

Faculty of physics

Advanced Quantum Field Theory

Notes of lectures by Prof. Martin Beneke, Ph.D.

Academic year 2025/2026
Summer semester

Author:

Simon Bukovšek

Last change: November 26, 2026

Contents

1 Non-Abelian gauge symmetry and quantization of gauge theories **4**
- 1.1 Non-Abelian gauge symmetry . . . 5
  - 1.1.1 Construction of the invariant Lagrangian . . . 5
  - 1.1.2 Some facts from group theory . . . 7
  - 1.1.3 Final form of the Lagrangian, self interaction and field equation . . . 10
- 1.2 Faddeev–Popov method . . . 13
  - 1.2.1 Path integral in axial gauge . . . 13
  - 1.2.2 Faddeev–Popov method and ghosts . . . 16
- 1.3 Renormalization of non-Abelian gauge theories . . . 19
  - 1.3.1 BRST symmetry and Ward identities . . . 19
  - 1.3.2 Renormalization – general considerations . . . 23
  - 1.3.3 Renormalization at the one-loop order . . . 26
- 1.4 The physical Hilbert space . . . 31
  - 1.4.1 Quantum electrodynamics in covariant gauge . . . 31
  - 1.4.2 Canonical quantization of non-Abelian gauge theory in covariant gauge . . . 32
  - 1.4.3 Gauge invariance of the S-matrix and conditions on the physical Hilbert space . . . 35
  - 1.4.4 Unitarity of the S-matrix on $\mathcal{H}_{\text{phys}}$ . . . 37
  - 1.4.5 Summary . . . 37

2 Spin-2 fields **39**
- 2.1 The massive spin-2 field . . . 39
- 2.2 The massless spin-2 field . . . 41
- 2.3 GR as an effective field theory of gravitation . . . 43
  - 2.3.1 Classical general relativity . . . 43
  - 2.3.2 Quantization of weak gravitational fields . . . 44
  - 2.3.3 Interaction with matter . . . 47
  - 2.3.4 Renormalization . . . 49

3 Spontaneous Symmetry Breaking **51**
- 3.1 Spontaneous breaking of global symmetries . . . 52
  - 3.1.1 Spontaneous symmetry breaking and the vacuum . . . 52
  - 3.1.2 Field-theoretical characterization of the vacuum through the effective action . . . 54
  - 3.1.3 Goldstone bosons . . . 59
  - 3.1.4 The linear σ-model . . . 64

3.1.5 Spontaneous symmetry breaking of approximate symmetries 67
3.2 Goldstone boson interactions 69
3.2.1 Non-linear symmetry representations 69
3.2.2 Goldstone boson effective Lagrangian 71
3.3 Spontaneously broken chiral symmetry of QCD 74
3.3.1 Approximate global symmetries of QCD 75
3.3.2 Masses of the QCD Goldstone bosons 76
3.3.3 The effective chiral Lagrangian 78

4 Electroweak symmetry breaking 85
4.1 Higgs mechanism 87
4.1.1 Unitary gauge 87
4.1.2 Renormalizable gauge 89
4.2 Scattering of longitudinal vector bosons 92
4.2.1 Partial wave unitarity 93
4.2.2 Unitarization by Higgs exchange 95
4.2.3 Goldstone boson equivalence theorem 96
4.3 Aspects of the Standard model Higgs sector 96
4.3.1 Higgs couplings in the SM (and beyond) 96
4.3.2 Effective potential and vacuum stability 99
4.3.3 Triviality 104

5 Infrared divergences and the strong interaction in high-energy processes 107
5.1 Soft and collinear divergences in QED 108
5.1.1 Soft photon emission amplitude 108
5.1.2 Virtual corrections and the Bloch–Nordsieck cancellation 109
5.2 IR divergences in QCD and jets 111
5.2.1 Hadron production at high energies 111
5.2.2 Calculation of jet production 117
5.3 Deep inelastic scattering (DIS) 120
5.3.1 Factorization and DGLAP evolution 126
5.3.2 Proton-proton collisions 132

6 Anomalies 136
6.1 $\pi^0 \to \gamma\gamma$ 136
6.2 Anomalous symmetries in the PI formalism 137
6.2.1 Computation of the chiral symmetry 138
6.2.2 Chiral anomaly and $\pi^0 \to \gamma\gamma$ 140
6.3 Diagrammatic derivation of the anomaly 141
6.3.1 The Abelian anomaly 145
6.3.2 Chiral anomaly in dimensional regularization 146
6.3.3 $U(1)_A$ and flavour anomaly in QCD 148
6.4 Consistency of chiral gauge theories 149
6.4.1 General results 150
6.4.2 Anomaly freedom of the Standard Model 151
6.4.3 Violation of baryon and lepton number conservation 153

Index 155

Disclaimer: these notes are not official and were not endorsed by professor. The content is almost entirely based on the lectures, but occasionally the notes were modified or extended slightly to improve clarity or add extra information. In particular, the mistakes are almost surely mine. If you find any, please let me know at simon.bukovsek@tum.de or Simon.Bukovsek@campus.lmu.de.

Chapter 1

Non-Abelian gauge symmetry and quantization of gauge theories

One of the basic lessons is that the construction of theories for the fundamental forces in the Nature is based on the concept of quantized *fields*, *symmetries* (in particular, Lorentz invariance) and the *action principle*. The action $S[\Psi]$ should be a functional of field configuration and invariant under the postulated symmetries. The dynamics follows from $\delta S = 0$ (*Euler–Lagrange equations*). The particle content follows from the poles of Green functions of the fields.

The concepts of fields, symmetries and particles are tied together, when one attempts to formulate quantum theories of massless spin-1 particles (more precisely, helicity ± 1). Recall the case of the *quantum electrodynamics* (QED):

$$\begin{aligned}\mathcal{L} &= -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \overline{\Psi}(i\gamma^\mu \mathrm{D}_\mu - m)\Psi + \mathcal{L}_{\text{gauge-fix}} \\ &= -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} - J_\mu A^\mu + \overline{\Psi}(i\partial\!\!\!/ - m)\Psi + \mathcal{L}_{\text{gauge-fix}},\end{aligned}$$

with $J_\mu = -e\overline{\Psi}\gamma_\mu\Psi$ called the *electric charge current* and $\mathrm{D}_\mu = \partial_\mu - ieA_\mu$ *covariant derivative*. The Lagrangian $\mathcal{L}$ is invariant under the *gauge symmetry*

$$\begin{aligned}A^\mu(x) &\to A^\mu(x) + \frac{1}{e}\partial^\mu\alpha(x) \\ \Psi(x) &\to e^{i\alpha(x)}\Psi(x).\end{aligned}$$

The gauge parameter $\alpha(x)$ depends on x, hence this is *local* (gauge) symmetry. Why is gauge symmetry necessary?

- It is not possible to describe helicity ± 1 particle states by a field $A^\mu(x)$, which transforms as a Lorentz vector (in $(1/2, 1/2)$ representation of the Lorentz group). If one nevertheless insists on this, then the field actually transforms as

  $$\mathcal{U}(\Lambda)A^\mu(x)\mathcal{U}^{-1}(\Lambda) = (\Lambda^-1)^\mu{}_\nu A^\nu(\Lambda x) + \partial^\mu\alpha(\Lambda x).$$

  To render the extra term irrelevant, $\mathcal{L}$ must be invariant under the gauge transformation of A^μ above and A^μ must couple to a conserved current J_μ, which is the Noether current of the gauge transformation of the Ψ field.

- We cannot use the antisymmetric tensor $F_{\mu\nu}$ (in the $(1,0)\oplus(0,1)$ representation of the Lorentz group) as the fundamental field, although this does describe helicity ± 1 particles, because then there would be no long-range Coulomb force.

Hence, the relativistic quantum field theory of massless spin-1 particles with long-range forces requires gauge symmetry.

But QED is not the most general case. Fields can transform under internal symmetry group G. In this case, the massless vector field can carry an additional internal index, i.e. $A^A_\mu(x)$ and similarly $\Psi(x)$. In this chapter we discuss this generalization, which is relevant in the Nature to describe the strong and electroweak force (eve though neither of them is long-ranged – we will discuss this extra twist in due time). We also supply the path integral quantization of QED, which was left out in the previous chapter, as a special case of the general treatment.

1.1 Non-Abelian gauge symmetry and invariant Lagrangians

We want to construct a theory of "matter fields", which transform under an internal symmetry, whose Noether current is the conserved current to which massless vector fields can couple, which generate the long-range force between matter particles:

$$\Psi_{\alpha a}(x) \xrightarrow{g\in G} \Psi'_{\alpha a}(x) = D_{ab}(g^{-1})(x)\Psi_{\alpha b}(x).$$

Here, α denotes the index with respect to the Lorentz transformations (scalar, Weyl, Majorana, Dirac fermion), a and b are internal symmetry indices, $g \in G$ is an element of some Lie group and $D(g)$ is a finite-dimensional representation R of the group element. We can anticipate that the symmetry is local, i.e. $D(g^{-1})$ can be different at different space-time points:

$$D(g)(x) = \exp(-i\varepsilon^A(x)T^A)$$

where $A = 1, \ldots, \dim G$ and T^A is the representation of the generator t^A of the Lie algebra. The latter is a $\dim R \times \dim R$ matrix. The T^A satisfy commutation relations

$$[T^A, T^B] = if^{ABC}T^C$$

where f^{ABC} are called the *structure constants* of the Lie algebra. They are independent of the representation R. Under an infinitesimal transformation, the field changes as

$$\delta\Psi_a(x) = \Psi'_a(x) - \Psi_a(x) = i\varepsilon^A(x)T^A_{ab}\Psi_b(x).$$

The parameters ε^A are real. QED is the special case of the Abelian group U(1) for which $T^A = 1$ and $f^{ABC} = 0$.

1.1.1 Construction of the invariant Lagrangian

Expressions such as $\partial^\mu \dagger^+ \partial_\mu\phi - m^2\phi^\dagger\phi$ and $\overline{\Psi}(i\gamma^\mu\partial_\mu - m)\Psi$ are invariant only under global, x-independent transformations, but not when $\varepsilon^A(x) \neq$ const. The problem is that $\partial_\mu\Psi_{\alpha a}$ foes not transform in the same way as $\Psi_{\alpha a}$:

$$\delta(\partial_\mu\Psi) = \partial_\mu(i\varepsilon^A T^A\Psi) = i\varepsilon^A T^A\partial_\mu\Psi + i(\partial_\mu\varepsilon^A)T^A\Psi.$$

The unwanted term looks like the extra term in the transformation of a massless vector field $A^A_\mu(x)$, so we introduce the *covariant derivative*

$$\mathrm{D}_{\mu,ab} = \delta_{ab}\partial_\mu - iA^A_\mu(x)T^A_{ab}.$$

This implies that there are always $\dim G$ gauge fields, independent of the representation R. If G is a non-Abelian group, these real vector fields are called non-Abelian gauge fields. The transformation of A^A_μ is determined by requiring

$$\delta(\mathrm{D}_\mu\Psi) \stackrel{!}{=} i\varepsilon^A T^A \mathrm{D}_\mu\Psi.$$

Then $(\mathrm{D}_\mu\phi)^\dagger(\mathrm{D}^\mu\phi) - m^2\phi^\dagger\phi$ and $\overline{\Psi}(i\gamma^\mu\mathrm{D}_\mu - m)\Psi$ are invariant under the gauge symmetry (provided $D(g)$ is unitary, e.g. R is a unitary representation of G. We will show later that this choice is always possible). The transformation of field A^A_μ can be obtained from the requirement for transformation of $\mathrm{D}_\mu\Psi$:

$$\begin{aligned}\delta(\mathrm{D}_\mu\Psi) &= \delta(\partial_\mu\Psi) - i\delta A^A_\mu T^A\Psi - iA^A_\mu T^A\delta\Psi \\ &= i\varepsilon^A T^A\mathrm{D}_\mu\Psi + i\varepsilon^A T^A iA^A_\mu T^B\Psi - i\delta A^B_\mu T^B\Psi - iA^A_\mu T^A i\varepsilon^B T^B\Psi + i\partial_\mu\varepsilon^A T^A\Psi \\ &\stackrel{!}{=} i\varepsilon^A T^A\mathrm{D}_\mu\Psi.\end{aligned}$$

From this it follows

$$\delta A^A_\mu T^A = i\varepsilon^A[T^A, T^B]A^B_\mu + \partial_\mu\varepsilon^A T^A$$

or, equivalently,

$$A^A_\mu \longrightarrow A^A_\mu + f^{ABC}\varepsilon^C A^B_\mu + \partial_\mu\varepsilon^A.$$

The second term can be written as $i\varepsilon^C T^C_{AB}A^B_\mu$ if we interpret T^C_{AB} as a generator. I.e. A_μ transforms itself under the internal symmetry, if G is non-Abelian. The last term is the inhomogeneous term which must be present to render the non-Lorentz covariant transformation irrelevant. The above holds for infinitesimal transformations. For large transformations $\mathcal{U} = D(g^{-1}) = e^{i\varepsilon^A T^A}$, this implies

$$\begin{aligned}\Psi &\longrightarrow \mathcal{U}\Psi, \\ \mathbf{A}_\mu \equiv A^A_\mu T^A &\longrightarrow \mathcal{U}\mathbf{A}_\mu\mathcal{U}^\dagger + i\mathcal{U}\partial_\mu\mathcal{U}^\dagger, \\ \mathrm{D}_\mu &\longrightarrow \mathcal{U}\mathrm{D}_\mu\mathcal{U}^\dagger.\end{aligned}$$

From here on, we will use bold letters to denote fields as elements of the representation, and capital Latin indices to denote their components in the chosen base of the representation. Remember from QED that the field strength tensor was defined as

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu = i[\mathrm{D}_\mu, \mathrm{D}_\nu].$$

The commutator looks like a differential operator, but is actually a function of x, since $[\partial_\mu, \partial_\nu] = 0$. The field strength tensor is gauge invariant and contains derivations of A_μ. Hence, it is a good building block for a gauge field kinetic term. For the general case define

$$\mathbf{G}_{\mu\nu} \equiv G^A_{\mu\nu}T^A = i[\mathrm{D}_\mu, \mathrm{D}_\nu] = \partial_\mu\mathbf{A}_\nu - \partial_\nu\mathbf{A}_\mu - i[\mathbf{A}_\mu, \mathbf{A}_\nu] = (\partial_\mu A^A_\nu - \partial_\nu A^A_\mu + f^{ABC}A^B_\mu A^C_\nu)T^A.$$

This object transforms as

$$\mathbf{G}_{\mu\nu} \longrightarrow \mathcal{U}\mathbf{G}_{\mu\nu}\mathcal{U}^\dagger, \quad \delta G^A_{\mu\nu} = f^{ABC}\varepsilon^C G^B_{\mu\nu}.$$

This is covariant but, in general, not invariant. However, $\mathrm{tr}(G_{\mu\nu}G^{\mu\nu})$ or $G^A_{\mu\nu}G^{A,\mu\nu}$ is invariant.

The most general gauge-invariant Lagrangian can be constructed from the covariant objects Ψ, $G_{\mu\nu}$ and D_μ and contracting the Lorentz indices to form a scalar. A_μ does not appear explicitly because it does not have a covariant transformation. Up to dimension 4, the possible terms are (considering $[A_\mu] = 1$, $[\mathrm{D}_\mu] = 1$, $G_{\mu\nu} = 2$)

$$\mathcal{L} = -\frac{1}{4} g_{AB} G^A_{\mu\nu} G^{B,\mu\nu} - \frac{1}{2}\theta_{AB}\varepsilon_{\mu\nu\rho\sigma} G^{A,\mu\nu} G^{B,\rho\sigma} + \mathcal{L}_{\mathrm{M}}(\text{matter fields}). \tag{1.1}$$

The Lagrangian containing matter fields is $\overline{\Psi}(i - \not{\mathrm{D}} - m)\Psi$ for Dirac spinor and $(\mathrm{D}_\mu\phi)^\dagger(\mathrm{D}^\mu\phi) - m^2\phi^\dagger\phi$ for a complex scalar. g_{AB} and θ_{AB} are matrices of constant (x-independent) parameters. The second term violates parity symmetry (and, because CPT conservation, also time reversal). Experimentally, pure gauge interactions are not observed to violate parity, so from now on we set $\theta_{AB} = 0$. For the strong force (quantum chromodynamics) it is not understood what is the reason for this. Explanations motivate the existence of a new particle – the axion, but this has not been observed either, though also not ruled out. Even if $\theta_{AB} \neq 0$, the second term is rather special, since it can be shown to be a total derivative. Hence, it contributes to the action only through certain topological effects. We will not discuss this in detail here.

1.1.2 Some facts from group theory

The matrix g_{AB} can be chosen to be symmetric and must be positive definite in order that the kinetic term for the gauge field is positive. Also, eq. 1.1 is gauge invariant only if

$$g_{AB} f^{ABC} + g_{CB} f^{BAD} = 0.$$

This follows from

$$\begin{aligned}\delta(g_{AB} G^A_{\mu\nu} G^{B,\mu\nu}) &= g_{AB} f^{ACD} \varepsilon^D G^C_{\mu\nu} G^{B,\mu\nu} + g_{AB} G^A_{\mu\nu} f^{ABC} \varepsilon^D G^{C,\mu\nu} \\ &= (g_{AB} f^{ABC} + g_{CB} f^{BAD}) \varepsilon^D G^A_{\mu\nu} G^{C,\mu\nu} \overset{!}{=} 0.\end{aligned}$$

This has important implications (for proof, see Weinberg, Vol. II, Ch. 15, App. A):

1. there sexists a basis for the generators t^A of the Lie algebra, such that the structure constants f^{ABC} are *totally antisymmetric*;
2. the Lie group G is a direct product of compact, simple groups and U(1) groups

   $$G = G_1 \times \cdots \times G_m \times \mathrm{U}(1) \times \cdots \times \mathrm{U}(1).$$

   Simple group means that there exist no non-trivial invariant subgroups H ($ghg^{-1} \in H$ for all $h \in H, g \in G$). Compact group means that the group is compact as a manifold.

Using a further basis transformation one van then bring g_{AB} into the diagonal form

$$g_{AB} = \begin{pmatrix} 1/g_1^2 & & & & & & & & & & & & \\ & \ddots & & & & & & & & & & & \\ & & 1/g_1^2 & & & & & & & & & & \\ & & & 1/g_2^2 & & & & & & & & & \\ & & & & \ddots & & & & & & & & \\ & & & & & 1/g_2^2 & & & & & & & \\ & & & & & & \ddots & & & & & & \\ & & & & & & & 1/g_m^2 & & & & & \\ & & & & & & & & \ddots & & & & \\ & & & & & & & & & 1/g_m^2 & & & \\ & & & & & & & & & & 1/g_1'^2 & & \\ & & & & & & & & & & & 1/g_2'^2 & \\ & & & & & & & & & & & & \ddots \end{pmatrix}.$$

Hence, the Lagrangian 1.1 falls apart into a sum of separate terms of the form

$$-\frac{1}{4g^2} G^A_{\mu\nu} G^{A,\mu\nu}$$

for every subgroup. Here, A is summed only from 1 to $\dim G_i$ for every i, or absent for U(1) factors. Without loss of generality we may assume that either G is compact and simple or $G = \mathrm{U}(1)$. The general case follows by adding several gauge field Lagrangians of the non-Abelian or U(1) type.

A result form representation theory states that the finite-dimensional representations of compact Lie groups are unitary. Hence, we can assume that $\mathcal{U} = D(g^{-1})$ is a unitary matrix, as was necessary. Then the generators T^A are hermitian matrices, $T^{A\dagger} = T^A$ (since $\mathcal{U} = e^{i\varepsilon^A T^A}$).

Classification of compact, simple Lie groups by Cartan

There is only a limited number of compact, simple Lie groups, consisting of the three infinite series SU(N), SO(N), Sp($2N$) and a few "special groups" $\mathrm{G}_2, \mathrm{F}_4, \mathrm{E}_6, \mathrm{E}_7, \mathrm{E}_8$.

- SU(N) for $N \geq 2$ are unitary matrices with $\det \mathcal{U} = 1$. It is called *special unitary group*. Since T^A are hermitian, $\operatorname{tr} T^A = 0$. The dimension of the group is $N^2 - 1$. (A hermitian matrix has N real diagonal elements and $N(N-1)/2$ complex elements above the diagonal. The ones below are then fixed. Accounting for the non-vanishing trace gives

  $$2 \times \frac{N(N-1)}{2} + N - 1 = N^2 - 1$$

  real parameters.)
- SO(N) are unitary matrices with $\det \mathcal{U} = 1$, which leave the scalar product $x_a \delta_{ab} x_b$ for real N-component vectors $\mathbf{x}$ and $\mathbf{y}$ invariant. They correspond to rotations in N-dimensional space and the matrices are real. The generators T^A are hermitian, traceless and purely imaginary, i.e. $T^{A*} = -T^A$. Its dimension is $N(N-1)/2$. The group is called *special orthogonal group*.

- $\mathrm{Sp}(2N)$ are unitary $2N \times 2N$ matrices which leave the scalar product $x_a \varepsilon_{ab} y_b$ invariant, where
$$\varepsilon_{ab} = \begin{pmatrix} & \mathbb{1}_{N\times N} \\ \mathbb{1}_{N\times N} & \end{pmatrix}.$$
It is called *symplectic group* and has dimension $2N(2N+1)/2$.

Since we usually consider complex fields, $\mathrm{SU}(N)$ and U_1 are the most important gauge groups.

Group elements can be represented as transformations of vector spaces:
$$g \in G \longrightarrow D(g) = e^{-i\varepsilon^A(g)T^A}.$$
Here, $D(g)$ is an $n \times n$ matrix, where n is the dimension of the representation R and T^A are generators. Since generators depend on the representation, one should write T^A_R for representation R. The group product has to be conserved on the representation, $D(g_1 g_2) = D(g_1)D(g_2)$. A representation is *irreducible* (often denoted as irrep) if the $D(g)$ operate on a vector space such that there is no subspace left invariant under the operation of all $D(g)$. In field theory, the vectors of the vector space are the field operators $\psi_a(x)$ where a is the vector's index. Since reducible representations can always be decomposed as a direct sum of irreducible ones, one usually considers only irreducible representations. These contain the minimal set of objects, which are necessarily linked by symmetry transformations. The two most important representations are:

- *fundamental representation*:
$$g = \mathcal{U} \longrightarrow D(g) = \mathcal{U}.$$
It is N dimensional for $\mathrm{SU}(N)$ and $\mathrm{SO}(N)$. As a convention, we choose the following normalization
$$\operatorname{tr} T^A_{\mathrm{F}} T^B_{\mathrm{F}} = \frac{1}{2}\delta^{AB}.$$
This can always be achieved because the left-hand side is a symmetric, positive definite matrix. (The transformation which diagonalizes $\operatorname{tr}(T^A T^B)$ is also the one that diagonalizes g_{AB}.) Then $\operatorname{tr}(T^A_R T^B_R) = T_R \delta^{AB}$ in *any* representation R. The symbol T_R is called the *Dynkin index* of the representation;

- *adjoint representation*: from the *Jacobi identity*
$$[[t^A, t^B], t^C] + [[t^C, t^A], t^B] + [[t^B, t^C], t^A] = 0$$
we can construct a similar identity for the structure constants
$$f^{DAB} f^{EDC} + f^{DCA} f^{EDB} + f^{DBC} f^{EDA} = 0.$$
From here, define
$$(T^A_{\mathrm{ad}})_{BC} \equiv -i f^{BCA}.$$
Then $[T^A_{\mathrm{ad}}, T^B_{\mathrm{ad}}] = i f^{CAB} T^C_{\mathrm{ad}}$ holds, since inserting the definition gives the Jacobi identity. Hence, T^A_{ad} are the generators of a $\dim(G)$-dimensional representation of G with $\mathcal{U}(g) = [\exp i\varepsilon^A(g) T^A_{\mathrm{ad}}]_{BC}$. *The adjoin representation is real*, because $i\varepsilon^A T^A_{\mathrm{ad}}$ is real.

Suppose ϕ^B is a set of $\dim G$ fields which transform in the adjoint representation, i.e. $\phi^B \to [\exp i\varepsilon^A T^A_{\mathrm{ad}}]_{BC}\phi^C$. Assume T^A_{ab} are the generators of another representation of dimension n. Then $\phi_{ab} \equiv \phi^A T^A_{ab}$ transforms as
$$\phi_{ab} \to \mathcal{U}_{aa'} \phi_{a'b'} \mathcal{U}^\dagger_{bb'}, \quad \mathcal{U} = e^{i\varepsilon^A T^A}$$
where T^A are the generators of this other representation.

Proof. Writing the infinitesimal transformation, we have

$$\begin{aligned} e^{i\varepsilon^A T^A}\phi^A T^A r^{-i\varepsilon^A T^A} &\approx \phi^A T^A + i\varepsilon^A [T^A, T^B]\phi^B + \dots \\ &= \phi^A T^A + i\varepsilon^A f^{ABC} T^C \phi^B \\ &= (\delta^{|}CB + i\varepsilon^A (T^A_{\text{ad}})_{CB})\phi^B T^C \\ &\approx (e^{i\varepsilon^A T^A_{\text{ad}}})_{CB}\phi^B T^C = \phi'^C T^C = \phi'_{ab}. \end{aligned}$$

□

The field strength tensor $G^A_{\mu\nu}$ transforms in the adjoint representation (see p. 6), but not A^A_μ due to the inhomogeneous term. This is consistent with A^A_μ and $G^A_{\mu\nu}$ being real fields.

Results for SU(N)

For SU(2) the basis of fundamental transformation are *Pauli matrices*

$$T^A_{\text{F}} = \frac{\sigma^A}{2}, \quad A = 1, 2, 3.$$

The structure constant is totally antisymmetric Levi–Civita symbol. There are also explicit expressions for T^A_{F} of SU(3) (called the *Gell-Mann matrices*). However, one rarely needs these explicit expressions. More important are the following sums and traces.

$$\begin{aligned} \operatorname{tr}(T^A_{\text{F}} T^B_{\text{F}}) &= \frac{1}{2}\delta^{AB}, \qquad (T^A_{\text{F}} T^A_{\text{F}})_{ab} = \frac{N^2-1}{2N}\delta_{ab} \\ \operatorname{tr}(T^A_{\text{ad}} T^B_{\text{ad}}) &= (T^A_{\text{ad}})_{CD}(T^B_{\text{ad}})_{DC} = -f^{CDA} f^{DCB} = f^{ACD} f^{BCD} = N\delta^{AB} \\ (T^A_{\text{ad}} T^A_{\text{ad}})_{\text{BC}} &= (T^A_{\text{ad}})_{BD}(T^A_{\text{ad}})_{DC} = -f^{BDA} f^{DCA} = N\delta^{BC}. \end{aligned}$$

1.1.3 Final form of the Lagrangian, self interaction and field equation

As we have shown before, the Lagrangian reads

$$\mathcal{L} = -\frac{1}{4g^2} G^A_{\mu\nu} G^{A,\mu\nu} + \mathcal{L}_{\text{M}}(\text{matter fields}).$$

It is more common to rescale the field $g\tilde{A}^A_\mu \equiv A^A_\mu$ and then rename $\tilde{A} \to A$. With this convention, we have

$$\mathcal{L} = -\frac{1}{4} G^A_{\mu\nu} G^{A,\mu\nu} + \mathcal{L}_{\text{M}}$$

and

$$\begin{aligned} &\text{D}_\mu = \partial_\mu - igA^A_\mu T^A \\ &\mathbf{A}_\mu \to \mathcal{U}\mathbf{A}_\mu\mathcal{U}^\dagger + \frac{i}{g}\mathcal{U}\partial_\mu\mathcal{U}^\dagger \qquad\qquad A^A_\mu \to A^A_\mu + f^{ABC}\varepsilon^C A^B_\mu + \frac{1}{g}\partial_\mu\varepsilon^A \\ &\mathbf{G}_{\mu\nu} = \frac{i}{g}[\text{D}_\mu, \text{D}_\nu], \qquad\qquad G^A_{\mu\nu} = \partial_\mu A^A_\nu - \partial_\nu A^A_\mu + g f^{ABC} A^B_\mu A^C_\nu. \end{aligned}$$

In this form the bilinear terms are independent of g, while g multiplies interaction vertices. We therefore interpret g as the coupling strength of the non-Abelian gauge interaction, similar to the electromagnetic coupling e.

Note the following important point. There is only one g whatever the matter field that transforms under the compact simple group. Also, the normalization of T_R^A is fixed by $\mathrm{tr}(T_R^A T_R^B) = \delta^{AB}/2$ for any R. There is no freedom for rescaling due to $[T_R^A, T_R^B] = if^{ABC}T_R^C$ once the basis for f^{ABC} is chosen. Hence, the coupling to matter fields is determined solely by specifying the representation. There is no freedom in choosing g. This is different from QED or U(1) Abelian gauge theory where the generator $T = -e_\psi$ can be chosen with a free value of e_ψ dor every different matter field.

Examples of gauge theories relevant to Nature are

- $\mathrm{U}(1)_{\mathrm{em}}$: Quantum Electrodynamics (QED), electromagnetic interaction. The field $A_\mu(x)$ represents the photon and the electron matter field transforms as $\Psi(x) \to e^{-i\varepsilon(x)e_\Psi}\Psi(x)$.
- $\mathrm{SU}(3)_{\mathrm{c}}$: Quantum Chromodynamics (QCD), strong interaction. The fields $A_\mu^A(x)$ represent gluons with $A = 1, \ldots, 8$ and the quarks transform in fundamental representation $q_a(x) \to (e^{i\varepsilon^A(x)T_{\mathrm{F}}^A})_{ab}q_b(x)$ with $a = 1, 2, 3$. Indices A and a are called *colour indices*.
- SU(2): weak interaction. The fields W_μ^A for $A = 1, 2, 3$ represent W-bosons. There are also left-handed quark and lepton fields and the Higgs field which transforms in the fundamental representation.

The standard model is completely determined by specifying which matter fields exist, by specifying their representation under $\mathrm{SU}(3)_{\mathrm{c}} \times \mathrm{SU}(2) \times \mathcal{U}(1)_Y$ and demanding gauge invariance. The subscript Y refers to the weak hypercharge and will be explained in due time.

Feynman rules and interaction vertices

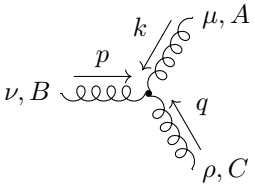
The gauge symmetry can be either U(1) or SU(N). The following holds for both.

$$\begin{aligned}\mathcal{L} &= -\frac{1}{4}G_{\mu\nu}^A G^{A,\mu\nu} + \overline{\Psi}(i\not{D} - m)\Psi \\ &= -\frac{1}{2}(\partial_\mu A_\nu^A \partial^\nu A^{A,\mu} - \partial_\mu A_\nu^A \partial^\nu A^{A,\mu}) + \overline{\Psi}(i\not{\partial} - m)\Psi \\ &\quad \underbrace{+ g\overline{\Psi}\gamma^\mu T^A \Psi A_\mu^A - gf^{ABC}(\partial_\mu A_\nu^A)A^{B,\mu}A^{C,\nu} - \frac{1}{4}g^2 f^{ABC}f^{ADE}A_\mu^B A_\nu^C A^{D,\mu}A^{E,\nu}}_{\mathcal{L}_{\mathrm{int}}}.\end{aligned}$$

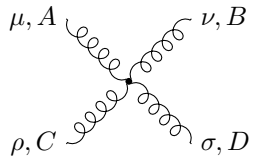
The Dirac fermion is in fundamental representation, so in notation $\Psi_{\alpha,a}$ the Dirac index α runs from 1 to 4 and the representation index runs from 1 to $n = \dim R = 3$.

- The term $m^2 A_\mu^A A^{A,\mu}/2$ is not allowed by gauge invariance. The vector particles must be massless.
- For non-Abelian gauge fields, $f^{ABC} \neq 0$ implies *gauge boson self-interaction*. They are completely determined by gauge invariance, because the A_μ^A field transforms in the adjoint representation (not trivially as the photon).

The interaction vertices in the momentum space are listed below.



$$= g f^{ABC} \left(g^{\mu\nu}(k-p)^\rho + g^{\nu\rho}(p-q)^\mu + g^{\rho\mu}(q-k)^\nu \right),$$



$$= \begin{aligned} &-ig^2[f^{ABE}f^{CDE}(g^{\mu\rho}g^{\nu\sigma} - g^{\mu\sigma}g^{\nu\rho}) + f^{ACE}f^{BDE}(g^{\mu\nu}g^{\rho\sigma} - g^{\mu\sigma}g^{\nu\rho}) \\ &\qquad + f^{ADE}f^{BCE}(g^{\mu\nu}g^{\rho\sigma} - g^{\mu\rho}g^{\nu\sigma})]. \end{aligned}$$

The fermion propagator is as usual

$$\overbrace{\Psi_{\alpha a}\overline{\Psi}_{\beta b}} \longrightarrow \frac{i(\not{p} + m)_{\alpha\beta}}{p^2 - m^2 + i\varepsilon}\delta_{ab}.$$

The propagator is diagonal in internal symmetry index. For the free gauge boson Lagrangian the standard procedure of inverting differential operator of the quadratic terms fails. The inverse does not exist. This is remedied by gauge-fixing and subsequent proper quantization (see Sec. 1.2). This will also lead to further vertices.

Field equations before gauge fixing

Start with Lagrangian

$$\mathcal{L} = -\frac{1}{4}G^A_{\mu\nu}G^{A,\mu\nu} + \mathcal{L}_{\mathrm{M}}(\Psi, \mathrm{D}_\mu\Psi).$$

To get the equation of motion, we first calculate the derivative with respect to $\partial_\mu A^A_\nu$:

$$\frac{\partial\mathcal{L}}{\partial(\partial_\mu A^A_\nu)} = -\frac{1}{2}G^B_{\rho\sigma}\frac{\partial G^{B,\rho\sigma}}{\partial(\partial_\mu A^A_\nu)} = -G^{A,\mu\nu}.$$

From this, Euler–Lagrange equations imply

$$\begin{aligned} \partial_\mu G^{A,\mu\nu} = -\frac{\partial\mathcal{L}}{\partial A^A_\nu} &= \frac{1}{2}G^B_{\rho\sigma}\frac{\partial G^{B,\rho\sigma}}{\partial A^A_\nu} - \frac{\partial\mathcal{L}_{\mathrm{M}}}{\partial(\mathrm{D}_\rho\Psi)}\frac{\partial(\mathrm{D}_\rho\Psi)}{\partial A^A_\nu} \\ &= gG^{B,\nu\rho}f^{BAC}A^C_\rho - ig\frac{\partial\mathcal{L}_{\mathrm{M}}}{\partial(\mathrm{D}_\nu\Psi)}T^A\Psi \equiv -gj^{A,\nu}. \end{aligned}$$

The quantity $-j^{A,\nu}$ is indeed the Noether current for the global gauge transformation $\delta\Psi = i\varepsilon^A T^A\Psi$ and $\delta A^C_\mu = f^{CBA}\varepsilon^A A^B_\mu$:

$$-j^{A,\nu} = \frac{\partial\mathcal{L}}{\partial(\partial_\nu A^C_\mu)}f^{CBA}A^B_\mu + \frac{\partial\mathcal{L}}{\partial(\partial_\nu\Psi)}iT^A\Psi = -G^{B,\nu\rho}f^{BAC}A^C_\rho + i\frac{\partial\mathcal{L}_{\mathrm{M}}}{\partial(\mathrm{D}_\nu\Psi)}T^A\Psi.$$

Hence, $\partial_\mu j^{A,\mu} = 0$, which can also be inferred because $\partial_\mu\partial_\nu G^{A,\mu\nu} = 0$ by antisymmetry. Note that $j^{A,\nu}$ and $\partial_\mu G^{A,\mu\nu}$ are not gauge invariant. This is because the gauge fields are self interacting

and therefore contribute to their source current. We can rewrite the current in a covariant form by defining the matter current

$$j_{\mathrm{M}}^{A,\mu} \equiv -i\frac{\partial_{\mathrm{M}}}{\partial(\mathrm{D}_\mu\Psi)}T^A\Psi.$$

Then $\mathrm{D}_\mu^{AB}G^{B,\mu\nu} = -gj_{\mathrm{M}}^{A,\nu}$ and $\mathrm{D}_\mu^{AB}j_{\mathrm{M}}^{B,\mu} = 0$. The latter states that the matter is "covariantly conserved". The first equation follows by rearranging the Noether current

$$-gj_{\mathrm{M}}^{A,\mu} = (\delta^{AB}\partial_\mu + gf^{ABC}A_\mu^C)G^{B,\mu\nu} = \mathrm{D}_\mu^{AB}G^{B,\mu\nu}.$$

The second equation follows from

$$\begin{aligned}\mathrm{D}_\mu^{AB}j_{\mathrm{M}}^{B,\mu} &= -\frac{1}{g}\mathrm{D}_\mu^{AB}\mathrm{D}_\nu^{BC}G^{C,\nu\mu} = -\frac{1}{2g}[\mathrm{D}_\mu,\mathrm{D}_\nu]_{AC}G^{C,\nu\mu}\\ &= \frac{i}{2}G_{\mu\nu}^B(T_{\mathrm{ad}}^B)_{AC}G^{C,\nu\mu} = \frac{1}{2}f^{ACB}G_{\mu\nu}^BG^{C,\nu\mu} = 0\end{aligned}$$

by antisymmetry of the adjoint indices in f^{ABC}. For Abelian U(1) gauge theory $f^{ABC} = 0$, hence $\mathrm{D}_\mu^{AB} \to \partial_\mu$ and the above equations reduce to the usual Maxwell equations with j_{M}^A being electromagnetic charge current.

1.2 Quantization of gauge theories with the Faddeev–Popov method

We will follow the same strategy as for the massive vector field:

1. start with Lagrangian,
2. identify the canonical variables and momenta,
3. construct the Hamiltonian,
4. write the Hamiltonian path integral,
5. reintroduce A^0 as an auxiliary integration variable,
6. integrate over canonical momenta,
7. obtain the Lagrange path integral.

In the course of this procedure, additional complications will arise from the massless vector field.

1.2.1 Path integral in axial gauge

The canonical momentum conjugate to A_μ^A is

$$\Pi_\mu^A = \frac{\partial\mathcal{L}}{\partial(\partial^0A^{A,\mu})} = G_{\mu0}^A.$$

By antisymmetry of $G_{\mu\nu}$ it follows $\Pi_0^A = 0$ so A_0^A is not a canonical variable and Π_0^A only acts as a constraint. This has to hold at all times, hence also $\partial^0\Pi^{A,0} = 0$ and the Euler–Lagrange equation for A_0^A is

$$\underbrace{\partial^0\Pi^{A,0}}_{0} - \partial^i\Pi^{A,i} + g\Pi^{B,i}f^{ABC}A^{C,i} + gj_{\mathrm{M}}^{A,0} = 0.$$

For the massive vector field the corresponding equation (p.??) contains an $m^2 A^0$ term and can be solved for A^0, which is the eliminated. But here it is independent of A^0 and therefore yields another constraint

$$\nabla \cdot \mathbf{\Pi}^A + g f^{ABC} \mathbf{\Pi}^B \cdot \mathbf{A}^C + g j_{\mathrm{M}}^{A,0} = 0$$

which is non-Abelian generalization of the Gauss law $\nabla \cdot \mathbf{E} = \rho$ in electrodynamics (there, $f^{ABC} = 0$, $\mathbf{\Pi} = \mathbf{E}$ and $ej^0 = \rho$). All bolded quantities here represent vectors in $\mathbb{R}^3$. The dynamics of $A^{A,0}$, however, is not determined by the equations. This always happens for gauge symmetries and is understandable, since the fact that the action is invariant under $A_\mu^A \to A_\mu^A + f^{ABC}\varepsilon^C A_\mu^B + \partial_\mu \varepsilon^A / g$ implies that there is a field redundancy in the description. The converse also holds. If there are constraints which cannot be solved for the canonical variable, then the system possesses a gauge symmetry.

Gauge fixing

The solution consists of fixing the gauge. That is, one imposes an additional condition, which removes the redundancy by *not* being gauge invariant, such that the system of constraints can be solved. Physical quantities should not (and do not, as will be seen) depend on how the redundancy is removed/gauge is fixed.

A standard gauge fixing in QED is Coulomb gauge $A^0 = 0$, $\nabla \cdot \mathbf{A} = 0$, which yields $\mathbf{A}$ as the three canonical coordinate fields subject to the one condition $\nabla \cdot \mathbf{A} = 0$ (transversality). Here we choose instead axial gauge $A^3 = 0$ (or more generally, $n^\mu A_\mu = 0$ for a spatial unit vector with $n^2 = -1$). This condition can always be fulfilled. If $A^3(x) \neq 0$, find a gauge transformation $\mathcal{U} = e^{i\varepsilon^A T^A}$ such that

$$0 = A'^3 = \mathcal{U} A^3 \mathcal{U}^\dagger + \frac{i}{g} \mathcal{U} \partial^3 \mathcal{U}^\dagger$$

or, equivalently,

$$\partial^3 e^{-i\varepsilon^A(x) T^A} = i g A^3(x) e^{-i\varepsilon^A(x) T^A}$$

which can be solved for $\varepsilon^A(x)$, given $A^3(x) \neq 0$. In axial gauge we immediately have only two canonical coordinate fields $A^{A,\mu}(x)$ with $\mu = 1, 2$ the number of which corresponds to the number of helicity states, ± 1. Since $\Pi^{A,3}$ is not a canonical momentum any more, we can determine A^{A0} from the Gauss constraint $\Pi^{A,3} = G^{A,30} = \partial^3 A^{A,0}$. This implies

$$-(\nabla^3)^2 A^{A,0} + \nabla^i \Pi^{A,i} + g f^{ABC} \Pi^{B,i} A^{C,i} + g j_{\mathrm{M}}^{A,0} = 0.$$

The index i sums only over 1 and 2. This is a second-order differential equation for $A^{A,0}(x)$ which can be solved. After these preliminaries, we can write down the standard Hamiltonian path integral formula for Green functions (p.??)

$$\begin{aligned} &\langle \Omega | T(\mathcal{O}_a(x_a)\mathcal{O}_b(x_b)\dots)|\Omega\rangle = \\ &\qquad |N|^2 \int \mathcal{D}[A^{A,i}]\, \mathcal{D}[\Pi^{A,i}]\, \mathcal{D}[\Psi, \Pi_\Psi] \\ &\qquad\qquad \exp\left(i \int \mathrm{d}^4 x \, (-\Pi^{A,i}(x) \dot{A}^{A,i}(x) + \Pi_\Psi \dot{\Psi} - \mathcal{H} + i\varepsilon) \right) \mathcal{O}_a(x_a)\mathcal{O}_b(x_b)\dots \end{aligned} \tag{1.2}$$

Here $i = 1, 2$ and Hamiltonian depends on $\Pi^{A,i}$ and $A^{A,i}$.

