The interaction strength, also known as the running coupling, can be related at two different scales as

$$e^2(\Lambda') = \frac{e^2(\Lambda)}{\varepsilon(\Lambda', \Lambda)}. \quad (\text{B.3})$$

where $\varepsilon(\Lambda', \Lambda)$ is the contribution to the vacuum dielectric constant from field modes between Λ' and Λ . Taking $\Lambda' < \Lambda$ i.e. $r' > r$, charge screening means that

$$\varepsilon(\Lambda', \Lambda) > 1. \quad (\text{B.4})$$

However, the dielectric properties of the quantum vacuum are related to its magnetic properties by Lorentz invariance (constant speed of light):

$$\varepsilon(\Lambda', \Lambda)\mu(\Lambda', \Lambda) = 1 \quad (\text{B.5})$$

⁵⁶The derivation in this section is adapted from the lecture notes by A. Grozin entitled ‘Asymptotic freedom: history and interpretation’, available as arXiv:0803.2589.

⁵⁷Note that here the energy scale is denoted by Λ for instead of μ as in Sections 12.3.2 and 12.3.3 of the main text; the reason is that in the following μ denotes the magnetic permeability of the quantum vacuum.

and so screening entails that the vacuum is diamagnetic:

$$\mu(\Lambda', \Lambda) < 1. \quad (\text{B.6})$$

This can be related to vacuum energy in an external magnetic field, which results in a polarisation contribution of the form

$$\Delta E_{\text{vac}} = \left(\frac{1}{\mu} - 1 \right) \frac{B^2}{2} V \quad (\text{B.7})$$

which can be interpreted as a change in vacuum energy. Therefore the magnetisation property of the vacuum can be determined by evaluating the shift of vacuum energy in an external magnetic field. To lowest order in the coupling constant $e(\Lambda)$ it will turn out that for a small change in scale

$$1 - \Lambda'/\Lambda \ll 1 \quad (\text{B.8})$$

the vacuum energy shift can be written in the form

$$\Delta E_{\text{vac}} = +\beta_0 \frac{e^2(\Lambda)}{(4\pi)^2} \log \frac{\Lambda^2}{\Lambda'^2} \frac{B^2}{2} V \quad (\text{B.9})$$

which results in

$$\mu = \epsilon^{-1} = 1 - \beta_0 \frac{e^2(\Lambda)}{(4\pi)^2} \log \frac{\Lambda^2}{\Lambda'^2}. \quad (\text{B.10})$$

Therefore we obtain

$$e^2(\Lambda') = \left[1 - \beta_0 \frac{e^2(\Lambda)}{(4\pi)^2} \log \frac{\Lambda^2}{\Lambda'^2} \right] e^2(\Lambda) \quad (\text{B.11})$$

or

$$\alpha(\Lambda) - \alpha(\Lambda') = +\beta_0 \frac{\alpha(\Lambda)^2}{2\pi} \log \frac{\Lambda}{\Lambda'} \quad (\text{B.12})$$

resulting in the differential equation

$$\frac{\partial \alpha(\Lambda)}{\partial \log \Lambda} = +\beta_0 \frac{\alpha(\Lambda)^2}{2\pi} \quad (\text{B.13})$$

which can be rewritten as

$$\Lambda \frac{\partial \alpha(\Lambda)}{\partial \Lambda} = +\beta_0 \frac{\alpha(\Lambda)^2}{2\pi}. \quad (\text{B.14})$$

Charge screening corresponds to $\beta_0 > 0$ which means the running coupling increases with energy, equivalently it decreases with distance.

The above equation can be solved explicitly with the result

$$\alpha(\Lambda) = \frac{\alpha(\Lambda_0)}{1 - \beta_0 \log(\Lambda/\Lambda_0)/2\pi} \quad (\text{B.15})$$

where $\alpha(\Lambda_0)$ is the initial condition given as the coupling at some reference scale Λ_0 . For asymptotically free theories with $\beta_0 < 0$ it is customary to rewrite the solution as

$$\alpha(\Lambda) = \frac{1}{|\beta_0| \log(\Lambda/\Lambda_0)/2\pi} \quad (\text{B.16})$$

which is the same as Eq. (12.108) with

$$b_0 = \frac{|\beta_0|}{2\pi}.$$

B.2 Charged scalar: Landau diamagnetism

Let us assume we have a charged spinless (scalar) particle of mass m with the Lagrangian density

$$\mathcal{L} = (D_\mu \varphi)^\dagger D^\mu \varphi - m^2 \varphi^\dagger \varphi \quad D_\mu \varphi = \partial_\mu \varphi - ie A_\mu \varphi. \quad (\text{B.17})$$