Introduction of $A^{A,0}$ as auxiliary integration variable

The Hamiltonian is obtained from $\mathcal{L}$ as usual by Legendre transformation

$$\mathcal{H} = \frac{1}{2}\Pi^{A,i}\Pi^{A,i} - \Pi^{A,i}(\partial^i A^{A,0} - g f^{ABC} A^{B,0} A^{C,i}) + \frac{1}{4} G^{A,ij} G^{A,ij} - \frac{1}{2}\partial^3 A^{A,0} \partial^3 A^{A,0} + \mathcal{H}_{\mathrm{M}} \quad (1.3)$$

where $A^{\mu,0}$ should be eliminated by the solution of the Gauss constraint. Consider now the following expression

$$\int \mathcal{D}[A^{A,0}, A^{A,i}, \Pi^{A,i}, \Psi, \Pi_\Psi] \, \exp\left(i \int \mathrm{d}^4x (-\Pi^{A,i}(x)\dot{A}^{A,i}(x) + \Pi_\Psi \dot{\Psi} - \mathcal{H}' + i\varepsilon) \right) \mathcal{O}_a(x_a)\mathcal{O}_b(x_b)\dots \quad (1.4)$$

where $\mathcal{H}'$ is the expression 1.3 with $A^{A,0}$ *not* fixed by the Gauss constraint, since it is now an integration variable. The expression 1.3 is a quadratic form in $A^{A,0}$, so we can perform the path integral over $A^{A,0}$ in 1.4. This replaces $A^{A,0}$ in the exponent by the solution to (p.??)

$$\frac{\delta}{\delta A^{A,0}}(\text{exponent}) = -\frac{\partial \mathcal{H}'}{\partial A^{A,0}} = 0$$

and adds the determinant of $(\partial^3)^2\delta^{(4)}(x-y)$ according to the standard formula for the Gaussian path integral. Since this determinant is field independent it cancels with the denominator (or can be absorbed in the normalization factor $|N|^2$). Now,

$$\begin{aligned} \frac{\partial \mathcal{H}'}{\partial A^{A,0}} &= \partial^i \Pi^{A,i} + g f^{BAC} \Pi^{B,i} A^{C,i} + (\partial^3)^2 A^{A,0} - g j^{A,0}_{\mathrm{M}} \\ &= -\nabla^i \Pi^{A,i} - g f^{ABC} \Pi^{B,i} A^{C,i} - g j^{A,0}_{\mathrm{M}} \end{aligned}$$

is exactly the Gauss constraint that determines $A^{A,0}$, so $\mathcal{H}'$ goes to $\mathcal{H}$ after integration over $A^{A,0}$ and 1.4 equals 1.2.

At this point, we have reintroduced $A^{A,0}$. We assume that $\mathcal{H}_{\mathrm{M}}$ is quadratic in Π_Ψ. Then $\mathcal{H}'$ is quadratic in the canonical momenta, and we can perform the corresponding Gaussian path integrals. Again, this sets Π_Ψ and $\Pi^{A,i}$ to their value at the stationary point

$$0 = \frac{\delta}{\delta \Pi_\Psi}(\text{exponent}) = \dot{\Psi} - \frac{\partial \mathcal{H}_{\mathrm{M}}}{\partial \Pi_\Psi} \quad \Longrightarrow \quad \mathcal{L}_{\mathrm{M}} = \Pi_\Psi \dot{\Psi} - \mathcal{H}_{\mathrm{M}}.$$

For fermionic Ψ, $\Pi_\Psi \sim \overline{\Psi}$ and we do not perform the integral. The other stationary point is

$$0 = \frac{\delta}{\delta \Pi^{A,i}}(\text{exponent}) = -\partial^0 A^{A,i} - \Pi^{A,i} + \partial^i A^{A,0} - g f^{ABC} A^{B,0} A^{C,i} = G^{A,i0} - \Pi^{A,i}.$$

From this it follows that the exponent becomes

$$\text{exponent} = -\frac{1}{4} G^{A,\mu\nu} G^A_{\mu\nu}|_{A^A_3 = 0} + \mathcal{L}_{\mathrm{M}}.$$

Hence,

$$\langle \Omega | T(\mathcal{O}_a(x_a)\mathcal{O}_b(x_b)\dots)|\Omega\rangle = |N|^2 \int \mathcal{D}[A^{A,\mu}]\,\mathcal{D}[\Psi]\,\delta(A^{A3}) \exp\left(i \int \mathrm{d}^4x\, \mathcal{L} \right) \mathcal{O}_a(x_a)\mathcal{O}_b(x_b)\dots \quad (1.5)$$

Note that for fermionic Ψ the measure is $\mathcal{D}[\Psi, \overline{\Psi}]$. We integrate over all $A^{A,\mu}$ for $\mu = 0, 1, 2, 3$ and implement $A^{A,3}(x) = 0$ via a functional δ-function

$$\delta(A^{A,3}) = \prod_A \prod_{\mathbf{x}} \delta(A^{A,3}(\mathbf{x})).$$

The above is the standard form of the Lagrangian path integral except for the gauge-fixing δ-functional. How do we interpret such expression? If $\delta(A^{A,3})$ were not present, the path integral would sum over all gauge field configurations, including those related by

$$\mathbf{A}'_\mu = \mathcal{U}\mathbf{A}_\mu\mathcal{U}^\dagger + i\mathcal{U}\partial_\mu\mathcal{U}^\dagger \quad \text{for some } \mathcal{U},$$

which do on represent physically distinct configurations. The gauge-fixing condition selects one number of the gauge orbit (the set of all $\mathbf{A}'$ obtained from some $\mathbf{A}$ by any $\mathcal{U}$) of $\mathbf{A}_\mu$ and eliminates the redundancy/over-counting of gauge-equivalent field configurations.

1.2.2 Faddeev–Popov method and ghosts

The non-covariant axial gauge is not convenient in practice. The above interpretation suggests that any other condition that picks a member of the gauge orbit could be chosen. This will be implemented now in several steps. First, parametrize the gauge transformation by $\mathcal{U}(\varepsilon)$. Then perform a change of variables in the path integral $\mathbf{A}^\mu, \Psi \to \mathbf{A}^\mu_\varepsilon, \Psi_\varepsilon$. The index ε denotes the gauge transformation of the fields for the given ε. If the path integral measure is invariant, and the operators $\mathcal{O}_a$ and $\mathcal{O}_b$ are invariant, both of which will be assumed (see remark below), the only change in 1.5 is $\delta(A^{A,3}) \to \delta(A^{A,3}_\varepsilon)$ since $\mathcal{L}$ is gauge invariant.

Remark. The measure for the gauge field *is* invariant. Under an infinitesimal gauge transformation

$$A'^A_\mu = A^A_\mu + f^{ABC}\varepsilon^C A^B_\mu + \frac{1}{g}\partial_\mu\varepsilon^A$$

the variation of the transformed field with respect to the original field is

$$\frac{\delta A'^A_\mu(x)}{\delta A^B_\nu(y)} = \delta^{AB}\delta^\nu_\mu\delta^{(4)}(x-y) + \delta^\nu_\mu\delta^{(4)}(x-y)f^{ABC}\varepsilon^C(x).$$

The Jacobi determinant can be calculated as follows

$$\begin{aligned}\det\left|\frac{\delta A'^A_\mu(x)}{\delta A^B_\nu(y)}\right| &= \exp\operatorname{tr}\ln\left(\delta^{AB}\delta^\nu_\mu\delta^{(4)}(x-y) + \delta^\nu_\mu\delta^{(4)}(x-y)f^{ABC}\varepsilon^C(x)\right)\\ &\approx \exp\operatorname{tr}(\delta^\nu_\mu\delta^{(4)}(x-y)f^{ABC}\varepsilon^C(x)) = 0\end{aligned}$$

since the trace has to be taken in $\mu\nu$, AB and xy and $f^{AAC} = 0$. The fermion measure, however, might be anomalous; this will be discussed in chapter ??. We must assume here that it is anomaly-free.

Now let $f^A[A]$ be a gauge-fixing functional and let the gauge-fixing condition be $f^A[\mathbf{A}(x)] = F^A(x)$. Define

$$M^{AB}[\mathbf{A}_\mu](x,y) = \left.\frac{\delta f^A[\mathbf{A}_\varepsilon(x)]}{\delta\varepsilon^B(y)}\right|_{\varepsilon=0}.$$

This can replace $\delta(A^3_\varepsilon) \to \delta(f^A[\mathbf{A}^\mu_\varepsilon] - F^A) \det M[\mathbf{A}]$ in 1.5, i.e. the path integral formula now reads

$$\begin{aligned}\langle\Omega|T(\mathcal{O}_a(x_a)\mathcal{O}_b(x_b)\dots)|\Omega\rangle = |N|^2 \int \mathcal{D}[A^{A,\mu}]\,\mathcal{D}[\Psi]\delta(f^A(A^{A',\mu}_\varepsilon) - F^A) \\ \det(M[A^{A'}_\mu]) \exp\left(i\int \mathrm{d}^4x\,\mathcal{L}\right)\mathcal{O}_a(x_a)\mathcal{O}_b(x_b)\dots\end{aligned} \tag{1.6}$$

This says that 1.5 is equivalent to a path integral with any other gauge fixing condition, since $f^A[A^{A'}_\mu]$ and $F^A(x)$ are arbitrary.

Proof. First, 1.5 is a special case $f^A[\mathbf{A}_\mu] = A^3_\mu$ and $F^A = 0$. In this case

$$\begin{aligned}M^{AB}[\mathbf{A}](x,y) &= \frac{\delta}{\delta\varepsilon^B(y)}(A^{A,3}(x) + f^{ADC}\varepsilon^C(x)A^{D,3}(x) + \partial^3\varepsilon^A(x)/g) \\ &= f^{ADB}\delta^{(4)}(x-y)A^{D,3}(x) + \frac{1}{g}\partial^3\delta^{(4)}(x-y)\delta^{AB}.\end{aligned}$$

The first term is zero, because $A^{A,3} = 0$ and thus $\det M^{AB}[\mathbf{A}_\mu]$ is independent of $\mathbf{A}_\mu$ and can be taken out of the path integral where it cancels with the denominator in $|N|^2$. This shows that $1.5 = 1.6$ in the special case of f^A and F^A. Second, we show that 1.6 is actually independent of f^A and F^A, which proves the assertion for general f^A, F^A. To see this, we multiply the numerator and denominator of the path integral with the constant factor

$$\int \mathcal{D}[\varepsilon^A]\,1 = \mathrm{vol}(G)^{\#N},$$

where $\mathrm{vol}(G)$ is the volume of the gauge group G (which is compact and thus has finite volume) and $\#N$ is the number of points in the space-time. This is well-defined in the discretized definition of the path integral. Since 1.6 is independent of ε, we can pull $\mathcal{D}[\varepsilon^A]$ into the integral. Then we get

$$\int \mathcal{D}[\varepsilon^A]\,\delta(f^A[\mathbf{A}^\mu_\varepsilon] - F^A) \det \left.\frac{\delta f^A[\mathbf{A}^\mu_\varepsilon]}{\delta\varepsilon^B}\right|_{\varepsilon=0} \overset{\varepsilon^A \to f^a}{=} \int \mathcal{D}[f^A]\,\delta(f^A - F^A) = 1.$$

This is independent of f^A and F^A and so is 1.6. □

We use the same trick again to bring 1.6 into yet another form. The purpose is to get rid of the δ-functions, which is difficult to implement in practice. Let $G[f^A]$ be a functional of functions f^A. Since F^A in 1.6 is arbitrary and 1.6 is actually independent of F^A as shown above, we can multiply the numerator and denominator of 1.6 with the constant $\mathcal{D}[F^A]\,G[F^A]$ which gives

$$\int \mathcal{D}[F^A]\,G[F^A] = \delta(f^A[\mathbf{A}^\mu_\varepsilon] - F^A) = G[f^A[\mathbf{A}^\mu_\varepsilon]].$$

This means we can perform a replacement $\delta(f^A[\mathbf{A}^\mu_\varepsilon] - F^A) \to G[f^A[\mathbf{A}^\mu_\varepsilon]]$ in 1.6. Then finally choose

$$G[f^A] \equiv \exp\left(i\int \mathrm{d}^4x\,\left(-\frac{1}{2\xi}f^A(x)f^A(x)\right)\right)$$

where ξ is an arbitrary real parameter. Then

$$\begin{aligned}\langle\Omega|T(\mathcal{O}_a(x_a)\mathcal{O}_b(x_b)\dots)|\Omega\rangle = |N|^2 \int \mathcal{D}[A^{A,\mu}]\,\mathcal{D}[\Psi]\det(M^{AB}[\mathbf{A}_\mu]) \\ \exp\left[i\int \mathrm{d}^4x\,\left(\mathcal{L} - \frac{1}{2\xi}f^Af^A\right)\right]\mathcal{O}_a(x_a)\mathcal{O}_b(x_b)\dots\end{aligned}$$

In this form the *gauge-fixing term* can be interpreted as an additional term in the Lagrangian

$$\mathcal{L}_{\mathrm{GF}} = -\frac{1}{2\xi} f^A f^A.$$

In a final step one would like to write $\det(M^{AB}[\mathbf{A}_\mu])$ as an exponential, too. This can be done with the Gaussian path integral for Grassmann fields (see p.??)

$$\det \frac{\delta f^A[\mathbf{A}^\mu_\varepsilon(x)]}{\delta \varepsilon^B}\bigg|_{\varepsilon=0} = \int \mathcal{D}[c^A, \bar{c}^A] \, \exp\left(\int \mathrm{d}^4x \, \mathrm{d}^4y \, \bar{c}^A(x) \frac{\delta f^A[\mathbf{A}^\mu_\varepsilon](x)}{\delta \varepsilon^B(y)}\bigg|_{\varepsilon=0} \right).$$

We introduced two new fields, a *ghost field* $c^A(x)$ and *anti-ghost field* $\bar{c}^A(x)$. We consider $c^A, \bar{c}^A$ as two independent, real Grassmann fields – the bar on $\bar{c}^A$ is only a piece of notation. They are scalar fields under Lorentz transformations (no spinor index), but are anti-commuting. They transform in the adjoint representation of the gauge group, i.e. $A = 1, \ldots, \dim G$. For a local gauge-fixing conditions $f^A(x)$ depends only on $A^A(x)$, so

$$\frac{\delta f^A[\mathbf{A}^\mu_\varepsilon](x)}{\delta \varepsilon^B(y)}\bigg|_{\varepsilon^A=0} = M^{AB}[\mathbf{A}_\mu](x)\delta^{(4)}(x-y)$$

and hence the path integral is

$$\langle \Omega | T(\mathcal{O}_a(x_a)\mathcal{O}_b(x_b)\ldots)|\Omega\rangle = |N|^2 \int \mathcal{D}[A^{A,\mu}, \Psi, c^A, \bar{c}^A] \, e^{i}S, \quad S = \int \mathrm{d}^4x \, (\mathcal{L} + \mathcal{L}_{\mathrm{GF}} + \mathcal{L}_{FP})$$

where

$$\mathcal{L}_{\mathrm{FP}} = \bar{c}^A(-iM^{AB}[\mathbf{A}_\mu])c^B$$

is the *ghost Lagrangian*. A common choice of gauge fixing is the *covariant gauge*, $f^A(x) = \partial_\mu A^{A,\mu}(x)$ where Lorentz invariance remains manifest. In this case it holds

$$\begin{aligned} \frac{\delta f^A[\mathbf{A}^\mu_\varepsilon(x)]}{\delta \varepsilon^B(y)} &= \frac{\delta(\partial_\mu(f^{ADC}\varepsilon^C A^{D,\mu} + \partial^\mu \varepsilon^A/g)(x))}{\delta \varepsilon^B(y)}\bigg|_{\varepsilon=0} \\ &= (\partial^{(x)}_\mu \partial^\mu_{(y)}\delta^{AB} + g f^{ABC}\partial^{(x)}_\mu A^{C,\mu})\frac{1}{g}\delta^{(4)}(x-y). \end{aligned} \tag{1.7}$$

Integrating the ghost term in the Lagrangian gives

$$\begin{aligned} \frac{1}{g}\int \mathrm{d}^4x \int \mathrm{d}^4y \, \bar{c}^A(x) \left(\partial^2_{(x)}\delta^{AB}\delta^{(4)}(x-y) + g f^{ABC}\partial^{(x)}_\mu (A^{C,\mu}(x)\delta^{(4)}(x-y))\right) c^B(y) = \\ = \frac{i}{g} i \int \mathrm{d}^4x \, \left(\partial_\mu \bar{c}^A \partial^\mu c^A - g f^{ABC}(\partial_\mu \bar{c}^A) c^B A^{C,\mu}\right). \end{aligned}$$

Redefine the ghost field $c^A \to ic^A$ to absorb the i and drop the $1/g$ because it gives an irrelevant constant (or alternatively redefine $c^A \to -igc^A$). Then

$$\begin{aligned} \tilde{L} &= \mathcal{L} + \mathcal{L}_{\mathrm{GF}} + \mathcal{L}_{\mathrm{FP}} \\ &= -\frac{1}{4}G^A_{\mu\nu}G^{A,\mu\nu} - \frac{1}{2\xi}(\partial_\mu A^{A,\mu})(\partial_\nu A^{A,\nu}) + \partial_\mu \bar{c}^A \partial^\mu c^A - g f^{ABC}(\partial_\mu \bar{c}^A) c^B A^{C,\mu} + \mathcal{L}_{\mathrm{M}}(\Psi, \mathrm{D}_\mu \Psi). \end{aligned}$$

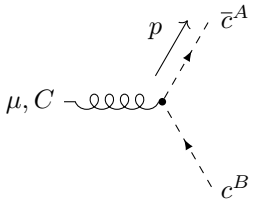
Note that c^A is now purely imaginary ($c^{A\dagger} = -c^A$) while $\bar{c}^A$ remains real. This is consistent with $\tilde{\mathcal{L}}$ being real/hermitian:

$$(\partial_\mu \bar{c}^A \partial^\mu c^A)^\dagger = \partial^\mu c^{A\dagger} \partial_\mu \bar{c}^{A\dagger} = -\partial^\mu c^A \partial_\mu \bar{c}^A = \partial_\mu \bar{c}^A \partial^\mu c^A.$$

Formally, the ghost kinetic term looks like that of a complex scalar field, but it is Grassmannian. It interacts with the gauge field. We can now complete the Feynman rules of non-Abelian gauge theory.

$$\nu, B \sim\sim\sim \mu, A \quad = \delta^{AB} \frac{i}{k^2 + i\varepsilon} \left(-g_{\mu\nu} + (1-\xi) \frac{k_\mu k_\nu}{k^2} \right),$$

$$B \text{ - - - - - } A \quad = \delta^{AB} \frac{i}{k^2 + i\varepsilon},$$



$$= +g f^{ABC} p^\mu.$$

(1) The derivation of the path integral assumed $\mathcal{O}_a(x_a)\ldots$ are gauge-invariant. In this case the derivation shows that the Green functions $\langle\Omega|T(\mathcal{O}_a(x_a)...)|\Omega\rangle$ are also gauge-invariant. In particular, they are independent of the gauge-parameter ξ of the covariant gauge. We will also use the formula to compute Green functions of the elementary fields $A^{A,\mu}, \Psi, c^A, \overline{c}^A$ which transform non-trivially. These Green functions are not gauge-independent. This makes us wonder whether the S-matrix elements computed via the LSZ procedure are gauge-invariant. We come back to this at the end of this chapter.

(2) We motivated non-Abelian gauge symmetry by the desire to have long-range interactions mediated by spin-1 fields for "matter fields" transforming under an internal symmetry. Now we see that the non-Abelian gauge fields have a life of their own, because they are self-interacting.

The theory of only non-Abelian gauge and ghost fields is called *pure Yang–Mills theory*. Its generating functional is

$$Z[J^A_\mu, \eta^A, \overline{\eta}^A] \equiv N \int \mathcal{D}[A^A_\mu, c^A, \overline{c}^A] \exp\left(i \int \mathrm{d}^4x \, (\mathcal{L} + \mathcal{L}_{\text{GF}} + \mathcal{L}_{\text{FP}} + J^A_\mu A^{A,\mu} + \overline{\eta}^A c^A + \overline{c}^A \eta^A) \right).$$

This generates the n-point functions of gauge and ghost fields.

1.3 Renormalization of non-Abelian gauge theories

1.3.1 BRST symmetry and Ward identities

The gauge-fixed Lagrangian $\tilde{\mathcal{L}} = \mathcal{L} - f^A f^A/(2\xi) + \overline{c}^A \Delta^A$ with

$$\Delta^A(x) = -g\alpha \int \mathrm{d}^4y \, \frac{\delta f^A[\mathbf{A}^\mu_\varepsilon(x)]}{\delta \varepsilon^B(y)}\bigg|_{\varepsilon=0} c^B(y) \overset{\text{covariant gauge}}{=} -\alpha \partial_\mu (\mathrm{D}^\mu c)^A(x) \tag{1.8}$$

is no longer gauge invariant (α is an arbitrary parameter, previously set to 1 to make the ghost kinetic term canonically normalized). However, it turns out to be invariant under a global, fermionic symmetry. The Ward identities of this symmetry are useful in the proof of renormalizability of non-Abelian gauge theories and the gauge-invariance of the S-matrix. The symmetry

is most easily expressed by introducing a bosonic auxiliary field (the *Nakanishi–Lautrup field*) $B^A(x)$ as follows:

$$G[f^A] = \exp\left(i\int \mathrm{d}^4x\, -\frac{f^A f^A}{2\xi}\right) = \int \mathcal{D}[B^A]\exp\left(i\int \mathrm{d}^4x\, B^A(x)(f^A(x)+\xi B^A(x)/2)\right).$$

Integrating over B^A sets $B^A \to -f^a/\xi$ in the exponent. An irrelevant constant is dropped when writing "=". Then the path integral formula involves $\mathcal{D}[B^A]$ as well and $\tilde{\mathcal{L}}$ reads instead

$$\mathcal{L} = \mathcal{L} + B^A f^A + \frac{1}{2}\xi B^A B^A + \bar{c}^A \Delta^A.$$

Definition of the infinitesimal BRST transformation

For a field $\phi \in \{A, \Psi, c, \bar{c}, B\}$ define the *BRST transformation* (named after Becchi, Rouet, Stora and Tyutin) as

$$\delta_\lambda \phi(x) \equiv \lambda s\phi(x)$$

where s is the BRST operator acting on field ϕ and λ is infinitesimal Grassmann parameter. On matter and gauge fields δ_λ acts like a gauge transformation with parameter $\varepsilon^A(x) = g\lambda c^A(x)$ (not that this is bosonic), hence $\delta_\lambda \mathcal{L} = 0$, because $\mathcal{L}$ is gauge invariant. On the other fields define δ_λ such that $\delta_\lambda \tilde{L} = 0$:

$$\begin{aligned}
\delta_\lambda \Psi &= ig\lambda c^A T^A \Psi,\\
\delta_\lambda A^A_\mu &= \lambda(\partial_\mu \delta^{AC} + g f^{ABC} A^B_\mu)c^C = \lambda \mathrm{D}^{AC}_\mu c^C,\\
\delta_\lambda c^A &= -\frac{1}{2} g\lambda f^{ABC} c^B c^C,\\
\delta_\lambda \bar{c}^A &= \frac{1}{\alpha}\lambda B^A\\
\delta_\lambda B^A &= 0.
\end{aligned}$$

This is a non-linear symmetry and is consistent with the antihermitian nature of c^A. The BRST transformation has the following properties:

(1) it is nilpotent, i.e. $s^2\mathcal{O}(x) = 0$ for any operator $\mathcal{O}(x)$,

(2) $\delta_\lambda \tilde{\mathcal{L}} = 0$.

Proof.

(1) Show this first for the fundamental fields. Using $c^A \lambda c^B = -\lambda c^A c^B$ and $T^A c^A c^B T^B = [T^A, T^B]c^A c^B/2 = i f^{ABC} T^C c^A c^B/2$ we first write the fermion transformation:

$$\begin{aligned}
\delta_\lambda(s\Psi) &= igT^A \delta_\lambda(c^A \Psi)\\
&= -\frac{i}{2} g^2 \lambda f^{ABC} T^A c^B c^C \Psi + g^2 \lambda T^A c^A c^B T^B \Psi\\
&= -\frac{i}{2} g^2 \lambda f^{ABC} T^A c^B c^C \Psi + \frac{i}{2} g^2 \lambda f^{ABC} T^A c^B c^C \Psi = 0.
\end{aligned}$$

The gauge field transformation vanishes due to Jacobi identity in the last step

$$\begin{aligned}\delta_\lambda(sA^A_\mu) &= \delta_\lambda(\partial_\mu c^A) + gf^{ABC}\delta_\lambda(A^C_\mu c^B)\\ &= -\frac{1}{2}g\lambda f^{ABC}\partial_\mu(c^B c^C) - gf^{CAB}\lambda(\mathrm{D}^{CD}_\mu c^D)c^B - gf^{CAB}A^C_\mu\frac{(-g)}{2}\lambda f^{BDE}c^D c^E\\ &= \lambda g^2(f^{CAB}f^{CDE}c^D c^B A^E_\mu + f^{CAB}f^{BDE}c^D c^E A^C_\mu/2)\\ &= \frac{1}{2}\lambda g^2 A^C_\mu c^D c^E(2f^{BAE}f^{BDC} + f^{CAB}f^{BDE}) = 0.\end{aligned}$$

The ghost field cancels by the antisymmetry of the structure constants

$$\begin{aligned}\delta_\lambda(sc^A) &= \frac{1}{2}gf^{ABC}\delta_\lambda(c^B c^C)\\ &= \frac{1}{4}g^2 f^{ABC}(\lambda f^{BDE}c^D c^E c^C + c^B\lambda f^{CDE}c^D c^E)\\ &= \frac{1}{2}g^2\lambda f^{ABC}f^{BDE}c^D c^E c^C = 0\end{aligned}$$

Nilpotency on $\overline{c}^A$ and B^A is trivial due to $\delta_\lambda B^A = 0$. Hence, s is nilpotent on arbitrary products of fields (not necessary at the same point):

$$\begin{aligned}\delta_\lambda(s(\phi_1\phi_2)) &= \delta_\lambda(s(\phi_1)\phi_2 + (-1)^{|\phi_1|}\phi_1 s(\phi_2))\\ &= (-1)^{|\phi_1|+1}s(\phi_1)s(\phi_2) + (-1)^{|\phi_1|}s(\phi_1)s(\phi_2) = 0.\end{aligned}$$

We used $s^2(\phi_{1,2}) = 0$ and $|s\phi_1| = |\phi_1| \pm 1$ (since $s(\phi_1)$ is bosonic/fermionic when ϕ_1 is fermionic/bosonic). The arbitrary product of fields can be proven by induction.

(2) We only need to check the last three terms in $\tilde{\mathcal{L}}$ as $\delta_\lambda\mathcal{L} = 0$,

$$\delta_\lambda\tilde{\mathcal{L}} = \delta_\lambda(B^A f^A + \xi B^A B^A/2 + \overline{c}^A\Delta^A).$$

First we calculate the transformation on f^A. We know that BRST changes A as a gauge transformation with $\varepsilon^A = g\lambda c^A$

$$\delta_\lambda f^A[\mathbf{A}_\mu] = \int \mathrm{d}^4 y\,\frac{\delta f^A[\mathbf{A}^\mu_\varepsilon(x)]}{\delta\varepsilon^B(y)}\bigg|_{\varepsilon=0} g\lambda c^B(y) = -\frac{1}{\alpha}\lambda\Delta^A.$$

The last step was a definition from page 19. Instead of evaluating $\delta_\lambda\tilde{\mathcal{L}}$ directly, we use this to note that

$$\delta_\lambda(\overline{c}^A f^A + \xi\overline{c}^A B^A/2) = (\delta_\lambda\overline{c}^A)(f^A + \xi B^A/2) + \overline{c}^A\delta_\lambda f^A = \frac{1}{\alpha}\lambda(B^A f^A + \xi B^A B^A/2 + \overline{c}^A\Delta^A).$$

Hence,

$$\tilde{L} = \mathcal{L} + \alpha s(\overline{c}^A f^A + \xi\overline{c}^A B^A/2).$$

Because of the nilpotency of s, it holds $\delta_\lambda\tilde{\mathcal{L}} = 0$.

□

This gives a more general formulation of gauge fixing. We can modify $\mathcal{L}$ to

$$\tilde{L} = \mathcal{L}[A^A_\mu, \Psi] + sF[A^A_\mu, \Psi, c^A, \overline{c}^A, B^A]$$

for any F. The sum of the gauge-fixing and ghost Lagrangians is the image of the BRST operator and $\tilde{\mathcal{L}}$ is the kernel. Then $\tilde{\mathcal{L}}$ is automatically BRST invariant and physical observations should not depend on whether one adds an arbitrary term in the image of the BRST operator.

BRST charge and Ward identities

For the remainder of this section, we will use the covariant gauge fixing. For $f^A(x) = \partial^\mu A^A_\mu$ the gauge-fixed Lagrangian reads

$$\tilde{\mathcal{L}} = -\frac{1}{4} G^A_{\mu\nu} G^{A,\mu\nu} + \mathcal{L}_{\mathrm{M}}(\Psi, \mathrm{D}_\mu \Psi) + (\partial^\mu \bar{c}^A)(\mathrm{D}^{AB}_\mu c^B) + B^A(\partial^\mu A^A_\mu + \xi B^A/2).$$

As for any global symmetry, the BRST symmetry $\delta_\lambda \phi = \lambda s\phi$ gives rise to a global conserved Noether current j^{BRST}_μ and conserved charge

$$Q_{\mathrm{BRST}} = \int \mathrm{d}^3\mathbf{x} j^{\mathrm{BSRT}}_0(x)$$

such that $\delta_\lambda \phi = \lambda s\phi = i[\lambda Q_{\mathrm{BRST}}, \phi]$. Because s is nilpotent, it holds $(Q_{\mathrm{BRST}})^2 = 0$. We will not the explicit expression for Q_{BRST}. The Green functions of non-Abelian gauge theory satisfy the usual equation-of-motion identities and the Ward identities of BRST symmetry (p.??). This special case is called the *Slavnov–Taylor formula*

$$\sum_{k=1}^{m} \langle \Omega | T(\phi_{n_1}(x_1) \dots \phi_{n_{k-1}}(x_{k-1}) \lambda s(\phi_{n_k}(x_k)) \phi_{n_{k+1}}(x_{k+1}) \dots \phi_{n_m}(x_m)) | \Omega \rangle = 0 \tag{1.9}$$

for $\phi_{n_i} \in \{\Psi, A^A_\mu, c^A, \bar{c}^A, B^A\}$. This assumes that BRST symmetry is non-anomalous. One can show that this is true whenever the gauge symmetry is non-anomalous. For the proof of renormalizability one uses the corresponding identity for the generating functional. The equation of motion gives

$$\begin{aligned} \langle \Omega | T\Big(\frac{\delta S[\phi_{n'}]}{\delta \phi_n} \phi_{n_1}(x_1) \dots \phi_{n_m}(x_m)\Big) | \Omega \rangle = \\ = i \sum_{i=1}^{m} \delta^{(4)}(x - x_i) \delta_{n n_i} \langle \Omega | T(\phi_{n_1}(x_1) \dots \widehat{\phi_{n_i}(x_i)} \dots \phi_{n_m}(x_m)) | \Omega \rangle \end{aligned} \tag{1.10}$$

where hat denotes the omission of the term. If the operator is fermionic, the formula gets an additional minus sign.

Transversality of the gauge boson self-energy

One of the first applications of Slavnov–Taylor formula is the proof that the gauge boson self-energy is transverse to all orders in perturbation theory. Consider the two-point function

$$\langle \Omega | T(A^A_\mu(x) A^B_\nu(y)) | \Omega \rangle \equiv \int \frac{\mathrm{d}^4 k}{(2\pi)^4} e^{-ik(x-y)} G^{AB}_{\mu\nu}(k).$$

The most general form of the Green function is

$$G^{AB}_{\mu\nu} = \delta^{AB} \frac{i}{k^2 + i\varepsilon} \left(\left(-g_{\mu\nu} + \frac{k_\mu k_\nu}{k^2} \right) A(k^2) - \frac{k_\mu k_\nu}{k^2} B(k^2) \right).$$

We split the function into transverse part $A(k^2)$ and longitudinal part $B(k^2)$. In the lowest order, this is the gauge boson propagator, hence

$$A(k^2) = 1 + \mathcal{O}(g^2), \quad B(k^2) = \xi + \mathcal{O}(g)^2.$$

We will now show that $B(k^2) = \xi$ is exact, i.e. no correction to any order in perturbation theory. We start with the equation of motion 1.10 with no contact terms

$$0 = \langle\Omega|T(\frac{\delta S}{\delta B^A(x)}\partial^\mu A^B_\mu(y))|\Omega\rangle = \langle\Omega|T((\partial^\nu A^A_\nu(x) + \xi B^A(x))\partial^\mu A^B_\mu(x))|\Omega\rangle\,.$$

Now take the second derivative of the two-point function

$$\begin{aligned}
\partial^\nu_{(x)}\partial^\mu_{(y)}\,\langle\Omega|T(A^A_\nu(x)A^B_\mu(x))|\Omega\rangle &\overset{1}{=} -\xi\,\langle\Omega|T(s(\overline{c}^A(x))\partial^\mu A^B_\mu(y))|\Omega\rangle \\
&\overset{2}{=} -\xi\,\langle\Omega|T(\overline{c}^A(x)s(\partial^\mu A^V_\mu(y)))|\Omega\rangle \\
&\overset{3}{=} -\xi\,\langle\Omega|T(\overline{c}^A\partial^\mu(\mathrm{D}^{BC}_\mu c^C(y)))|\Omega\rangle \\
&\overset{4}{=} \xi\,\langle\Omega|T(\frac{\delta S}{\delta\overline{c}^B(y)}\overline{c}^A(x))|\Omega\rangle \\
&\overset{5}{=} -i\xi\delta^{AB}\delta^{(4)}(x-y).
\end{aligned}$$

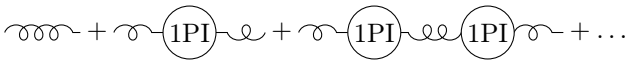
Each step of this derivation is worth commenting.

1. We used $B^A = s(\overline{c}^A)$ with $\alpha = 1$ and $\xi B^A = -\partial^\mu A^A_\mu$ in covariant gauge.
2. Slavnov–Taylor identity 1.9 only has two terms, so we can move the operator to the other field. The minus sign comes from commuting λ and $\overline{c}^A$.
3. Used the BRST transformation of the gauge field.
4. We have rewritten $\partial^\mu(\mathrm{D}^{BC}_\mu c^C)$ as a derivative of the action since this is exactly the term that appears with c^C.
5. We used the equation of motion again, this time the contact term is relevant. The minus sign comes from $\overline{c^A}$ being fermionic.

The Fourier transformation of the Green function gives

$$(-ik^\nu)(+ik^\mu)G^{AB}_{\mu\nu}(k) = -iG(k^2)\delta^A B = -i\xi\delta^{AB}.$$

Hence, $B(k^2) = \xi$ holds exactly. If one writes $G^{AB}_{\mu\nu}$ in terms of one-particle irreducible self-energy,



with

$$\Pi_{\mu\nu}(k) = (k^2 g_{\mu\nu} - k_\mu k_\nu)\Pi(k^2) + k_\mu k_\nu \Pi_2(k^2)$$

then the above result is equivalent to $A(k^2) = (1 - \Pi(k^2))^{-1}$ and $\Pi_2(k^2) = 0$. The self-energy is transverse: $k^\mu\Pi_{\mu\nu} = 0$. We used this in QED (p.??) to express the on-shell field renormalization constant $Z_3 = (1 - \Pi_{\mathrm{bare}}(k^2 = 0))^{-1}$.

1.3.2 Renormalization – general considerations

For definiteness, we assume that the matter fields consist of Dirac fermions in the fundamental representation

$$\tilde{\mathcal{L}} = -\frac{1}{4}G^A_{\mu\nu}G^{A,\mu\nu} + \overline{\Psi}(i\not{D} - m)\Psi + \mathcal{L}_{\mathrm{GF}} + \mathcal{L}_{\mathrm{FP}}.$$

We use the version where the B^A field is integrated out (p. 18). For any vertex define (p.??)

$$\Delta_i = d - a_i - \sum_f n_{if}(d/2 - 1 + s_f)$$

where d is the number of space-time dimensions, a_i is the number of derivatives, n_{fi} is the number of fields f at vertex i and $s_f = 0$ for bosons and $s_f = 1/2$ for fermions. One can easily check that all vertices in $\tilde{\mathcal{L}}$ have $\Delta_i = 0$, i.e. they are renormalizable. The *superficial degree of divergence* of Green functions is then

$$D(\gamma) = 4 - \frac{3}{2}E_\Psi - E_A - E_{c,\bar{c}}$$

where each E_i denotes the number of external fields of type i. Hence, the 1PI Green functions given in table 1.1 can be divergent.

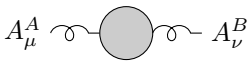
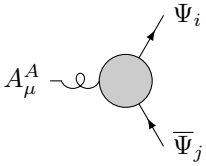
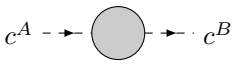
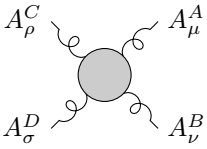
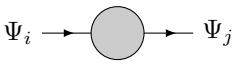
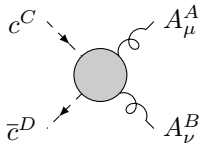
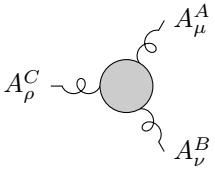
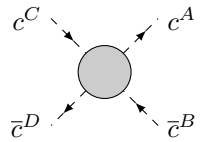
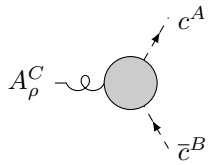
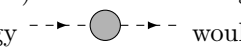
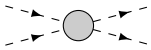
Diagram	Superf. div. D	Actual div.	Diagram	Superf. div. D	Actual div.
	2	0		0	0
	2	0		0	0
	1	0		0	finite
	1	0		0	finite
	1	0	anything else	finite	finite

Table 1.1: Divergent diagrams in the Yang-Mills theory.

Since all Δ_i are renormalizable, the theory can be made finite by adding counter-terms for the above vertices. However, the real point of the proof of renormalizability is to show that the

result is compatible with gauge invariance and can be obtained by rescaling the *single* coupling g in $\tilde{\mathcal{L}}$ and the fields. The proof uses the Ward identities and the BRST symmetry of $\tilde{\mathcal{L}}$. It is technical and will not be exposed here, except for a few remarks.

- In the first place, one needs a regulator that does not break the BRST/gauge symmetry, so that the regularized theory still has the symmetry. Dimensional regularization is the regulator of choice.
- The actual degree of divergence is often less than the superficial one and at most logarithmic. For example,  could be quadratically divergent, but we know that $\Pi_{\mu\nu}(k) \propto k^\mu k^\nu - g^{\mu\nu}k^2$ and not Λ^2 for a momentum cutoff Λ, so $\Pi(k^2)$ can be at most log-divergent. Similarly, a quadratic divergence is the ghost self energy  would imply a ghost mass counter-term, which would violate BRST invariance. However, the ghost self-energy must be proportional to the external momentum, since the ghost vertex is proportional to p^μ. Lorentz invariance then implies p^2, so  is only log-divergent. For the same reason are  and  finite. These vertices would be compatible with BRST symmetry, but they cannot be obtained by rescaling of the parameters and fields of $\tilde{\mathcal{L}}$ in covariant gauge.

Start from $\tilde{\mathcal{L}}$ and introduce a counter-term for every term present in the original Lagrangian (this implies the use of renormalized perturbation theory, p.??). Then

$$
\begin{aligned}
\tilde{\mathcal{L}} = & -\frac{1}{4}(\partial_\mu A^A_\nu - \partial_\nu A^A_\mu)^2 - \frac{1}{2\xi}(\partial_\mu A^{A,\mu})^2 + \overline{\Psi}\partial\!\!\!/\Psi - m\overline{\Psi}\Psi + g\overline{\Psi}\gamma^\mu T^A \Psi A^A_\mu + \\
& - g f^{ABC}(\partial_\mu A^A_\nu) A^{B,\mu} A^{C,\nu} - \frac{1}{4} g^2 f^{ABC} f^{ADE} A^B_\mu A^C_\nu A^{D,\mu} A^{E,\nu} + \\
& + (\partial_\mu \bar{c}^A)(\partial^\mu c^A) - g f^{ABC}(\partial_\mu \bar{c}^A) c^B A^C_\mu \\
& - \frac{\delta_3}{4}(\partial_\mu A^A_\nu - \partial_\nu A^A_\mu)^2 - \frac{\delta_\xi}{2\xi}(\partial_\mu A^{A,\mu})^2 + \delta_2 \overline{\Psi}\partial\!\!\!/\Psi - \delta_m m\overline{\Psi}\Psi + \delta_1 g\overline{\Psi}\gamma^\mu T^A \Psi A^A_\mu + \\
& - \delta_1^{3g} g f^{ABC}(\partial_\mu A^A_\nu) A^{B,\mu} A^{C,\nu} - \frac{\delta_1^{4g}}{4} g^2 f^{ABC} f^{ADE} A^B_\mu A^C_\nu A^{D,\mu} A^{E,\nu} + \\
& + \delta_3^c (\partial_\mu \bar{c}^A)(\partial^\mu c^A) - \delta_1^c g f^{ABC}(\partial_\mu \bar{c}^A) c^B A^C_\mu ,
\end{aligned}
$$

where all quantities (fields, m, g) are renormalized ones. The terms marked with blue are part of the renormalized Lagrangian $\mathcal{L}_\mathrm{r}$ and the terms in red are in the counter-term Lagrangian $\mathcal{L}_\mathrm{ct}$. If this Lagrangian is to follow from a rescaling of fields and parameters (which can be shown) of the original Lagrangian in terms of bare quantities, then the counter-terms must be related in a certain way.