We compute its vacuum energy in a box $L_x \times L_y \times L_z$, for a magnetic field B along z . We can choose the vector potential $A_y = Bx$ and $A_x = A_z = 0$ giving the equation of motion

$$\begin{aligned} & [(\partial^\mu - ie A^\mu)(\partial_\mu - ie A_\mu) + m^2] \varphi = 0 \\ \text{i.e. } & (\square + e^2 B^2 x^2 + 2ie B x \partial_y + m^2) = 0 \end{aligned} \quad (\text{B.18})$$

We look for the field modes in the form

$$\varphi = e^{-iEt + ik_y y + ik_z z} \phi_n(x) \quad (\text{B.19})$$

which results in

$$\left(-\partial_x^2 + e^2 B^2 \left(x - \frac{k_y}{eB} \right)^2 \right) \phi_n(x) = (E^2 - m^2 - k_z^2) \phi_n(x).$$

Using the solutions of the usual harmonic oscillator

$$\left(-\frac{1}{2} \partial_x^2 + \frac{\omega^2}{2} x^2 - \mathcal{E}_n \right) \psi_n(x) = 0, \quad \mathcal{E}_n = \omega \left(n + \frac{1}{2} \right) \quad (\text{B.20})$$

the energy levels of a relativistic charged particle in a magnetic field are given by the Landau levels

$$E_n(k_z) = \pm \sqrt{m^2 + k_z^2 + 2eB \left(n + \frac{1}{2} \right)} \quad (\text{B.21})$$

and

$$\phi_n(x) = \psi_n \left(x - \frac{k_y}{eB} \right) \quad (\text{B.22})$$

Henceforth we neglect the mass m for simplicity. As usual, the degeneracy of these levels is given by

$$\frac{eBL_x L_y}{2\pi} \quad (\text{B.23})$$

and summation over the z momenta is given by the integral

$$\frac{L_z}{2\pi} \int dk_z. \quad (\text{B.24})$$

The vacuum energy is then

$$E_{\text{vac}} = \frac{L_z}{2\pi} \int dk_z \frac{eBL_x L_y}{2\pi} \sum_n \sqrt{k_z^2 + 2eB \left(n + \frac{1}{2} \right)} = \frac{eBV}{(2\pi)^2} \int dk_z \sum_n \sqrt{k_z^2 + 2eB \left(n + \frac{1}{2} \right)}. \quad (\text{B.25})$$

The sum can be performed using the Euler-Maclaurin formula

$$\sum_{n=0}^N f\left(n + \frac{1}{2}\right) = \int_0^{N+1} f(x)dx - \frac{1}{24} (f'(N+1) - f'(0)) + \dots \quad (\text{B.26})$$

Let us consider the integral term first:

$$\int_0^{N+1} f(x)dx = \frac{eBV}{(2\pi)^2} \int dk_z \int_0^{N+1} dx \sum_n \sqrt{k_z^2 + 2eBx} \quad (\text{B.27})$$

and defining

$$eBx = \frac{k_\perp^2}{2} \quad (\text{B.28})$$

it can be rewritten as

$$\frac{V}{(2\pi)^2} \int dk_z \int_0^{k_\perp^{\max}} k_\perp dk_\perp \sqrt{k_z^2 + k_\perp^2} = \frac{V}{(2\pi)^3} \int dk_z dk_x dk_y \sqrt{k_z^2 + k_x^2 + k_y^2} \quad (\text{B.29})$$

which is just the vacuum energy without a magnetic field. So the shift is

$$\Delta E_{\text{Landau}} = \frac{1}{6} \frac{eBV}{(4\pi)^2} \int dk_z \left| \frac{d}{dx} \sqrt{k_z^2 + 2eBx} \right|_{x=0} = \frac{e^2 B^2 V}{6 (4\pi)^2} \int \frac{dk_z}{|k_z|}. \quad (\text{B.30})$$

Recalling that we only need the contribution of modes $\Lambda' < |k_z| < \Lambda$

$$\int \frac{dk_z}{|k_z|} \rightarrow 2 \int_{\Lambda'}^{\Lambda} \frac{dk_z}{k_z} = \log \frac{\Lambda^2}{\Lambda'^2} \quad (\text{B.31})$$

we obtain the result

$$\Delta E_{\text{vac}} = \frac{1}{3} \frac{e^2(\Lambda)}{(4\pi)^2} \log \frac{\Lambda^2}{\Lambda'^2} \frac{B^2}{2} V. \quad (\text{B.32})$$

Note that here we specifically substituted e with its value at the scale Λ since we compute the contribution for a small interval between Λ' and Λ . Therefore for a complex scalar we obtain

$$\beta_0 = \frac{1}{3} \quad (\text{B.33})$$

which is positive, leading to charge screening. As a result, the coupling in scalar quantum electrodynamics increases with energy or, equivalently, the effective charge decreases with distance.