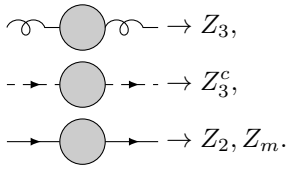
$$
\begin{aligned}
\Psi_0 &= \sqrt{Z_2}\Psi, & m_0 &= Z_m m, \\
A_0^{A,\mu} &= \sqrt{Z_3} A^{A,\mu}, & \xi_0 &= Z_\xi \xi, \\
c_0^A &= \sqrt{Z_3^c} c^A, & g_0 &= Z_g \tilde{\mu}^\varepsilon g.
\end{aligned}
$$

Due to Ward identity (p.??) it holds $Z_\xi = Z_3$. Also keep in mind that in dimensional regularization (with dimreg scale $\tilde{\mu}$), the bare coupling receives small dimension $[g_0] = \varepsilon$ (while g is still dimensionless). The relations between renormalization factors and prefactors in the counter-term

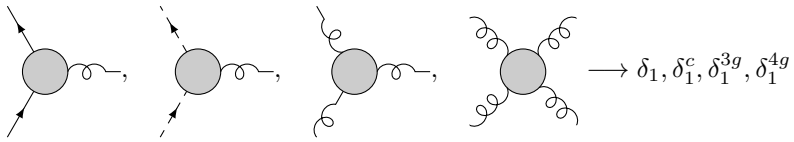
Lagrangian are as follows

$$\delta_2 = Z_2 - 1, \qquad \delta_3 = Z_3 - 1, \qquad \delta_3^c = Z_3^c - 1,$$
$$\delta_m = Z_2 Z_m - 1, \qquad \delta_\xi = 0, \qquad \delta_1 = Z_g Z_2 Z_3^{1/2} - 1,$$
$$\delta_1^{3g} = Z_g Z_3^{3/2} - 1, \qquad \delta_1^{4g} = Z_g^2 Z_3^2 - 1, \qquad \delta_1^c = Z_g Z_3^c Z_3^{1/2} - 1.$$

Most renormalization factors can be obtained from self-energies of the fields:



After the corresponding counter-terms are determined by suitable renormalization conditions, *all* vertex counter-terms



must be related to a *single* renormalization factor Z_g because there is only one coupling constant. This is precisely guaranteed by the Ward identity for the gauge symmetry and the Slavnov–Taylor identity.

Note. In the non-Abelian case, there is no Ward identity that relates Z_g and Z_3, since $A^{A,\mu}$ cot only couples to $\overline{\Psi}\gamma_\mu\Psi$, but also to itself. The derivation for QED (p.??) does not work, because the Noether current is no longer simply $\overline{\Psi}\gamma_\mu\Psi$.

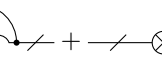
1.3.3 Renormalization at the one-loop order

In this subsection we will compute the renormalization constants Z_2, Z_m, Z_3, Z_g. The result is directly applicable to quantum chromodynamics (QCD), the theory of quarks Ψ_a and gluons A_μ^A. The choice of the renormalization scheme is important. Z-factors are not unique, they are determined by renormalization conditions, such that 1PI Green functions are finite. Here, we use dimensional regularization (p.??) and the $\overline{\text{MS}}$scheme (p.??). This scheme is almost universally used for QCD (while for QED and electroweak interaction, both $\overline{\text{MS}}$and on-shell scheme are in use). Taylor expansion of factors Z_i in terms of scale-dependent renormalized coupling in the $\overline{\text{MS}}$scheme $\alpha_\text{s}(\mu) = g(\mu)^2/(4\pi)$ is generically given by

$$Z_i = 1 - \frac{\alpha_\text{s}(\mu)}{4\pi\varepsilon} c_i + \mathcal{O}(\alpha_\text{s}^2).$$

The numerical factor c_i is different for each renormalization constant.

A) First, we compute Z_2 and Z_m. They are obtained from the fermion self-energy:

$$-i\Sigma =$$  $$+$$  $$+ \mathcal{O}(g^4)$$

Slashed propagator means that it is amputated and $\otimes$ denotes the counterterm vertex from $\mathcal{L}_\text{ct}$. We will work in renormalized perturbation theory. We need only the $1/\varepsilon$ so we will

introduce a simple method to determine the pole part. The symbolic expression for the above diagrams is

$$-i\Sigma = (ig)^2\tilde{\mu}^{2\varepsilon}\int\frac{\mathrm{d}^d k}{(2\pi)^d}\frac{\gamma^\mu T^A i(\not p+\not k+m)\gamma^\nu T^B}{(p+k)^2-m^2}\frac{(-i)g_{\mu\nu}\delta^{AB}}{k^2}+i\delta_{ab}(\delta_2\not p-\delta_m m)+\mathcal{O}(g^4).$$

Since Σ is gauge invariant, we can use any gauge – we choose *Feynman gauge* ($\xi=1$). We work in $d=4-2\varepsilon$ dimensions and in $\overline{\text{MS}}$scheme it holds

$$\tilde{\mu}=\mu\sqrt{\frac{e^{\gamma_\text{E}}}{4\pi}}.$$

The $1/\varepsilon$ pole is a UV divergence, which arises from the region $k\gg p,m$. Thus, we expand the integrand in terms of k and only keep the highest powers:

$$\frac{1}{k^4}\gamma^\mu(\not k 2(p\cdot k)\not k/k^2+\not p+m+\dots)\gamma_\mu$$

All omitted terms are no UV longer divergent. By Lorentz invariance we can replace $k^\mu k^\nu\to k^2 g^{\mu\nu}/d$ for any integrand of form $f(k^2)k^\mu k^\nu$. If the integrand is proportional to $f(k^2)k^\mu$, then the integral is zero (also due to Lorentz invariance). Thus, one obtains

$$-i\Sigma=-g^2T^A_{ac}T^A_{cb}\left((2-d)\frac{-2}{d}g^{\mu\nu}p_\mu\gamma_\nu+(2-d)\not p+dm\right)\tilde{\mu}^{2\varepsilon}\int\frac{\mathrm{d}^d}{(2\pi)^d}\frac{1}{k^4}+i\delta_{ab}(\delta_2\not p-\delta_m m).$$

The product of representation matrices is $(T^AT^A)_{ab}=(N^2-1)/(2N)\delta_{ab}=C_\text{F}\delta_{ab}$, where C_F is the Casimir of the fundamental representation. Now, the remaining integral is infrared divergent for $k\to 0$, but this is an artifact of the expansion for large k. The UV divergence is logarithmic and therefore remains the same when we replace $k^2\to k^2-m^2$. The result for the integral is $A(2,\Delta)$ (p.??), hence we can use

$$\tilde{\mu}^{2\varepsilon}\int\frac{\mathrm{d}^d k}{(2\pi)^d}\frac{1}{k^4}\to\frac{i}{(4\pi)^2}\frac{1}{\varepsilon}$$

for the purpose of extracting the $1/\varepsilon$ UV pole. The result reads

$$-i\Sigma=-i\delta_{ab}\left(\frac{g^2C_\text{F}}{(4\pi)^2}\frac{1}{\varepsilon}(-\not p+4m)-\delta_2\not p+\delta_m m\right).$$

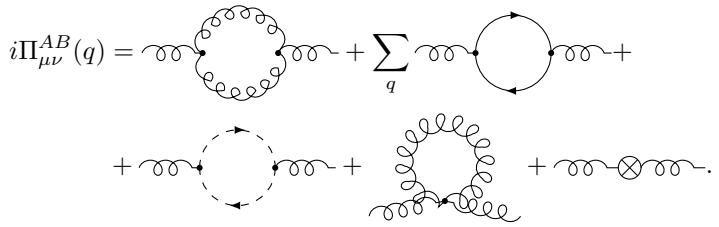
From $\delta_2=Z_2-1$ and $\delta_m=Z_mZ_2-1$, we get

$$Z_2=1-\frac{\alpha_\text{s}C_\text{F}}{4\pi}\frac{1}{\varepsilon},\quad Z_m=1-\frac{\alpha_\text{S}C_\text{F}}{4\pi}\frac{3}{\varepsilon}.$$

These are Feynman gauge quark field and mass renormalizations. In QED the result is the same with $C_\text{F}\to 1$ and $\alpha_\text{s}\to\alpha_\text{em}$.

Note. The above form holds when one uses $\tilde{\mu}$. Another convention in the literature defines $g_0=Z_g\mu^\varepsilon g$. Then the $\overline{\text{MS}}$scheme is defined by the above expression with replacement $1/\varepsilon\to 1/\varepsilon-\gamma_\text{E}+\ln(4\pi)$. The rule is that $\overline{\text{MS}}$renormalized Green functions should not contain γ or $\ln(4\pi)$ terms.

B) To calculate Z_3 we need to determine the gluon self-energy. At one-loop order, there is the bubble for gluon, ghost and each fermion, the tadpole (snail) diagram and the counter-term. Let the gluon momentum be q and external indices ν, B and μ, A:



All diagrams are amputated. The gluon bubble and snail have symmetry factor $1/2$, the ghost and fermion loops have -1 because of anti-commuting fields. The group theory factors are $f^{ACD}f^{BCD} = N\delta^{AB}$ for the gluon bubble and $\mathrm{tr}(T^AT^B) = \delta^{AB}/2$ for fermion bubble. Assume there are n_f fermion fields and use Feynman gauge. The counter-term has the form $-i\delta_3(g_{\mu\nu}q^2 - q_\mu q_\nu)\delta^{AB}$. Here, the expansion model pays off, because it renders the computation algebraic. An elementary calculation gives

$$\begin{aligned} i\Pi^{AB}_{\mu\nu}(q) &= -g^2\delta^{ab}\int \frac{\mathrm{d}^dk}{(2\pi)^d}\frac{1}{k^4}\left[\left(\frac{(d-2)^2}{d}N + \frac{2(d-2)}{d}n_\mathrm{f}\right)k^2 g_{\mu\nu} + \right. \\ &\qquad \left. + \left(-\frac{5}{3}N + \frac{2}{3}n_\mathrm{f}\right)(q^2 g_{\mu\nu} - q_\mu q_\nu) + \text{finite}\right] - i\delta_3(g_{\mu\nu}q^2 - q_\mu q_\nu)\delta^{AB} \\ &= -i\delta^{AB}(g_{\mu\nu}q^2 - q_\mu q_\nu)\left(\frac{g^2}{(4\pi)^2\varepsilon}\left[-\frac{5}{3}N + \frac{2}{3}n_\mathrm{f}\right] + \delta_3\right) + \text{finite}. \end{aligned}$$

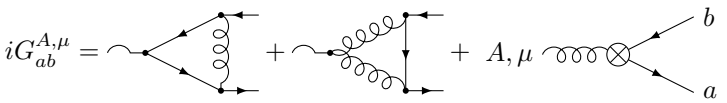
The longitudinal part in the integral expression looks like a quadratic divergence, which is not transverse and cannot be removed by a counter-term. In dimreg, a quadratic divergence would give a pole in *two* dimensions, since

$$\int \frac{\mathrm{d}^d}{(2\pi)^d}\frac{1}{k^2} \propto \frac{1}{2-d}.$$

This pole here is removed by prefactor which is proportional to $(2-d)$. Hence, the gauge boson two-point function is only logarithmically divergent. The renormalization factor in Feynman gauge is then

$$Z_3 = 1 - \frac{\alpha_\mathrm{s}}{4\pi\varepsilon}\left(-\frac{5}{3}N + \frac{2}{3}n_\mathrm{f}\right).$$

C) To calculate Z_g we can choose any of the above listed three- or four-point functions that give logarithmic divergence. We will choose gauge-boson–fermion vertex, but any other would also be possible. Let the gluon have indices A, μ and fermions a and b:



All diagrams are amputated. The first diagram is as in QED, except for the SU(N) group factor

$$T^BT^AT^B = T^BT^BT^A + T^B[T^A, T^B] = C_\mathrm{F}T^A + if^{ABC}T^BT^C = -\frac{1}{2N}T^A.$$

The second diagram has colour factor

$$(-i)f^{ABC}T^BT^C = -\frac{i}{2}f^{ABC}[T^B,T^C] = \frac{1}{2}f^{ABC}f^{BCD}T^D = \frac{N}{2}T^A.$$

The counter-term is $i\delta_1 g\gamma^\mu T^A_{ab}$. The expansion for large k is simple here, because the divergence is only logarithmic by power counting, so we can drop all masses and external momenta in the integrand. The result is

$$iG^{A,\mu}_{ab} = igT^A\gamma^\mu\left(\frac{\alpha_\mathrm{s}}{4\pi\varepsilon}(C_\mathrm{F}+N)+\delta_1\right).$$

Using $\delta_1 = Z_gZ_2Z_3^{1/2}-1$ one finds

$$Z_g = 1-\frac{\alpha_\mathrm{s}}{4\pi\varepsilon}\left(\frac{11}{6}N-\frac{n_\mathrm{f}}{6}\right).$$

This is in fact independent of the gauge parameter ξ.

Scale dependence of the non-Abelian gauge coupling

Recall how the scale-dependence of $\overline{\mathrm{MS}}$couplings is obtained (p.??):

$$\mu^2\frac{\mathrm{d}\alpha_\mathrm{s}(\mu)}{\mathrm{d}\mu^2} = \mu^2\frac{\mathrm{d}}{\mathrm{d}\mu^2}\left(\tilde{\mu}^{-2\varepsilon}Z_g^{-2}\alpha_0\right) = -\varepsilon\alpha_\mathrm{s}(\mu)-\frac{1}{Z_g^2}\mu^2\frac{\mathrm{d}Z_g^2}{\mathrm{d}\mu^2}\alpha_\mathrm{s}(\mu) = \beta(\alpha_\mathrm{s}).$$

We used α_0 for the bare coupling. It holds

$$\mu^2\frac{\mathrm{d}Z_i}{\mathrm{d}\mu^2} = -\frac{c_i}{4\pi\varepsilon}\mu^2\frac{\mathrm{d}\alpha_\mathrm{s}(\mu)}{\mathrm{d}\mu^2} = \frac{\alpha_\mathrm{s}(\mu)}{4\pi}c_i+\mathcal{O}(\alpha_\mathrm{s}^2),$$

which in our case becomes

$$\beta(\alpha_\mathrm{s}) = -\varepsilon\alpha_\mathrm{s}+\beta_0\alpha_\mathrm{s}^2+\mathcal{O}(\alpha_\mathrm{s}^3),\quad \beta_0 = -\frac{1}{4\pi}\left(\frac{11}{3}N-\frac{2}{3}n_\mathrm{f}\right).$$

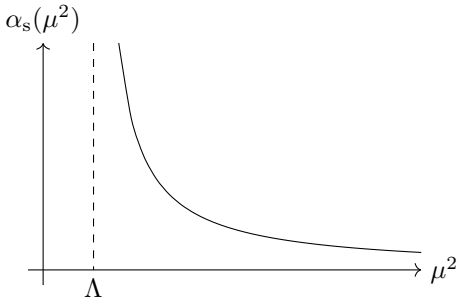
This holds for SU(N) gauge theory with n_f Dirac fermions. The β-function for QED is $\beta_0^{\mathrm{QED}} = -(4n_\mathrm{f}/3)/(4\pi)$, which is a limit for n_f electron species and μ much bigger than masses of these electrons. The coefficient in front of n_f is twice as large as in non-Abelian theories, because there the trace of two generators is normalized by 1/2 ($\mathrm{tr}(T^AT^B) = \delta^{AB}/2$). In QED, $\beta_0^{\mathrm{QED}} > 0$ and the QED coupling *increases* with μ ($\mu = 0$ is an IR fixed point, see p.??). In non-Abelian gauge theory, the gauge boson self energy interaction causes the negative $11/N/3$ term, i.e. the self interaction anti-screens the charge. When n_f is not too large, which is the case relevant to Nature, then $\beta_0 < 0$. Hence, the non-Abelian gauge coupling decreases at higher energies (trivial UV fixed point, p.??):

$$\mu^2\frac{\mathrm{d}\alpha_\mathrm{s}}{\mathrm{d}\mu^2} = \beta_0\alpha_\mathrm{s}^2 \quad\Longrightarrow\quad \alpha_\mathrm{s}(\mu) = \frac{\alpha_\mathrm{s}(\mu_0)}{1+(-\beta_0)\alpha_\mathrm{s}(\mu_0)\ln(\mu^2/\mu_0^2)}.$$

For $\alpha_\mathrm{s}\to 0$ the energy scale grows to infinity, which is called *asymptotic freedom*. At

$$\Lambda = \mu_0 e^{1/(2\beta_0\alpha_\mathrm{s}(\mu_0))}$$

the coupling grows to infinity. When the coupling is strong, bound states can form and perturbation theory breaks down.



Dimensional transmutation

Note that we can also write

$$\alpha_s(\mu) = \frac{1}{-\beta_0 \ln(\mu^2/\Lambda^2)}.$$

Suppose now the Lagrangian contains two massless fermions, up and down,

$$\mathcal{L} = -\frac{1}{4} G^A_{\mu\nu} G^{A,\mu\nu} + \sum_{u,d} \overline{\Psi}_f i \not{D} \Psi_f.$$

This theory appears to have one dimensionless parameter g. Thus, the theory is naively scale-invariant. The above is a good approximation for a theory of nuclei. But in a scale invariant theory there could be no massive states like protons, pions or nuclei. Regularization breaks the scale invariance, and it is not recovered in the renormalized theory. It is *anomalous*. We see that dimensionless $\alpha_s(\mu)$ can be traded for the dimensionful parameter Λ. This is called *dimensional transmutation*. In fact, if $\mathcal{L}$ described the entire world, it is a theory without any parameter, since we need one parameter to fix units. That is, we can arbitrarily fix units such that, say, $\Lambda = 1\,\text{GeV}$ and measure any other dimensionful number in units of Λ. In this sense, QCD with massless quarks is perhaps the most perfect, non-trivial theory relevant to Nature. But there is more in Nature.

Scale-dependent fermion masses

Because $Z_m \neq 1$, $m(\mu)$ in the $\overline{\text{MS}}$scheme is scale dependent with *anomalous dimension*

$$\gamma_m(\alpha_s) = -\frac{1}{Z_m} \mu^2 \frac{d}{d\mu^2} Z_m = -\frac{3 C_F \alpha_s}{4\pi} + \mathcal{O}(\alpha_s^2).$$

Then solving $\mu^2 dm/d\mu^2 = \gamma_m m$ gives (p.??)

$$m(\mu) = m(\mu_0) \exp\left(\int_{\alpha_s(\mu_0)}^{\alpha_s(\mu)} d\alpha_s \, \frac{\gamma_m(\alpha_s)}{\beta(\alpha_s)} \right) = m(\mu_0) \exp\left(\int_{\alpha_s(\mu_0)}^{\alpha_s(\mu)} d\alpha_s \, \frac{\gamma_0 \alpha_s}{\beta_0 \alpha_s^2} + \dots \right) =$$

$$= m(\mu_0) \exp\left(\frac{\gamma_0}{\beta_0} \ln\left(\frac{\alpha_s(\mu)}{\alpha_s(\mu_0)} \right) + \dots \right) = m(\mu_0) \left(\frac{\alpha_s(\mu)}{\alpha_s(\mu_0)} \right)^{\gamma_0/\beta_0} (1 + \mathcal{O}(\alpha_s)).$$

Since $\gamma_0/\beta_0 > 0$, fermion masses decrease as μ increases in non-Abelian theories. This effect has been observed for the bottom quark:

$$m_b(m_b) \approx 4.2\,\text{GeV} \to m_b(M_Z) \approx 2.8\,\text{GeV}.$$

1.4 The physical Hilbert space; unitarity and gauge invariance of the S-matrix

We started with the observation that the gauge field A^A_μ has only two "physical components" (for every A), corresponding to the helicity ± 1 massless particles. But after quantization we ended up with a path integral with six fields (for every A), A^A_μ, c^A and $\overline{c}^A$. The generating functional allows us to compute Green functions of these fields, and they were used to renormalize the theory. However, the scattering amplitudes obtained by LSZ are not all physical. They are also not gauge invariant. The gauge invariance held only for gauge invariant $\mathcal{O}_a$, not the gauge-dependent elementary fields. In this section we discuss how to identify the physical states and scattering amplitudes in the larger Hilbert space of unphysical and physical gauge bosons, and of ghost states.

1.4.1 Quantum electrodynamics in covariant gauge

We first give an elementary discussion for the Abelian U(1), which exposes some of the relevant issues. Take the covariant gauge $\xi = 1$ (Feynman gauge). The gauge field now only has one component of the four space-time dimensions, $A^A_\mu \to A_\mu$, there are no ghosts and $f^{ABC} = 0$:

$$\mathcal{L} = -\frac{1}{2} g_{\mu\nu} (\partial_\rho A^\mu)(\partial^\rho A^\nu) + \mathcal{L}_{\rm M}(\Psi, \mathrm{D}_\mu \Psi).$$

The kinetic term looks like four independent, real, scalar fields, except that the kinetic term of A^0 has the wrong sign. The canonical quantization means a free field of the form

$$A_\mu(\mathbf{x}, t) = \sum_{\lambda=0}^{3} \int \frac{\mathrm{d}^3\mathbf{k}}{(2\pi)^3} \frac{1}{2k_0} \left(e^{-ikx} \varepsilon^\mu(k,\lambda) a(k,\lambda) + \text{h.c.} \right)$$

with canonical commutation relations of $A^\mu, \dot{A}^\mu$. There are two transverse photon polarization states ($\lambda = 1, 2$), and a longitudinal ($\lambda = 3$) and scalar ($\lambda = 0$) polarizations, which are unphysical. It holds

$$\langle k'\lambda' | k\lambda \rangle = 2k^0 (2\pi)^3 \delta^{(3)}(k - k')(-g_{\lambda'\lambda}).$$

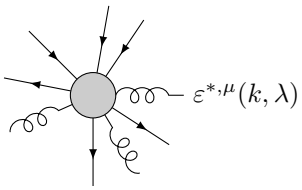
The metric at the end of the expression comes from $\varepsilon^{*,\mu}(k,\lambda')\varepsilon_\mu(k,\lambda) = +g_{\lambda'\lambda}$. The scalar mode has states with negative norm. Compare this to Coulomb gauge or axial gauge where only the transversely polarized states exist. In covariant gauge, canonical quantization leads to an indefinite Hilbert space. To rescue the probabilistic interpretation of quantum theory one needs to single out a *physical Hilbert space*, in which all states except for the unique null vector have positive norm:

phyiscal Hilbert space = Fock space constructed from the transverse photon states.

For this to be consistent, it must not be possible to produce the unphysical polarizations in a scattering process. For example, in $e^+e^- \to \gamma\gamma$, the photons must be transverse. Then the S-operator restricted to the physical subspace is unitary. (It is unitary on the full space already by definition.) Denote by $|i, \mathrm{tr}\rangle$ the transverse part of the state i and by $[S^\dagger S]_{\rm res}$ the matrix element restricted to the physical Hilbert space:

$$\begin{aligned} \langle f, \mathrm{tr} | [S^\dagger S]_{\rm res} | i, \mathrm{tr} \rangle &= \sum_{n,\mathrm{tr}} \langle f, \mathrm{tr} || S^\dagger | n, \mathrm{tr} \rangle \langle n, \mathrm{tr} | S | i, \mathrm{tr} \rangle = \sum_{\text{all } n} \langle t, \mathrm{tr} | S^\dagger | n \rangle \langle n | S | i, \mathrm{tr} \rangle \\ &= \langle f, \mathrm{tr} | S^\dagger S | i, \mathrm{tr} \rangle = \langle f, \mathrm{tr} | i, \mathrm{tr} \rangle . \end{aligned}$$

In short, $[S^\dagger S]_{\text{res}} = \mathbb{1}_{\text{res}}$. Consider the following scattering process.



This contains $\varepsilon^{*,\mu}(k,\lambda)M_\mu(k)$ where M_μ is defined as everything else. Upon squaring one gets

$$\sum_{\lambda=0}^{3}(-1)^{\delta_{\lambda 0}}\varepsilon^{*,\mu'}(k,\lambda)\varepsilon^{\mu}(k,\lambda) = -g^{\mu'\mu}.$$

Keep in mind that $\lambda = 0$ gets negative weight due to negative norm. This is the polarization sum including unphysical states. For the transverse polarizations only it holds

$$\sum_{\lambda=1,2}\varepsilon^*_{\mu'}(k,\lambda)\varepsilon_\mu(k,\lambda) = -g_{\mu'\mu} + \frac{(n\cdot k)(k_\mu n_{\mu'} + k_{\mu'}n_\mu) - k_{\mu'}k_\mu}{|\mathbf{k}|^2} = -g_{\mu'\mu} + \frac{k_\mu \overline{k}_{\mu'} + k_{\mu'}\overline{k}_\mu}{k\cdot\overline{k}},$$

where $n^\mu = (1,\mathbf{0})$ and $\overline{k}^\mu = (k^0, -k^i)$. The unphysical polarizations do not contribute (that is, $\lambda = 0$ and $\lambda = 3$ cancel each other), if

$$k^\mu M_\mu(k) = 0. \tag{1.11}$$

Exactly the same condition guarantees that the scattering matrix is gauge invariant. In fact, we motivated the introduction of gauge symmetry for massless vector fields by the requirement that there is an invariance under $\varepsilon^\mu(k) \to \varepsilon^\mu(k) + \alpha(k)k^\mu$ (p.??) in order to allow the massless limit in the polarization sum of a massive particle. The formal proof that 1.11 holds for any scattering process can be found, for instance, in Peskin and Schröder (Ch. 4.7) using combinatorial methods. Or one derives the Ward identity for the generating functional of Green functions (connected, $W[J]$). The proof does not work for the non-Abelian theory. The process $f + \overline{f} \to g + g$ *does* produce unphysical gauge boson polarizations. To completely cancel their contributions one has to add the production of further unphysical states, the ghosts.

1.4.2 Canonical quantization of non-Abelian gauge theory in covariant gauge

For this section, we choose $\xi = 1$ (Feynman gauge) which makes the gauge boson propagator simple

$$-\frac{ig^{\mu\nu}\delta^{AB}}{k^2 + i\varepsilon}.$$

We are working with Lagrangian

$$\tilde{\mathcal{L}} = -\frac{1}{4}G^A_{\mu\nu}G^{A,\mu\nu} + \mathcal{L}_{\text{M}}(\Psi, \mathrm{D}_\mu\Psi) - (\partial^\mu B^A)A^A_\mu + \frac{1}{2}B^A B^A + (\partial_\mu\overline{c}^A)(\mathrm{D}_\mu c)^A.$$

The canonical momenta can be readily obtained as

$$\Pi^{A,\mu} = G^{A,\mu 0}, \quad \Pi_\Psi = \frac{\partial\mathcal{L}_{\text{M}}}{\partial(\mathrm{D}_0\Psi)}, \quad \Pi_B = -A^A_0, \quad \Pi^A_c = \partial^0\overline{c}^A, \quad \Pi^A_{\overline{c}} = -(\mathrm{D}_0 c)^A.$$

Note that $\Pi^{A,0} = 0$ as usual and $A^{A,0}$ is not a canonical variable, but is the field canonically conjugate to B^A. From the canonical momenta, the canonical (anti-)commutation relations follow immediately

$$\begin{aligned}
[A_i^A(t,\mathbf{x}), G^{B,j0}(t,\mathbf{y})] &= i\delta^{AB}\delta^{(3)}(\mathbf{x}-\mathbf{y}),\\
[B^A(t,\mathbf{x}), -A^{B,0}(t,\mathbf{y})] &= i\delta^{AB}\delta^{(3)}(\mathbf{x}-\mathbf{y}),\\
\{c^A(t,\mathbf{x}), \partial^0\overline{c}^B(t,\mathbf{y})\} &= i\delta^{AB}\delta^{(3)}(\mathbf{x}-\mathbf{y}),\\
\{\overline{c}^A(t,\mathbf{x}), -(\mathrm{D}^0 c)^B(z,\mathbf{y})\} &= i\delta^{AB}\delta^{(3)}(\mathbf{x}-\mathbf{y}).
\end{aligned}$$

Assume first that the system can be treated perturbatively. We want to show that the Fock space of the free theory (and hence, presumably, the theory for small g) contains negative-norm states. For $g = 0$, the equations of motion are (in Feynman gauge)

$$\partial^2 A_\mu^A = \partial^2 c^A = \partial^2 \overline{c}^A = 0, \quad B^A = -\partial^\mu A_\mu^A. \tag{1.12}$$

Hence,

$$\begin{aligned}
[A_i^A(t,\mathbf{x}), \dot{A}_j^B(t,\mathbf{y})] &= i\delta^{AB}\delta^{(3)}(\mathbf{x}-\mathbf{y}),\\
[\partial^\mu A_\mu^A(t,\mathbf{x}), A^{B,0}(t,\mathbf{y})] &= i\delta^{AB}\delta^{(3)}(\mathbf{x}-\mathbf{y}),\\
[A_\mu^A(t,\mathbf{x}), \dot{A}_\nu^B(t,\mathbf{y})] &= -ig_{\mu\nu}\delta^{AB}\delta^{(3)}(\mathbf{x}-\mathbf{y}),\\
\{c^A(t,\mathbf{x}), \dot{\overline{c}}^B(t,\mathbf{y})\} = -\{\overline{c}^A(t,\mathbf{x}), \dot{c}^B(t,\mathbf{y})\} &= i\delta^{AB}\delta^{(3)}(\mathbf{x}-\mathbf{y}).
\end{aligned}$$

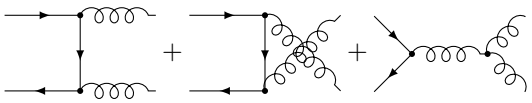
The free fields can be expressed as

$$\begin{aligned}
A_\mu^A(x) &= \sum_{\lambda=0}^{3}\int \frac{\mathrm{d}^3\mathbf{k}}{(2\pi)^3}\frac{1}{2k_0}\left(e^{-ikx}\varepsilon_\mu^A(k,\lambda)a(k,\lambda,A) + \text{h.c.}\right),\\
c^A(x) &= (-i)\int \frac{\mathrm{d}^3\mathbf{k}}{(2\pi)^3}\frac{1}{2k_0}\left(e^{-ikx}c(k,A) + \text{h.c.}\right),\\
\overline{c}^A(x) &= \int \frac{\mathrm{d}^3\mathbf{k}}{(2\pi)^3}\frac{1}{2k_0}\left(e^{-ikx}\overline{c}(k,A) + \text{h.c.}\right).
\end{aligned}$$

Recall that c^A is antihermitian while $\overline{c}^A$ is hermitian. The creation and annihilation operators obey the following commutation relations

$$\begin{aligned}
[a(k,\lambda,A), a^\dagger(k',\lambda',B)] &= -g_{\lambda\lambda'}\delta^{AB}2k_0(2\pi)^3\delta^{(3)}(\mathbf{k}-\mathbf{k}'),\\
\{c(k,A), c^\dagger(k',B)\} = \{\overline{c}(k,A), \overline{c}(k',B)\} &= i\delta^{AB}2k_0(2\pi)^3\delta^{(3)}(\mathbf{k}-\mathbf{k}').
\end{aligned}$$

Thus, the states $a^\dagger(k,0,A)\,|0\rangle$ and $(c^\dagger(k,A) + i\overline{c}^\dagger(k,A))\,|0\rangle$ and many-particle states containing them have *negative norm*. As expected, a restriction to a positive definite physical Hilbert space is necessary, and we expect it to contain only the transverse polarizations and no ghosts. Suppose we compute $f + \overline{f} \to g + g$:

$T(f\overline{f} \to gg) =$ 

One finds

$$\sum_{\lambda,\lambda'=0}^{3} |T(f\overline{f} \to g(\lambda)g(\lambda'))|^2 = \sum_{\lambda,\lambda'=1}^{2} |T(f\overline{f} \to g(\lambda)g(\lambda'))|^2 +$$
$$+ 2\left(ig\overline{v}(p_2)\gamma^\alpha T^C u(p_1)\frac{-i}{(p_1+p_2)^2}gf^{ABC}k_{1\alpha}\right)\left(-ig\overline{u}(p_1)\gamma^\beta T^D v(p_2)\frac{i}{(p_1+p_2)^2}gf^{ABD}k_{1\beta}\right).$$

The second line shows that unphysical states are produced. But if $f\overline{f} \to \mathrm{gh}\,\overline{\mathrm{gh}}$ is also non-zero and one finds

$$|T(f\overline{f} \to \mathrm{gh}\,\overline{\mathrm{gh}})|^2 = \left|\right| =$$
$$= -\left(ig\overline{v}(p_2)\gamma^\alpha T^C u(p_1)\frac{-i}{(p_1+p_2)^2}gf^{ABC}k_{1\alpha}\right)\left(-ig\overline{u}(p_1)\gamma^\beta T^D v(p_2)\frac{i}{(p_1+p_2)^2}gf^{ABD}k_{1\beta}\right). \tag{1.13}$$

In the phase space integration, the $f\overline{f} \to gg$ is multiplied by $1/2$, because there are two identical particles in the final state, hence

$$\frac{1}{2}|T(f\overline{f} \to gg)|^2_{\text{full}} + |T(f\overline{f} \to \mathrm{gh}\,\overline{\mathrm{gh}})|^2 = \frac{1}{2}|T(f\overline{f} \to gg)|^2_{\text{transverse}}.$$

That is, the hypothesis that the physical Hilbert space should consist only of transverse polarizations while the unphysical gauge boson and ghost states cancel works. But the cancellation is much more complicated than in QED, *because the unphysical states have interactions.*

A comment is necessary on the negative weight, i.e. the factor (-1) in 1.13, of the ghost production process. If the metric of states on the Fock space (single-particle) is $\langle n|n'\rangle = G_{nn'}$ (usually would be $\delta_{nn'}$), then the completeness relation reads

$$\sum_{n,n'} |n\rangle\,(G^{-1})_{nn'}\,\langle n'| = \mathbb{1}.$$

For gauge bosons we had $G_{\lambda\lambda'} = -g_{\lambda\lambda'}$ which yielded

$$\sum_{\lambda=0}^{3}(-1)^{\delta_{\lambda 0}}\varepsilon^{*,\mu}(k,\lambda)\varepsilon^{\nu}(k,\lambda).$$

For ghosts, we have

$$\begin{pmatrix}\langle \mathrm{gh}|\mathrm{gh}\rangle & \langle \mathrm{gh}|\overline{\mathrm{gh}}\rangle \\ \langle \overline{\mathrm{gh}}|\mathrm{gh}\rangle & \langle \overline{\mathrm{gh}}|\overline{\mathrm{gh}}\rangle\end{pmatrix} = \begin{pmatrix}0 & (-i)i \\ (-i)(-i) & 0\end{pmatrix} = \begin{pmatrix}0 & 1 \\ -1 & 0\end{pmatrix}.$$

The first factors $(-i)$ come from the definition of the field c^A (remember that $\overline{c}^A$ has no i factors in front) and the second factors of $\pm i$ come from canonical anti-commutation relations. Hence, $\mathrm{gh}\,\overline{\mathrm{gh}}$ production gets the weight

$$|\mathrm{gh}\rangle\,(G^{-1})_{12}\,\langle\overline{\mathrm{gh}}| \quad\otimes\quad |\overline{\mathrm{gh}}\rangle\,(G^{-1})_{21}\,\langle \mathrm{gh}|\,,$$

i.e. the factor $(-1)\cdot 1 = -1$.

1.4.3 Gauge invariance of the S-matrix and conditions on the physical Hilbert space

We use the requirement that the S-matrix must be gauge-invariant to define the physical Hilbert space. Recall the general gauge fixing Lagrangian (p. 21) $\tilde{L} = \mathcal{L} + sF$ with $s\tilde{L} = 0$ where s is the BRST operator. For physical in/out states we want

$$\delta_F S_{\alpha\beta} = \delta_F(\langle\beta;\text{out}|\alpha;\text{in}\rangle) = 0.$$

Here, δ_F means a variation of the gauge fixing functional F. A change of $F \to F + \delta F$ changes $\tilde{L}$ and therefore also the Hamiltonian. We interpret it as a charge of H_{int}. The in/out states were defined as a solution of the *Lippmann–Schwinger equation* (p.??):

$$\ker \Psi_\alpha^{\text{in/out}} = |\phi_\alpha\rangle + \frac{1}{E_0 - \hat{H}_0 \pm i\varepsilon} \hat{H}_{\text{int}} \, |\Psi_\alpha^{\text{in/out}}\rangle \,.$$

Hence, the variation with respect to F becomes is

$$\delta_F(|\alpha;\text{in/out}\rangle) = \frac{1}{E_\alpha - \hat{H}_0 \pm i\varepsilon} \delta_F H \, |\alpha;\text{in/out}\rangle \,.$$

Multiplying by arbitrary $\langle\beta;\text{in}|$ gives

$$\begin{aligned}
\langle\beta;\text{in}|\, \delta_F(|\alpha;\text{in/out}\rangle) &= \frac{1}{E_\alpha - E_\beta \pm i\varepsilon} \, \langle\beta;\text{in}|\delta_F H|\alpha;\text{in/out}\rangle \\
&= \mp i \int_{t,-\infty}^{\infty,t} \mathrm{d}t \, e^{i(E_\alpha - E_\beta \pm i\varepsilon)t} \, \langle\beta;\text{in}|\delta_F H|\alpha;\text{in/out}\rangle \\
&= \mp i \int_{t,-\infty}^{\infty,t} \mathrm{d}t \, \langle\beta;\text{in}|e^{-iHt}\delta_F H r^{iHt}|\alpha;\text{in/out}\rangle \,.
\end{aligned}$$

Hence,

$$\delta_F(|\alpha,\text{in/out}\rangle) = \mp i \int_{t,-\infty}^{t,\infty} \mathrm{d}t \, e^{-iHt}\delta_F H e^{iHt} \, |\alpha;\text{in/out}\rangle \,.$$

Now we write a long calculation that will show that BRST charge annihilates physical states:

$$\begin{aligned}
0 &= \delta_F(\langle\beta;\text{out}|\alpha;\text{in}\rangle) = (\delta_F \, \langle\beta;\text{out}|) \, |\alpha;\text{in}\rangle + \langle\beta;\text{in}|\, \delta_F(|\alpha;\text{out}\rangle) = \\
&= -i \int_{-\infty}^{t} \mathrm{d}t \, \langle\beta;\text{out}|e^{-iHt}\delta_F H e^{iHt}|\alpha;\text{in}\rangle - i \int_{t}^{\infty} \mathrm{d}t \, \langle\beta;\text{out}|e^{-iHt}\delta_F H e^{iHt}|\alpha;\text{in}\rangle \\
&= -i \int_{-\infty}^{\infty} \mathrm{d}t \, \langle\beta;\text{out}|\delta_F H|\alpha;\text{in}\rangle = i \int \mathrm{d}^4x \, \langle\beta;\text{out}|\delta_F\tilde{\mathcal{L}}|\alpha;\text{in}\rangle = \\
&= i \, \langle\beta;\text{out}|\delta_F S|\alpha;\text{in}\rangle = i \, \langle\beta;\text{out}|sX|\alpha;\text{in}\rangle = \\
&= - \, \langle\beta;\text{out}|[Q_{\text{BRST}}, X]|\alpha;\text{in}\rangle \,.
\end{aligned}$$

In the derivation we used that $|\alpha;\text{in/out}\rangle$ are eigenstates of H with eigenvalue E_α, and we used energy conservation $E_\alpha = E_\beta$. We also introduced the variation of the gauge fixing Lagrangian

$$X = \int \mathrm{d}^4x \, \delta F.$$

In the last step, it was used that $sX = i[Q_{\mathrm{BRST}}, X]$ (p. 22). A sufficient condition for the gauge-invariance of the S-matrix of *physical states* is

$$Q_{\mathrm{BRST}} \,|\alpha; \mathrm{in}\rangle$$

for $|\alpha; \mathrm{in}\rangle$ a physical state. Similarly holds for the out-states, since Q_{BRST} is a conserved charge. Thus, physical states lie in the kernel of the BRST symmetry charge operator ($\ker Q_{\mathrm{BRST}}$). However, $\ker Q_{\mathrm{BRST}}$ is not the physical Hilbert space $\mathcal{H}_{\mathrm{phys}}$. For any $|\Psi\rangle \in \mathcal{H}$, $Q_{\mathrm{BRST}}\,|\Psi\rangle$ is in $\ker Q_{\mathrm{BRST}}$ because $Q^2_{\mathrm{BRST}} = 0$, but is a *zero-norm state*. One can show, however, that $\ker Q_{\mathrm{BRST}}$ does not contain negative-norm states. (We shall do this below for the free theory. The general proof is given in Kugo, "Eichtheorie", Ch. 5.8/5.9).

We can now construct the physical Hilbert space as a quotient space of equivalence classes of states that differ only by zero-norm states. Let $\mathrm{im}\, Q_{\mathrm{BRST}}$ be the set of all zero-norm states in $\ker Q_{\mathrm{BRST}}$. Then we define

$$\mathcal{H}_{\mathrm{phys}} = \frac{\ker Q_{\mathrm{BRST}}}{\mathrm{im}\, Q_{\mathrm{BRST}}}.$$

This is the *cohomology* of the operator Q_{BRST}. The states $|\hat{\Psi}\rangle$ in $\mathcal{H}_{\mathrm{phys}}$ are equivalence classes (for a given $|\Psi\rangle \in \ker Q_{\mathrm{BRST}}$)

$$|\hat{\Psi}\rangle = \{|\Psi'\rangle \in \ker Q_{\mathrm{BRST}};\ |\Psi'\rangle - |\Psi\rangle \in \mathrm{im}\, Q_{\mathrm{BRST}}\}.$$

The scalar product in $\mathcal{H}_{\mathrm{phys}}$ is defined as

$$\langle\hat{\Psi}|\hat{\varphi}\rangle = \langle\Psi|\varphi\rangle$$

where $|\Psi\rangle$ and $|\varphi\rangle$ are any representatives from the respective equivalence classes. For this to be well-defined, the scalar product must be independent of the chosen representative. This follows, since

$$\langle\Psi|f\rangle = 0 \quad \forall\ |\Psi\rangle \in \ker Q_{\mathrm{BRST}}, \forall\ |f\rangle \in \mathrm{im}\, Q_{\mathrm{BRST}}. \tag{1.14}$$

Proof. Since both $|\Psi\rangle$ and $|f\rangle$ are elements $|Q\rangle_{\mathrm{BRST}}$, any linear combination will also be in this kernel. In particular, $|\Psi\rangle + \alpha\,|f\rangle \in \ker Q_{\mathrm{BRST}}$ for all $\alpha \in \mathbb{C}$. The norm of this state is $2\mathrm{Re}\,(\alpha\,\langle f|\Psi\rangle) + \langle\Psi|\Psi\rangle \geq 0$ since $\ker Q_{\mathrm{BRST}}$ does not contain negative norm states. For arbitrary complex α this can hold only if $\langle f|\Psi\rangle = 0$. □

The above defined scalar product on $\mathcal{H}_{\mathrm{phys}}$ is positive definite by construction. The following two observations make the abstract condition $Q_{\mathrm{BRST}}\,|\Psi\rangle > 0$ for physical states plausible.

1) Equation of motion – the Euler–Lagrange equations for A^A_μ for the gauge-fixed Lagrangian $\tilde{L}$ (p. 22 and 32) is

$$\mathrm{D}^{AB}_\mu G^{B,\mu\nu} + g j^{A,\nu}_{\mathrm{M}} = \partial^\nu B^A + g f^{ABC}(\partial^\nu \bar{c}^B) c^C.$$

$\tilde{\mathcal{L}}$ is also invariant under global gauge transformations (with $B^A, c^A, \bar{c}^A$ in the adjoint representation). In terms of the Noether current of this symmetry, the field equation can be written as (the derivation is left as an exercise to the reader)

$$\partial_\mu G^{A,\mu\nu} + g j^{A,\nu} = i\{Q_{\mathrm{BRST}}, (\mathrm{D}^\nu \bar{c})^A\}.$$

The right-hand side vanishes between physical states, hence on $\mathcal{H}_{\mathrm{phys}}$ the field equation retains the form of the Maxwell equation.

2) Asymptotic states: these are free states, so consider the BRST transformation for $g = 0$:

$$\delta_\lambda \Psi = 0, \quad \delta_\lambda c^A = 0, \quad \delta_\lambda A^A_\mu = \lambda \partial_\mu c^A, \quad \delta_\lambda \bar{c}^A = \lambda B^A, \quad \delta_\lambda B^A = 0.$$

The gauge boson polarization vectors have two physical (transverse) polarizations and two unphysical ones

$$\varepsilon^\pm_\mu(k) = \frac{1}{\sqrt{2}}(\varepsilon_\mu(k,0) \mp \varepsilon_\mu(k,3)) = \frac{1}{\sqrt{2}}(1, \mp \mathbf{k}/|\mathbf{k}|).$$

Hence, in momentum space it holds

$$\delta_\lambda(\varepsilon^\mu(k,\lambda) A^A_\mu(k)) = \lambda \varepsilon^\mu(k,\lambda) k_\mu c^A.$$

For $\lambda = \pm$ this expression is not zero and the ε^- gauge polarization and the anti-ghost have $\delta_\lambda \neq 0$ (they are not in $\ker Q_{\mathrm{BRST}}$ and hence not physical). On the other hand $B^A \propto \partial_\mu A^{A,\mu}$ (from the equation of motion, eq. 1.12) is the BRST variation of the antighost, and $\partial_\mu c^A$ of A^A_μ. For ε^+ polarization $\partial_\mu A^{A,\mu} \neq 0$ and c^A is the BRST variation of the ε^- polarization (due to ∂_μ). Ghost states and ε^+ polarization states can be represented as $Q_{\mathrm{BRST}} |\Psi\rangle$ for some $|\Psi\rangle$ (antighost for ε^- and ε^+ for ghost) and are zero-norm states which are eliminated in the quotient space construction.

$\mathcal{H}_{\mathrm{phys}}$ is built from equivalence classes corresponding to the transverse gauge boson states. This is exactly what was expected on physical grounds.

1.4.4 Unitarity of the S-matrix on $\mathcal{H}_{\mathrm{phys}}$

The previous discussion guarantees a consistent quantum theory of the space of physical states. We start with both $\tilde{\mathcal{L}} = \mathcal{L} + sF$ and $\tilde{\mathcal{H}}$ hermitian. The scattering operator is unitary on the Hilbert space including unphysical states. We want to show that S is unitary also on $\mathcal{H}_{\mathrm{phys}} = \ker Q_{\mathrm{BRST}} / \operatorname{im} Q_{\mathrm{BRST}} \subset \mathcal{H}$. Since Q_{BRST} is a conserved charge, it commutes with S, so

$$Q_{\mathrm{BRST}}(S|\Psi\rangle) = S Q_{\mathrm{BRST}} |\Psi\rangle = 0, \quad \text{for } |\Psi\rangle \in \ker Q_{\mathrm{BRST}}.$$

Thus, S maps from $\ker Q_{\mathrm{BRST}}$ to $\ker Q_{\mathrm{BRST}}$, as well as $S^\dagger$. But it also maps $\operatorname{im} Q_{\mathrm{BRST}}$ into $\operatorname{im} Q_{\mathrm{BRST}}$, since for all $|\Psi\rangle \in \ker Q_{\mathrm{BRST}}$ and $|f\rangle \in \operatorname{im} Q_{\mathrm{BRST}}$ it holds

$$\langle\Psi| (S|f\rangle) = \langle\Psi|S|f\rangle = (\langle\Psi| S)|f\rangle = 0.$$

This is because $\langle\Psi| S \in \ker Q_{\mathrm{BRST}}$ and the scalar product between an element of the image and an element in the kernel is zero (eq. 1.14). In other words, $S|f\rangle$ also has zero norm. Hence, we can define the scattering operator $\hat{S}$ on $\mathcal{H}_{\mathrm{phys}}$ as $S_{\mathrm{phys}}|\hat{\Psi}\rangle = \widehat{S|\Psi\rangle}$ since all $S|\Psi\rangle$ lie in the same equivalence class if $|\Psi\rangle$ do, so the definition is unique. Then for all $|\hat{\Psi}\rangle$ and $|\hat{\varphi}\rangle$ in $\mathcal{H}_{\mathrm{phys}}$ it holds

$$\langle\hat{\varphi}|S^\dagger_{\mathrm{phys}} S_{\mathrm{phys}}|\hat{\Psi}\rangle = \langle\widehat{S\varphi}|\widehat{S\Psi}\rangle = \langle S\varphi|S\Psi\rangle = \langle\varphi|S^\dagger S|\Psi\rangle = \langle\varphi|\Psi\rangle = \langle\hat{\varphi}|\hat{\Psi}\rangle.$$

Thus, S_{phys} is unitary on $\mathcal{H}_{\mathrm{phys}}$.

1.4.5 Summary

We start with gauge invariant Lagrangian $\mathcal{L}$. We have two options:

- non-covariant quantization (for instance in axial gauge), no ghosts, only two transverse polarizations, physical Hilbert space from the beginning;

- gauge fixing, Faddeev–Popov method, ghosts, BRST formalism.

In the latter case, one can maintain Lorentz invariance but:

- one obtains ghosts, unphysical gauge boson components, Fock space with indefinite metric;
- definition of physical states via $\ker Q_{\mathrm{BRST}}(\mathrm{im}\, Q_{\mathrm{BRST}}$ (positive definite);
- gauge invariance and unitary S-matrix for the scattering of states from $\mathcal{H}_{\mathrm{phys}}$.

In practice this means the scattering of transverse gauge bosons, since it is simplest to choose those as representatives of the equivalence class.

Chapter 2

Spin-2 fields

We have already discussed spin-0, spin-1/2 and spin-1 fields at length, and now we will skip spin-3/2 (applicable to gravitinos, described by Rarita–Schwinger equation). This chapter will be dedicated to spin-2 fields and the general relativity as an effective quantum field theory of weak gravitational fields. This will conclude the analysis of elementary relativistic quantum fields, as fields with higher spin cannot be consistently defined or lack any applicability.

2.1 The massive spin-2 field

Remember that vector fields were built from the $(1/2, 1/2)$ representation of the Lorentz group. This representation is equivalent to the vector representation A_μ (which is 4-dimensional). This representation can be reducible as a representation of the little group SO(3): $\mathbf{4} = \mathbf{1} \oplus \mathbf{3}$. Similarly, we will get the spin-2 representation by multiplying two vector representations:

$$(1/2, 1/2) \otimes (1/2, 1/2) = \underbrace{(1,0) \oplus (0,1)}_{\text{antisymmetric tensor} F_{\mu\nu}} \oplus \underbrace{(1,1) \oplus (0,0)}_{\text{symmetric tensor} h_{\mu\nu}} .$$

However, this representation is not irreducible as it can be represented by any Lorentz tensor of rank 2 $B_{\mu\nu}$. It can be split into symmetric (10 dimensions) and antisymmetric parts (6 dimensions, a sum of self-dual and anti-self-dual form). The antisymmetric part is associated with spin-1 objects, hence the spin-2 objects have to be represented by symmetric tensor fields. (An alternative would be $(2,0) \oplus (0,2)$ representation – similar to $(1,0) \oplus (0,1)$ for spin-1 particles. The reason for continuing with $h_{\mu\nu}$ are also similar. In the end we want to be able to describe long-range forces generated by the exchange of massless spin-2 particles.) Under the little group SO(3), the symmetric tensor decomposes as

$$\mathbf{10} = \mathbf{1} \oplus \mathbf{1} \oplus \mathbf{3} \oplus \mathbf{5}.$$

The parts are h_{00}, h_{ii}, h_{0i} and $h_{ij} - \delta_{ij} h_{kk}/3$. Since a massive spin-$j$ representation is $2j + 1$ dimensional, we would like h_{ij} to be the canonical coordinates. In the analogous case of the vector field A_μ we ensured that only A_i are degrees of freedom by ensuring that the conjugated momentum Π_0 vanishes identically. This determined the free Lagrangian for the massive vector field uniquely (p.??). We now rederive this result differently and then generalize it to the spin-2 field. With the spin-1 field, we demanded that only the non-trivial part of the little group

representation remains dynamical (the **3** part). We separate the field into the "transverse" and "longitudinal" parts as follows:

$$A_\mu = A^{\mathrm{T}}_\mu + \partial_\mu \pi, \quad \text{with } \partial^\mu A^{\mathrm{T}}_\mu = 0.$$

Then $\partial^\mu A^{\mathrm{T}}_\mu = 0$ provides the one condition that reduces the degrees of freedom in A^{T}_μ from four to three. The most general bilinear terms in the Lagrangian for A_μ are (p.??)

$$\begin{aligned}\mathcal{L}_0 &= -\frac{a}{4}F_{\mu\nu}F^{\mu\nu} + \frac{b}{2}(\partial_\mu A^\nu)(\partial^\mu A_\nu) + \frac{c}{2}(\partial_\mu A^\mu)^2 + \frac{m^2}{2}A_\mu A^\mu \\ &= -\frac{a}{4}F^{\mathrm{T}}_{\mu\nu}F^{\mathrm{T},\mu\nu} + \frac{b}{2}(\partial_\mu A^{\mathrm{T},\nu})(\partial^\mu A^{\mathrm{T}}_\nu) + \frac{1}{2}m^2 A^{\mathrm{T}}_\mu A^{\mathrm{T},\mu} + \\ &\quad + \frac{b}{2}(\partial_\mu\partial_\nu\pi)(\partial^\mu\partial^\nu\pi) + \frac{c}{2}(\partial^2\pi)^2 + \frac{1}{2}m^2(\partial_\mu\pi)(\partial^\mu\pi).\end{aligned}$$

We omitted any terms that vanish due to $\partial^\mu A^{\mathrm{T}}_\mu = 0$ after integration by parts. Note that the first and the second term are not independent. Hence, we can drop one of them and set $b = 0$. The propagator for the longitudinal mode π contains

$$\frac{i}{m^2k^2 - (b+c)k^4} = \frac{1}{m^2}\left(\frac{i}{k^2} - \frac{i}{k^2 - m^2/(b+c)}\right).$$

The negative sign in the second term indicates a negative norm state (ghost). To avoid this, we set $b = -c$. Together with the dependence of the first two terms, we can set $b = c = 0$ and rescale the fields to get $a = 1$ (canonical normalization):

$$\mathcal{L}_0 = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \frac{1}{2}m^2 A_\mu A^\mu.$$

If we calculate the conjugate momentum, we see that its time component vanishes identically.

Now, we will repeat the same procedure for the symmetric tensor $h_{\mu\nu}$. The most general bilinear terms are (using the shorthand $h = h^\mu{}_\mu$)

$$\mathcal{L}_0 = \frac{1}{2}ah^{\mu\nu}\Box h_{\mu\nu} - \frac{1}{2}bh^{\mu\nu}\partial_\mu\partial^\alpha h_{\nu\alpha} + \frac{1}{2}ch\Box h + \frac{1}{2}dh\partial_\mu\partial_\nu h^{\mu\nu} - \frac{1}{2}m^2h^{\mu\nu}h_{\mu\nu} - \frac{1}{2}n^2h^2. \quad (2.1)$$

The terms were already collected together as much as possible using integration by parts. We decompose the field into irreducible representations of the little group:

$$h_{\mu\nu} = h^{\mathrm{T}}_{\mu\nu} + \partial_\mu\pi_\nu + \partial_\nu\pi_\mu = h^{\mathrm{T}}_{\mu\nu} + \partial_\mu\pi^{\mathrm{T}}_\nu + \partial_\nu\pi^{\mathrm{T}}_\mu + 2\partial_\mu\partial_\nu\pi^{\mathrm{L}},$$

with

$$\partial^\mu h^{\mathrm{T}}_{\mu\nu} = 0, \quad \partial^\mu\pi^{\mathrm{T}}_\mu = 0, \quad \pi_\mu = \pi^{\mathrm{T}}_\mu + \partial_\mu\pi^{\mathrm{L}}.$$

After inserting this into the Lagrangian 2.1, we want to avoid terms that have more than two derivatives on π^{L} or π^{L} (that would imply ghosts). The problematic terms are:

$$\mathcal{L}_0 \supset -\frac{1}{2}\pi^{\mathrm{L}}\partial^4\pi^{\mathrm{L}}(4m^2+4n^2) + \frac{1}{2}\pi^{\mathrm{L}}\partial^6\pi^{\mathrm{L}}4(a+b+c+d) + \frac{1}{2}\pi^{\mathrm{T},\mu}\partial^4\pi^{\mathrm{T}}_\mu(-2a-b) + \frac{1}{2}h\partial^4\pi^{\mathrm{L}}(8c+4d).$$

In order to avoid ghosts, we set

$$m^2 + n^2 = 0, \quad b = -2a, \quad c = -a, \quad d = 2a$$

and $a = -1$ for the canonical normalization. The Lagrangian for the free massive spin-2 field becomes

$$\mathcal{L}_0 = \frac{1}{2}\partial^\lambda h_{\mu\nu}\partial_\lambda h^{\mu\nu} - \partial_\mu h^{\mu\lambda}\partial^\nu h_{\nu\lambda} - \frac{1}{2}(\partial^\lambda h)(\partial_\lambda h) + \partial^\mu h \partial^\nu h_{\mu\nu} - \frac{1}{2}m^2(h^{\mu\nu}h_{\mu\nu} - h^2).$$

This is called the *Fierz–Pauli Lagrangian* (free theory of a massive spin-2 particle). The tensor $h^{\mathrm{T}}_{\mu\nu}$ has six components (not 5), however, the equations of motion give the constraint $h = 0$ (requiring the tensor to be transverse and traceless). The Euler–Lagrange equations give

$$\partial^2 h_{\mu\nu} - \partial_\mu\partial^\lambda h_{\lambda\nu} - \partial_\nu\partial^\lambda h_{\lambda\mu} - \eta_{\mu\nu}\partial^2 h + \eta_{\mu\nu}\partial^\lambda\partial^\omega h_{\lambda\omega} + \partial_\mu\partial_\nu h = -m^2(h_{\mu\nu} - \eta_{\mu\nu}h).$$

When we differentiate with ∂^μ, the left-hand side vanishes, and it follows (for $m \neq 0$) $\partial^\mu h_{\mu\nu} = \partial_\nu h$. We insert this into the Euler–Lagrange equation and get

$$\partial^2 h_{\mu\nu} - \partial_\mu\partial_\nu h = -m^2(h_{\mu\nu} - \eta_{\mu\nu}h).$$

The trace of this is zero on the left-hand side, so we get $h = 0$. Moreover, we have $\partial^\mu h_{\mu\nu} = 0$ and $(\partial^2 + m^2)h_{\mu\nu} = 0$ for each component. This means that every component of $h_{\mu\nu}$ satisfies the Klein–Gordon equation and that $h_{\mu\nu}$ carries only 5 dynamical degrees of freedom (the **5** representation of the little group). In Nature, however, there are no known massive spin-2 particles. We will thus move to the massless case.

2.2 The massless spin-2 field

There are several problems that we encounter in the limit $m \to 0$. The issues are exactly the same as those in the massless spin-1 case:

- We cannot construct a Lorentz covariant field theory for helicity ± 2 particles from the symmetric tensor representation. We would need to have a representation (A, B) such that $|A - B| = 2$. The simplest choice is $(2, 0) \oplus (0, 2)$, but this does not generate a $1/r$ static force. This is the same reason why we didn't choose $(1, 0) \oplus (0, 1)$ for the photon.
- The polarization sum
$$\sum_{\lambda=0,\pm1,\pm2} \varepsilon^{\mu\nu}(p,\lambda)\varepsilon^{*\rho\sigma}(p,\lambda)$$
contains singular $1/m^2$ terms (from $\lambda = 0, \pm 2$). If we direct the momentum along the z-axis, the $\lambda = \pm 2$ polarizations are of the form
$$\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & \pm i & 0 \\ 0 & \pm i & -1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$
- We can try to manually remove the unphysical polarizations and construct the free field
$$h_{\mu\nu}(x) = \sum_{\lambda=\pm2} \int \frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3 2E_{\mathbf{p}}} \left(e^{-ipx}\varepsilon_{\mu\nu}(p,\lambda)a(p,\lambda) + \text{h.c.}\right).$$
Under a Lorentz transformation, we get additional terms that do not vanish:
$$\mathcal{U}(\Lambda)h^{\mu\nu}(x)\mathcal{U}^{-1}(\Lambda) = (\Lambda^{-1})^\mu{}_\rho(\Lambda^{-1})^\nu{}_\sigma h^{\rho\sigma}(\Lambda x) + \partial^\mu\xi^\nu(\Lambda x) + \partial^\nu\xi^\mu(\Lambda x),$$

where ξ^μ is some function involving the parameters of the little group transformation. For the non-covariance to be irrelevant for the dynamics, the action must be invariant under the gauge symmetry:

$$h_{\mu\nu}(x) \to h_{\mu\nu}(x) + \partial_\mu \xi_\nu(x) + \partial_\nu \xi_\mu(x).$$

- The Fierz–Pauli Lagrangian has a smooth limit $m \to 0$:

$$\mathcal{L}_0 = \frac{1}{2}\partial^\lambda h_{\mu\nu}\partial_\lambda h^{\mu\nu} - \partial_\mu h^{\mu\lambda}\partial^\nu h_{\nu\lambda} - \frac{1}{2}\partial^\lambda h \partial_\lambda h + \partial^\mu h \partial^\nu h_{\mu\nu}.$$

  It is invariant (up to a total derivative) under the gauge transformation. Note that we explicitly constructed the Lagrangian to be ghost-free in the massive case, so we can choose $\pi^\mu = \xi^\mu$ and the gauge invariance is guaranteed.

Suppose we want to couple $h_{\mu\nu}$ to some "matter current" by adding the term

$$-\int \mathrm{d}^4x \; h^{\mu\nu} T_{\mu\nu}$$

to the action. Then the gauge invariance requires $\partial_\mu T^{\mu\nu} = 0$. This gives four Noether currents and corresponding conserved charges, which need to be a 4-vector

$$P^\nu = \int \mathrm{d}^3\mathbf{x}\, T^{0\nu}(x).$$

We only know one conserved 4-vector in nature: the energy-momentum 4-vector. This implies that $T^{\mu\nu}$ is the energy-momentum tensor and the gauge symmetry is local space-time translations $x^\mu \to x^\mu + \xi^\mu$. This also makes clear why one cannot have massless fields of spin more than 2 with long-range interactions. They would have to couple to a current associated with a tensorial charge, but there is no symmetry that can give such a charge.

Let's count the degrees of freedom. Naively, the symmetric tensor $h_{\mu\nu}$ has 10 components. The gauge symmetry removes 4 components and the equations of motion (the constraints $\partial^\mu h_{\mu\nu} - \partial_\nu h = 0$) remove another 4 components. We are left with 2 dynamical degrees of freedom, which correspond to the helicity ± 2 states. The remaining dynamical states reside in the transverse-traceless part (s^{T}_{ij}) of the tensor h^{T}_{ij}:

$$h_{ij} = \frac{1}{3}\delta_{ij}h_{kk} + s_{ij} = \frac{1}{3}\delta_{ij}h_{kk} + s^{\mathrm{T}}_{ij} + s^{\mathrm{V}}_{ij} + s^{\mathrm{S}}_{ij}.$$

Up to now, we have discussed the free field. The translation gauge symmetry suggests a relation to local general coordinate transformations, which is the symmetry underlying general relativity. In the non-relativistic limit, it holds $T^{00} \propto m$, so we conclude that this force has to couple to mass. Indeed, general relativity is built on a symmetric tensor field $g_{\mu\nu}(x)$. It is usually given a geometric interpretation as the metric of a curved space-time manifold. With this we can construct GR from the bottom up from $h_{\mu\nu}$. Remember first how the non-Abelian gauge transformations were constructed:

$$A^A_\mu \to A^A_\mu + \partial_\mu \varepsilon^A + f^{ABC}\varepsilon^C A^B_\mu.$$

The other index of the gauge field A^A_μ can be replaced by a space-time index for $h_{\mu\nu}$:

$$h_{\mu\nu} \to h_{\mu\nu} + \partial_\mu \varepsilon_\nu + \partial_\nu \varepsilon_\mu + \text{terms like } \varepsilon \times h.$$

However, the extra terms at the end can make the Lagrangian non-Lorentz invariant. To avoid this, we have to add new terms of higher order in $h_{\mu\nu}$ to the Lagrangian. Repeating this procedure infinitely many times, we get the full Einstein–Hilbert action of general relativity:

$$S_{\mathrm{GR}} = \frac{1}{16\pi G_{\mathrm{N}}} \int \mathrm{d}^4x \, \sqrt{-g} R,$$

where $g = \det(g_{\mu\nu})$ and R is the Ricci scalar.

2.3 GR as an effective field theory of gravitation

We have shown that GR can be constructed from the massless spin-2 field $h_{\mu\nu}$. Now, we will treat GR as an effective quantum field theory of weak gravitational fields.

2.3.1 Classical general relativity

GR is a classical relativistic field theory with the fundamental field being $g_{\mu\nu}(x)$. Why do we choose this field?

- The charge is P^μ, the current is $T^{\mu\nu}$, so the field should be a symmetric tensor like $T^{\mu\nu}$.
- Gravity is a long-range force and always attractive (couples to mass). Non-relativistically, the Higgs field can couple to mass as $V(r) = -2m_1m_2/v^2 \cdot r^{-1}$. However, the Higgs doesn't couple to photons, to which the gravity does couple (bending of light). This leaves a tensor field as the simplest (and in fact only) option.

The Einstein–Hilbert action follows almost uniquely from the nature of the fundamental field, the requirements of covariance and charge conservation and the Newtonian limit, $g_{\mu\nu}$ as a field and second order in derivative.

From the metric, one constructs the affine connection (*Christoffel symbols*)

$$\Gamma^\lambda_{\mu\nu} = \frac{1}{2} g^{\lambda\sigma} (\partial_\mu g_{\nu\sigma} + \partial_\nu g_{\mu\sigma} - \partial_\sigma g_{\mu\nu}).$$

The symmetry of the theory is the group of general coordinate transformations (diffeomorphisms). The covariant derivative for a general tensor is defined with the Christoffel symbols

$$\nabla_\lambda T^{\mu\nu\ldots}{}_{\rho\sigma\ldots} = \partial_\lambda T^{\mu\nu\ldots}{}_{\rho\sigma\ldots} + \Gamma^\mu{}_{\lambda\alpha} T^{\alpha\nu\ldots}{}_{\rho\sigma\ldots} + \Gamma^\nu{}_{\alpha\lambda} T^{\mu\alpha\ldots}{}_{\rho\sigma\ldots} + \cdots - \Gamma^\alpha{}_{\lambda\rho} T^{\mu\nu\ldots}{}_{\alpha\sigma\ldots} - \ldots$$

From the connection, one constructs the *Riemann curvature tensor*

$$R^\rho{}_{\sigma\mu\nu} = \partial_\mu \Gamma^\rho{}_{\nu\sigma} - \partial_\nu \Gamma^\rho{}_{\mu\sigma} + \Gamma^\rho{}_{\mu\lambda} \Gamma^\lambda{}_{\nu\sigma} - \Gamma^\rho{}_{\nu\lambda} \Gamma^\lambda{}_{\mu\sigma},$$

and further the *Ricci tensor* and the *Ricci scalar*

$$R_{\mu\nu} = R^\lambda{}_{\mu\lambda\nu}, \quad R = g^{\mu\nu} R_{\mu\nu}.$$

The general relativity has a nice geometric interpretation. The line element can in general be written as $\mathrm{d}s^2 = g_{\mu\nu}(x) \mathrm{d}x^\mu \, \mathrm{d}x^\nu$. The theory is invariant under the general coordinate transformation

$$x^\mu \to x'^\mu(x), \quad \mathrm{d}x'^\mu = \frac{\partial x'^\mu}{\partial x^\nu}(x) \, \mathrm{d}x^\nu.$$

The Jacobian of the transformation is x-dependent. There is also the inverse of the metric $g^{\mu\nu}$, which satisfies $g^{\mu\lambda}g_{\lambda\nu} = \delta^{\mu}{}_{\nu}$. The Christoffel symbol is not a tensor, but the covariant derivative is, i.e. it converts a tensor to a tensor with rank increased by one. The Riemann tensor is a tensor of rank four and contains two derivatives on the fundamental field. It is therefore a candidate to build the field equation and the action. Using the appropriate Newtonian limit and the second order in derivatives, we get the *Einstein field equations*:

$$R_{\mu\nu} - \frac{1}{2}g_{\mu\nu}R + \Lambda g_{\mu\nu} = -8\pi G_{\mathrm{N}} T_{\mu\nu}.$$

This is the equation of motion for the *Einstein–Hilbert action*

$$S_{\mathrm{EH}} = \frac{1}{16\pi G_{\mathrm{N}}} \int \mathrm{d}^4x \, \sqrt{-g}(R - 2\Lambda)$$

after varying with respect to $g_{\mu\nu}$ (without matter coupling). The full action is $S = S_{\mathrm{EH}}[g_{\mu\nu}] + S_{\mathrm{M}}[g_{\mu\nu}, \phi, \dots]$, where S_{M} is the matter action. The energy-momentum tensor is *defined* as

$$\delta S_{\mathrm{M}} = -\frac{1}{2} \int \mathrm{d}^4x \, \sqrt{-g} T^{\mu\nu}(x) \, \delta g_{\mu\nu}.$$

General relativity can be viewed as a theory of a dynamical space-time, where energy-momentum as the source covers spaces. Or it can be viewed as the most general theory of an interacting, massless spin-2 field – up to terms of higher mass dimension that one could add to the action.

The mass dimensions of these objects are $[g_{\mu\nu}] = 0$, $[R] = 2$, $[\Lambda] = 2$, $[G_{\mathrm{N}}] = -2$ and $[T_{\mu\nu}] = 4$. The action is a scalar and is thus invariant under general coordinate transformations. The metric transforms as

$$g_{\mu\nu}(x) \to g'_{\mu\nu}(x') = \frac{\partial x^{\alpha}}{\partial x'^{\mu}} \frac{\partial x^{\beta}}{\partial x'^{\nu}} g_{\alpha\beta}(x).$$

Infinitesimally ($x'^{\mu} = x^{\mu} + \varepsilon^{\mu}(x)$), this becomes

$$\delta g_{\mu\nu}(x) = g'_{\mu\nu}(x) - g_{\mu\nu}(x) = -\partial_{\mu}\varepsilon^{\rho} g_{\rho\nu} - \partial_{\nu}\varepsilon^{\rho} g_{\mu\rho} - \varepsilon^{\rho}\partial_{\rho} g_{\mu\nu}.$$

If $g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$ with $|h_{\mu\nu}| \ll 1$, we get $\delta h_{\mu\nu} = -\partial_{\mu}\varepsilon_{\nu} - \partial_{\nu}\varepsilon_{\mu}$, which is exactly the gauge transformation with a homogeneous and inhomogeneous terms. We will therefore quantize this theory by using the methods introduced for non-Abelian gauge theories of vector fields. The complications are mainly of technical nature.

2.3.2 Quantization of weak gravitational fields

We will quantize fluctuations around the flat Minkowski space-time: $g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu}$. The Minkowski metric is the vacuum expectation value of the metric field

$$\langle \Omega | g_{\mu\nu} | \Omega \rangle = \eta_{\mu\nu}.$$

This also means that $\eta_{\mu\nu}$ has to be a solution of the classical field equations in vacuum, which is the case if we set $\Lambda = 0$ (if $\Lambda \neq 0$ we would need to quantize around de Sitter or anti-de Sitter background, which is quite difficult). The Lagrangian we are left with is

$$\mathcal{L}_{\mathrm{EH}} = \frac{2}{\kappa^2}\sqrt{-g}R, \qquad \kappa = \sqrt{32\pi G_{\mathrm{N}}}.$$

We want to expand this Lagrangian in powers of $h_{\mu\nu}$:

$$-g = -\det(\eta_{\mu\rho}(\delta^\rho{}_\nu + h^\rho{}_\nu)) = -\det(\eta_{\mu\nu})\exp(\mathrm{tr}(\ln(\mathbb{1}+h))) = \exp(h^\alpha{}_\alpha - h^\alpha{}_\beta h^\beta{}_\alpha/2 + \dots).$$

Including also the square root, we get

$$\sqrt{-g} = 1 + \frac{1}{2}h^\alpha{}_\alpha + \frac{1}{8}(h^\alpha{}_\alpha)^2 - \frac{1}{4}h^\alpha{}_\beta h^\beta{}_\alpha + \dots$$

The expansion for R is

$$\begin{aligned} R = g^{\mu\nu}R_{\mu\nu} &= (\eta^{\mu\nu} - h^{\mu\nu})(R^{(1)}_{\mu\nu} + R^{(2)}_{\mu\nu} + \dots) \\ &= (\eta^{\mu\nu} - h^{\mu\nu})(\frac{1}{2}\partial^\rho\partial_\mu h_{\nu\rho} + \frac{1}{2}\partial^\rho\partial_\nu h_{\mu\rho} - \frac{1}{2}\partial^2 h_{\mu\nu} - \frac{1}{2}\partial_\mu\partial_\nu h^\rho{}_\rho + \dots). \end{aligned}$$

The easiest approach is to expand the Lagrangian in powers of h:

$$\mathcal{L}_{\mathrm{EH}} = \mathcal{L}^{(1)} + \mathcal{L}^{(2)} + \dots$$

The first order term is (using $h = h^\alpha{}_\alpha$)

$$\frac{\kappa^2}{2}\mathcal{L}^{(1)} = \partial^\rho\partial^\mu h_{\mu\rho} - \partial^2 h,$$

and it is a total derivative, so it can be dropped. The second order term is exactly the same as the kinetic term of the massless spin-2 field:

$$\mathcal{L}^{(2)} = \frac{1}{\kappa^2}\left(\frac{1}{2}\partial^\lambda h_{\mu\nu}\partial_\lambda h^{\mu\nu} - \partial_\mu h^{\mu\lambda}\partial^\nu h_{\nu\lambda} - \frac{1}{2}\partial^\lambda h\partial_\lambda h + \partial^\mu h\partial^\nu h_{\mu\nu}\right).$$

In order to do quantum field theory, we need to understand the propagators and vertices. For the propagator, we need to write the Lagrangian in a suitable form

$$\mathcal{L}_{\mathrm{EH}} = \frac{1}{\kappa^2}\left(-\frac{1}{2}h_{\mu\nu}\mathrm{D}^{\mu\nu\rho\sigma}h_{\rho\sigma} + \mathcal{O}(h^3)\right) = -\frac{1}{2}\hat{h}_{\mu\nu}\mathrm{D}^{\mu\nu\rho\sigma}\hat{h}_{\rho\sigma} + \kappa\mathcal{O}(\hat{h}^3) + \kappa^2\mathcal{O}(\hat{h}^4), \qquad (2.2)$$

where $\mathrm{D}^{\mu\nu\rho\sigma}$ is some second-order differential operator. We defined the canonically normalized field $\hat{h}_{\mu\nu} = h_{\mu\nu}/\kappa$ with $[\hat{h}] = 1$. As long as $|\kappa\hat{h}| \ll 1$, the expansion is valid. Sadly, the operator $\mathrm{D}^{\mu\nu\rho\sigma}$ is not invertible because of the gauge symmetry. So, let's fix the gauge! We need to make a choice, whether to break covariance or gauge invariance. Most of the time, the covariance is preferred, so we choose the physical gauge. This is the analogue of transverse gauge in QED, $\nabla\cdot\mathbf{A} = 0$, where the two degrees of freedom are transverse to the direction of motion. The electric potential A^0 vanishes in the vacuum, however in the presence of charges, it obeys the Poisson equation $\nabla^2 A^0 = -4\pi\rho$. With the spin-2 field, the transverse gauge (4 constraints) reads

$$\partial_i h_{i0} = 0, \quad \partial_i s_{ij} = 0$$

and the equations of motion give the additional 4 constraints (in the absence of sources)

$$\partial^\mu h_{\mu\nu} = \partial_\nu h, \quad h = 0, \quad \partial^\mu h_{\mu\nu} = 0, \quad \partial^2 h_{\mu\nu} = 0.$$

The first two constraints give $\dot{h}_{00} =$ and then $\nabla^2 h_{00} = 0$. The vacuum initial conditions give $h_{00} = 0$. The spatial trace also vanishes $h^i{}_i = h - h_{00} = 0$. The third constraint gives for $\nu = i$

$$0 = \dot{h}_{0i} - \partial_j h_{ji} = \dot{h}_{0i} - \partial_j(\delta_{ij}h^k_k + s_{ij}) = \dot{h}_{0i}.$$

With the initial condition and $\nabla^2 h_{0i} = 0$, the vacuum solution is $h_{0i} = 0$. To recap, we have

$$h_{00} = h_{0i} = h_{0\mu} = 0$$

in the absence of sources. If there are sources, these components are non-zero, but non-dynamical, determined from

$$\nabla^2 h_{0\mu} \propto T_{0\mu}.$$

The only dynamical components are in the transverse-traceless part s^{T}_{ij}, which, in the absence of sources ($\partial^2 s^{\mathrm{T}}_{ij} = 0$), admit plane wave solutions with helicity ± 2. When we quantize them, we get operators in the Fock space of gravitons:

$$s^{\mathrm{T}}_{ij}(x) = \sum_{\lambda=\pm 2} \int \frac{\mathrm{d}^3\mathbf{k}}{(2\pi)^3} \frac{1}{2E_{\mathbf{k}}} \left(e^{-ikx} \varepsilon_{ij}(p,\lambda) a(p,\lambda) + \text{h.c.} \right).$$

The operators a and $a^\dagger$ destroy and create *gravitons* with helicity ± 2.

Equally, we could use the *de Donder gauge* (harmonic gauge):

$$\partial^\nu h_{\mu\nu} - \frac{1}{2} \partial_\mu h^\nu{}_\nu = 0.$$

Applying the *Faddeev–Popov method* gives the generating functional integral

$$Z[\mathcal{J}_{\mu\nu}] = N \int \mathcal{D}[h_{\mu\nu}]\, G[f_\mu[h]] \det(M) \exp\left(iS_{\mathrm{EH}} + i \int \mathrm{d}^4x\, \mathcal{J}_{\mu\nu} h^{\mu\nu} \right).$$

The gauge-fixing functional is usually taken to be Gaussian

$$G[f_\mu] = \exp\left(+ \frac{i}{\xi} \int \mathrm{d}^4x\, f_\mu(x) f^\mu(x) \right),$$

which is an equivalent of the R_ξ gauge in non-Abelian gauge theories. The Faddeev–Popov determinant $\det(M)$ is

$$M_{\mu\nu}(x,y) = \left. \frac{\delta f_\mu[h^\varepsilon_{\mu\nu}(x)]}{\delta \varepsilon^\nu(y)} \right|_{\varepsilon^\nu = 0} = M_{\mu\nu}[h](x) \delta^{(4)}(x-y).$$

The determinant can be written as a path integral over Grassmann fields η_μ and $\overline{\eta}_\mu$ (ghost fields)

$$\det M = N \int \mathcal{D}[\eta_\mu, \overline{\eta}_\mu] \exp\left(i \int \mathrm{d}^4x\, \overline{\eta}^\mu M_{\mu\nu}[h](x) \eta^\nu \right).$$

In summary the full Lagrangian consists of the graviton self-interactions, the gauge-fixing term, the ghost term and the matter coupling:

$$\mathcal{L} = \mathcal{L}_{\mathrm{EH}} + \mathcal{L}_{\mathrm{gf}} + \mathcal{L}_{\mathrm{gh}} + \mathcal{L}_M.$$

The bilinear terms in the Lagrangian can be taken together to get the graviton propagator

$$\mathcal{L}_{\mathrm{bilin}} = -\frac{1}{2} h_{\mu\nu} M^{\mu\nu,\rho\sigma} \partial^2 h_{\rho\sigma},$$

where

$$M_{[\mu\nu][\rho\sigma]} = \frac{1}{2} \left(\eta_{\mu\rho}\eta_{\nu\sigma} + \eta_{\mu\sigma}\eta_{\nu\rho} - \eta_{\mu\nu}\eta_{\rho\sigma} \right).$$

The propagator will be

$$\frac{i}{k^2+i\varepsilon}N_{[\mu\nu][\rho\sigma]}$$

where the tensor $N_{[\mu\nu][\rho\sigma]}$ satisfies

$$M^{[\mu\nu][\alpha\beta]}N_{[\alpha\beta][\rho\sigma]} = \frac{1}{2}(\delta^{\mu}{}_{\rho}\delta^{\nu}{}_{\sigma} + \delta^{\mu}{}_{\sigma}\delta^{\nu}{}_{\rho}).$$

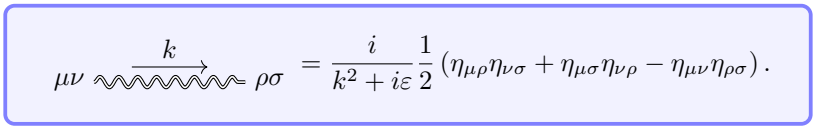
We can think of $[\mu\nu]$ as a single index running from 1 to 10. The general form for $N_{[\alpha\beta][\rho\sigma]}$ is

$$N_{[\alpha\beta][\rho\sigma]} = a(\eta_{\alpha\rho}\eta_{\beta\sigma} + \eta_{\alpha\sigma}\eta_{\beta\rho}) - b\eta_{\alpha\beta}\eta_{\rho\sigma}.$$

To find a and b we write

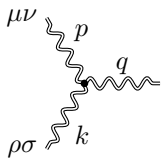
$$\begin{aligned}\frac{1}{2}(\delta^{\mu}{}_{\rho}\delta^{\nu}{}_{\sigma} + \delta^{\mu}{}_{\sigma}\delta^{\nu}{}_{\rho}) &= \frac{1}{2}(\eta^{\mu\alpha}\eta^{\nu\beta} + \eta^{\mu\beta}\eta^{\nu\alpha})a(\eta_{\alpha\rho}\eta_{\beta\sigma} + \eta_{\alpha\sigma}\eta_{\beta\rho}) - \frac{b}{2}(\eta^{\mu\alpha}\eta^{\nu\beta} + \eta^{\mu\beta}\eta^{\nu\alpha})\eta_{\alpha\beta}\eta_{\rho\sigma} + \\ &\quad - \frac{1}{2}\eta^{\mu\nu}\eta^{\alpha\beta}a(\eta_{\alpha\rho}\eta_{\beta\sigma} + \eta_{\alpha\sigma}\eta_{\beta\rho}) - \frac{1}{2}\eta^{\mu\nu}\eta^{\alpha\beta}(-b)\eta_{\alpha\beta}\eta_{\rho\sigma} \\ &= a(\delta^{\mu}{}_{\rho}\delta^{\nu}{}_{\sigma} + \delta^{\mu}{}_{\sigma}\delta^{\nu}{}_{\rho}) - b\eta^{\mu\nu}\eta_{\rho\sigma} - a\eta^{\mu\nu}\eta_{\rho\sigma} + 2b\eta^{\mu\nu}\eta_{\rho\sigma}.\end{aligned}$$

It follows $a = 1/2$ and $-b - a + 2b = 0$, hence $a = b = 1/2$ and $M^{[\mu\nu][\rho\sigma]} = N^{[\mu\nu][\rho\sigma]}$. The graviton propagator in the de Donder gauge is



$$= \frac{i}{k^2+i\varepsilon}\frac{1}{2}\left(\eta_{\mu\rho}\eta_{\nu\sigma} + \eta_{\mu\sigma}\eta_{\nu\rho} - \eta_{\mu\nu}\eta_{\rho\sigma}\right).$$

The vertices of the graviton self-interactions go to arbitrary high order in κ (and the number of legs). For n legs, the prefactor is κ^{n-2} and there are always two derivatives acting on the fields (from Riemann tensor). The three point self-interaction vertex is (roughly)



$$\sim \kappa\left(pp\eta\eta + p^2\eta\eta\eta\right).$$

Similarly, the 4-point vertex has the form $\kappa^2(pp\eta\eta\eta + p^2\eta\eta\eta\eta)$ and this generalizes to n-graviton vertices. However, we will see below that after renormalization, there are further vertices with more factors of momentum. In addition, there are the graviton-ghost interactions.