B.3 Dirac fermion: Landau paramagnetism and Pauli diamagnetism

For a Dirac fermion the vacuum energy is given by

$$E_{\text{vac}} = - \sum_i \omega_i, \quad (\text{B.34})$$

where the negative sign is due to Fermi statistics. Surprisingly, this means that the contribution of Landau levels is paramagnetic, since the only change from the above derivation for the charged scalar is a relative sign and a factor of 2 accounting for the spin degeneracy:

$$\Delta E_{\text{Landau}} = -2 \frac{1}{3} \frac{e^2}{(4\pi)^2} \log \frac{\Lambda^2}{\Lambda'^2} \frac{B^2}{2} V. \quad (\text{B.35})$$

Why is the sign change compared to the usual Landau diamagnetism? The reason is that the usual condensed matter derivation is performed in the presence of a Fermi sea consisting of positive energy electrons, instead of the Dirac vacuum considered here.

However, for a fermion we also have a spin contribution, which in this case leads to diamagnetism (again contrary to the usual result). Neglecting the mass, the energy levels are given by the Zeeman effect as

$$\omega = \sqrt{k^2 - g_s s_z e B} \quad (\text{B.36})$$

with g_s denoting the gyromagnetic factor for a particle of spin s . For a massless particle the spin projection s_z can only take values $\pm s$, therefore

$$\begin{aligned} \Delta E_{\text{Pauli}} &= -V \int \frac{d^3 k}{(2\pi)^3} \left(\sqrt{k^2 - g_s s e B} + \sqrt{k^2 + g_s s e B} - 2k \right) \\ &= V \frac{(g_s s e B)^2}{4} \int \frac{d^3 k}{(2\pi)^3} \frac{1}{k^3} + O((eB)^4) = 2(g_s s)^2 \frac{e^2}{(4\pi)^2} B^2 \int \frac{dk}{k} \\ &= 2(g_s s)^2 \frac{e^2}{(4\pi)^2} \log \frac{\Lambda^2}{\Lambda'^2} \frac{B^2}{2} V. \end{aligned} \quad (\text{B.37})$$

For a Dirac fermion one can obtain the gyromagnetic factor as follows. The Dirac equation can be written as

$$\gamma^\mu D_\mu \psi = 0 \quad D_\mu = \partial_\mu - ie A_\mu \quad (\text{B.38})$$

which implies

$$\begin{aligned} 0 &= \gamma^\lambda \gamma^\mu D_\lambda D_\mu \psi = \frac{1}{2} \{ \gamma^\lambda, \gamma^\mu \} D_\lambda D_\mu \psi + \frac{1}{2} [\gamma^\lambda, \gamma^\mu] D_\lambda D_\mu \psi \\ &= D^2 \psi + \frac{1}{4} [\gamma^\lambda, \gamma^\mu] [D_\lambda, D_\mu] \psi \end{aligned} \quad (\text{B.39})$$

The first term is the same as the Klein-Gordon equation (B.18) in external field, and the second term describes the spin Zeeman coupling. Using

$$[D_\lambda, D_\mu] = -ie F_{\lambda\mu} \quad (\text{B.40})$$

and $F_{xy} = -F_{yx} = -B$, this term can be written as

$$\frac{1}{4} [\gamma^\lambda, \gamma^\mu] [D_\lambda, D_\mu] \psi \rightarrow 2eB \Sigma_3 \psi \quad (\text{B.41})$$

where

$$\Sigma_3 = \frac{1}{2} \sigma^{12} \quad \sigma^{12} = \frac{i}{4} [\gamma^1, \gamma^2] \quad (\text{B.42})$$

is the operator for the z component of the spin, which is the same as the generator of rotations in the rest frame in Eq. (5.139). Therefore the gyromagnetic factor is $g_s = 2$ for a particle with spin $s = 1/2$, giving

$$\Delta E_{\text{Pauli}} = 2 \frac{e^2}{(4\pi)^2} \log \frac{\Lambda^2}{\Lambda'^2} \frac{B^2}{2} V \quad (\text{B.43})$$

and so

$$\Delta E_{\text{vac}} = \Delta E_{\text{Landau}} + \Delta E_{\text{Pauli}} = \frac{4}{3} \frac{e^2}{(4\pi)^2} \log \frac{\Lambda^2}{\Lambda'^2} \frac{B^2}{2} V \quad (\text{B.44})$$

giving

$$\beta_0 = \frac{4}{3}. \quad (\text{B.45})$$

The upshot is that coupling the electromagnetic field to a Dirac fermion field also leads to charge screening. Substituting this result into the running coupling equation (B.15) gives equation (12.97).