2.3.3 Interaction with matter

The graviton also couples to matter:

$$S_M = \int \mathrm{d}^4x\,\sqrt{-g}\mathcal{L}_M(\psi, \nabla_\lambda\psi).$$

It is easiest to start with a scalar field ϕ (because the covariant derivative is just the partial derivative)

$$\mathcal{L}_\phi = \frac{1}{2}(\partial_\mu\phi\partial^\mu\phi - m^2\phi^2) = \frac{1}{2}g^{\mu\nu}\partial_\mu\phi\partial_\nu\phi - \frac{1}{2}m^2\phi^2.$$

The coupling goes via the energy-momentum tensor

$$S_M = S_{\phi,\text{free}} - \kappa \int \mathrm{d}^4x\, h^{\mu\nu} T_{\mu\nu} + \mathcal{O}(h^2),$$

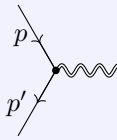
where

$$T_{\mu\nu} = \frac{1}{2}\left(\partial_\mu\phi\partial_\nu\phi - \frac{1}{2}\eta_{\mu\nu}\partial_\rho\phi\partial^\rho\phi + \frac{1}{2}m^2\eta_{\mu\nu}\phi^2\right).$$

This is called the *minimal coupling* of matter to gravity. There could also exist an extra term

$$S_M \supset -\frac{1}{2}\int \mathrm{d}^4x\, \sqrt{-g}\xi R\phi^2$$

where ξ is dimensionless coupling constant. However, we won't examine this interaction in this course. For the minimally coupled scalar vertex (for $\xi = 0$) we get



$$= -\frac{i\kappa}{2}(p_\mu p'_\nu + p_\nu p'_\mu - \eta_{\mu\nu}(p\cdot p' - m^2)).$$

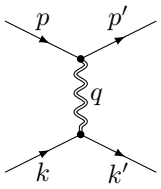
The arrows on the propagators only denote the direction of the momentum and not the fermion nature of the particle. Coupling to the electromagnetic field is

$$S_A = \int \mathrm{d}^4x\, \sqrt{-g}\left(-\frac{1}{4}F_{\mu\nu}F^{\mu\nu}\right) = -\frac{1}{4}\int \mathrm{d}^4x\, \sqrt{-g}g^{\mu\alpha}g^{\nu\beta}F_{\mu\nu}F_{\alpha\beta},$$

where

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

The coupling of spinor/fermionic fields is more complicated, because they do not transform as Lorentz tensors, but rather under representations of the universal covering group $\mathrm{SL}(2,\mathbb{C})$ of $\mathrm{SO}(1,3)$. This requires *vierbein formalism*. Now, we will consider the scattering of non-relativistic scalar particles (aka planets) through graviton exchange (note that the scalars are denoted by fermion lines in order to signal the direction of momentum flow):



$$= \frac{-i\kappa}{2}(p_\mu p'_\nu + p_\nu p'_\mu - \eta_{\mu\nu}(p\cdot p' - m^2))\frac{i(\eta^{\mu\rho}\eta^{\nu\sigma} + \eta^{\mu\sigma}\eta^{\nu\rho} - \eta^{\mu\nu}\eta^{\rho\sigma})/2}{q^2 + i\varepsilon}\cdot$$
$$\cdot\frac{-i\kappa}{2}(k_\rho k'_\sigma + k_\sigma k'_\rho - \eta_{\rho\sigma}(k\cdot k' - M^2)).$$

In the non-relativistic limit, $p^\mu \to mv^\mu$ and $v^\mu = (1,\mathbf{0})$. Both vertex factors become $2m^2v^\mu v^\nu$ and after the contractions, we get

$$\mathcal{M} = \frac{(-i\kappa)^2}{4}2m^2 2M^2\frac{i}{q^2+i\varepsilon}\frac{1}{2} \approx \frac{i\kappa^2}{2}\frac{m^2M^2}{\mathbf{q}^2}.$$

We can compare this result to the non-relativistic exchange of photons:

$$\mathcal{M} = ie(p+p')^\mu\frac{-i\eta_{\mu\nu}}{q^2+i\varepsilon}ie(k+k')^\nu \approx -ie^2\frac{4mM}{\mathbf{q}^2}.$$

As we can see, the coupling e is replaced by κm. It is more inspiring to compare the potentials in position space. The Fourier transform of $1/\mathbf{q}^2$ is

$$\int \frac{\mathrm{d}^3\mathbf{q}}{(2\pi)^3} e^{i\mathbf{q}\cdot\mathbf{r}} \frac{1}{\mathbf{q}^2} = -\frac{1}{4\pi r}.$$

The factors $2m$ and $2M$ are absorbed as the relativistic normalization of states. The potentials are thus

$$V_{\mathrm{em}}(r) = \frac{e^2}{4\pi r}, \quad V_{\mathrm{gr}}(r) = -\frac{\kappa^2}{32\pi}\frac{Mm}{r} = -\frac{GMm}{r}.$$

The electromagnetic potential is repulsive for like charges, while the gravitational potential is always attractive (this happens because the electromagnetic interaction is linear in p, while the graviton one is even). We should not be surprised by the familiar result, since we have constructed the theory to have the correct non-relativistic limit. Another important point is the mass in the Feynman diagram is inertial mass, and exactly this mass comes into the gravitational potential, which is a manifestation of the equivalence principle.

2.3.4 Renormalization

As we discussed before, the Lagrangian is (eq. 2.2)

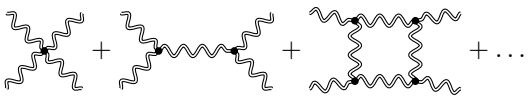
$$\mathcal{L} = \mathcal{L}_{\mathrm{EH}} + L_{\mathrm{GF+ghost}} = -\frac{1}{2} h_{\mu\nu} D^{\mu\nu\rho\sigma} h_{\rho\sigma} + \mathcal{L}_{\mathrm{ghost}} + \kappa\mathcal{O}(\partial^2 h^3) + \kappa^2\mathcal{O}(\partial^2 h^4).$$

Since $[h] = 1$ and $[\kappa] = -1$ *all interactions are non-renormalizable* (starting at triple graviton self-interaction at dimension 5), including the matter interactions! Einstein–Hilbert action defines an effective field theory of graviton for vector fields, valid up to the Planck scale

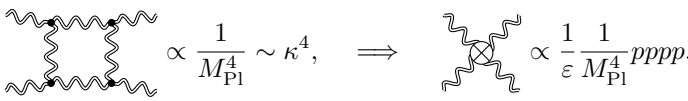
$$\frac{1}{\kappa} = \frac{1}{\sqrt{32\pi G_{\mathrm{N}}}} = \frac{M_{\mathrm{Pl}}}{2} = 1.218 \cdot 10^{18}\,\mathrm{GeV}.$$

This is 15 orders of magnitude larger than any other known microscopic scale (such as vev of the Higgs boson). This is the reason why graviton is relevant only for macroscopic bodies with very large "charge" (mass). For scattering energy $E \ll M_{\mathrm{Pl}}$, the contribution from n-graviton vertex is suppressed by $(\kappa E)^{n-2} = (E/M_{\mathrm{pl}})^{n-2}$ (the factors of energy just have to match the factors of κ). As for any effective (non-renormalizable) QFT, quantum correction can be consistently computed as an expansion in E/M_{Pl} by adding all higher-dimensional operators (up to the dimension required by the desired accuracy) compatible with the symmetries of the action.

Example 1. Consider the graviton-graviton scattering



The $2 \to 2$ scattering is dimensionless, and the loops are UV-divergent. We need to add a counter-term to cancel the divergence (since it has to be dimensionless, and it comes with four factors of κ, it also has to come with four derivatives).



This means that we have to add a local operator of the form $\partial\partial\partial\partial hhhh/M_{\mathrm{Pl}}^4$ to the Einstein–Hilbert action as a counter-term. We cannot absorb this into a redefinition of the parameters in the original Lagrangian (like gravitational constant). The terms that are fourth order in derivatives and h are R^2, $R_{\mu\nu}R^{\mu\nu}$ and $R_{\mu\nu\rho\sigma}R^{\mu\nu\rho\sigma}$. However, the last term can be written as a combination of the first two terms (Gauss–Bonnet theorem) and is not independent up to a topological term.

In pure gravity, the effective action is

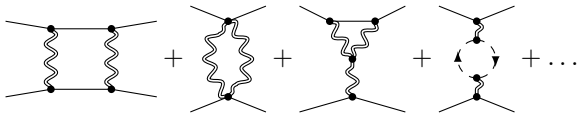
$$S_{\mathrm{eff}} = \int \mathrm{d}^4x\, \sqrt{-g}\left(-\tilde{\Lambda} + \frac{1}{\kappa^2}R + c_1R^2 + c_2R_{\mu\nu}R^{\mu\nu} + \dots\right).$$

As a QFT, EH action has to be extended to the general EFT. The divergences at 1-loop are absorbed into renormalized c_1 and c_2 and so on.

Note. For $S_M = 0$, we have $R = R_{\mu\nu} = 0$ by equations of motion. So the one-loop, on-shell graviton scattering amplitude is not divergent. But for $S_M \neq 0$, the terms R^2, $R_{\mu\nu}R^{\mu\nu}$ must appear.

Quantum corrections to the Newtonian potential

We can return to the scattering of two massive scalars via graviton exchange and calculate the quantum corrections to the Newtonian potential.



When the momentum transfer is small, $|\mathbf{q}| \ll m, M$ (and $\mathbf{r} \gg 1/m_1, 1/m_2$) we can expand the scattering amplitude as

$$\mathcal{M} = -i\frac{G_{\mathrm{N}}m_1m_2}{q^2}\left(\tilde{a} + \tilde{b}G_{\mathrm{N}}(m_1+m_2)\sqrt{-q^2} + \tilde{c}G_{\mathrm{N}}q^2\left(\ln\frac{-q^2}{\mu^2} + \tilde{d}\right) + \dots\right).$$

The last constant $\tilde{d}$ is dependent on the renormalization scheme. In the position space, the potential reads

$$V(r) = -\frac{G_{\mathrm{N}}m_1m_2}{r}\left(1 + b\frac{G_{\mathrm{N}}(m_1+m_2)}{r} + c\frac{G_{\mathrm{N}}}{r^2} + \dots\right) + d\delta^{(3)}(\mathbf{r}).$$

In 2017, the values of b and c were calculated to be $b = 3$ and $c = 41/(10\pi)$. To restore the units, we have

$$[G_{\mathrm{N}}] = \frac{\mathrm{m}^3}{\mathrm{kg\,s}^2}, \quad [\hbar] = \mathrm{J\,s} = \frac{\mathrm{kg\,m}^2}{\mathrm{s}}, \quad [c] = \frac{\mathrm{m}}{\mathrm{s}}.$$

The potential in physical units is thus

$$V(r) = -\frac{G_{\mathrm{N}}m_1m_2}{r}\left(1 + 3\frac{G_{\mathrm{N}}(m_1+m_2)}{rc^2} + \frac{41}{10\pi}\frac{G_{\mathrm{N}}\hbar}{r^2c^3} + \dots\right)$$

We can see that the first correction is classical (doesn't depend on $\hbar$, sometimes called "classical loop") and the second correction is quantum (depends on $\hbar$). The smallest distances in terrestrial high energy physics are of order 1 TeV, which amounts to $G_{\mathrm{N}}/r^2 \sim 10^{-34}$. Experimentally, we are very far from testing these corrections. However, in the everyday life the gravity still plays the most important role, because the gravitational charges exist in exorbitant amounts (like the Earth mass).

Chapter 3

Spontaneous Symmetry Breaking

Let's first recapitulate some basic notions. Assume we have a scalar field ϕ with action

$$S = \int \mathrm{d}^4x \; \mathcal{L}(\phi_n, \partial_\mu \phi_n)$$

that is invariant under a group G with symmetry transformation $\phi'_n(x) = \mathcal{U}(g)\phi_n(x)\mathcal{U}(g)^{-1}$. For continuous groups and infinitesimal transformations, parametrized by $\varepsilon^A = \varepsilon^A(g)$, $A = 1, \ldots, \dim G$, the field transforms as

$$\delta\phi_n(x) = \varepsilon^A \delta^A \phi_n(x) \equiv \phi'_n(x) - \phi_n(x) = \varepsilon^A F_n^A[\phi_m, \partial_\mu \phi_m] = i[\varepsilon^A Q^A, \phi_n(x)],$$

where Q^A are the conserved charges associated with the symmetry. The last equality holds if F_n^A does not depend on $\dot{\phi}_n$ and if $K^0(x)$ that appears in the Noether theorem vanishes. Then $j^{0,A} = F_n^A \partial\mathcal{L}/\partial\dot{\phi}_n$ and F_n commutes with $\phi_{n'}$ at equal times, which proves the equality. We distinguish several types of symmetries.

- **Global vs. local symmetry**: global means that $\mathcal{U}(g)$ is the same for all x^μ and $\varepsilon^A F^A$ does not depend explicitly on x^μ.
- **Internal vs. spacetime symmetry**: for internal symmetries, $\phi'_n(x)$, when expressed in terms of fields, depends only on fields at the same point x^μ.
- **Linear vs. non-linear symmetry**: for linear symmetries, $\delta\phi_n(x)$ is linear in fields $\phi_m(x)$, that is

  $$F_n^A[\phi_m, \partial_\mu \phi_m] = i \int \mathrm{d}^4y \, \hat{L}_{nm}^A(x-y)\phi_m(y) = iL_{nm}^A(x, \partial^\mu)\phi_m(x),$$

  where $L_{nm}^A(x, \partial^\mu)$ is a linear operator. For instance, for translations it holds $\varepsilon^A \to \varepsilon^\mu$ and $L_{nm}^\mu = \delta_{nm}\partial^\mu$. For an internal symmetry, it holds $L_{nm}^A = \delta^{(4)}(x-y)T_{nm}^A$, hence L_{nm}^A can only be a space-time independent matrix without differential operators,

  $$\delta\phi_n(x) = i\varepsilon^A L_{nm}^A \phi_m(x).$$

  L_{nm}^A is the generator of the matrix (finite-dimensional) representation of the Lie algebra of the symmetry group G. Poincaré and gauge symmetries have so far been linearly represented on fields, but the BRST symmetry is an example of a non-linear symmetry, since $\delta_\lambda c^A \propto f^{ABC} c^B c^C$.

3.1 Spontaneous breaking of global symmetries

3.1.1 Spontaneous symmetry breaking and the vacuum

Vacuum $|\Omega\rangle$ is the lowest energy state of the theory. It can be a degenerate state, but we assume it corresponds to an eigenvalue in the discrete spectrum of the Hamiltonian H (contrary to single-particle states with eigenvalues $\sqrt{m^2+\mathbf{p}^2}$).

Definition 3.1.1 (Spontaneous symmetry breaking)**.** A symmetry associated with variations δ^A is spontaneously broken, if there exists an operator $A(\phi)$, for which $\langle\Omega|\delta^A A(\phi)(x)|\Omega\rangle \neq 0$. The term $\delta^A A(\phi)$ is then called an *order parameter*.

Remark.

a) In principle, Poincaré symmetry could be spontaneously broken. This case, for instance, the speed of light in vacuum could be anisotropic. This is not observed in Nature and does not happen in the so-called Standard Model. In the following we assume that translation and Lorentz symmetry is unbroken, and we consider only internal symmetries. More precisely, we assume $\mathcal{U}(\Lambda,a)\,|\Omega\rangle = |\Omega\rangle$ for all $\Lambda \in \mathrm{SO}(1,3)$ and $a \in \mathbb{R}^4$.

b) Let $b(x) \equiv \langle\Omega|\delta^A A(x)|\Omega\rangle \neq 0$ for some internal symmetry. Then $B(x) \equiv \delta^A A(x)$ must be a scalar under Lorentz transformations and $b(x)$ independent of x.

Proof. Consider a translation (which is unbroken by assumption)

$$0 = \langle\Omega|\delta^\mu B(x)|\Omega\rangle = \langle\Omega|\partial^\mu B(x)|\Omega\rangle = \partial^\mu \langle\Omega|B(x)|\Omega\rangle = \partial^\mu b(x),$$

so b has to be a constant. Next, consider a Lorentz transformation

$$B'(x) = \mathcal{U}(\Lambda)B(x)\mathcal{U}(\Lambda)^{-1} = \exp(i\omega_{\mu\nu}\Sigma^{\mu\nu}/2)B(\Lambda x),$$

where $\Sigma^{\mu\nu}$ depends on the Lorentz representation of B (0 for scalar, $i[\gamma^\mu,\gamma^\nu]/4$ for Dirac spinor, etc.). Then, after expanding the exponent and Λx for $\Lambda = \mathbb{1} + \omega$, we get

$$\delta^{\mu\nu}B(x) = \frac{i}{2}(i(x^\mu\partial^\nu - x^\nu\partial^\mu) + \Sigma^{\mu\nu})B(x)$$

and, since Lorentz symmetry is unbroken,

$$0 = \langle\Omega|\delta^{\mu\nu}B(x)|\Omega\rangle = \frac{i}{2}\Sigma^{\mu\nu}b.$$

From this it follows that $\Sigma^{\mu\nu}b = 0$ for all μ,ν, which is only possible if B is a scalar under Lorentz transformations. □

c) Spontaneous symmetry breaking can only occur if the vacuum is degenerate, i.e., if it is not, all symmetries are unbroken.

Proof. For an internal symmetry it holds

$$\delta\phi = i[\varepsilon^A Q^A,\phi], \quad \phi'(x) = e^{i\varepsilon^A Q^A}\phi(x)e^{-i\varepsilon^A Q^A}.$$

Since Q^A is hermitian, $\mathcal{U}(g)$ is unitary, and we could write it as an exponential of Q^A. Suppose now that $|\Omega\rangle$ is non-degenerate. Since Q^A is a conserved charge, it holds $[H,\mathcal{U}(g)] = 0$

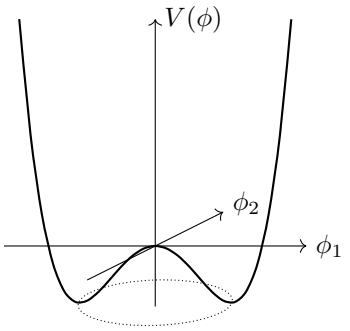
for all $g \in G$, so $\mathcal{U}(g)\,|\Omega\rangle$ has the same energy as $|\Omega\rangle$. Since $|\Omega\rangle$ is non-degenerate, it follows that $\mathcal{U}(g)\,|\Omega\rangle = e^{i\varepsilon^A Q^A}\,|\Omega\rangle \propto |\Omega\rangle$. Thus, $|\Omega\rangle$ is an eigenstate of every Q^A with some real eigenvalue q^A (real because Q^A is hermitian). Hence, for an arbitrary operator $A(x)$, we have

$$\langle\Omega|\delta^A A(x)|\Omega\rangle = i\,\langle\Omega|Q^A A(x) - A(x)Q^A|\Omega\rangle = 0$$

and all symmetries are unbroken. □

On the other hand, if the vacuum is degenerate, $\mathcal{U}(g)\,|\Omega\rangle$ can be proportional to a different vacuum $|\Omega'\rangle$. Spontaneous symmetry breaking is then associated with the fact that due to some prior external influence (for example), the system is in a particular vacuum state which is not invariant under G.

An example from classical mechanics is a sombrero potential. Here, the symmetry group G is the group of rotations of a vector $\boldsymbol{\phi} = (\phi_1, \phi_2)$ in the plane, i.e., $G = \mathrm{O}(2) = \mathrm{U}(1)$.

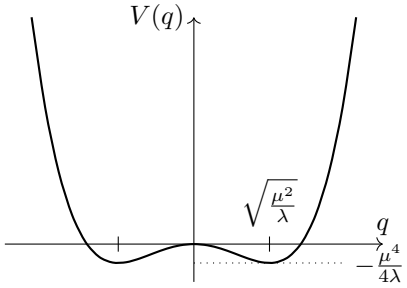


This description carries over to fields, since $\phi_i = \langle\Omega|\phi_i(x)|\Omega\rangle$ is x-independent.

We will consider spontaneous symmetry breaking classically, quantum-mechanically and in field theory.

Classical level. Consider a one-dimensional system with a double-well potential

$$L = \frac{1}{2}\dot{q}^2 - V(q), \qquad V(q) = -\frac{1}{2}\mu^2 q^2 + \frac{\lambda}{4}q^4.$$



The Lagrangian L is invariant under the $\mathbb{Z}_2$ parity transformation $P : q \to -q$. However, the classical ground state is either $+\sqrt{\mu^2/\lambda}$ or $-\sqrt{\mu^2/\lambda}$, and parity is spontaneously broken.

Quantum-mechanical level. The spectrum of the low energy states is close to a doubly degenerate harmonic oscillator with

$$E_n \approx -\frac{\mu^4}{4\lambda} + \omega^2(n + 1/2), \quad \omega^2 = V''(\pm\sqrt{\mu^2/\lambda}) = 2\mu^2.$$

The wave-functions are centred around $\pm\sqrt{\mu^2/\lambda}$ for the $|n,\pm\rangle$ states. However, $|n,\pm\rangle$ are not even approximate eigenstates of the Hamiltonian, because of quantum-mechanical tunnelling. This lifts the degeneracy by a small amount and (to a good approximation), the energy and wave-function of the two lowest-lying states $|\Omega,\pm\rangle$ are now

$$|\Omega,\pm\rangle = \frac{1}{\sqrt{2}}(|0,+\rangle \pm |0,-\rangle)$$

with energy difference

$$E_+ - E_- = -E_0 e^{-\Delta}, \quad \Delta = 2\int_{-\sqrt{\mu^2/\lambda}}^{\sqrt{\mu^2/\lambda}} \mathrm{d}q\, \sqrt{2(V(q)-E_0)}.$$

The difference is exponentially suppressed, nut non-vanishing due to the tunnelling effect. The parity symmetry is not broken quantum-mechanically. The ground state $|\Omega,+\rangle$ is parity even ($P|\Omega,\pm\rangle = \pm|\Omega,\pm\rangle$).

Field-theoretical level. Promote a generalized coordinate $q(t)$ to a field $\phi_{\mathbf{x}} \equiv \phi(\mathbf{x})$. The Lagrangian becomes

$$L = \sum_{\mathbf{x}} \frac{1}{2}\left(\dot{\phi}_{\mathbf{x}}^2 - \sum_{\mathbf{n}}\left[\frac{\phi_{\mathbf{x}+a\mathbf{n}} - \phi_{\mathbf{x}}}{a}\right]^2\right) - \left(-\frac{1}{2}\mu^2\phi_{\mathbf{x}}^2 + \frac{\lambda}{4}\phi_{\mathbf{x}}^4\right).$$

The sum over $\mathbf{n}$ vanishes for constant fields. Now, for many coordinates, the system has to tunnel between the degenerate ground states $\phi_{\mathbf{x}} = \pm\sqrt{\mu^2/\lambda}$ for all $\mathbf{x}$ and the tunnelling amplitude is proportional to $e^{-N\Delta}$, where N is the number of space points. The parity symmetry is not restored by tunnelling in an infinite (continuous) system and the degeneracy is not lifted. There are two degenerate, unequivalent ground states and the symmetry is spontaneously broken when the system has settled into one of the ground states.

3.1.2 Field-theoretical characterization of the vacuum through the effective action

Recall that the generating functional of correlation functions is defined as

$$Z[J] \equiv \frac{\displaystyle\int \mathcal{D}[\phi]\, \exp\left(iS[\phi] + i\int \mathrm{d}^4x\, J(x)\phi(x)\right)}{\displaystyle\int \mathcal{D}[\phi] \exp\left(iS[\phi]\right)}.$$

The n-point function is then calculated by taking functional derivatives with respect to J and setting $J = 0$,

$$\langle\Omega|T(\phi(x_1)\dots\phi(x_n))|\Omega\rangle = \prod_{i=1}^{n}\frac{1}{i}\frac{\delta}{\delta J(x_i)} Z[J]\big|_{J=0}.$$

The generating functional can also be expressed as an exponential, $Z[J] = e^{iW[J]}$, where $W[J]$ is the generating functional of connected Green functions,

$$G^{\mathrm{c}}(x_1,\dots,x_n) = \prod_{i=1}^{n}\frac{1}{i}\frac{\delta}{\delta J(x_i)} iW[J]\big|_{J=0}.$$

For notational simplicity, we will consider a single real scalar field.

- We can take functional derivatives at $J \neq 0$. This gives the Green functions G^J and $G^{c,J}$ in the presence of external source $J(x)$.
- We can also generate Green functions of arbitrary operators $A(x)$ by adding

$$+i \int \mathrm{d}^4x \; J_A(x) A(x)$$

to the exponent in the definition of $Z[J]$ and taking functional derivatives with respect to J_A. We will not use this in the following, but the generalization should become clear.

The *effective action* $\Gamma[\varphi]$ is defined as

$$\Gamma[\varphi] \equiv W[J] - \int \mathrm{d}^4x \; \varphi(x) J(x).$$

This is a functional of the classical field $\varphi(x)$ and not an operator. It can be thought of as the functional Legendre transform of $W[J]$, i.e., for given $\varphi(x)$, $J(x)$ is determined by solving

$$\varphi(x) = \frac{\delta W[J]}{\delta J[x]}$$

for $J(x)$ and inserting the solution on the right-hand side of the definition of $\Gamma[\varphi]$. Then define

$$\Gamma^{(n)}_{\varphi_0}(x_1, \ldots, x_n) = \frac{\delta}{\delta\varphi(x_1)} \cdots \frac{\delta}{\delta\varphi(x_n)} \Gamma[\varphi]\big|_{\varphi=\varphi_0}$$

and

$$\Gamma[\varphi] = \sum_{n=0}^{\infty} \frac{1}{n!} \int \mathrm{d}^4x_1 \ldots \mathrm{d}^4x_n \, \Gamma^{(n)}_{\varphi_0}(x_1, \ldots, x_n) \hat{\varphi}(x_1) \ldots \hat{\varphi}(x_n)$$

where $\hat{\varphi} = \varphi - \varphi_0$. This is the functional Taylor expansion around φ_0. For slowly varying field configurations $\varphi(x)$ one can perform a *gradient expansion*

$$\Gamma[\varphi] = \int \mathrm{d}^4x \; \left(-V_{\text{eff}}(\varphi) + \frac{1}{2} Z(\varphi) \partial_\mu \varphi \partial^\mu \varphi + \ldots \right).$$

This defines the *effective potential* $V_{\text{eff}}(\varphi)$ as a function of the classical field φ. Here are listed some properties of the effective action and potential.

a) The possible vacuum expectation values

$$\varphi = \langle \Omega | \phi(x) | \Omega \rangle = \frac{1}{i} \frac{\delta}{\delta J(x)} i W[J]\big|_{J=0}$$

in the absence of external sources are the field configurations where $\Gamma[\varphi]$ is stationary, i.e.,

$$\frac{\delta \Gamma[\varphi]}{\delta \varphi(x)} = 0.$$

- Since φ must be constant in Poincaré-invariant theories, it follows

$$\frac{\delta \Gamma}{\delta \varphi(x)} = -\frac{\partial V_{\text{eff}}}{\partial \varphi(x)},$$

so it is the stationary points of the effective potential that determine the possible vacuum expectation values of the field.

- A linear symmetry $\delta^A\phi \propto \varphi$ must be spontaneously broken, if the stationary points of $\Gamma[\varphi]$ (or $V_{\text{eff}}(\varphi)$) occur only for $\varphi(x) \neq 0$, since then also $\langle\Omega|\delta^A\phi|\Omega\rangle \neq 0$.

b) $\Gamma[\varphi]$ is the generating functional for the 1PI (one-particle irreducible) Green functions. More precisely, for $n \geq 3$, $i\Gamma_\varphi^{(n)}(x_1,\ldots,x_n)$ is the amputated 1PI n-point function in the presence of external source J, such that $\langle\Omega|\phi(x)|\Omega\rangle = \varphi(x)$. For $n = 2$, it holds $i\Gamma_\varphi^{(2)}(x_1,x_2) = (-G^{\text{c},J}(x_1,x_2))^{-1}$ (the inverse of the exact, connected 2-point function in the presence of J).

c) In the classical limit $\hbar \to 0$, $\Gamma[\varphi]$ reduces to the classical action $S[\varphi]$. That is, $\Gamma[\varphi]$ is the classical action and includes all quantum corrections when $\hbar \neq 0$. This means that $S[\varphi]$ is the generating functional for 1PI tree-level diagrams — indeed these are nothing but the vertices following from $\mathcal{L}_{\text{int}}$.

d) V_{eff} is the minimum of the energy density of the system under the constraint that the field $\phi(x)$ takes the constant vacuum expectation value (vev) $\varphi = \langle\Omega|\phi(x)|\Omega\rangle$. The vev of $\phi(x)$ is determined by the absolute minima of V_{eff}. If these do not occur at $\varphi = 0$, linear symmetries under which ϕ is not invariant are spontaneously broken, since $\delta^A\phi \propto \phi \neq 0$.

From Item c) it follows that for small quantum corrections (weak coupling) $V_{\text{eff}} = -\mathcal{L}\big|_{\partial\phi\to 0}$, so the effective potential is approximately given by the negative of the classical potential. In the Standard Model, $V_{\text{eff}} = -\mu^2\varphi^\dagger\varphi + \lambda(\varphi^\dagger)^2/4$ and small corrections. Here φ is related to the Higgs field. Hence, all symmetries under which the Higgs field is not invariant are spontaneously broken. The exact vacuum expectation value is given by the minimum of the effective potential including quantum corrections.

Proof.

a) We have

$$\frac{\delta\Gamma}{\delta\varphi(x)} = \frac{\delta W}{\delta\varphi(x)} - \int \mathrm{d}^4y\, \varphi(y)\frac{\delta J(x)}{\delta\varphi(x)} = \int \mathrm{d}^4y\, \frac{\delta W[J]}{\delta J(y)}\frac{\delta J(y)}{\delta\varphi(x)} - J(x) - \int \mathrm{d}^4y\, \varphi(y)\frac{\delta J(y)}{\delta\varphi(x)} = -J(x).$$

In the first equality we used the definition of $\Gamma[\varphi]$ and in the second equality we used the functional chain rule and the definition of $\varphi(x)$ as the functional derivative of $W[J]$. Hence, for $J = 0$, we have $\delta\Gamma/\delta\varphi(x) = 0$. Note that this has the same form as the classical equation of motion, $\delta S/\delta\phi(x) = 0$, which is consistent with property Item c) for $\hbar \to 0$.

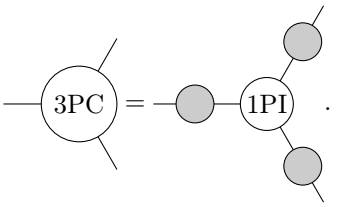
b) We will not prove the entire statement completely, but will first focus on the $n = 2$ case. We write

$$\begin{aligned}\delta^{(4)}(x-y) = \frac{\delta\varphi(x)}{\delta\varphi(y)} &= \frac{\delta}{\delta\varphi(y)}\frac{\delta W[J]}{\delta J(x)} \\ &= i\int \mathrm{d}^4z\, \frac{\delta^2 W[J]}{i\delta J(x) i\delta J(z)}\frac{\delta J(z)}{\delta\varphi(y)} = -\int \mathrm{d}^4z\, G^{\text{c},J}(x,z)\frac{i\delta^2\Gamma}{\delta\varphi(z)\delta\varphi(y)}.\end{aligned}$$

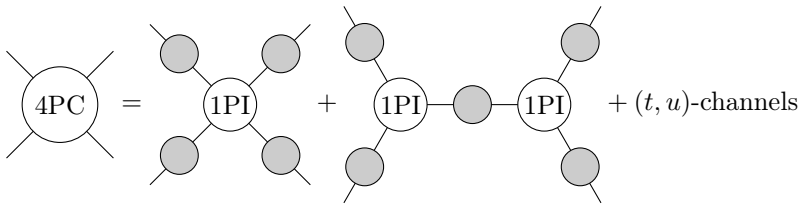
In the third equality we used the chain rule and in the last equality we used the definition of G^{c} and $J = -\delta\Gamma/\delta\varphi$. From the result above, it follows that $-i\Gamma_\varphi^{(2)} = (G^2(2)_{\text{c},J})^{-1}$. In the free theory, $G^{(2)}_{\text{c},J}(x,z) = \Delta_{\text{F}}(x-z)$ and $\Gamma^{(2)}(x,z) = (\partial^2_{(z)} + m^2)\delta^{(4)}(y-z)$, so the relation corresponds to the usual definition of the Feynman propagator, $(\partial^2_{(y)} + m^2)\Delta_{\text{F}}(y,x) = \delta^{(4)}(x-y)$. In the same way, one can show

$$G^{(3)}_{\text{c},J}(x_1,x_2,x_3) = \int \mathrm{d}^4z_1\, \mathrm{d}^4z_2\, \mathrm{d}^4z_3 \frac{i\delta^3\Gamma}{\delta\varphi(z_1)\delta\varphi(z_2)\delta\varphi(z_3)} G^{(2)}_{\text{c},J}(x_1,z_1)G^{(2)}_{\text{c},J}(x_2,z_2)G^{(2)}_{\text{c},J}(x_3,z_3).$$

Diagrammatically, this corresponds to



Every connected diagram contributing to the three-point function can be written as in the right diagram, which confirms that $\Gamma^{(3)}_\varphi$ is the 1PI three-point function. In the next order, one derives a relation between $G^{(4)}_{\mathrm{c},J}$, $\Gamma^{(4)}_\varphi$, $\Gamma^{(3)}_\varphi$ and $G^{(2)}_{\mathrm{c},J}$, which corresponds to the following diagrammatic relation:

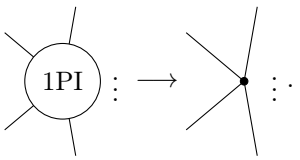


so that $\Gamma^{(4)}_\varphi$ is the 1PI four-point function. The general proof can be found in Weinberg, Vol. 2, Chapter 16.

c) Let's reintroduce $\hbar$ in the definition of generating functional,

$$Z[J] = N \int \mathcal{D}[\phi] \exp\left(\frac{i}{\hbar} S[\phi] + \frac{i}{\hbar} \int \mathrm{d}^4x\ \phi(x) J(x)\right) = e^{iW[J]}.$$

The propagator is the inverse of the quadratic terms in S, hence counts as $\hbar$. Every vertex has a factor of $1/\hbar$. A connected diagram with P propagators and V vertices has a factor of $\hbar^{P-V} = \hbar^{L-1}$ (because $L = P - V + 1$ for connected diagrams). Hence, in the limit $\hbar \to 0$, only tree-level diagrams dominate. The tree diagrams are built from the vertices of S and correspond precisely to the diagram topologies above with replacement



On the right-hand side is the vertex corresponding to the interaction term in S (provided it exists). Hence, $\Gamma = S$. Alternatively, one can perform the functional integral for $\hbar \to 0$ by the stationary phase approximation. Let φ_J be the stationary point, i.e.,

$$\frac{\delta}{\delta\phi(x)} \left(S[\phi] + \int \mathrm{d}^4x\ \phi(x) J(x)\right)\Big|_{\phi=\varphi_J} = 0$$

for a given J. Then

$$e^{iW[J]} \propto e^{(i/\hbar)(S[\phi_J] + \int \mathrm{d}^4x\ \phi_J(x) J(x))}.$$

Writing it differently, we have $W[J] = S[\varphi_J] + \int \mathrm{d}^4x\ \varphi_J(x) J(x)$. Since $\delta S[\varphi]/\delta\varphi(x)|_{\varphi=\varphi_J} = -J(x)$, by comparing with the definition of $\Gamma[\varphi]$, we get $\Gamma[\varphi] = S[\varphi]$ for $\hbar \to 0$.

d) Finally, we will prove the interpretation of V_{eff} as the minimum of the energy density. Recall that the starting point of the path integral formalism is the amplitude $\langle f|e^{-iHT}|i\rangle$ for the transition $\langle i\rangle \to \langle f\rangle$ in the time interval T. A special case is

$$1 = \langle\Omega|e^{iHT}|\Omega\rangle \overset{T\to\infty}{=} N\int \mathcal{D}[\phi]e^{iS[\phi]}.$$

Now, add to H the source term

$$H \to H[J] \equiv H - \int \mathrm{d}^3\mathbf{x}\, J(\mathbf{x})\phi(\mathbf{x})$$

and accordingly

$$S[\phi] \to S_J[\phi] + \int \mathrm{d}^4x\; J(x)\phi(x).$$

Assume the source J is turned on adiabatically, remains constant in the interval T, and is then adiabatically turned off. The system goes from $|\Omega\rangle$ into the ground state $|\Omega_J\rangle$ in the presence of J and back to $|\Omega\rangle$. $|\Omega_J\rangle$ is the ground state of $H[J]$ and has energy E_J. Hence,

$$e^{-iE_JT} = \langle\Omega_J|e^{-iH[J]T}|\Omega_J\rangle = N\int \mathcal{D}[\phi]e^{iS_J[\phi]} = Z[J] = e^{iW[J]}.$$

For this particular source configuration, it holds $W[J] = -E_JT$ and from $H[J]\,|\Omega_J\rangle = E_J\,|\Omega_J\rangle$ one concludes

$$H\,|\Omega_J\rangle = E_J\,|\Omega_J\rangle + \int \mathrm{d}^3\mathbf{x}\, J(\mathbf{x})\phi(\mathbf{x})\,|\Omega_J\rangle\,.$$

We are now seeking the normalized state $|\psi\rangle$ which minimizes the energy expectation value $\langle\psi|H|\psi\rangle$ under the constraint $\langle\psi|\phi(\mathbf{x})|\psi\rangle = \varphi$. Introducing Lagrange multipliers λ_1 and $\lambda_2(\mathbf{x})$, one minimizes

$$\langle\psi|H|\psi\rangle - \lambda_1\,\langle\psi|\psi\rangle - \int \mathrm{d}^3\mathbf{x}\,\lambda_2(\mathbf{x})\,\langle\psi|\phi(\mathbf{x})|\psi\rangle\,.$$

Since $|\Omega_J\rangle$ was the lowest energy state in the presence of J, it must be the solution to the minimization problem above, so $|\psi\rangle = |\Omega_J\rangle$ and $\lambda_1 = E_J$, $\lambda_2(\mathbf{x}) = J(\mathbf{x})$. The minimal energy is

$$E_{\min} = \langle\Omega_J|H|\Omega_J\rangle = E_J + \int \mathrm{d}^3\mathbf{x}\, J(\mathbf{x})\varphi(\mathbf{x}) = \frac{1}{T}\left(-W[J] + \int \mathrm{d}^4x\; J(x)\varphi(x)\right) = -\frac{1}{T}\Gamma[\varphi].$$

Since $J = \lambda_2$ is determined (as a Lagrange multiplier) such that $\langle\Omega_J|\phi(\mathbf{x})|\Omega_J\rangle = \varphi(x)$, this precisely corresponds to the definition of $\Gamma[\varphi]$ as the Legendre transform of $W[J]$. Choosing now $\varphi(\mathbf{x})$ (and therefore $J(\mathbf{x})$) to be spatially constant, we find

$$\mathcal{E}_{\min} = \frac{E_{\min}}{V} = -\frac{1}{VT}\Gamma[\varphi] = V_{\text{eff}}(\varphi).$$

V_{eff} is thus the minimal energy density of the system for given $\varphi = \langle\Omega_J|\phi(x)|\Omega_J\rangle$.

□

Proposition 3.1.1. *If $S[\phi_n]$ is invariant under a non-anomalous linear symmetry, then the effective action $\Gamma[\varphi]$ is also invariant.*

Proof. First, derive the *Ward identity*

$$\int \mathcal{D}[\phi_n] e^{iS[\phi_n]+i\int \mathrm{d}^4x\ J_n(x)\phi_n(x)} \int \mathrm{d}^4x\ J_n(x) F_n^A[\phi_m(x)] = 0$$

for the symmetry $\phi_n(x) \to \phi_n(x) + \varepsilon^A F_n^A[\phi_m; x]$. Write

$$\begin{aligned}
Z[J_n] &= N \int \mathcal{D}[\phi_n] e^{iS[\phi_n]+i\int \mathrm{d}^4x\ J_n(x)\phi_n(x)} \\
&= N \int \mathcal{D}[\phi'_n] e^{iS[\phi'_n]+i\int \mathrm{d}^4x\ J_n(x)\phi'_n(x)} \\
&= N \int \mathcal{D}[\phi_n] \left| \det \frac{\delta \phi'_{n'}}{\delta \phi_n} \right| \exp\left(iS[\phi_n] + i\int \mathrm{d}^4x\ \frac{\delta S}{\delta \phi_n}\delta\phi_n + i\int \mathrm{d}^4x\ J_n(x)(\phi_n(x)+\delta\phi_n(x)) \right) \\
&= N \int \mathcal{D}[\phi_n] e^{iS[\phi_n]+i\int \mathrm{d}^4x\ J_n(x)\phi_n(x)} \left(1 + i\int \mathrm{d}^4x\ J_n(x) F_n^A[\phi_m(x)] + \dots \right).
\end{aligned}$$

In the first equality, we simply renamed the integration variable $\phi \to \phi'$. In the second equality, we changed the variables $\phi'_n = \phi_n + \delta\phi_n$ with $\delta\phi_n = \varepsilon^A F_n^A[\phi_m; x]$. For the last equality, we used the invariance of S under the symmetry transformation and the fact that the Jacobian is trivial for a non-anomalous symmetry. Then we expanded the exponential to first order in ε^A. The first term in the expansion gives back $Z[J_n]$ and the second term must vanish, which gives the Ward identity by independence of ε^A. For a linear symmetry, it holds $F_n^A[\phi_m; x] = iT^A_{nm}\phi_m(x)$, (where T^A is the generator of the symmetry). Hence, the Ward identity implies

$$\begin{aligned}
0 &= \int \mathrm{d}^4x\ J_n(x) iT^A_{nm} N \int \mathcal{D}[\phi_n] e^{iS[\phi_n]+i\int \mathrm{d}^4x\ J_n(x)\phi_n(x)} \phi_m(x) \\
&= \int \mathrm{d}^4x\ J_n(x) iT^A_{nm} \langle \Omega | \phi_m(x) | \Omega \rangle_J \\
&= (-i) \int \mathrm{d}^4x\ \frac{\delta \Gamma[\varphi]}{\delta \varphi_n(x)} T^A_{nm} \varphi_m(x).
\end{aligned}$$

But the right-hand side is precisely the variation of $\Gamma[\varphi]$ under the symmetry transformation

$$\Gamma[\varphi'] = \Gamma[\varphi + \delta\varphi] = \Gamma[\varphi] + \int \mathrm{d}^4x\ \frac{\delta\Gamma[\varphi]}{\delta\varphi_n(x)}\delta\varphi_n(x).$$

Hence, $\Gamma[\varphi] = \Gamma[\varphi']$ is invariant under the symmetry transformation. □

3.1.3 Goldstone bosons

Theorem 3.1.2 (*Goldstone's theorem*)**.** *Assume a theory specified by $\mathcal{L}(\phi_n, \partial_\mu\phi_n)$ is*

a) invariant under Poincaré transformations (Lorentz + translations) and these are not spontaneously broken,

b) invariant under the linear internal symmetry $\delta\phi_n = \varepsilon^A \delta^A \phi_n = i\varepsilon^A T^A_{nm}\phi_m$.

Then there exists a massless particle for every internal symmetry that is spontaneously broken.

To specify the characterization of the broken symmetry, take the internal symmetry group G and let $A = 1, \ldots, \dim G$. By taking linear combinations find the maximal number of linearly independent generators T^i for which $\langle\Omega|\delta^i\phi_n|\Omega\rangle \neq 0$. They obviously generate a subgroup $H \subset G$. So T^i, $i = 1, \ldots, \dim H$ are the generators of the unbroken subgroup H and X^A, $A = 1, \ldots, \dim G - \dim H$ are the generators of the broken symmetries. The generators X^A do not firm a Lie algebra.

We shall prove the Goldstone theorem in two ways.

First proof. The assumption that the theory is invariant under G includes that the symmetry is not anomalous. Hence, $\Gamma[\varphi]$ is invariant as well:

$$\int \mathrm{d}^4x \, \frac{\delta\Gamma[\varphi]}{\delta\varphi(x)} T^A_{nm}\phi_m(x) = 0.$$

By Poincaré invariance, φ_m is constant in the absence of an external source. But for constant φ_m, the functional derivative reduces to an ordinary derivative, so

$$\frac{\delta\Gamma}{\delta\varphi_n} = \frac{\delta}{\delta\varphi_n}\left(-\int \mathrm{d}^4y \, V_{\mathrm{eff}}(\varphi_n)\right) = -\frac{\partial V_{\mathrm{eff}}}{\partial\varphi_n}.$$

First we get

$$\frac{\partial V_{\mathrm{eff}}(\varphi)}{\partial\varphi_n} T^A_{nm}\varphi_m = 0,$$

and after another derivative with respect to φ_k,

$$\frac{\partial^2 V_{\mathrm{eff}}}{\partial\varphi_k\partial\varphi_n} T^A_{nm}\varphi_m + \frac{\partial V_{\mathrm{eff}}}{\partial\varphi_n} T^A_{nk} = 0.$$

Now assume $\overline{\varphi}_n$ is the (constant) field configuration that minimizes V_{eff}, that is, the vacuum expectation value of ϕ_n in the ground state. We consider the case $\overline{\varphi}_n \neq 0$, since otherwise the symmetry would be unbroken. In particular, this means that ϕ_n is a scalar field (Lorentz symmetry unbroken). Then

$$\frac{\partial^2 V_{\mathrm{eff}}}{\partial\varphi_k\partial\varphi_n} T^A_{nm}\varphi_m|_{\varphi=\overline{\varphi}} = \frac{1}{i}\frac{\partial^2 V_{\mathrm{eff}}}{\partial\varphi_k\partial\varphi_n}\delta^A\varphi_n|_{\varphi=\overline{\varphi}} = 0. \tag{3.1}$$

We now show that the second derivative of V_{eff} at $\overline{\varphi}$ is the mass matrix. First, define the Fourier transform of 1PI n-point functions of fields $\phi_{k_1}(x_1), \ldots \phi_{k_n}(x_n)$,

$$\int \mathrm{d}^4x_1 \ldots \mathrm{d}^4x_n e^{ip_1x_1+\cdots+ip_nx_n}\Gamma^{(n)}_{k_1\ldots k_n}(x_1, \ldots, x_n) \equiv (2\pi)^4\delta^{(4)}(p_1 + \cdots + p_n)\tilde{\Gamma}^{(n)}_{k_1\ldots k_n}(p_1, \ldots, p_n).$$

For constant fields,

$$\begin{aligned} -\int \mathrm{d}^4x \, V_{\mathrm{eff}}(\varphi) = \Gamma[\varphi] &= \sum_{n=0}^{\infty}\frac{1}{n!}\int \mathrm{d}^4x_1 \ldots \mathrm{d}^4x_n \, \Gamma^{(n)}_{k_1\ldots k_n}(x_1, \ldots, x_n)\hat{\varphi}_{k_1}(x_1)\ldots\hat{\varphi}_{k_n}(x_n) \\ &= \sum_{n=0}^{\infty}\frac{1}{n!}\hat{\varphi}_{k_1}\ldots\hat{\varphi}_{k_n}(2\pi)^4\delta^{(4)}(0)\tilde{\Gamma}^{(n)}_{k_1\ldots k_n}(p_1 = 0, \ldots, p_n = 0), \end{aligned}$$

hence

$$V_{\mathrm{eff}}(\varphi) = -\sum_{n=0}^{\infty}\frac{1}{n!}\hat{\varphi}_{k_1}\ldots\hat{\varphi}_{k_n}\tilde{\Gamma}^{(n)}_{k_1\ldots k_n}(p_i = 0).$$

We are using summation convention for the indices k_i. The effective potential has been expressed in terms of 1PI n-point functions in momentum space at zero external momentum. In particular,

$$\left.\frac{\partial^2 V_{\text{eff}}}{\partial\varphi_k\partial\varphi_n}\right|_{\varphi=\overline{\varphi}} = -\tilde{\Gamma}^{(2)}_{kn,\overline{\varphi}}(p=0) = (-i)(\tilde{G}^{\text{c}}_{kn}(p=0)|_{J=0})^{-1}.$$

The right-hand side is the inverse of the exact, connected two-point function $\langle\Omega|T(\phi_k\phi_n)|\Omega\rangle$ in momentum space at $p=0$,

$$\tilde{G}^{-1}_{\text{c},kn} = p^2 - m_{kn} - \Pi_{kn}(p) \overset{p\to 0}{=} -[m^2+\Pi(0)]_{kn} \equiv M^2_{kn}.$$

Here, m^2_{kn} is the tree level mass matrix ($\mathcal{L} \supset -m^2_{kn}\phi_k\phi_n/2$) and $\Pi_{kn}(p)$ is the self-energy quantum correction. Hence,

$$M^2_{kn}\delta^A\overline{\varphi}_n = 0.$$

For the $N = \dim G - \dim H$ linearly independent, broken symmetries $\delta^A\overline{\varphi}_n \neq 0$, hence $\delta^A\overline{\varphi}_n$ are N linearly independent eigenvectors of the mass matrix M^2_{kn} to the eigenvalue zero. Since poles in Green functions correspond to one-particle states, with $p^2 = m^2$ corresponding to mass m of the particle, and since φ has to be a scalar field, we conclude that there must be (at least) N massless, spin-0 particles, the *Goldstone bosons*. □

Second proof. Consider the vacuum expectation value of the Noether current $j^\mu(x)$ of the symmetry together with the field $\phi_n(y)$,

$$I = \langle\Omega|j^\mu(x)\phi_n(y)|\Omega\rangle\,.$$

First insert a complete set of intermediate states with p_N being the total momentum of the N-particle state $|N\rangle$,

$$I = \sum_N \int \frac{\mathrm{d}^3\mathbf{p}_N}{(2\pi)^3}\frac{1}{2E_{\mathbf{P}_N}} \langle\Omega|j^\mu(x)|N(p_N)\rangle\,\langle N(p_N)|\phi_n(y)|\Omega\rangle\,.$$

The second bra-ket can be shifted $y\to 0$, to yield $e^{ip_N y}\langle N(p_N)|\phi_n(0)|\Omega\rangle$. Similar happens to the first bra-ket. Convert the integral measure to a Lorentz-invariant one and add another integral over p,

$$I = \int \frac{\mathrm{d}^4p}{(2\pi)^4}e^{-ip(x-y)}\sum_N(2\pi)^4\delta^{(4)}(p-p_N)\cdot$$
$$\cdot\int\frac{\mathrm{d}^4p_N}{(2\pi^4)}(2\pi)\theta(p^0_N)\delta(p^2_N-m^2_N)\,\langle\Omega|j^\mu(0)|N(p_N)\rangle\,\langle N(p_N)|\phi_n(0)|\Omega\rangle\,.$$

By the Lorentz structure, the first bra-ket must be proportional to p^μ_N, and the second one can only depend on $p^2_N = m^2_N$. Now, write the expression in terms of the *spectral density function* $\rho_n(p^2)$,

$$I = i\frac{\partial}{\partial x_\mu}\int\frac{\mathrm{d}^4p}{(2\pi)^4}e^{-ip(x-y)}2\pi\theta(p^0)\rho_n(p^2).$$

This defines the spectral function by comparison with the previous expression. Note that $\rho_n(p^2)\neq 0$ only for $p^2\geq 0$, since $m^2_N\geq 0$. The derivative comes from the p^μ_N factor in the first bra-ket. The preceding expression can be rewritten as

$$I = \int_0^\infty \mathrm{d}s\,\rho_n(s)i\frac{\partial}{\partial x_\mu}\int\frac{\mathrm{d}^4p}{(2\pi)^4}e^{-ip(x-y)}2\pi\theta(p^0)\delta(p^2-s) \equiv \int_0^i nfty\mathrm{d}s\,\rho_n(s)i\frac{\partial}{\partial x_\mu}\Delta_+(x-y;s),$$

where $\Delta_+(x-y;s)$ is the Wightman function of a free scalar field with mass $\sqrt{s}$. If we take the commutator between $j^\mu(x)$ and $\phi_n(y)$, we get instead

$$\langle\Omega|[j^\mu(x),\phi_n(y)]|\Omega\rangle = \int_0^\infty \mathrm{d}s\, \left(\rho_n(s) i\partial_\mu^{(x)}\Delta_+(x-y;s) - \overline{\rho}_n(s) i\partial_\mu^{(x)}\Delta_+(y-x;s)\right).$$

The function $\Delta_+(x-y;s)$ in non-vanishing for space-like distances and is an even function for $(x-y)^2<0$. Hence, $\partial_\mu^{(x)}\Delta_+(x-y)$ is odd and thus $\partial_\mu^{(x)}\Delta_+(x-y) = \partial_\mu^{(x)}\Delta_+(y-x)$. But for space-like distances, causality requires $[j^\mu(x),\phi_n(y)]=0$ and since x and y are arbitrary, it must hold $\rho_n(s)=\overline{\rho}_n(s)$. Now apply $\partial_\mu^{(x)}$ to the Noether current and use the Ward identity $\partial_\mu j^\mu(x)=0$ to get

$$\begin{aligned}
0 &= \langle\Omega|\partial_\mu j^\mu(x),\phi_n(y)|\Omega\rangle \\
&= i\partial_{(x)}^2 \int_0^\infty \mathrm{d}s\, \rho_n(s)(\Delta_+(x-y;s)-\Delta_+(y-x;s)) \\
&= -i\int_0^\infty \mathrm{d}s\, s\rho_n(s)(\Delta_+(x-y;s)-\Delta_+(y-x;s)).
\end{aligned}$$

In the last equality we used that $(\partial^2+s)\Delta_+(x-y;s)=0$. The term in parentheses is non-vanishing for time-like and light-like distances, $(x-y)^2\geq 0$, so the only way to satisfy the equation above is that $s\rho_n(s)=0$ for $s>0$. Hence, $\rho_n(s)=K_n\delta(s)$ for some constant K_n. To determine K_n, we use the equal-time commutator

$$\langle\Omega|[j^0(x),\varphi_n(y)]|\Omega\rangle_{x^0=y^0} = iK_n\partial_0^{(x)}(\delta_+(x-y;0)-\Delta_+(y-x;0))|_{x_0=y^0} = K_n\delta^{(3)}(\mathbf{x}-\mathbf{y}).$$

The delta function was obtained from the definition of $\Delta_+(x-y;s)$,

$$\begin{aligned}
i\partial_0^{(x)}\Delta_+(x-y;0) &= i\partial_0^{(x)}\int\frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3}\frac{1}{2p_0}e^{-ip(x-y)} \\
&= \frac{i}{2p_0}(-ip^0)e^{-ip_0(x_0-y_0)}\int\frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3}e^{-i\mathbf{p}\cdot(\mathbf{x}-\mathbf{y})} \\
&\overset{x^0=y^0}{=} \frac{1}{2}\delta^{(3)}(\mathbf{x}-\mathbf{y}).
\end{aligned}$$

To extract K_n, we integrate the equal-time commutator over $\mathbf{x}$ and use the charge operator Q:

$$K_n = \langle\Omega|[Q,\phi_n(x)]|\Omega\rangle = (-i)\langle\Omega|\delta\phi_n(x)|\Omega\rangle = T_{nm}\langle\Omega|\phi_m(x)|\Omega\rangle\,.$$

If the symmetry is spontaneously broken, then $K_n\neq 0$ and the spectral function has a $\delta(s)$ contribution. For the time-ordered two-point function, this implies

$$\begin{aligned}
\langle\Omega|T(j^\mu(x)\phi_n(0))|\Omega\rangle &= \int_0^\infty \mathrm{d}s\,\rho_n(s)i\partial_\mu^{(x)}\int\frac{\mathrm{d}^3\mathbf{p}}{(2\pi)^3}\frac{1}{2p_0}(\theta(x^0)e^{-ipx}+\theta(-x^0)e^{ipx})|_{p^0=\sqrt{\mathbf{p}^2+s}} \\
&= \int_0^\infty \mathrm{d}s\,\rho_n(s)i\partial_\mu^{(x)}\int\frac{\mathrm{d}^4p}{(2\pi)^4}\frac{ie^{-ipx}}{p^2-s+i\epsilon} \\
&= \int_0^\infty \mathrm{d}s\,\rho_n(s)i\partial_\mu^{(x)}\Delta_{\mathrm{F}}(x;s) \\
&= K_n i\partial_\mu^{(x)}\Delta_{\mathrm{F}}(x;0) = \int\frac{\mathrm{d}^4p}{(2\pi)^4}e^{-ipx}\frac{iK_np_\mu}{p^2+i\epsilon}.
\end{aligned}$$

The existence of a pole at $p^2=0$ implies the existence of a massless particle state $|\pi(\mathbf{p})\rangle$, the Goldstone boson. □

The second proof gives insight on certain matrix elements with Goldstone boson states. From the general analysis of two-point functions, we infer

$$\int \mathrm{d}^4x\; e^{ipx}\,\langle\Omega|T(j^\mu(x)\phi_n(0))|\Omega\rangle \xrightarrow{p^2\to 0} \frac{i}{p^2+i\varepsilon} M_j^\mu(p)\overline{M}_{\phi_n}(p),$$

where

$$\int \mathrm{d}^4x\; e^{iqx}\,\langle\Omega|j^\mu(x)|\pi(\mathbf{p})\rangle = (2\pi)^4\delta^{(4)}(p-q)M_j^\mu(q),$$
$$\int \mathrm{d}^4x\; e^{iqx}\,\langle\pi(\mathbf{p})|\phi_n(0)|\Omega\rangle = (2\pi)^4\delta^{(4)}(p-q)\overline{M}_{\phi_n}(q).$$

This follows from inserting a complete set of states. Only $|\pi(p)\rangle$ contributes to the singular piece for $p^2\to 0$. Since $\phi_n(x)$ is a scalar field, the π-particle must also have spin 0 and the same parity and internal symmetry properties as ϕ_n or j^μ. Otherwise, M_j^μ or $\overline{M}_{\phi_n}$ would be zero and there would be no pole in the Green function. Define

$$\langle\Omega|j^\mu(x)|\pi(\mathbf{p})\rangle = if_\pi e^{-ipx}p^\mu,$$
$$\langle\pi(\mathbf{p})|\phi_n(x)|\Omega\rangle = \sqrt{Z_n}e^{ipx}.$$

The form is completely fixed by Lorentz and translation invariance, the only free parameters are Z_n and f_π. With this, we match $\overline{M}_{\phi_n}(p)=\sqrt{Z_n}$ and $M_j^\mu(p)=if_\pi p^\mu$, as well as

$$K_n = if_\pi\sqrt{Z_n} = T_{nm}\,\langle\Omega|\phi_m(x)|\Omega\rangle\,.$$

If there is more than one broken symmetry, consider the corresponding currents j_μ^A, $A=1,\ldots,\dim G-\dim H$ and generalize as follows: there are now N linearly independent Goldstone bosons π^A and the matrix elements are

$$\langle\Omega|j_\mu^A(x)|\pi^B(\mathbf{p})\rangle = if_\pi^{AB}e^{-ipx}p_\mu,$$
$$\langle\pi^A(\mathbf{p})|\phi_n(x)|\Omega\rangle = \sqrt{Z_n^A}e^{ipx},$$

where f_π^{AB} is the matrix of Goldstone decay constants. The other relations generalize to

$$K_n\to K_n^A,\quad M_j^\mu\to M_j^{AB,\mu},\quad \overline{M}_{\phi_n}\to\overline{M}^B_{\phi_n}$$

and

$$if_\pi^{AB}\sqrt{Z_n^B} = -i\,\langle\Omega|\delta^A\phi_n(x)|\Omega\rangle = T_{nm}^A\,\langle\Omega|\phi_m(x)|\Omega\rangle\,.$$

Next, we define an asymptotic field which creates and annihilates Goldstone bosons π^A with $\langle\pi^A(p)|\pi^B(x)|\Omega\rangle=\delta^{AB}e^{ipx}$. Then for scattering amplitudes of Goldstone bosons

$$\phi_n(x)\xrightarrow{x^0\to\pm\infty}\sqrt{Z_n^A}\pi^A(x),\qquad j_\mu^A(x)\xrightarrow{x^0\to\pm\infty}-f_\pi^{AB}\partial_\mu\pi^B(x).$$

Then, for example, the following holds: $\langle\Omega|\partial_\mu\pi^A(x)|\pi^B(p)\rangle=-ip_\mu\delta^{AB}e^{-ipx}$.

3.1.4 The linear σ-model

This provides a concrete example of the general results derived so far. The *linear σ-model* is a theory of N real scalar fields ϕ_n, $n = 1, \ldots, N$ with Lagrangian

$$\mathcal{L} = \frac{1}{2}\partial_\mu\phi_n\partial^\mu\phi_n - \frac{\lambda}{4}(\phi_n\phi_n - v^2)^2.$$

It is invariant under internal symmetry $\mathrm{O}(N)$, $\phi_n \to e^{i\varepsilon^A T^A_{nm}}\phi_m$. The exponential has to be real, orthogonal matrix, hence T^A is hermitian and purely imaginary. There are $\dim \mathrm{O}(N) = N(N-1)/2$ independent generators. The continuous piece of the symmetry group is $\mathrm{SO}(N)$, with the full group being $\mathrm{O}(N) = \mathrm{SO}(N) \otimes \mathbb{Z}_2$. For small coupling (weak interaction) $\lambda \ll 1$ the effective potential equals the classical potential $V_{\text{eff}}(\varphi) \approx V(\phi) = \lambda(\varphi_n\varphi_n - v^2)^2/4$.

- If $v^2 < 0$, the minimum of $V(\phi)$ lies at $\phi_n = 0$, i.e., $\langle\Omega|\phi_n(x)|\Omega\rangle = 0$. Then also $\langle\Omega|\delta^A\phi_n(x)|\Omega\rangle = 0$ for all A and $\mathrm{O}(N)$ symmetry is unbroken.
- If $v^2 > 0$, then the minima re degenerate with $\varphi_n\varphi_n = v^2$ ($\varphi_n = \langle\Omega|\phi_n(x)|\Omega\rangle$). Without loss of generality, we can parametrize the fields such that

$$\langle\Omega|\phi_N(x)|\Omega\rangle = v, \quad \langle\Omega|\phi_i(x)|\Omega\rangle = 0, \quad i = 1, \ldots, N-1.$$

Hence, the symmetry is broken according to $\mathrm{O}(N) \to \mathrm{O}(N-1)$ where the unbroken subgroup $\mathrm{O}(N-1)$ corresponds to transformations that leave $\langle\Omega|\phi_N(x)|\Omega\rangle$ unchanged, i.e., rotations around the N-axis in the field space. The number of broken generators is $\dim \mathrm{O}(N) - \dim \mathrm{O}(N-1) = N-1$, hence there are $N-1$ Goldstone bosons. The mass matrix (ignoring quantum corrections) is

$$M^2_{nm} = \frac{\partial^2 V(\phi)}{\partial\phi_n\partial\phi_m}\bigg|_{\phi_i=0,\phi_N=v} = \frac{\lambda}{2}(\delta_{nm}(\phi_k\phi_k - v^2) + 4\phi_n\phi_k)\bigg|_{\phi_i=0,\phi_N=v} = 2\lambda v^2 \begin{pmatrix} 0 & & & \\ & \ddots & & \\ & & 0 & \\ & & & 1 \end{pmatrix}.$$

The $N-1$ eigenvalues 0 correspond to $N-1$ Goldstone modes and the eigenvalue $2\lambda v^2$ corresponds to the massive σ-mode. The excitations around the vacuum can be parametrized as

$$\phi_i(x) \equiv \pi_i(x), \quad \phi_N(x) \equiv v + \sigma(x).$$

The resulting Lagrangian is

$$\mathcal{L} = \frac{1}{2}\left(\partial_\mu\sigma\partial^\mu\sigma - 2\lambda v^2\sigma^2 + \partial_\mu\pi_i\partial^\mu\pi_i\right) - \lambda v\sigma(\sigma^2 + \pi_i\pi_i) - \frac{\lambda}{4}(\sigma^2 + \pi_i\pi_i)^2.$$

The existence of $N-1$ massless Goldstone will remain true even after including quantum corrections, but the mass of the σ-particle will be corrected by mass renormalization.

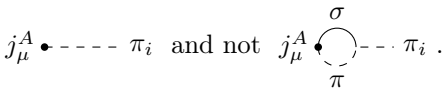
To get the Goldstone decay constant, we first calculate the Noether current of the $\mathrm{O}(N)$ symmetry

$$j^A_\mu = i(\partial^\mu\phi_n)T^A_{nm}\phi_m = i(\partial^\mu\pi_i)T^A_{ij}\pi_j + i(\partial^\mu\sigma)T^A_{Ni}\pi_i + i(\partial^\mu\pi_i)T^A_{iN}(v+\sigma).$$

There is no T^A_{NN} term, because T^A is antisymmetric (traceless). At tree level, we find

$$\langle\Omega|j^A_\mu(0)|\pi_i(p)\rangle = iT^A_{jN}v\,\langle\Omega|\partial_\mu\pi_j|\pi_i(p)\rangle = iT^A_{jN}v(-i)\delta_{ij}p_\mu.$$

At tree level, we only include



The previous calculation shows $f_\pi^{A,i} = (-i)T_{iN}^A v$. The unbroken generators are of the form

$$\begin{pmatrix} T_{ij} & 0 \\ 0 & 0 \end{pmatrix}$$

where T_{ij} are antisymmetric $(N-1)\times(N-1)$ matrices generating $\mathrm{O}(N-1)$. Hence, $T_{iN}^A = 0$ for $i = 1, \ldots, N-1$ when A refers to an unbroken generator (as it should be). The $N-1$ broken symmetry generators can be chosen such that $T_{AN}^A = i$ (no sum) and $T_{BN}^A = 0$ for $B \neq A$,

$$\begin{pmatrix} & & \vdots \\ & 0 & i \\ & & \vdots \\ \ldots & -i \quad \ldots & 0 \end{pmatrix}.$$

Then $f_\pi^{AB} = v\delta^{AB}$ (A and B referring to broken generators), i.e., the Goldstone decay constant is equal to the vacuum expectation value of the field ϕ_N.

Special case $N = 4$

Note that $\mathrm{O}(4) \cong \mathrm{SU}(2) \otimes \mathrm{SU}(2)$ (as Lie algebras). This is analogous to the Lorentz group $\mathrm{SO}(1,3)$. Instead of representing ϕ_n as a four-component vector under $\mathrm{O}(4)$ rotations, we can map them to 2×2 matrices according to

$$\Sigma(x) = \phi_4(x)\mathbb{1} + i\tau^k\phi_k(x).$$

Here, τ^k, $k = 1, 2, 3$ represent *Pauli matrices*. Then

$$\operatorname{tr}\Sigma^\dagger\Sigma = \operatorname{tr}(\sigma_4^2\mathbb{1} + \tau^k\phi_k\tau^l\phi_l) = 2\phi_n\phi_n$$

and

$$\operatorname{tr}(\partial_\mu\Sigma^\dagger\partial^\mu\sigma) = 2\partial_\mu\phi_n\partial^\nu\phi_n.$$

With this, the Lagrangian can be rewritten as

$$\mathcal{L} = \frac{1}{4}\operatorname{tr}(\partial_\mu\Sigma^\dagger\partial^\mu\Sigma) - \frac{\lambda}{16}(\operatorname{tr}(\Sigma^\dagger\Sigma) - 2v^2)^2.$$

The $\mathrm{O}(4)$ symmetry is now represented as $\mathrm{SU}(2)_\mathrm{L} \otimes \mathrm{SU}(2)_\mathrm{R}$ where $\Sigma \to D_\mathrm{R}\Sigma D_\mathrm{L}^\dagger$ with $D_{\mathrm{L,R}} \in \mathrm{SU}(2)_{\mathrm{L,R}}$ two independent $\mathrm{SU}(2)$ transformations. It is clear that $\mathcal{L}$ is invariant under $\mathrm{SU}(2)_\mathrm{L} \otimes \mathrm{SU}(2)_\mathrm{R}$. For $v^2 > 0$, the symmetry is broken since $V(\Sigma)$ is minimized for $\operatorname{tr}(\Sigma^\dagger\Sigma) = 2v^2$. Without loss of generality, we can parametrize the fields such that

$$\langle\Omega|\Sigma(x)|\Omega\rangle = v\mathbb{1},$$

which exactly corresponds to the previous parametrization $\langle\Omega|\phi_4(x)|\Omega\rangle = v$ and $\langle\Omega|\phi_k(x)|\Omega\rangle = 0$. This is invariant only when $D_\mathrm{L} = D_\mathrm{R}$, hence the symmetry is broken to the "diagonal" subgroup $\mathrm{SU}(2)_\mathrm{L} \otimes \mathrm{SU}(2)_\mathrm{R} \to \mathrm{SU}(2)_\mathrm{v}$. In order to make the Goldstone fields explicit, write

$$\Sigma(x) = (v + \sigma(x))e^{i\boldsymbol{\tau}\cdot\boldsymbol{\pi}/v} \equiv (v + \sigma(x))\mathcal{U}(\boldsymbol{\pi}).$$

Note that σ and π_i are not the same as in the previous parametrization. Under $\mathrm{SU}(2)_\mathrm{L}\otimes\mathrm{SU}(2)_\mathrm{R}$, the new fields transform as

$$\sigma \to \sigma \quad \mathcal{U}(\boldsymbol{\pi}) \to D_\mathrm{R}\mathcal{U}(\boldsymbol{\pi})D_\mathrm{L}^\dagger.$$

The transformation of $\mathcal{U}(\boldsymbol{\pi})$ is simple, but the π^i transform in a complicated, non-linear way. In terms of new fields, the terms in the Lagrangian are

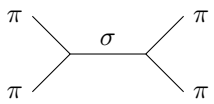
$$\begin{aligned}\mathrm{tr}(\partial_\mu\Sigma^\dagger\partial^\mu\Sigma) &= 2\partial_\mu\sigma\partial^\mu\sigma + (v+\sigma)^2\,\mathrm{tr}(\partial_\mu\mathcal{U}(\boldsymbol{\pi})^\dagger\partial^\mu\mathcal{U}(\boldsymbol{\pi}))\\ \mathrm{tr}(\Sigma^\dagger\Sigma) &= 2(v+\sigma)^2.\end{aligned}$$

Many terms got cancelled, partly because $\mathcal{U}(\boldsymbol{\pi})$ is unitary and partly because $\mathcal{U}^\dagger\partial^\mu\mathcal{U} + \partial^\mu\mathcal{U}^\dagger\mathcal{U} = \partial^\mu(\mathcal{U}^\dagger\mathcal{U}) = 0$. Constructing the Lagrangian in terms of σ and $\mathcal{U}(\boldsymbol{\pi})$ gives

$$\mathcal{L} = \frac{v^2}{2}\,\mathrm{tr}(\partial_\mu\mathcal{U}^\dagger\partial^\mu\mathcal{U})(1+\sigma/v)^2 + \frac{1}{2}(\partial_\mu\sigma\partial^\mu\sigma - 2\lambda v^2\sigma^2) - \lambda v\sigma^3 - \frac{\lambda}{4}\sigma^4.$$

The important difference compared to the Lagrangian with the linear parametrization is that the π^i fields only appear with derivatives. All interactions of Goldstone bosons are proportional to their momentum and become weak at low energies. When $E \ll m_\sigma$ one can integrate out the σ field to obtain an effective Lagrangian. Since σ does not transform under $\mathrm{SU}(2)_\mathrm{L}\otimes\mathrm{SU}(2)_\mathrm{R}$, the effective Lagrangian retains the same symmetry. This is the advantage of the non-linear representation relative to the previous linear representation.

Consider scattering for $E \ll m_\sigma = \sqrt{2\lambda}v$. Then σ-field can only appear as an internal line in Feynman diagrams.



The vertex $\sigma\pi\pi$ has the form

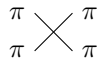
$$\frac{v^2}{4}2\partial_\mu\boldsymbol{\pi}\partial^\mu\boldsymbol{\pi}\frac{1}{v^2}\frac{2\sigma}{v} = \frac{1}{v}\sigma\partial^\mu\boldsymbol{\pi}\partial_\mu\boldsymbol{\pi}.$$

The two-to-two scattering amplitude is



$$\propto (E/v)^2\frac{1}{m_\sigma^2} \sim \frac{E^2}{m_\sigma}^2 \times \frac{E^2}{v^2}.$$

Compare this to



$$\sim \frac{E^2}{v^2}$$

from the $\partial^\mu\pi\partial_\mu pi\,\pi^2/v^2$ term in $\mathrm{tr}(\partial_\mu\mathcal{U}^\dagger\partial^\mu\mathcal{U})$. The leading Goldstone interactions are given by

$$\mathcal{L}^{(0)} = \frac{v^2}{4}\,\mathrm{tr}(\partial_\mu\mathcal{U}^\dagger(\boldsymbol{\pi})\partial^\mu\mathcal{U}(\boldsymbol{\pi}))$$

which remains $\mathrm{SU}(2)_\mathrm{L}\otimes\mathrm{SU}(2)_\mathrm{R}$ invariant. For $E \ll m_\sigma$, the Lagrangian can be systematically improved by matching diagrams with σ-fields, $\mathcal{L}_\mathrm{eff} = \mathcal{L}^{(0)} + \mathcal{L}^{(2)} + \ldots$. It only contains π-fields and is still invariant under $\mathrm{SU}(2)_\mathrm{L}\otimes\mathrm{SU}(2)_\mathrm{R}$. For example, the previously shown process with σ-exchange can be reproduced by a four-π interaction in $\mathcal{L}^{(2)}$ with four derivatives (since it contains E^4),

$$\mathcal{L}^{(2)} \supset \frac{v^2}{8m_\sigma}\,\mathrm{tr}(\partial_\mu\mathcal{U}^\dagger(\boldsymbol{\pi})\partial^\mu\mathcal{U}(\boldsymbol{\pi}))^2.$$

3.1.5 Spontaneous symmetry breaking of approximate symmetries

Consider a theory whose (effective) action is invariant under a linear symmetry (with generators T^A). Assume that the symmetry is explicitly broken by a small perturbation. We consider the effective potential as function of fields φ_n, which correspond to the expectation values of the field operators $\phi_n(x)$, which can either be elementary scalar fields or Lorentz-invariant combinations of fields (e.g. $\overline{\Psi}\Psi$, $A_\mu A^\mu$):

$$V(\varphi) = V_0(\varphi) + V_1(\varphi).$$

All the potentials are effective, but we omit $_{\text{eff}}$ subscript for simplicity. V_0 is invariant under the symmetry, $(\partial V_0/\partial\varphi_n)T^A_{nm}\varphi_m = 0$ and V_1 is a small perturbation that breaks the symmetry. Let φ_0 be the minimum of V_0 and $\varphi_0 + \varphi_1$ of $V_0 + V_1$. Assume φ_1 is small compared to φ_0. Expand the potential around $\varphi = \varphi_0 + \varphi_1$ to the first order

$$\left.\frac{\partial V}{\partial\varphi_n}\right|_{\varphi_0+\varphi_1} = 0, \quad \left.\frac{\partial^2 V_0}{\partial\varphi_n\partial\varphi_n}\right|_{\varphi_0}\varphi_{1m} + \left.\frac{\partial V_1}{\partial\varphi_n}\right|_{\varphi_0} = 0. \tag{3.2}$$

The zeroth order term vanishes because φ_0 is the minimum of V_0. Since V_0 is invariant under the symmetry, the first term vanishes when contracted with the symmetry generator $T^A_{nm}\varphi_{0m}$. So, multiply the second equation with $T^A_{nk}\varphi_k$ to get

$$\frac{\partial V_1}{\partial\varphi_n}T^A_{nk}\varphi_k\big|_{\varphi=\varphi_0} = 0. \tag{3.3}$$

The derivation of the last expression assumed that φ_1 is small, since V_1 is small. However, the minima of $V_0(\varphi)$ are degenerate (must be, since φ transforms under the symmetry). φ_1 is small only for particular choice of φ_0. To understand this, consider the following example:

$$V_0(\phi_1,\phi_2) = (\phi_1^2 + \phi_2^2 - 2a^2)^2, \quad V_1(\phi_1,\phi_2) = \phi_1.$$

The unperturbed potential V_0 is SO(2) symmetric. The family of minima of V_0 is parametrized by

$$\varphi_0 = \mathcal{U}\begin{pmatrix}a\\a\end{pmatrix} = \begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}\begin{pmatrix}a\\a\end{pmatrix}.$$

Then the Eq. (3.3) reads

$$\left.\frac{\partial V_1}{\partial\varphi_n}\right|_{\varphi_0} T^A\varphi_0 = \begin{pmatrix}1 & 0\end{pmatrix}\begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix}\begin{pmatrix}cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}\begin{pmatrix}a\\a\end{pmatrix} = (\cos\theta - \sin\theta)a = 0.$$

This can be satisfied only for $\sin\theta = \cos\theta$, so

$$\varphi_0' = \frac{1}{\sqrt{2}}\begin{pmatrix}1 & 1\\ -1 & 1\end{pmatrix}\begin{pmatrix}a\\a\end{pmatrix} = \begin{pmatrix}\sqrt{2}a\\0\end{pmatrix}.$$

That is, if φ_0' is chosen to point in the direction of the explicit symmetry breaking, then φ_1 is small and the expansion is valid.

Claim 3.1.3. *It is always possible to choose the minimum φ_0 such that Eq. (3.3) is satisfied, namely by choosing φ_0 in the direction of the symmetry breaking term.*

Proof. If φ_0 is a minimum of V_0, then so is $D(g)\varphi_0$ where $D(g) = e^{i\varepsilon^A(g)T^A}$ is an element of the representation of the symmetry group G on the fields φ_n. For internal symmetries, G is a compact manifold (Lie group), hence $D(g)\varphi_0$ defines a compact manifold in the field space, when g runs over all $g \in G$. Because of compactness, the function $V_a(D(g)\varphi_0)$ of G has a minimum at some $g_0 \in G$. For this choice,

$$0 = \frac{\partial V_1(D(g)\varphi_0)}{\partial \varepsilon^A}\bigg|_{\varepsilon^A(g_0)} = \frac{\partial V_1}{\partial \varphi_n}\bigg|_{D(g_0)\varphi_0} \cdot \frac{\partial (D(g)\varphi_0)}{\partial \varepsilon^A}\bigg|_{\varepsilon^A(g_0)} = \frac{\partial V_1}{\partial \varphi_n}\bigg|_{D(g_0)\varphi_0} iT^A_{nm} D(g_0)\varphi_0 = 0.$$

This proves the claim. □

We can now compute the masses of the approximate Goldstone bosons. The connected two-point Green function behaves in the vicinity of the mass pole as

$$G_{\mathrm{c}}^{nm} \overset{p^2 \to m^2}{\sim} \frac{i\sqrt{Z_n^A}\sqrt{Z_m^B}}{p^2 - m_{AB}^2},$$

where Z_n^A is defined via $\langle \pi^A(p)|\phi_n(x)|\Omega\rangle = \sqrt{Z_n^A} e^{ipx}$. The value of the second derivative of the potential at the minimum gives the inverse two-point connected Green function at zero momentum, so we can extract the mass matrix as

$$-iG^{-1}_{\mathrm{c},nm}(p=0) = \frac{\partial^2 V}{\partial\varphi_n \partial\varphi_n}\bigg|_{\varphi_0+\varphi_1}, \quad M^2_{AB} = \sqrt{Z_n^A}\sqrt{Z_m^B}(-i)G^{-1}_{\mathrm{c},nm}(p=0).$$

Expand M^2_{AB} in the small symmetry-breaking terms

$$M^2_{AB} = \sqrt{Z_n^A}\sqrt{Z_m^B}\left(\frac{\partial^2 V_0}{\partial\varphi_n\partial\varphi_m}\bigg|_{\varphi_0} + \frac{\partial^3 V_0}{\partial\varphi_n\partial\varphi_m\partial\varphi_p}\bigg|_{\varphi_0}\varphi_1 + \frac{\partial^2 V_1}{\partial\varphi_n\partial\varphi_m}\bigg|_{\varphi_0} + \dots\right).$$

The first term vanishes for the broken symmetries. The $\sqrt{Z_n^A}$ correspond to the eigenvectors of $\delta^A\overline{\varphi}_n$ with zero eigenvalue, up to f_π^{AB}. We want to express this entirely in terms of V_1 and φ_0. From $(\partial V_0/\partial\varphi_n)T^A_{nm}\varphi_m = 0$ (symmetry of V_0) it follows by differentiating with respect to φ_k and φ_l that

$$\frac{\partial^2 V_0}{\partial\varphi_n\partial\varphi_k}T^A_{nl} + \frac{\partial^2 V_0}{\partial\varphi_n\varphi_l}T^A_{nm} + \frac{\partial^3 V_0}{\partial\varphi_n\partial\varphi_k\partial\varphi_l}T^A_{nm}\varphi_m = 0.$$

Multiply this with $(T^B\varphi_0)_k\varphi_{1l}$ and evaluate the derivatives at $\varphi = \varphi_0$:

$$\frac{\partial^2 V_0}{\partial\varphi_n\partial\varphi_k}\bigg|_{\varphi_0}(T^A\varphi_1)_n(T^B\varphi_0)_k + \frac{\partial^2 V_0}{\partial\varphi_n\partial\varphi_l}\bigg|_{\varphi_0}(T^AT^B\varphi_0)_n\varphi_{1l} + \frac{\partial^3 V_0}{\partial\varphi_n\partial\varphi_k\partial\varphi_l}\bigg|_{\varphi_0}\varphi_{1l}(T^A\varphi_0)_n(T^B\varphi_0)_k = 0.$$

We are interested in the case where T^B belongs to a broken symmetry, so the first term vanishes. According to Eq. (3.2), the second term is $-(\partial V_1/\partial\varphi_m)|_{\varphi_0}(T^AT^B\varphi_0)_n$ and the expression reduces to

$$\frac{\partial^3 V_0}{\partial\varphi_n\partial\varphi_k\partial\varphi_l}\bigg|_{\varphi_0}\varphi_{1l}(T^A\varphi_0)_n(T^B\varphi_0)_k = \frac{\partial V_1}{\partial\varphi_m}\bigg|_{\varphi_0}(T^AT^B\varphi_0).$$

Now, multiply the previous expansion of M^2_{AB} in terms of the derivatives of V_0 and V_1 by $if_\pi^{A'A}if_\pi^{B'B}$ and use $if_\pi^{A'A}\sqrt{Z_n^A} = T^{A'}_{nm}\varphi_{0m}$. Then multiply by the inverse of f_n to get

$$\begin{aligned} M^2_{AB} &= -f^{-1}_{\pi,AA'}f^{-1}_{\pi,BB'}\left(\frac{\partial^3 V_0}{\partial\varphi_n\partial\varphi_k\partial\varphi_l}\bigg|_{\varphi_0}\varphi_{1l}(T^{A'}\varphi_0)_n(T^{B'}\varphi_0)_k + \frac{\partial^2 V_1}{\partial\varphi_n\partial\varphi_n}\bigg|_{\varphi_0}(T^{A'}\varphi_0)_n(T^{B'}\varphi_0)_m\right) \\ &= -(f_\pi^{-1})_{AA'}(f_\pi^{-1})_{BB'}\left(\frac{\partial V_1}{\partial\varphi_n}\bigg|_{\varphi_0}(T^{A'}T^{B'}\varphi_0)_n + \frac{\partial^2 V_1}{\partial\varphi_n\partial\varphi_m}\bigg|_{\varphi_0}(T^{A'}\varphi_0)_n(T^{B'}\varphi_0)_m\right). \end{aligned}$$

This is the general formula for the mass matrix of approximate Goldstone bosons. It can be slightly simplified if we assume linear symmetry breaking, i.e.,

$$\mathcal{L} \supset -\lambda_n \varphi_n \equiv -\mathcal{H}_1.$$

Equivalently, $V_1(\varphi) = \lambda_n \varphi_n$ for $\varphi_n = \langle\Omega|\phi_n(x)|\Omega\rangle$. Note again that φ_n does not have to be an elementary field. We can still consider an effective action $\Gamma[\varphi_n]$ by adding a source term $J_n\phi_n$ to the action where ϕ_n is a composite operator or function of fields that are integrated in the path integral. With

$$\delta\phi_n = i\varepsilon^A T^A_{nm}\phi_m = i\varepsilon^A [Q^A, \phi_n]$$

it follows from Eq. (3.3) that $\lambda_n T^A_{nm}\varphi_{0m} = 0$. Moreover,

$$\langle\Omega_0|[Q^A, \mathcal{H}_1]|\Omega_0\rangle = \lambda_n \langle\Omega_0|[Q^A, \phi_n]|\Omega_0\rangle = \lambda_n T^A_{nm} \langle\Omega_0|\phi_n|\Omega_0\rangle = \lambda_n T^A_{nm}\varphi_{0m} = 0,$$

where Ω_0 is the vacuum of the unperturbed theory. Since for linear symmetry breaking, the second derivative of V_1 vanishes, the mass matrix simplifies to

$$M^2_{AB} = -(f_\pi^{-1})_{AA'}(f_\pi^{-1})_{BB'}(T^{A'}T^{B'}\varphi_0)_n = -(f_\pi^{-1})_{AA'}(f_\pi^{-1})_{BB'} \langle\Omega_0|[Q^{B'}, [Q^{A'}, \mathcal{H}_1]]|\Omega_0\rangle \tag{3.4}$$

which is, in general, non-vanishing, but small (proportional to $\mathcal{H}_1$). In the presence of small explicit symmetry breaking, the Goldstone bosons are no longer massless, but acquire small masses proportional to the symmetry breaking couplings.

3.2 Goldstone boson interactions

We consider the interactions of Goldstone bosons for small energies $E \ll \Lambda$, where Λ is the mass scale of the non-Goldstone modes. The idea is that the Goldstone particles are the lightest particles in the spectrum, so one can write an effective theory for Goldstone modes alone. We want to constrain this Lagrangian as much as possible by the underlying (spontaneously broken) symmetry. This generalizes the discussion of the $O(4)$ σ-model.