B.4 Non-Abelian gauge fields

For simplicity, let us first consider the $SU(2)$ case. We use the Lagrangian defined as in (12.69), which for a pure gauge theory without matter fields is

$$\mathcal{L} = -\frac{1}{4}G_{\mu\nu}^a G^{a\mu\nu} \quad G_{\mu\nu}^a = \partial_\mu A_\nu^a - \partial_\nu A_\mu^a + g\varepsilon^{abc}A_\mu^b A_\nu^c \quad (\text{B.46})$$

The field equations are

$$\partial^\mu G_{\mu\nu}^a + g\varepsilon^{abc}A^{\mu\nu}G_{\mu\nu}^c = 0 \quad (\text{B.47})$$

Consider the field equation of the charged field

$$A_\mu^- = A_\mu^1 - iA_\mu^2 \quad (\text{B.48})$$

in the external field A_μ^3 to linear order, which can be written as

$$\begin{aligned} D^\nu G_{\mu\nu}^- + igG_{\mu\nu}^3 A^{-\nu} &= 0 \\ G_{\mu\nu}^- &= D_\mu A_\nu^- - D_\nu A_\mu^- \quad D_\mu = \partial_\mu - igA_\mu^3 \end{aligned} \quad (\text{B.49})$$

In the gauge $D^\mu A_\mu^- = 0$ and using $[D_\mu, D_\nu] = -igG_{\mu\nu}^3$, we can write the equations of motion as

$$D^2 A_\mu^- - 2igG_{\mu\nu}^3 A^{-\nu} = 0 \quad (\text{B.50})$$

We take a magnetic field of colour 3 and oriented along z , so $G_{xy}^3 = -G_{yx}^3 = -B$. The two spin polarisations of the A^- field are given by $A^{-y} = is_z A^{-x}$ with $s_z = \pm 1$, so finally

$$(D^2 + 2igBs_z) A^{-x} = 0, \quad (\text{B.51})$$

showing that the gyromagnetic factor of a spin-1 field is $g_1 = 2$.

Therefore, in this case the Landau diamagnetic contribution is

$$\Delta E_{\text{vac}} = 2 \frac{1}{3} \frac{e^2(\Lambda)}{(4\pi)^2} \log \frac{\Lambda^2}{\Lambda'^2} \frac{B^2}{2} V, \quad (\text{B.52})$$

with a degeneracy factor 2 due to the two polarisations of massless spin-1 particle compared to the scalar result (B.32). The Pauli paramagnetic contribution is given by

$$-2(g_s s)^2 \frac{e^2}{(4\pi)^2} \log \frac{\Lambda^2}{\Lambda'^2} \frac{B^2}{2} V, \quad (\text{B.53})$$

with a sign opposite to the fermionic result (B.37) due to the difference in the sign of the vacuum energy between bosons and fermions.

Note that the Landau diamagnetic contribution with a coefficient of $1/3$ is overcompensated by Pauli paramagnetism with the coefficient $-(g_1 1)^2 = -4$, arriving at the net coefficient

$$\frac{1}{3} - 4 = -\frac{11}{3} \quad (\text{B.54})$$

i.e. we obtain antiscreening!

We can now consider a general non-Abelian field theory corresponding to a simple gauge group G . The coupling strength of a field multiplet to any of the gluon fields is given by

$$gt_R^a \quad (\text{B.55})$$

where t_R^a is the generator of the representation of G in which the field multiplet transforms. To evaluate the total contribution, we need to replace e^2 by the sum of the multiplet charges squared, which can be computed from

$$\text{Tr} gt_R^a gt_R^a \quad (\text{B.56})$$

for any fixed a . Here no summation is meant and all choices of a give the same result due to group symmetry. Introducing the quadratic Casimir

$$\text{Tr} t_R^a t_R^b = C_R \delta^{ab} \quad (\text{B.57})$$

we obtain

$$\beta_0 = -\frac{11}{3}C_{\text{adj}} + \frac{4}{3}C_f + \frac{1}{3}C_s \quad (\text{B.58})$$

where C_{adj} is the quadratic Casimir of the adjoint representation (the gluon contribution), while C_f and C_s are Casimirs for the Dirac fermions and complex scalar fields, respectively. For an $SU(N_c)$ gauge group $C_{\text{adj}} = N_c$ and with n_f fermions in the adjoint representation with $C_f = 1/2$ the result is

$$\beta_0 = -\frac{11}{3}N_c + \frac{2}{3}n_f \quad (\text{B.59})$$

which is the result quoted in Eq. (12.106). In particular, QCD with $N_c = 3$ is asymptotically free as long as $n_f \leq 16$.