3.2.1 Non-linear symmetry representations

A linear symmetry operates as $\Psi \to D(g)\Psi$ for $g \in G$ in some field operators ψ_n collected in a vector Ψ. The ψ_n do not necessarily have to be elementary fields. Assume that the symmetry G is spontaneously broken to a subgroup $H \subset G$ by vacuum expectation values of some G-non-invariant operators (these operators must be Lorentz scalars, because we assume that Lorentz symmetry is unbroken). Since H is unbroken,

$$D(h)_{nm} \langle\Omega|\psi_m|\Omega\rangle = \langle\Omega|\psi_n|\Omega\rangle, \quad \text{or} \quad \langle\Omega|\delta_h\Psi|\Omega\rangle = 0$$

for all $h \in H$. We split the generators of G into unbroken generators T^i $(i = 1, \ldots, \dim H)$ and the remaining broken generators X^A. The following commutation relations hold:

$$[T^i, T^j] = if_{ijk}T^k, \quad [T^i, X^A] = if_{iAB}T^B, \quad [X^A, X^B] = if_{ABi}T^i + if_{ABC}X^C.$$

The first set of relations holds because T^i close an algebra, and the second set holds because the first set implies $f_{ijA} = 0$ and hence by antisymmetry $f_{iAB} = 0$. Assume the standard normalization $\operatorname{tr} T^A T^B = \delta^{AB}/2$ in G. Thus,

$$\operatorname{tr} T^i T^j = \frac{1}{2}\delta^{ij}, \quad \operatorname{tr} T^i X^A = 0, \quad \operatorname{tr} X^A X^B = \frac{1}{2}\delta^{AB}.$$

Since H is unbroken, $T^i_{mn}\langle\Omega|\psi_m|\Omega\rangle = 0$ for all i and ψ_m. There exist $\dim G - \dim H$ Goldstone bosons. Define new fields $\tilde{\psi}_n$, which are "orthogonal" to the Goldstone modes in the sense

$$\psi_n(x) \equiv \sum_m \Sigma_{nm}(x)\tilde{\psi}_m(x), \quad \Sigma(x) \in \{D(g)\,|\,g \in G\}. \tag{3.5}$$

Goldstone modes are parametrized by $\Sigma(x)$, and $\tilde{\psi}_m$ are the remaining fields. This form can always be achieved. Assume ψ_n is a real field and $D(g)$ a real representation of G (without loss of generality, since we can always split a complex field in its real and imaginary part). The Goldstone modes are contained in the linearly independent vectors

$$\delta\psi^A_n = iX^A_{nm}\langle\Omega|\psi_m|\Omega\rangle\,,$$

hence the orthogonality condition for $\tilde{\psi}$ is

$$0 = \tilde{\psi}_n(x)T^A_{nm}\langle\Omega|\psi_m|\Omega\rangle\,. \tag{3.6}$$

This is trivially fulfilled for $A = i$, since $T^i_{nm}\langle\Omega|\psi_m|\Omega\rangle = 0$ for unbroken H. This gives $\dim G - \dim H$ conditions. Hence, $\dim G - \dim H$ degrees of freedom are removed from ψ_n by the definition of $\tilde{\psi}_n$ and those are now parametrized by $\Sigma(x)$. This coincides with the number of Goldstone bosons. The parametrization is not unique, though. We could replace Σ by $\Sigma D(h)$ for any $h \in H$,

$$\psi_n = \sum_m \Sigma_{nm}\tilde{\psi}_m = \sum_m (\Sigma D(h))_{nm}(D(h)^{-1}\tilde{\Psi})_m.$$

So we have to show that if $\tilde{\Psi}$ satisfies the orthogonality condition, then so does $D(h)^{-1}\tilde{\Psi}$. Since $T^i_{nm}\langle\Omega|\psi_m|\Omega\rangle$ is trivially zero, it remains to check

$$(\tilde{\Psi}T^i)_n X^A_{nm}\langle\Omega|\psi_m|\Omega\rangle = (\tilde{\Psi}X^A)_n T^i_{nm}\langle\Omega|\psi_m|\Omega\rangle + if_{iAB}\tilde{\psi}_n X^B_{nm}\langle\Omega|\psi_m|\Omega\rangle = 0.$$

In the first equality, we used $T^iX^A = X^AT^i + if_{iAB}X^B$ and in the second equality we again used the invariance of the vacuum under H for the first part and orthogonality of $\tilde{\Psi}$ for the second part.

It remains to show that there exists a solution of the orthogonality condition such that $\Sigma(x)$ is a representation matrix $D(g)$ for some $g \in G$. Consider $f(g) = \psi_n D(g)_{nm}\langle\Omega|\psi_m|\Omega\rangle$ as a function of g on the compact group manifold. Then $f(g)$ has a maximum at some g_0. Define $\Sigma = D(g_0)$. At g_0 the variation of f with respect to g around g_0 gives

$$0 = \delta f = i\varepsilon^A\psi_n\Sigma_{nm}T^A_{mk}\langle\Omega|\psi_k|\Omega\rangle\,.$$

The orthogonality condition is satisfied for $\tilde{\psi} = (\Psi\Sigma)_m = \Sigma^{\mathrm{T}}_{mn}\psi_n = (\Sigma^{-1})_{mn}\psi_n$ ($D(g)$ is a real unitary representation, so Σ is an orthogonal matrix), i.e., after inverting we get $\psi_n = \Sigma_{nm}\tilde{\psi}_m$ with $\Sigma = D(g_0)$ a representation matrix.

We consider Σ and Σ' as equivalent, if there exists $h \in H$ such that $\Sigma' = \Sigma D(g)$, i.e., the Σ field parametrizes the $\dim G - \dim H$ dimensional quotient space G/H. We can write any $\Sigma(x)$ as

$$\Sigma(x) = e^{i\pi^A(x)X^A}e^{i\theta^i(x)T^i},$$

and we choose $\theta_i = 0$ as the standard representative of the equivalence class. The fields $\pi^A(x)$ are the Goldstone boson fields. We would like to understand how $\tilde{\Psi}$ and Σ transform under G and H:

$$\Psi(x) = \Sigma(\pi(x))\tilde{\Psi}(x) \xrightarrow{g\in G} D(g)\Psi(x) = D(g)\Sigma(\pi(x))\tilde{\Psi}(x) = \Sigma(\pi'(x))D(h)\tilde{\Psi}(x) = \Sigma(\pi'(x))\tilde{\Psi}(x).$$

In the second to last equality we used that $D(g)\Sigma(\pi)$ is an element of $D(g')|g' \in G$, hence there exists an $h \in H$ and an element from G/H such that

$$D(g)\Sigma(\pi) = \Sigma(\pi')D(h) \tag{3.7}$$

where h and π' depend on given g and π. The transformations are thus as follows

$$\Sigma(\pi) \xrightarrow{g} \Sigma(\pi'(g,\pi)) = D(g)\Sigma(\pi)D(h(g,\pi))^{-1},$$
$$\tilde{\Psi} \xrightarrow{g} D(h(g,\pi))\tilde{\Psi}.$$

Since π' and h usually depend in a non-linear way on g, this is called a *non-linear representation* of the symmetry g (as compared to $\Psi \to D(g)\Psi$ which is linear). However, the transformation of the unbroken subgroup H remains linear.

Proof. Start from $[T^i, X^A]?if_{iAB}X^B$. In the notation of group elements it reads

$$D(h)X^A D(h)^{-1} = \mathcal{D}_{AB}(h)X^B$$

where $D(h) = e^{i\varepsilon^i T^i}$ and $\mathcal{D}(h) = e^{f_{iAB}\varepsilon^i}$ a real representation of H. In terms of Goldstone fields, we have $D(h)\Sigma(\pi)D(h)^{-1} = \Sigma(\mathcal{D}(h)\pi)$ and hence

$$\pi^A \xrightarrow{h} \pi'^A = \mathcal{D}_{AB}(h)\pi^B,$$
$$\tilde{\Psi} \xrightarrow{h} \tilde{\Psi}' = D(h)\tilde{\Psi}.$$

Since $D(h)\Sigma(\pi) = (D(h)\Sigma(\pi)D(h)^{-1})D(h)$, $h(h\pi) = h$ solves 3.7. □

Example 2. Consider the O(N) σ-model. In this notation, we have $\psi_n(x) = \phi_n(x)$, $\tilde{\psi}_m = (0,\ldots,0,\sigma(x))^{\mathrm{T}}$ and $\Sigma_{nm}(x)$ is an O(N) rotation:

$$\phi_n(x) = \Sigma_{nm}(x)\begin{pmatrix} 0 \\ \vdots \\ 0 \\ \sigma(x) \end{pmatrix} = \Sigma_{nN}(x)\sigma(x).$$

When we break the symmetry O(N) $\to$ O($N-1$), $D(h)$ is an O($N-1$) matrix of the form

$$\begin{pmatrix} \mathrm{O}(N-1) & 0 \\ 0 & 1 \end{pmatrix}.$$

$\Sigma(nm)$ transforms in the fundamental representation from left and right, i.e., $D(h)\Sigma D(h)^{-1}$ while $\tilde{\psi}$ is a singlet under O($N-1$). Hence, the non-Goldstone mode is also a singlet under the full symmetry group O(N).

3.2.2 Goldstone boson effective Lagrangian

Our aim is to construct a general Lagrangian, which is invariant under the non-linearly represented symmetry. Consider the partial derivative of the field

$$\partial_\mu\Psi(x) = \partial_\mu(\Sigma(x)\tilde{\Psi}(x)) = \Sigma(x)\partial_\mu(\partial_\mu\tilde{\Psi}(c) + \Sigma^{-1}(x)(\partial_\mu\Sigma)(x)\Psi\tilde{(x)})$$
$$\equiv \Sigma(x)(\partial_\mu - iX^A\alpha^A_{\perp\mu}(x) - iT^i\alpha^i_{\parallel\mu}(x))\tilde{\Psi}(x).$$

Since $\Sigma^{-1}(x)\partial_\mu\Sigma(x)$ is an element of the Lie algebra, we wrote it as $-iX^A\alpha^A_{\perp\mu}(x) - iT^i\alpha^i_{\parallel\mu}(x)$, where $\alpha_\perp$ and $\alpha_\parallel$ depend on x through $\pi^A(x)$ in $\Sigma(x)$. For a G invariant $\mathcal{L}(\Psi, \partial_\mu\Psi)$ the overall $\Sigma(x)$ will cancel in the G-invariant product of field, thus the Lagrangian only contains terms with at least one derivative of the Goldstone fields $\pi^A(x)$. Since $D(g)\Sigma(\pi) = \Sigma(\pi')D(h(g,\pi))$, the derivative of this gives

$$D(g)\partial_\mu\Sigma(x) = \partial_\mu(\Sigma(\pi')D(h(g,\pi))).$$

Multiply the left-hand side by $\Sigma^{-1}(\pi)D(g)^{-1}$ and the right-hand side by $D(h)^{-1}\Sigma(\pi)^{-1}$ (both are equal),

$$\begin{aligned}\Sigma(\pi)^{-1}\Sigma(\pi) &= D(h(g,\pi))^{-1}\Sigma(\pi')^{-1}\partial_\mu(\Sigma(\pi')D(h(g,\pi))) \\ &= D(h(g,\pi))^{-1}(\Sigma(\pi')^{-1}\partial_\mu\Sigma(\pi'))D(h(f,\pi)) + D(h(g,\pi))^{-1}\partial_\mu D(h(g,\pi)).\end{aligned}$$

Now we invert the equation, to get

$$\Sigma(\pi')^{-1}\Sigma(\pi') = D(h(g,\pi))(\Sigma(\pi)^{-1}\partial_\mu\Sigma(x))D(h(g,\pi))^{-1} - (\partial_\mu D(h(g,\pi)))D(h(g,\pi))^{-1}.$$

The last term only contains T^i. From this we read how $\alpha_\perp$ and $\alpha_\parallel$ transform:

$$\begin{aligned}X^A\alpha^A_{\perp\mu}(\pi'(x)) &= D(h(g,\pi(x)))X^A\alpha^A_{\perp\mu}D(h(g,\pi(x)))^{-1}, \\ T^i\alpha^i_{\parallel\mu}(\pi'(x)) &= D(h(g,\pi(x)))T^i\alpha^i_{\parallel\mu}D(h(g,\pi(x)))^{-1} + iD(h(g,\pi(x)))\partial_\mu D(h(g,\pi(x)))^{-1}.\end{aligned}$$

This means that $\alpha^A_{\perp\mu}$ transforms covariantly under G, while $\alpha^i_{\parallel\mu}$ has an inhomogeneous transformation similar to a gauge field. $\pi^A(x)$ itself does not have a simple transformation. We can use $\alpha^i_{\parallel\mu}$ to construct a covariant derivative on the non-Goldstone ("matter") fields $\tilde\Psi$

$$\mathrm{D}_\mu \equiv \partial_\mu - iT^i\alpha^i_{\parallel\mu}(x).$$

Then $\mathrm{D}_\mu\tilde\Psi$ transforms under $g \in G$ as

$$\begin{aligned}(\partial_\mu - iD(h)T^i\alpha^i_{\parallel\mu}D(h)^{-1} - i(-i)(\partial_\mu D(h))D(h)^{-1})D(h)\tilde\psi = \\ = D(h)(\partial_\mu - iT^i\alpha^i_{\parallel\mu})\tilde\Psi + (\partial_\mu D(h))\tilde\Psi - (\partial_\mu D(h))\tilde\Psi \\ = D(h)\mathrm{D}_\mu\tilde\psi.\end{aligned}$$

Any Lagrangian, which is invariant under the global linear symmetry H and which consists only of products $\tilde\Psi$, $\mathrm{D}_\mu\tilde\Psi$ and $\alpha^A_{\perp\mu}$ is automatically invariant under the full symmetry G, where the fields π^A and $\tilde\Psi$ (or Σ and $\tilde\Psi$) transform non-linearly under G as specified above. To get the simplest Goldstone boson interaction first consider

$$X^A\alpha^A_{\perp\mu}(\pi') = \mathcal{D}_{AB}(h(g,\pi))\alpha^A_{\perp\mu}(\pi)X^B,$$

where, as we defined earlier, $D(h)X^AD(h)^{-1} = \mathcal{D}_{AB}X^B$. From this it follows $\alpha^A_{\perp\mu}(\pi') = \mathcal{D}_{AB}(h(g,\pi))\alpha^B_{\perp\mu}(\pi)$. Here, $\mathcal{D}$ is real and unitary matrix, hence $\mathcal{D}^{\mathrm{T}} = \mathcal{D}^\dagger = \mathcal{D}^{-1}$. Usually $\mathcal{D}(h)$ is a reducible representation of H, thus we decompose it into reducible subspaces by choosing a suitable basis. The matrix $\mathcal{D}$ becomes block diagonal. The minimal invariant Lagrangian reads

$$\mathcal{L}^{(2)} = \sum_i \frac{1}{2}f_i^2\alpha^{Ai}_{\perp\mu}(\pi(x))\alpha^{Ai\mu}_\perp(\pi(x)).$$

Index i labels the irreducible subspaces. For given i the sum over A runs only over the generators belonging to subspace i. To see the Goldstone interactions, express this in terms of the $\pi^A(x)$. First compute

$$\begin{aligned}\Sigma^{-1}\partial_\mu\Sigma &= e^{-i\pi^A X^A}\partial_\mu e^{i\pi^A X^A}\\ &= \left(1 - i\pi^A X^A + \dots\right)\partial_\mu\left(1 + i\pi^A X^A - \frac{1}{2}\pi^A X^A \pi^B X^B + \dots\right)\\ &= iX^A\partial_\mu\pi^A + \pi^A X^A\partial_\mu\pi^B X^B - \frac{1}{2}(\partial_\mu\pi^A)X^A\pi^B X^B - \frac{1}{2}\pi^A X^A(\partial_\mu\pi^B)X^B + \dots\\ &= -i\left(-X^A\partial_\mu\pi^A + \frac{i}{2}[\pi^A X^A, \partial_\mu\pi^B X^B] + \dots\right)\end{aligned}$$

The commutator contributes to $\alpha_\perp$ and $\alpha_\parallel$, so we pick the X^C piece of the commutator for $\alpha_\perp$,

$$\alpha^A_{\perp\mu} = -\partial_\mu\pi^A + \#f^{ABC}\pi^B\partial_\mu\pi^C.$$

The minimal Lagrangian then reads

$$\mathcal{L}^{(2)} = \sum_i \frac{1}{2}f_i^2\partial_\mu\pi^{Ai}\partial^\mu\pi^{Ai} + \text{interactions}.$$

To canonically normalize the kinetic term, relabel $\pi^{Ai} \to \pi^{Ai}/f_i$. Thus, the kinetic term only has a factor of $1/2$ and interaction terms come in powers of $1/f_i$. When $E \ll M$ (the mass of the non-Goldstone fields), only Goldstone particles can appear in a scattering process. One can then construct an effective Lagrangian for low-energy processes of Goldstone particles in an expansion in E/M. At what order E^P does a given diagram contribute? Generalize the discussion to the case where there are non-Goldstone fields (which of course cannot be produced or destroyed – like electrons in quantum mechanics interacting with long-wavelength photons with $E_\gamma \ll m_e$). For a given diagram we have

$$P = \sum_i d_i V_i + 4L - 2I_\pi - I_M$$

where

- P are the powers of E,
- V_i is the number of vertices of type i which contain d_i derivatives,
- L is the number of loops,
- I_π is the number of internal Goldstone lines with propagator $1/q^2$, and
- I_M us the number of internal non-Goldstone lines. Their propagator counts only as $1/q$, as

$$\frac{1}{(p_M+q)^2 - M^2} \approx \frac{1}{2p_M\cdot q}, \quad \frac{\not{p}_M + \not{q} + M}{(p_M+q)^2 - M^2} \approx \frac{M}{2p_M\cdot q}.$$

For connected diagrams, there are further relations,

$$L = I_\pi + I_M - \sum_i V_i + 1, \quad 2I_M + E_M = \sum_i V_i n_i,$$

where n_i is the number of non-Goldstone lines at vertex of type i. With these relations, we can eliminate I_π and I_M to get

$$P = \sum_i V_i(d_i + n_i/2 - 2) + 2L - E_M/2 + 2.$$

The largest contribution arises from the smallest P, because the expansion is in E/M. In the case of Goldstone scattering (no matter fields, $E_M = 0$, $n_i = 0$), the expression simplifies even further

$$P = \sum_i V_i(d_i - 2) + 2L + 2.$$

This is definitely positive, because each vertex contains at least two derivatives. The dominant contribution arises from tree diagrams and the terms from $\mathcal{L}^{(2)}$ since these are exactly the terms with two derivatives.

If we now add matter fields, it is possible to have $d_i = 1$, but then $n_i \geq 2$ (since matter fields cannot be destroyed or created). The leading terms are $n_i = 2$ and $d_i = 1$. The effective Lagrangian is then

$$\mathcal{L}_M = \overline{N}(i\gamma^\mu \mathrm{D}_\mu - M)N - \sum_i g_i \overline{N}\gamma^\mu X^{Ai} \alpha_{\perp\mu}^{Ai} N,$$

where N denotes the matter field (previously Ψ) and $\mathrm{D}_\mu = \partial_\mu - iT^i\alpha_\parallel^i(x)$. This is the most general H invariant Lagrangian and therefore also G invariant ($\alpha_{\perp,\parallel}$ contain a single derivative on π^A). By adding all possible terms with more derivatives one can improve the approximation systematically in E/M. The theory is in general non-renormalizable, but this is irrelevant, since from the beginning it is constructed as an effective theory for $E \ll M$ and contains all terms allowed by the symmetry. The only assumption is that there is a mass gap between the Goldstone bosons and the non-Goldstone particles, but this is guaranteed by the Goldstone theorem when the symmetry is spontaneously broken,

3.3 Spontaneously broken chiral symmetry of QCD

Quantum chromodynamics (QCD) is an SU(3) non-Abelian gauge theory with Dirac fermions (*quarks*) in the fundamental representation. Internal degree of freedom, on which SU(3) operates is called *colour*. In the context of the Standard model it is also referred to as the *strong interaction*. The fundamental fields of the theory are the *gluons* A_μ^A ($A = 1, \ldots, 8$) in adjoint representation and quarks $q_{\alpha a f}$, where

- index α denotes the Dirac spinor, $\alpha = 1, \ldots, 4$
- $a = 1, 2, 3$ is the colour index for the 3-dimensional fundamental representation, and
- f is the flavour index, denoting "up", "down", "strange", "charm", "bottom" and "top".

All flavours have exactly the same gauge interaction, only their mass is different. The Lagrangian is

$$\mathcal{L} = -\frac{1}{4} G_{\mu\nu}^A G^{A\mu\nu} + \sum_f \overline{q}_f (i\not{D} - m_f) q_f + \mathcal{L}_{\text{gauge-fix}} + \mathcal{L}_{\text{ghost}}.$$

The quark mass can be chosen diagonal in "flavour index" by field rotation.

The gauge interaction is scale-dependent and becomes non-perturbatively strong at a scale $Q \lesssim 1\,\text{GeV}$. Quarks form bound states called *hadrons*. *Confinement* means that all observed

bound states are SU(3) (colour-) singlets. That is, the colour charge is permanently confined in neutral, composite objects. There are no asymptotic, or free, quark or gluon states. That is on empirical fact, which has not yet been proven analytically. There are two types of particles.

- *Mesons* $q_a\overline{q}_a$, singlets in $\mathbf{3}\otimes\overline{\mathbf{3}}=\mathbf{1}\oplus\mathbf{8}$. The symbol $\mathbf{3}$ represents fundamental and $\overline{\mathbf{3}}$ the anti-fundamental representation.
- *Baryons* $\varepsilon_{abc}q_aq_bq_c$, singlets in $\mathbf{3}\otimes\mathbf{3}\otimes\mathbf{3}=\mathbf{1}\otimes\mathbf{6}\otimes\mathbf{8}$.

Proton is made of uud and held together by gluons. Its mass is $m_p=938\,\mathrm{MeV}$, which is much smaller than the sum of its quarks, $2m_u+m_d\approx 10\,\mathrm{MeV}$ (defined as $\overline{\mathrm{MS}}$-renormalized at scale $\mu=2\,\mathrm{GeV}$). Most of the proton's mass is binding energy. This is only possible for strong interactions.

3.3.1 Approximate global symmetries of QCD

We are interested in the low-energy regime $Q\lesssim 1\,\mathrm{GeV}$.

f	u	d	s	c	b	t
m_f	2.5 MeV	5 MeV	90 MeV	1.3 GeV	4.2 GeV	175 GeV

The lightest three quarks (called *light quarks*) have mass below QCD scale, the others (*heavy quarks*) have the mass above QCD scale. In this section, we will ignore the heavy quarks,

$$\mathcal{L}_q=\sum_{f=u,d,s}(\overline{q}_{\mathrm{L}f}i\not{D}q_{\mathrm{L}f}+\overline{q}_{\mathrm{R}f}i\not{D}q_{\mathrm{R}f}-m_f[\overline{q}_{\mathrm{L}f}q_{\mathrm{R}f}+\overline{q}_{\mathrm{R}f}q_{\mathrm{L}f}]).$$

In the approximation $m_f=0$ we can rotate $q_{\mathrm{L}f}$ and $q_{\mathrm{R}f}$ separately by unitary matrices in flavour space, $q_{\mathrm{L}f}\to\mathcal{U}_{\mathrm{L}fg}q_{\mathrm{L}g}$, $q_{\mathrm{R}f}\to\mathcal{U}_{\mathrm{R}fg}q_{\mathrm{R}g}$ for $\mathcal{U}_{\mathrm{L}},\mathcal{U}_{\mathrm{R}}\in\mathrm{U}(3)$. The global flavour symmetry can be decomposed as follows

$$\mathrm{U}(3)_{\mathrm{L}}\otimes\mathrm{U}(3)_{\mathrm{R}}=\mathrm{U}(1)_{\mathrm{B}}\otimes\mathrm{U}(1)_{\mathrm{A}}\otimes\mathrm{SU}(3)_{\mathrm{V}}\otimes\mathrm{SU}(3)_{\mathrm{A}}.$$

The transformations within individual subgroups are

$$\mathrm{U}(1)_{\mathrm{B}}:q_f\to e^{i\varepsilon}q_f,\quad \mathrm{U}(1)_{\mathrm{A}}:q_f\to e^{i\varepsilon\gamma_5}q_f,$$

$$\mathrm{SU}(3)_{\mathrm{V}}:q_f\to(e^{i\varepsilon^AT^A})_{fg}q_g,\quad \mathrm{SU}(3)_{\mathrm{A}}:q_f\to(e^{i\varepsilon^AT^A\gamma_5})_{fg}q_g.$$

Here we choose to replace the L and R rotations by vectorial and axial one, which is simply a different way of choosing the independent transformations.

- Vectorial transformations rotate L and R in the same way.
- Axial transformations rotate L and R in the opposite way. This follows from

$$\gamma_5=\begin{pmatrix}-\mathbb{1}&0\\0&\mathbb{1}\end{pmatrix},\quad q_f=\begin{pmatrix}q_{\mathrm{L}f}\\q_{\mathrm{R}f}\end{pmatrix}\to\begin{pmatrix}e^{-i\varepsilon^AT^A}&0\\0&e^{i\varepsilon^AT^A}\end{pmatrix}\begin{pmatrix}q_{\mathrm{L}f}\\q_{\mathrm{R}f}\end{pmatrix}=\begin{pmatrix}e^{-i\varepsilon^AT^A}q_{\mathrm{L}f}\\e^{i\varepsilon^AT^A}q_{\mathrm{R}f}\end{pmatrix}.$$

Now consider the various transformations.

- $\mathrm{U}(1)_{\mathrm{B}}$: the symmetry is present even in the presence of the quark mass term $m_f\overline{q}_fq_f$. It is an exact global symmetry. Noether charge is baryon or quark number (antiquark has quark number -1). Quarks are only produced as quark-antiquark pairs. It guarantees that the lightest baryon (proton) is stable.

- $\mathrm{U}(1)_\mathrm{A}$: this symmetry is anomalous, i.e., not a symmetry of the quantum theory, $\delta_{\mathrm{U}(1)_\mathrm{a}} S = 0$, $\delta_{\mathrm{U}(1)_\mathrm{A}} \Gamma \neq 0$.
- $\mathrm{SU}(3)_\mathrm{V}$: the symmetry holds even for massive quarks, provided all masses are equal. Then $m(\overline{q}_{\mathrm{L}f} q_{\mathrm{R}f} + \overline{q}_{\mathrm{R}f} q_{\mathrm{L}f})$ is invariant under SU(3) rotations of $q_{\mathrm{L}f}, q_{\mathrm{R}f}$ because L, R are rotated in the same way. This is an approximate global symmetry, i.e., there is small explicit breaking by quark mass differences. The $\mathrm{SU}(2)_\mathrm{V}$ subgroup operating only on u and d quarks is less strongly broken, since $m_u, m_d \ll m_s$. This is the SU(2) isospin symmetry of the strong interaction. It is also the $\mathrm{SU}(3)_\mathrm{V}$ which leads to the classification of hadrons in terms of (approximate) SU(3) representations (Gell-Mann, 1965, postulate of three quarks).
- $\mathrm{SU}(3)_\mathrm{A}$: this symmetry is broken by $m_f \neq 0$ even for equal quark masses. However, contrary to $\mathrm{SU}(3)_\mathrm{V}$, one does not observe this symmetry in the hadron spectrum – one does not observe $\mathrm{SU}(3)_\mathrm{A}$ multiplets which would imply a partner with opposite parity for every hadron.

Note. The Noether current of $\mathrm{SU}(3)_\mathrm{V}$ is $\overline{q}_f \gamma^\mu T^A_{fg} q_g$, for $\mathrm{SU}(3)_\mathrm{A}$ it is $\overline{q}_f \gamma^\mu \gamma_5 T^A_{fg} q_g$ so the $\mathrm{SU}(3)_\mathrm{A}$ Noether charge Q^A_A has negative parity. Hence, if $|p\rangle$ is a proton state, the state $Q^A_\mathrm{A}|p\rangle$ is also an eigenstate of the Hamiltonian, but has opposite parity,

$$\mathcal{U}_P(Q^A_\mathrm{A}|p\rangle) = \mathcal{U}_P Q^A_\mathrm{A} \mathcal{U}_P^{-1} \mathcal{U}_P |p\rangle = -Q^A_\mathrm{A}|p\rangle\,.$$

Such states are not observed, which means that $\mathrm{SU}(3)_\mathrm{A}$ is spontaneously broken through non-perturbative effects. This means $\delta_{\mathrm{SU}(3)_\mathrm{A}} \Gamma = 0$, but the vacuum is not invariant. In reality, it is spontaneously broken, approximate symmetry, since $m_f \neq 0$ but the explicit breaking is small, especially for the $\mathrm{SU}(2)_\mathrm{A}$ subgroup operating on q_u and q_d. We therefore expect 8 Goldstone bosons. Due to explicit breaking of the symmetry, they are not exactly massless.

3.3.2 Masses of the QCD Goldstone bosons

The chiral symmetry is spontaneously broken to the diagonal subgroup

$$\mathrm{SU}(3)_\mathrm{L} \otimes \mathrm{SU}(3)_\mathrm{R} \to \mathrm{SU}(3)_\mathrm{V}.$$

For the Noether charges we have the same commutation relations as for the Lie algebra generators,

$$\begin{aligned}
[L^A, L^B] &= i f^{ABC} L^C, \quad \mathrm{SU}(3)_\mathrm{L},\\
[R^A, R^B] &= i f^{ABC} R^C, \quad \mathrm{SU}(3)_\mathrm{R},\\
[L^A, R^B] &= 0, \quad \text{since tensor product.}
\end{aligned}$$

If we decompose the tensor product as $\mathrm{SU}(3)_\mathrm{L} \otimes \mathrm{SU}(3)_\mathrm{R} = \mathrm{SU}(3)_\mathrm{V} \otimes \mathrm{SU}(3)_\mathrm{A}$, the vectorial and axial charges are $Q^A_\mathrm{V} = R^A + L^A$ and $Q^A_\mathrm{A} = R^A - L^A$. Hence,

$$[Q^A_\mathrm{V}, Q^B_\mathrm{A}] = i f^{ABC} Q^C_\mathrm{A}.$$

Thus, it also holds

$$[Q^A_\mathrm{V}, j^{\mu B}_\mathrm{A}(x)] = i f^{ABC} j^{\mu B}_\mathrm{A}(x)$$

for the $\mathrm{SU}(3)_\mathrm{A}$ Noether current. The $\mathrm{SU}(3)_\mathrm{V}$ variation of this current is

$$\delta^A_\mathrm{V} j^{\mu B}_\mathrm{A}(x) = i[Q^A_\mathrm{V}, j^{\mu B}_\mathrm{A}(x)] = -f^{ABC} j^{\mu C}_\mathrm{A}(x) = i(T^A_\mathrm{ad})_{BC} j^{\mu C}_\mathrm{A}(x).$$

The axial current transforms in the adjoint (octet) representation of $\mathrm{SU}(3)_\mathrm{V}$. Hence, also the Goldstone bosons form an octet under $\mathrm{SU}(3)_\mathrm{V}$ (up to the small explicit breaking of $\mathrm{SU}(3)_\mathrm{V}$). Goldstone bosons have equal masses and decay constants (at least for $m_u = m_d = m_s$), so choose basis such that $f_\pi^{AB} = f_\pi \delta^{AB}$. For simplicity, consider in the following the $\mathrm{SU}(2)_\mathrm{L} \otimes \mathrm{SU}(2)_\mathrm{R}$ chiral symmetry for up and down quarks, neglecting the strange quark. Then there is a triplet of (approximate) Goldstone bosons. They can be identified with the three *pions* $\pi^\pm, \pi^0$ with masses $m_\pi \approx 140\,\mathrm{MeV}$, which is smaller than any other hadron made out of u and d quarks ($m_\rho = 770\,\mathrm{MeV}$). To compute the masses, use the result Eq. (3.4). The perturbation Hamiltonian is

$$\mathcal{H}_1 = m_u \bar{q}_u q_u + m_d \bar{q}_d q_d = M_{ij} (q\bar{q})_{ji}, \quad M = \begin{pmatrix} m_u & 0 \\ 0 & m_d \end{pmatrix}.$$

This is linear in the composite, scalar field $\phi_n = (q\bar{q})_{ij}$ where n corresponds to ij. The mass matrix of the Goldstone bosons reads

$$M^2_{AB} = -\frac{1}{f_\pi^2} \langle \Omega_0 | [Q^B_\mathrm{A}, [Q^A_\mathrm{A}, \bar{q} M q]] | \Omega_0 \rangle = -\frac{1}{f_\pi^2} M_n (T^A_{nm} T^B_{mk} \langle \Omega_0 | (q\bar{q})_k | \Omega_0 \rangle),$$

where T^A_{nm} is the generator of $\mathrm{SU}(2)_\mathrm{A}$ in the representation of the composite field ϕ_n. To get the commutator $[Q^A_\mathrm{A}, \bar{q} M q]$ we first compute the $\mathrm{SU}(2)_\mathrm{A}$ transformation of the $\bar{q} M q$:

$$\bar{q} M q \xrightarrow{\mathrm{SU}(2)_\mathrm{A}} \bar{q} e^{i\varepsilon^A T^A \gamma_5} M e^{i\varepsilon^A T^A \gamma_5} q = \bar{q} M q + i\varepsilon^A \bar{q} \{\gamma_5 T^A, M\} q.$$

Note the + sign in the exponential on both sides of M—this is because $e^{-i\varepsilon^A T^A \gamma_5} \gamma_0 = \gamma_0 e^{i\varepsilon^A T^A \gamma_5}$. From this we conclude

$$[Q^A_\mathrm{A}, \bar{q} M q] = \frac{1}{i} \delta^A_{\mathrm{SU}(2)_\mathrm{A}} (\bar{q} M q) = \bar{q} \{T^A, M\} \gamma_5 q.$$

Applying this twice gives

$$M^2_{AB} = -\frac{1}{f_\pi^2} \langle \Omega_0 | \bar{q} \{T^B, \{T^A, M\}\} | \Omega_0 \rangle = -\frac{1}{f_\pi^2} \langle \Omega_0 | \bar{u} u | \Omega_0 \rangle \operatorname{tr}(\{T^B, \{T^A, M\}\}) = M_\pi^2 \delta^{AB}.$$

We used that $|\Omega_0\rangle$ is $\mathrm{SU}(2)_\mathrm{V}$ invariant, hence

$$\langle \Omega_0 | \bar{q}_i q_j | \Omega_0 \rangle = \delta_{ij} \langle \Omega_0 | \bar{u} u | \Omega_0 \rangle = \delta_{ij} \langle \Omega_0 | \bar{d} d | \Omega_0 \rangle .$$

Inverting the relation gives

$$M_\pi^2 = -\frac{\langle \Omega_0 | \bar{u} u | \Omega_0 \rangle}{3 f_\pi^2} \operatorname{tr}(\{T^A, \{T^A, M\}\}) = -\frac{\langle \Omega_0 | \bar{u} u | \Omega_0 \rangle}{3 f_\pi^2} \operatorname{tr}(4 T^A T^A M) = -\frac{\langle \Omega_0 | \bar{u} u | \Omega_0 \rangle}{3 f_\pi^2} \operatorname{tr}(3M).$$

We used $T^A = \sigma^A/2$ in the fundamental representation. Now use $\operatorname{tr} M = m_u + m_d$, and we obtain the *Gell-Mann–Oakes–Renner relation*

$$M_\pi^2 = -\frac{m_u + m_d}{f_\pi^2} \langle \Omega_0 | \bar{u} u | \Omega_0 \rangle .$$

It establishes a relation between Goldstone properties (f_π, M_π) and the quark-antiquark vacuum expectation value (condensate). Note that $\bar{u} u = \bar{q}_u q_u$ is not $\mathrm{SU}(3)_\mathrm{A}$ invariant and is therefore an order parameter of the spontaneously broken chiral symmetry. Note also that $M_\pi^2 \to 0$ for $m_u, m_d \to 0$, as it should be.

Remark. The above relation does not demonstrate that $\langle\Omega_0|\overline{u}u|\Omega_0\rangle \neq 0$. If that were the case, M_π^2 would be determined by second-order terms in the symmetry breaking. However, numerical, non-perturbative studies of QCD give a non-vanishing quark condensate.

Let's now generalize to $\mathrm{SU}(3)_\mathrm{L}\otimes\mathrm{SU}(3)_\mathrm{R}$ with $m_u, m_d, m_s \neq 0$. The octet of Goldstone bosons consists of $\pi^\pm$, π^0, $K^\pm$, K^0, $\overline{K}^0$ and η. The charged *kaon* masses are

$$M_{K^\pm}^2 = -\frac{m_u + m_s}{f_\pi^2}\,\langle\Omega_0|\overline{u}u|\Omega_0\rangle$$

and for the neutral kaons

$$M_{K^0}^2 = -\frac{m_d + m_s}{f_\pi^2}\,\langle\Omega_0|\overline{u}u|\Omega_0\rangle\,.$$

From this, we can calculate the ratio of the quark masses (denote $\overline{m}_q = (m_u + m_d)/2$)

$$\frac{m_s + \overline{m}_q}{2\overline{m}_q} = \frac{M_{K^\pm}^2 + M_{K^0}^2}{2M_{\pi^\pm}^2} = \frac{(493.68\,\mathrm{MeV})^2 + (497.67\,\mathrm{MeV})^2}{2(139.57\,\mathrm{MeV})^2} = 12.61.$$

From this we determine $m_s/\overline{m}_q = 24.2$.

3.3.3 The effective chiral Lagrangian

Now, we will apply the general theory from Section 3.2.2 to chiral symmetry in QCD to describe pion interactions at low energy. The full symmetry group is $G = \mathrm{SU}(3)_\mathrm{L} \otimes \mathrm{SU}(3)_\mathrm{R}$ and the unbroken subgroup is $H = \mathrm{SU}(3)_\mathrm{V}$. The coset/quotient space $G/H = \mathrm{SU}(3)_\mathrm{A}$ is parametrized as $\Sigma(x) = e^{i\gamma_5\pi^A(x)T^A}$. Then

$$D(g)\Sigma(\pi) = \Sigma(\pi')D(h(g,\pi))$$
$$e^{i(\varepsilon_\mathrm{V}^A T^A + \varepsilon_\mathrm{A}^A T^A\gamma_5)}e^{i\gamma_5\pi^A(x)T^A} = e^{i\gamma_5\pi'^A(x)T'^A}e^{i\theta^i(\varepsilon_\mathrm{V},\varepsilon_\mathrm{A},\pi^A(x))T^i}.$$

Decompose this to $P_{\mathrm{L/R}} = (\mathbb{1} \mp \gamma_5)/2$ ($\mathbb{1} = P_\mathrm{L} + P_\mathrm{R}$, $\gamma_5 = P_\mathrm{R} - P_\mathrm{L}$)

$$e^{iT^A((\varepsilon_\mathrm{V}^A - \varepsilon_\mathrm{A}^A)P_\mathrm{L} + (\varepsilon_\mathrm{V}^A + \varepsilon_\mathrm{A}^A)P_\mathrm{R})}e^{i\pi^A T^A(P_\mathrm{R} - P_\mathrm{L})} = e^{i\pi'^A T'^A(P_\mathrm{R} - P_\mathrm{R})}e^{i\theta^A T^A(P_\mathrm{R} + P_\mathrm{L})}.$$

Use $\varepsilon_\mathrm{L}^A = \varepsilon_\mathrm{V}^A - \varepsilon_\mathrm{A}^A$, $\varepsilon_\mathrm{R}^A = \varepsilon_\mathrm{V}^A + \varepsilon_\mathrm{A}^A$ and some properties of the projectors ($P_{\mathrm{L/R}}^n = P_{\mathrm{L/R}}$, $P_\mathrm{R}P_\mathrm{L} = P_\mathrm{L}P_\mathrm{L} = 0$) to first show a general relation

$$e^{iT^A(a_1P_\mathrm{L} + a_2P_\mathrm{R})} = e^{iT^Aa_1}P_\mathrm{L} + e^{iT^Aa_2}P_\mathrm{R}$$

from which we can then derive

$$e^{i\varepsilon_\mathrm{L}T^A}e^{-i\pi^AT^A} = e^{-i\pi'^AT'^A}e^{i\theta^iT^i},$$
$$e^{i\varepsilon_\mathrm{R}T^A}e^{i\pi^AT^A} = e^{i\pi'^AT'^A}e^{i\theta^iT^i}.$$

At this point, one departs from the general formalism to exploit the particular structure of symmetry breaking of a tensor product to the diagonal subgroup. In this case, instead of $\Sigma(x) = e^{i\gamma_5\pi^A(x)T^A}$, we can choose a parametrization of the Goldstone fields in terms of

$$\mathcal{U}(x) \equiv e^{2i\pi^A(x)T^A}.$$

The advantage is that contrary to the transformation of $\Sigma(x)$, which contains the complicated expression $\theta^A(\varepsilon_\mathrm{L}, \varepsilon_\mathrm{R}, \pi^A(x))$, $\mathcal{U}(\pi(x))$ has the simple transformation

$$e^{i\varepsilon_\mathrm{R}^A T^A}\mathcal{U}(x)e^{-i\varepsilon_\mathrm{L}^A T^A} = \mathcal{U}(\pi'), \qquad D_\mathrm{R}(g)\mathcal{U}(\pi)D(g)^{-1} = \mathcal{U}(\pi')$$

for $g \in G = \mathrm{SU}(3)_\mathrm{L} \otimes \mathrm{SU}(3)_\mathrm{R}$. Object $\mathcal{U}(\pi)$ transforms *linearly* in the $(\mathbf{3}, \overline{\mathbf{3}})$ representation of $\mathrm{SU}(3) \otimes \mathrm{SU}(3)$. The field $\pi^A(x)$ itself still transforms non-linearly. Note the similar construction for the $\mathrm{O}(4) \cong \mathrm{SU}(2) \otimes \mathrm{SU}(2)$ σ-model.

We will now change the notation $\pi \to \phi/f$ (with f of mass dimension 1) so that the field will have canonical mass dimension, $\mathcal{U}(\phi) = e^{2i\phi^A T^A/f}$. Using the Gell-Mann matrices for the generators of the fundamental representation of SU(3), we can write

$$\sqrt{2}\phi^A T^A = \begin{pmatrix} \frac{1}{\sqrt{2}}\pi^0 + \frac{1}{\sqrt{6}}\eta_8 & \pi^+ & K^+ \\ \pi^- & -\frac{1}{\sqrt{2}}\pi^0 + \frac{1}{\sqrt{6}}\eta_8 & K^0 \\ K^- & \overline{K}^0 & -\frac{2}{\sqrt{6}}\eta_8 \end{pmatrix}.$$

The naming will become clear later. It is evident that $(K^0)^\dagger = \overline{K}^0$, $(\pi^+)^\dagger = \pi^-$ and $(\pi^0)^\dagger = \pi^0$. To construct the leading Goldstone Lagrangian $\mathcal{L}^{(2)}$ we need $\alpha^A_{\perp\mu}$. However, since $\mathcal{U}(x)$ transforms linearly and is unitary, one can write the most general Lagrangian directly

$$\mathcal{L}^{(2)} = \frac{f^2}{2}\,\mathrm{tr}(\partial_\mu \mathcal{U}^\dagger \partial^\mu \mathcal{U}) + \mathcal{O}(p^4).$$

The first term is trivial and is the only one with two derivatives. All consecutive terms have four or more derivatives (hence p^4). It is not possible to write down a term that would five Goldstone mass terms $m^2_{AB}\phi^A\phi^B$, as it should be. The effective theory corresponds to a derivative expansion, but Goldstone terms without derivatives do not appear in non-linear representations.

The discussion above can be applied to calculate, for example, the scattering amplitude for $\pi^+\pi^0 \to \pi^+\pi^0$ in leading order in p. In order to expand the fields in $\mathcal{U}$, we first denote $\phi = \phi^A T^A$ and calculate

$$\begin{aligned} \partial_\mu \mathcal{U} &= \frac{2i}{f}\partial_\mu\phi - \frac{2}{f^2}\partial_\mu(\phi\phi) - \frac{4i}{3f^3}\partial_\mu(\phi\phi\phi) + \ldots, \\ \partial_\mu \mathcal{U}^\dagger &= -\frac{2i}{f}\partial_\mu\phi - \frac{2}{f^2}\partial_\mu(\phi\phi) + \frac{4i}{3f^3}\partial_\mu(\phi\phi\phi) + \ldots \end{aligned}$$

Combining the various cross terms in the product gives

$$\begin{aligned} \partial_\mu(\phi\phi)\partial^\mu(\phi\phi) &= (\partial_\mu\phi)\phi(\partial^\mu\phi)\phi + \phi(\partial_\mu\phi)(\partial_\mu\phi)\phi + (\partial_\mu\phi)\phi\phi\partial^\mu\phi + \phi(\partial_\mu\phi)\phi(\partial_\mu\phi) \\ &\overset{\text{trace}}{=} 2\phi\phi(\partial_\mu\phi)(\partial^\mu\phi), \\ \partial_\mu\phi\partial^\mu(\phi\phi\phi) &= (\partial_\mu\phi)(\partial^\mu\phi)\phi\phi + (\partial_\mu\phi)\phi(\partial^\mu\phi)\phi + (\partial_\mu\phi)\phi\phi(\partial_\mu\phi) \\ &= 2\phi\phi(\partial_\mu\phi) + \phi(\partial_\mu\phi)\phi(\partial^\mu\phi). \end{aligned}$$

Hence, the Lagrangian up to four powers of ϕ and up to two derivatives reads

$$\begin{aligned} \mathcal{L}^{(2)} &= \mathrm{tr}(\partial_\mu\phi\partial^\mu\phi) + \frac{1}{f^2}\,\mathrm{tr}\left(\partial_\mu(\phi\phi)\partial^\mu(\phi\phi) - \frac{2}{3}\partial_\mu\phi\partial^\mu(\phi\phi\phi) - \frac{2}{3}\partial_\mu(\phi\phi\phi)\partial^\mu\phi\right) \\ &= \mathrm{tr}(\partial_\mu\phi\partial^\mu\phi) - \frac{2}{3f^2}\,\mathrm{tr}(\phi\phi\partial_\mu\phi\partial^\mu\phi) + \frac{2}{3f^2}(\phi(\partial_\mu\phi)\phi\partial^\mu\phi). \end{aligned}$$

The remaining terms are with 6 or more ϕ fields which do not contribute at tree level (and 1-loop diagrams are suppressed). The first component of the first term would be, for example,

$$(\partial_\mu\phi\partial^\mu\phi)_{11} = \frac{1}{2}\left[\left(\frac{1}{\sqrt{2}\partial_\mu\pi^0 + \frac{1}{\sqrt{6}}\partial_\mu\eta_8}\right)\left(\frac{1}{\sqrt{2}}\partial^\mu\pi^0 + \frac{1}{\sqrt{6}}\partial^\mu\eta_8\right) + \partial_\mu\pi^+\partial^\mu\pi^-\right].$$

Taking the trace of the whole matrix gives correctly normalized kinetic terms. The ones we are interested in are

$$\frac{1}{2}\partial_\mu\pi^0\partial^\mu\pi^0 + \partial_\mu\pi^+\partial^\mu\pi^-.$$

The neutral pion is a real scalar field, and the charged pion is a complex scalar field. In the following, we will only be interested in the pion interaction, so we can simplify

$$\sqrt{2}\phi^A T^A \to \sqrt{2}\pi^i\frac{\sigma^i}{2} = \begin{pmatrix}\pi^0/\sqrt{2} & \pi^+\\ \pi^- & -\pi^0/\sqrt{2}\end{pmatrix},$$

where $\pi^0 = \pi^3$ and $\pi^\pm = (\pi^1 \mp i\pi^2)/\sqrt{2}$. Using $\mathrm{tr}(\sigma^i\sigma^j) = 2\delta^{ij}$ and $\mathrm{tr}(\sigma^i\sigma^j\sigma^k\sigma^l) = 2(\delta^{ij}\delta^{kl} - \delta^{ik}\delta^{jl} + \delta^{il}\delta^{jk})$ we expand the Lagrangian as follows

$$\begin{aligned}
\mathcal{L}^{(2)} &= \frac{1}{2}\partial_\mu\pi^i\partial^\mu\pi^i - \frac{1}{6f^2}\pi^i\pi^i\partial_\mu\pi^j\partial^\mu\pi^j + \frac{1}{6f^2}(\pi^i\partial_\mu\pi^i)(\pi^j\partial^\mu\pi^j) + \dots\\
&= \frac{1}{2}\partial_\mu\pi^0\partial^\mu\pi^0 + \partial_\mu\pi^+\partial^\mu\pi^- +\\
&\quad + \frac{1}{3f^2}\left(-(\pi^0)^2\partial_\mu\pi^-\partial^\mu\pi^+ + \pi^0\pi^-\partial_\mu\pi^0\partial^\mu\pi^+ + \pi^0\pi^+\partial_\mu\pi^0\partial^\mu\pi^- +\right.\\
&\quad \left. -\pi^-\pi^+\partial_\mu\pi^0\partial^\mu\pi^0 + \frac{1}{2}(\pi^-)^2\partial_\mu\pi^+\partial^\mu\pi^+ + \frac{1}{2}(\pi^+)^2\partial_\mu\pi^-\partial^\mu\pi^- + \pi^-\pi^+\partial_\mu\pi^-\partial^\mu\pi^+\right) + \dots
\end{aligned}$$

For the scattering $\pi^+\pi^0 \to \pi^+\pi^0$ we only need the first four terms in the f^{-2} piece (note that outgoing π^+ is represented by π^-):

$$\begin{aligned}
A(\pi^+\pi^0 \to \pi^+\pi^0) &= \\
&= \frac{1}{3f^2}\left(-2ip_3(-ip_1) + (-ip_1)(ip_4 - ip_2) + ip_3(ip_4 - ip_2) - 2ip_4(-ip_2)\right)\\
&= -\frac{1}{3f^2}\left(2p_1\cdot p_3 + 2p_2\cdot p_4 + (p_3 - p_1)\cdot(p_4 - p_2)\right)\\
&= -\frac{1}{3f^2}\left(4p_1\cdot p_3 + 2p_1\cdot p_2 - 2p_1\cdot p_4\right)\\
&= \frac{t}{f^2}.
\end{aligned}$$

We used that the pions are massless $p_i^2 = 0$ and the momentum conservation ($p_1 + p_2 = p_3 + p_4$). In the last step we introduced the Mandelstam variable $t = (p_3 - p_4)^2 = -2p_1\cdot p_3$. We could follow the same steps to compute the other $\pi\pi \to \pi\pi$ or $\pi\pi \to 2n\pi$ processes. At order p^2 in momenta, the only parameter of the theory is f.

Explicit chiral symmetry breaking

The chiral symmetry is also explicitly broken by external fields coupling to the QCD Lagrangian. This breaking adds correction to the Goldstone boson masses. The most general way the fermions can couple to external scalar and vector fields is

$$\mathcal{L}'_{\mathrm{QCD}} = \mathcal{L}_{\mathrm{QCD}} + \overline{q}\gamma^{\mu}(v_{\mu} + \gamma^{5}a_{\mu})q - \overline{q}(s - i\gamma_5 p)q$$

where v_μ, a_μ, s and p are 3×3 hermitian matrices in the flavour space. We may consider them as sources of composite fields in the generating functional. This can be used to describe quark masses and the coupling to gauge bosons by setting the sources to specific values

$$s \to M = \begin{pmatrix} m_u & 0 & 0 \\ 0 & m_d & 0 \\ 0 & 0 & m_s \end{pmatrix}, \qquad v_\mu \to -eQA_\mu = -eA_\mu \begin{pmatrix} 2/3 & 0 & 0 \\ 0 & -1/3 & 0 \\ 0 & 0 & -1/3 \end{pmatrix}.$$

Analogously we could add electroweak gauge bosons by setting v_μ and a_μ to appropriate values. The new Lagrangian is invariant under local $\mathrm{SU}(3)_{\mathrm{L}} \otimes \mathrm{SU}(3)_{\mathrm{R}}$ transformations, $q_{\mathrm{L}} \to D_{\mathrm{L}} q_{\mathrm{L}}$, $q_{\mathrm{R}} \to D_{\mathrm{R}} q_{\mathrm{R}}$:

$$\begin{aligned} s + ip &\to D_{\mathrm{R}}(s + ip)D_{\mathrm{L}}^{\dagger}, \\ \ell_\mu &\to D_{\mathrm{L}}\ell_\mu D_{\mathrm{L}}^{\dagger} + iD_{\mathrm{L}}\partial_\mu D_{\mathrm{L}}^{\dagger}, \\ r_\mu &\to D_{\mathrm{R}} r_\mu D_{\mathrm{R}}^{\dagger} + iD_{\mathrm{R}}\partial_\mu D_{\mathrm{R}}^{\dagger}, \end{aligned}$$

where $\ell_\mu \equiv v_\mu - a_\mu$ and $r_\mu = v_\mu + a_\mu$. Therefore, the Goldstone effective Lagrangian in the presence of the external fields also has to be locally $\mathrm{SU}(3)_{\mathrm{L}} \otimes \mathrm{SU}(3)_{\mathrm{R}}$ invariant. The sources ℓ_μ and r_μ transform inhomogeneously, so must appear in covariant derivatives only

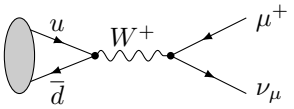
$$\mathrm{D}_\mu \mathcal{U} = \partial_\mu - ir_\mu \mathcal{U} + i\mathcal{U}\ell_\mu \to D_{\mathrm{R}} \mathrm{D}_\mu \mathcal{U} D_{\mathrm{L}}^{\dagger}.$$

The term $s + ip$ transforms homogeneously, so $\mathrm{tr}((s+ip)^{\dagger}\mathcal{U})$ is invariant. Introduce $\chi = 2B_0(s + ip)$, where B_0 is a new parameter whose interpretation will become clear. We will also show that χ counts as two derivatives, so $\mathcal{L}^{(2)}$ needs to only contain one χ:

$$\mathcal{L}^{(2)} \equiv \frac{f^2}{4}\,\mathrm{tr}((\mathrm{D}_\mu \mathcal{U})^{\dagger}(\mathrm{D}^\mu \mathcal{U}) + \mathcal{U}^{\dagger}\chi + \chi^{\dagger}\mathcal{U}).$$

Determination of the pion decay constant

The positive pion decays into anti-muon and corresponding neutrino.



According to the Fermi theory, the interaction between the quarks and leptons reads

$$\frac{G_{\mathrm{F}}}{\sqrt{2}} V_{ud}\overline{u}\gamma^{\mu}(1 - \gamma_5)d \cdot \overline{\mu}\gamma_\mu(1 - \gamma_5)\nu_\mu + \mathrm{h.c.}$$

The amplitude is

$$A(\pi^+ \to \mu^+\nu_\mu) = \frac{G_{\mathrm{F}}}{\sqrt{2}} V_{ud}^{*}\overline{u}_{\nu_\mu}\gamma^{\mu}(1 - \gamma_5)v_{\mu^+} \langle \Omega|\overline{d}\gamma_\mu(1 - \gamma_5)u|\pi^+\rangle\,.$$

Here, $|\Omega\rangle$ represents the QCD vacuum, and the expectation value of the quarks forming the pion $\langle\Omega|\overline{d}\gamma_\mu(1-\gamma_5)u|\pi^+\rangle$ can be determined from the pion lifetime $\tau_{\pi^+\to\mu^+\nu_\mu} \approx 2.6\cdot 10^{-8}\,\mathrm{s}$. Note that only $\gamma_\mu\gamma_5$ contributes to the decay. On the other hand, the $\langle\Omega|\overline{d}\gamma_\mu(1-\gamma_5)u|\pi^+\rangle$ is also the Goldstone matrix of the $\mathrm{SU(3)_A}$ Noether current, which defines the pion decay constant. The Noether current in terms of the Goldstone fields ca be obtained as follows:

$$\overline{q}_i\gamma^\mu\gamma_5 q_j \leftrightarrow \frac{\delta\Gamma}{\delta a_{\mu,ji}} = 2\left(\frac{\delta\Gamma}{\delta r_{\mu,ji}} - \frac{\delta\Gamma}{\delta \ell_{\mu,ji}}\right).$$

Here, Γ is the effective action for generating functional with sources v^μ, a^μ, s and p. For the quadratic terms in the Lagrangian, the latter expression becomes

$$\frac{\delta\Gamma}{\delta a_{\mu,ji}} = 2\frac{f^2}{4}\left((\mathrm{D}_\mu\mathcal{U})i\mathcal{U}^\dagger + (-i)\mathcal{U}\mathrm{D}_\mu\mathcal{U}^\dagger - (\mathrm{D}_\mu\mathcal{U}^\dagger)i\mathcal{U} - (-i)\mathcal{U}^\dagger\mathrm{D}_\mu\mathcal{U}\right)_{ji} = -2f\partial^\mu(\phi^A T^A)_{ji}.$$

The remaining terms have more than one ϕ or contain external currents. Continue with

$$\begin{aligned}\langle\Omega|\overline{d}\gamma_\mu(1-\gamma_5)u|\pi^+(p)\rangle &= -\langle\Omega|\overline{q}_2\gamma_\mu\gamma_5 q_1|\pi^+(p)\rangle = 2f\langle\Omega|\partial_\mu(\phi^A T^A)_{12}|\pi^+(p)\rangle \\ &= \sqrt{2}f\partial_\mu\langle\Omega|\pi^+(x)|\pi^+(p)\rangle = \sqrt{2}f\partial_\mu e^{-ipx} = -i\sqrt{2}fp^\mu.\end{aligned}$$

From the pion lifetime one then finds $f = 92.4\,\mathrm{MeV}$. This determines the strength of all pion interactions at $\mathcal{O}(p^2)$.

Symmetry breaking through quark masses

The following relation holds

$$\overline{q}_i q_j \overset{\mathcal{L}'_{\mathrm{QCD}}}{\leftrightarrow} -\frac{\Gamma}{\delta s_{ij}} \overset{\mathcal{L}^{(2)}}{\leftrightarrow} -\frac{f^2}{4}2B_0(\mathcal{U}+\mathcal{U}^\dagger)_{ji}.$$

From this, we obtain

$$\langle\Omega|\overline{q}_i q_j|\Omega\rangle = -\frac{f^2B_0}{2}\langle\Omega|(\mathcal{U}+\mathcal{U}^\dagger)_{ji}|\Omega\rangle = -f^2B_0\delta_{ij}.$$

Hence, B_0 is determined by the quark condensate

$$B_0 = -\frac{1}{f^2}\langle\Omega|\overline{u}u|\Omega\rangle.$$

Now put $s = M$ (quark mass matrix). Then $\mathcal{L}^{(2)}$ contains the following terms (denote $\phi = \phi^A T^A$)

$$\frac{f^2}{2}B_0\,\mathrm{tr}(M(\mathcal{U}+\mathcal{U}^\dagger)) = f^2B_0\,\mathrm{tr}(M\cos(2\phi^A T^A/f)) = -2B_0\,\mathrm{tr}(M\phi^2) + \frac{2}{3}\frac{B_0}{f^2}\,\mathrm{tr}(M\phi^4) + \ldots$$

The first term is the Goldstone mass term and the second gives a correction to $\pi\pi \to \pi\pi$ scattering. For $m_u \neq m_d$, the Goldstone mass matrix contains a mixed term $\propto \pi^0\eta_8$ so the physical π^0, η_8 states are linear combinations of π^0, η_8 appearing in ϕ. Diagonalizing the mass matrix gives

$$\begin{aligned} m^2_{\pi^\pm} &= (m_u+m_d)B_0 & m^2_{K^0} &= m^2_{\overline{K}^0} = (m_d+m_s)B_0 \\ m^2_{\pi^0} &= (m_u+m_d-\delta)B_0 & m^2_\eta &= \frac{4m_s+m_u+m_d}{3}B_0 + \delta B_0 \\ m^2_{K^\pm} &= (m_u+m_s)B & \delta &= \frac{1}{2}\frac{(m_d-m_u)^2}{2m_s-m_u-m_d}. \end{aligned}$$

The first equation agrees with the Gell-Mann–Oakes–Renner relation. Since B_0 is a priori unknown, this gives mass relations from which one can extract quark mass ratios from the known meson masses

$$m_u : m_d : m_s = 0.55 : 1 : 20.3.$$

If one interprets the difference between the vector meson masses $m_K - m_\rho \approx 120\,\text{MeV}$ as $m_s - m_d$, one gets $m_s = 120\,\text{MeV}$. In reality, the strange mass is scale and scheme dependent, so this estimate is not rigorous, but $\mathcal{O}(100\,\text{MeV})$ is correct. To get the mass correction to the $\pi^+\pi^0 \to \pi^+\pi^0$ scattering amplitude, we include the term

$$\mathcal{L}^{(2)} \supset \frac{(m_u + m_d)}{6f^2} B_0 \pi_0^2 \pi^+\pi^-$$

to the Lagrangian. Note that there are no derivatives in the explicit symmetry breaking terms. The amplitude reads

$$\begin{aligned} A(\pi^+\pi^0 \to \pi^+\pi^0) &= -\frac{1}{3f^2}(2p_1\cdot p_3 + 2p_2\cdot p_4 + (p_3 - p_1)\cdot(p_4 - p_2) - m_\pi^2) \\ &= -\frac{1}{3f^2}(4p_1\cdot p_3 + 2p_1(p_\cdot - p_4) - m_\pi^2) = -\frac{1}{3f^2}(6p_1\cdot p_2 - 3m_\pi^2) \\ &= \frac{t - m_\pi^2}{f^2}, \end{aligned}$$

where $t = (p_1 - p_3)^2 = 2m_\pi^2 - 2p_1\cdot p_3$. Note that m_π^2 naturally counts as t, i.e. as two derivatives and therefore appears in the $\mathcal{O}(p^2)$ Lagrangian.

Goldstone Lagrangian beyond leading order

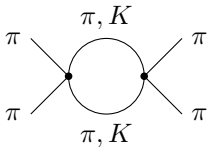
We will now sketch how the effective Lagrangian can by systematically improved in the sense of an effective field theory. In the context of the spontaneously broken SU(3) × SU(3) symmetry of QCD, the method is also known as *chiral perturbation theory* (χPT). We have already shown that the momentum power p of a diagram can be determined as

$$P = \sum_i v_i(d_i - 2) + 2L + 2.$$

At E^4 ($P = 4$), one-loop diagrams with vertices from $\mathcal{L}^{(2)}$ and the tree diagrams with one vertex $v_i = 1$ and four derivatives from $\mathcal{L}^{(4)}$. When quark masses are included, $B_0 M \sim m^2_{\text{Goldstone}} \sim \chi$ counts like two derivatives.

$$\begin{aligned} \mathcal{L}^{(4)} &= \text{most general SU(3) × SU(3) invariant built from } \mathcal{U} \text{ and } \chi = B_0 M \\ &= L_1(\operatorname{tr}((\mathrm{D}_\mu\mathcal{U})^\dagger \mathrm{D}^\mu\mathcal{U}))^2 + \cdots + L_5 \operatorname{tr}((\mathrm{D}_\mu\mathcal{U})^\dagger \mathrm{D}^\mu\mathcal{U}(\mathcal{U}^\dagger\chi + \chi^\dagger\mathcal{U})) + \\ &\quad + \text{temrs with no derivatives and two } \chi \end{aligned}$$

The $\mathcal{L}^{(4)}$ gives additional terms to the scattering amplitude. The tree diagrams are polynomial in the external momenta. The loop diagrams give non-analytic contributions:



$$\propto \frac{1}{f^4}\int \frac{\mathrm{d}^4k}{(2\pi)^4} g(k, p_i; m^2)$$

$$\propto \frac{1}{f^2}\frac{1}{16\pi^2 f^2}\left(h(p_i, m)\left[\frac{1}{\varepsilon} + \ln\frac{\mu^2}{p_i p_j}\right] + \text{finite}\right)$$

The divergent part can be removed by adding a counterterm proportional to one of the L_i in $\mathcal{L}^{(4)}$. The renormalized coupling $L_i(\mu)$ is now scale-dependent, but

$$L_i(\mu) \cdot \text{tree} + \log \frac{\mu^2}{p_i p_j} \times \text{one-loop}$$

is scale independent. The $\log(\mu^2/(p_i p_j))$ are called *chiral logarithms*. They are independent of the unknown L_i couplings and can be predicted/calculated unambiguously from $\mathcal{L}^{(2)}$. The $L_i(\mu)$ have to be determined from measurements. The effective expansion parameter is $(E/(4\pi f))^2$, i.e., chiral perturbation theory is valid for $E \ll 4\pi f \sim 1\,\text{GeV}$.

Note. At $\mathcal{O}(p^4)$ there is a subtlety. When one constructs the most general effective theory compatible with the symmetries, it is enough that the action is invariant. It can happen that there are terms that can be added to $\mathcal{L}$ which are not invariant, but rather their $\text{SU}(3) \times \text{SU}(3)$ variation is a total divergence and hence the action is invariant. Such terms are called *Wess–Zumino–Witten terms*. Here they are related to an anomaly of the $\text{SU}(3) \times \text{SU}(3)$ symmetry, when the Noether currents are coupled to the external v_μ and a_μ fields. We do not discuss this further here.

Chapter 4

Electroweak symmetry breaking and the Higgs boson

Electroweak symmetry refers to $\mathrm{SU}(2) \times \mathrm{U}(1)_Y$ gauge symmetry of the Standard Model. Since only one massless spin-1 particle is observed, the photon, this symmetry must be broken

$$\mathrm{SU}(2) \times \mathrm{U}(1)_Y \to \mathrm{U}(1)_{\mathrm{em}}, \quad e^{-i\varepsilon^A(x)T^A} e^{-i\varepsilon(x)Y} \to e^{-i\varepsilon(x)Q}, \quad Q = T^3 + Y.$$

Here, T^A is an SU(2) generator in any representation, Y is the *hypercharge*, which can be any real number, and Q is the electric charge in units of positron charge. It is easily seen that the spontaneous symmetry breaking of a gauge (local) symmetry does not lead to Goldstone bosons. Recall the parametrization (eq. 3.5)

$$\Psi_n(x) = \Sigma_{nm}(x)\tilde{\Psi}_m(x)$$

where $\Sigma(x) \in \{D(g)|g \in G\}$ are Goldstone modes and $\tilde{\Psi}$ are the non-Goldstone modes. We can look at the expression as Ψ_n a gauge transformation of $\tilde{\Psi}_n$. Since the theory is gauge-invariant, $\mathcal{L}$ expressed in terms of Σ and $\tilde{\Psi}$ does not contain Σ at all. The parametrization in terms of $\tilde{\Psi}$ instead of Ψ corresponds to choosing a particular gauge – *unitary gauge*. In unitary gauge, the particle content of the theory corresponds (at weak coupling) directly to physical particles.

We discuss the breaking of electroweak (EW) gauge symmetry in the SM for the case of the minimal *Higgs sector*, as usually implied by the definition of the SM. The sector consists of three fields

$$\begin{aligned}
&W_\mu^A, \quad A = 1, 2, 3, \quad \mathrm{SU}(2) \text{ gauge field,} \\
&B_\mu, \quad \mathrm{U}(1)_Y \text{ gauge field,} \\
&\boldsymbol{\phi} = \begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix}, \quad \mathrm{SU}(2) \text{ doublet of complex scalar fields with hypercharge } Y = 1/2.
\end{aligned}$$

Note that

$$\tilde{\boldsymbol{\phi}} \equiv i\sigma^2\boldsymbol{\phi}^* = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} \phi_1^* \\ \phi_2^* \end{pmatrix} = \begin{pmatrix} \phi_2^* \\ -\phi_1^* \end{pmatrix}$$

also transforms as an SU(2) doublet and has *hypercharge* $-1/2$. The latter is obvious from complex conjugation, the proof of the former is identical to showing that $\varepsilon\psi_{\mathrm{R}}^*$ transforms as a

left-handed spinor. Explicitly (using $\sigma^2\sigma^{A*}\sigma^2 = -\sigma^A$ in the last step),

$$\tilde{\boldsymbol{\phi}} \xrightarrow{\mathrm{SU}(2)} i\sigma^2 \left(e^{-i\varepsilon^A\sigma^A/2}\boldsymbol{\phi}\right)^* = i\sigma^2 e^{i\varepsilon^A\sigma^{A*}/2}\boldsymbol{\phi}^*$$
$$= i\sigma^2 e^{i\varepsilon^A\sigma^{A*}/2}\sigma^2\sigma^2\boldsymbol{\phi}^* = e^{-i\varepsilon^A\sigma^A/2} i\sigma^2\boldsymbol{\phi}^* = e^{-i\varepsilon^A\sigma^A/2}\tilde{\boldsymbol{\phi}}.$$

The Standard Model Lagrangian without strong interaction and fermions reads

$$\mathcal{L} = -\frac{1}{4}W^A_{\mu\nu}W^{A,\mu\nu} - \frac{1}{4}B_{\mu\nu}B^{\mu\nu} + (\mathrm{D}_\mu\boldsymbol{\phi})^\dagger(\mathrm{D}^\mu\boldsymbol{\phi}) + \mu^2\boldsymbol{\phi}^\dagger\boldsymbol{\phi} - \frac{\lambda}{4}(\boldsymbol{\phi}^\dagger\boldsymbol{\phi})^2$$

with $i\mathrm{D}_\mu = i\partial_\mu - gT^AW^A_\mu - g'YB_\mu$. This is the most general $\mathrm{SU}(2)\times\mathrm{U}(1)$ invariant renormalizable Lagrangian. The Higgs sector has the form of a linear σ-model (Sec. 3.1.4):

$$\Sigma = \begin{pmatrix} \phi_2^* & \phi_1^* \\ -\phi_1^* & \phi_2 \end{pmatrix} = (\tilde{\boldsymbol{\phi}}\,|\,\boldsymbol{\phi}).$$

The trace of two such objects is

$$\mathrm{tr}(\Sigma^\dagger\Sigma) = \mathrm{tr}\left(\begin{pmatrix} \phi_2 & -\phi_1 \\ \phi_1^* & \phi_2^* \end{pmatrix}\begin{pmatrix} \phi_2^* & \phi_1^* \\ -\phi_1^* & \phi_2 \end{pmatrix}\right) = \mathrm{tr}\begin{pmatrix} \phi_2\phi_2^* + \phi_1\phi_1^* & 0 \\ 0 & \phi_1\phi_1^* + \phi_2\phi_2^* \end{pmatrix} = 2\boldsymbol{\phi}^\dagger\boldsymbol{\phi}.$$

Hence, the part of the Lagrangian without gauge fields reads

$$\mathcal{L}_{\text{no gauge}} = \frac{1}{2}\,\mathrm{tr}(\partial_\mu\Sigma^\dagger\partial^\mu\Sigma) - \frac{\lambda}{16}\left(\mathrm{tr}(\Sigma^\dagger\Sigma) - \frac{4\mu^2}{\lambda}\right)^2 + \text{const.}$$

For future convenience, we will denote $v^2 \equiv 4\mu^2/\lambda$. Hence, the SM Higgs sector has a global, "chiral" $\mathrm{SU}(2)\times\mathrm{SU}(2)$ symmetry, under which

$$\Sigma \to e^{-i\varepsilon^A_{\mathrm{L}}T^A}\Sigma e^{i\varepsilon^A_{\mathrm{R}}T^A}.$$

Compare this to the operation of gauge transformations on Σ:

$$\Sigma = (\tilde{\boldsymbol{\phi}}\,|\,\boldsymbol{\phi}) \to \left(e^{-i\varepsilon^AT^A}e^{+\varepsilon/2}\tilde{\boldsymbol{\phi}}\,\middle|\,e^{-i\varepsilon^AT^A}e^{-i\varepsilon/2}\boldsymbol{\phi}\right) =$$
$$= e^{-i\varepsilon^AT^A}(\tilde{\boldsymbol{\phi}}\,|\,\boldsymbol{\phi})\begin{pmatrix} e^{i\varepsilon/2} & 0 \\ 0 & e^{-i\varepsilon/2} \end{pmatrix} = e^{-i\varepsilon^AT^A}\Sigma e^{i\varepsilon T^3}. \tag{4.1}$$

For $v^2 > 0$, $\boldsymbol{\phi}$ develops a vacuum expectation value $\boldsymbol{\phi}_0 = (0, v/\sqrt{2})^\intercal$ (without loss of generality, we can assume that the breaking happens in the second direction). Hence,

$$\Sigma = \frac{v}{\sqrt{2}}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

and the global $\mathrm{SU}(2)\times\mathrm{SU}(2)$ symmetry of the Higgs sector is spontaneously broken to the diagonal subgroup $\mathrm{SU}(2)$. In the absence of gauge symmetry, this would result in the Goldstone bosons. From eq. 4.1 we see that the first factor of $\mathrm{SU}(2)$ corresponds to the global $\mathrm{SU}(2)$ gauge transformations, while the second factor of $\mathrm{SU}(2)$ contains but is larger than $\mathrm{U}(1)_Y$. The extended symmetry is called the *custodial symmetry* of the SM Higgs sector. The covariant derivative acting on Σ in eq. 4.1 is

$$\mathrm{D}_\mu\Sigma = \partial_\mu\Sigma + igT^AW^A_\mu\Sigma - ig'\Sigma T^3B_\mu.$$

Note that $\mathrm{D}_\mu\Sigma$ is $\mathrm{SU}(2)\times\mathrm{U}(1)_Y$ covariant but is not even globally $\mathrm{SU}(2)\times\mathrm{SU}(2)$ covariant since

$$\Sigma T^3 B_\mu \xrightarrow[\text{global}]{\mathrm{SU}(2)\times\mathrm{SU}(2)} D_\mathrm{L}\sigma D_\mathrm{R}^\dagger T^3 B_\mu \neq D_\mathrm{L}\Sigma T^3 B_\mu D_\mathrm{R}^\dagger$$

since $[D_\mathrm{R}^\dagger, T^3] \neq 0$ in general, while

$$\mathbf{W}_\mu \xrightarrow[\text{global}]{\mathrm{SU}(2)\times\mathrm{SU}(2)} D_\mathrm{L}\mathbf{W}_\mu D_\mathrm{L}^\dagger D_\mathrm{L}\Sigma D_\mathrm{R}^\dagger = D_\mathrm{L}\mathbf{W}_\mu\Sigma D_\mathrm{R}^\dagger.$$

The custodial symmetry of the Higgs sector is weakly (since $g' \ll 1$) explicitly broken by the hypercharge interaction.

4.1 Higgs mechanism

4.1.1 Unitary gauge

As in the linear σ-model, we write

$$\Sigma(x) = \frac{1}{\sqrt{2}}(v + H(x))\underbrace{e^{iT^A\varphi^A(x)/v}}_{\mathcal{U}(\varphi)}.$$

The Lagrangian becomes

$$\mathcal{L} = -\frac{1}{4}W_{\mu\nu}^A W^{A,\mu\nu} - \frac{1}{4}B_{\mu\nu}B^{\mu\nu} + \frac{v^2}{4}\,\mathrm{tr}((\mathrm{D}_\mu\mathcal{U}(\varphi))^\dagger(\mathrm{D}^\mu\mathcal{U}(\varphi))) + \text{terms containing } H(x).$$

The mass of the H-boson is $m_H^2 = \lambda v^2/2$. For $m_H \gg v$ (large λ) and scattering energies $E \ll m_H$, one can write down an effective Lagrangian without H-field and the electroweak sector of the SM is described by an $\mathrm{SU}(2)\times\mathrm{SU}(2)$-invariant chiral Lagrangian coupled (weakly) to the $\mathrm{SU}(2)\times\mathrm{U}(1)_Y$ gauge fields

$$\mathcal{L} = -\frac{1}{4}W_{\mu\nu}^A W^{A,\mu\nu} - \frac{1}{4}B_{\mu\nu}B^{\mu\nu} + \mathcal{L}^{(2)} + \mathcal{L}^{(4)}.$$

The first term $\mathcal{L}^{(2)}$ has one single coupling v, while the second term $\mathcal{L}^{(4)}$ has couplings L_i which have to be computed for every specific model of EW symmetry breaking. A model-independent description of EW symmetry breaking based only on the pattern of symmetry breaking $(\mathrm{SU}(2)_\mathrm{L}\times\mathrm{SU}(2)_\mathrm{R} \to \mathrm{SU}(2)$ and consequently $\mathrm{SU}(2)_\mathrm{L}\times\mathrm{U}(1)_Y \to \mathrm{U}(1)_\mathrm{em}$ for the gauge symmetry) is valid only for scattering energies below $4\pi v$ or m_H. Meanwhile, we know that there exists a scalar particle with mass $m_H = 125\,\mathrm{GeV}$ not much larger than m_W. This could be a singlet under the $\mathrm{SU}(2)\times\mathrm{U}(1)$ symmetry (as is the non-Goldstone field $H(x)$ in the linear σ-model) different from $H(x)$, or it could be $H(x)$ field. The latter corresponds to the SM Higgs mechanism, but the first possibility is experimentally not completely excluded.

The $\mathcal{L}^{(2)}$ part of the Lagrangian in unitary gauge $\mathcal{U}(\varphi) \equiv 1$ is

$$\begin{aligned}
\mathcal{L}^{(2)} &= \frac{v^2}{4} \operatorname{tr}\left((gT^A W_\mu^A - g'T^3 B_\mu)(gT^B W^{B,\mu} - g'T^3 B^\mu)\right) \\
&= \frac{v^2}{4}\left(g^2 \sum_{A,B=1}^{2} W_\mu^A W^{B,\mu} \operatorname{tr}(T^A T^B) + (gW_\mu^3 - g'B_\mu)(gW^{3\mu} - g'B^\mu) \operatorname{tr}(T^3 T^3)\right) \\
&= \frac{v^2}{4}\left(\frac{g^2}{2} \sum_{A=1}^{2} W_\mu^A W^{A,\mu} + (gW_\mu^3 - g'B_\mu)(gW^{3\mu} - g'B^\mu)\frac{1}{2}\right) \\
&= \frac{g^2 v^2}{4} W_\mu^- W^{+\mu} + \frac{g^2 + g'^2}{8} v^2 Z_\mu Z^\mu.
\end{aligned}$$

We have introduced three new fields

$$\begin{aligned}
W_\mu^\pm &\equiv \frac{1}{\sqrt{2}}(W_\mu^1 \pm i W_\mu^2), \quad \text{complex vector field with mass } m_W^2 = \frac{g^2 v^2}{4} \\
Z_\mu &\equiv \frac{gW_\mu^3 - g'B_\mu}{\sqrt{g^2 + g'^2}}, \quad \text{real vector field with mass } m_Z^2 = \frac{v^2}{4}(g^2 + g'^2), \\
A_\mu &\equiv \frac{g'W_\mu^3 + gB_\mu}{\sqrt{g^2 + g'^2}}, \quad \text{massless real vector field corresponds to the unbroken } \mathrm{U}(1)_{\mathrm{em}}.
\end{aligned}$$

There are no Goldstone bosons. Instead, three of the four massless gauge fields acquired masses.

Note. Define the *Weinberg angle* (sometimes called weak mixing angle)

$$\cos\theta_{\mathrm{W}} \equiv \frac{g}{\sqrt{g^2 + g'^2}}.$$

With this, we define the *ρ parameter*

$$\rho \equiv \frac{m_W}{m_Z \cos\theta_{\mathrm{W}}} = \frac{\sqrt{g^2 + g'^2}}{g}\frac{m_W}{m_Z} = 1.$$

This relation follows from $\mathcal{L}^{(2)}$ alone and hence depends only on the $\mathrm{SU}(2) \times \mathrm{SU}(2)$ symmetry. In the SM the relation $\rho = 1$ is modified by quantum effects involving g' and the coupling of the Higgs field to fermions, since these break custodial SU(2) explicitly.

A general pattern for spontaneous symmetry breaking in unitary gauge goes as follows.

- We start with a gauge group G with $\dim G$ massless gauge bosons.
- The symmetry is broken to a subgroup H. There are $\dim H$ massless gauge bosons left, corresponding to the unbroken symmetries (generators). The remaining $\dim G - \dim H$ massless gauge bosons and some number of massless Goldstone bosons combine to $\dim G - \dim H$ *massive* gauge bosons. The number of degrees of freedom matters, because a massive spin-1 particle has there rather than two polarization states.

This can be explicitly seen when the symmetry breaking is caused by vacuum expectation values of elementary scalar fields as in the case for the standard Higgs mechanism. Imagine we have a real scalar field ϕ_n in some representation of G with generators T^A. After expanding around the vacuum expectation value $\phi_n = v_n + \phi'_n$ the kinetic term gives

$$\frac{1}{2}\mathrm{D}_\mu\phi_n \mathrm{D}^\mu\phi_n = \frac{1}{2}(\partial_\mu\phi'_n - igT^A_{nm}v_m A^A_\mu - igT^A_{nm}\phi'_m A^A_\mu)^2 \supset \frac{1}{2}g^2(iT^A v)(iT^B v)A^A_\mu A^{B,\mu}.$$

We read out the mass matrix $M^2_{AB} = g^2(iT^Av)(iT^Bv)$. For a real representation iT^A is real, so M^2_{AB} is also real. For arbitrary (real) λ^A it holds

$$M^2_{AB}\lambda^A\lambda^B = (\underbrace{i\lambda^AT^Av}_{w_n})(i\lambda^BT^Bv) = w_nw_n \geq 0,$$

so M^2_{AB} has non-negative eigenvalues. The zero eigenvalues correspond to vectors satisfying $(\lambda^AT^A)_{nm}v_m = 0$, i.e. the λ^AT^A with this property correspond to the unbroken gauge symmetries. In the SM, $\mathbf{v} = (0, v)^\intercal$, so

$$(\lambda^AT^A)\mathbf{v} = \sum_{A=1}^{3}\lambda^A\frac{\sigma^A}{2}\mathbf{v} + \lambda_Y\frac{Y}{2}\mathbf{v} = \frac{v}{2}\begin{pmatrix}\lambda^1 - i\lambda^2\\ -\lambda^3 + \lambda_Y\end{pmatrix}.$$

The unbroken symmetry corresponds to $\lambda^1 = \lambda^2 = 0$ and $\lambda_Y = \lambda^3$, i.e. $Q \equiv T^3 + Y$. Unitary gauge is useful to see the physical degrees of freedom, but not practical for quantization and perturbation theory, because the propagator of massive vector fields (ch.??) is

$$\frac{i}{k^2 - m^2}\left(-g_{\mu\nu} + \frac{k_\mu k_\nu}{m^2}\right).$$

Because this is constant for large k, every interaction is non-renormalizable, and the theory is only useful as an effective theory for $E \ll 4\pi m/g$.

4.1.2 Renormalizable gauge

Assume first (as above) a theory with gauge symmetry G and real, scalar field ϕ_n, which transforms as $\boldsymbol{\phi} \xrightarrow{g} D(g)\boldsymbol{\phi}$. The Lagrangian is again

$$\mathcal{L} = -\frac{1}{4}G^A_{\mu\nu}G^{A,\mu\nu} + \frac{1}{2}(\mathrm{D}_\mu\phi)_n(\mathrm{D}^\mu\phi)_n - V(\phi_n), \quad i\mathrm{D}_\mu = \partial_\mu + igT^AA^A_\mu.$$

Assume that the potential is such that some of the scalar fields get a vacuum expectation values, and define $\phi_n = v_n + \tilde{\phi}_n$. The quadratic terms in $\mathcal{L}$ then read

$$\mathcal{L} = -\frac{1}{4}G^A_{\mu\nu}G^{A,\mu\nu} + \frac{1}{2}\partial_\mu\tilde{\phi}_n\partial^\mu\tilde{\phi}_n - \underbrace{g\partial_\mu\tilde{\phi}_niT^A_{nm}v_mA^{A,\mu}}_{(1)} + \underbrace{\frac{1}{2}g^2iT^A_{nk}v_kiT^B_{nm}v_mA^A_\mu A^{B,\mu}}_{(2)} +$$
$$\underbrace{-\frac{1}{2}\frac{\partial^2V}{\partial\phi_n\partial\phi_m}\bigg|_{\boldsymbol{\phi}=\mathbf{v}}\tilde{\phi}_n\tilde{\phi}_m}_{(3)} + \text{interactions}.$$

We can drop $V(v_n)$ and linear terms in $\tilde{\phi}_n$, since v_n is determined by $\partial V/\partial\phi_m|_{\phi=v} = 0$. The three relevant parts in the Lagrangian are explained below.

(1) The $\tilde{\phi}$-A term which mixes $\tilde{\phi}$ and A fields contributes only for the Goldstone modes in $\tilde{\phi}$. The scalar fields without Goldstone modes satisfy the orthogonality condition $\tilde{\phi}_nT^A_{nm}v_m = 0$ (eq. 3.6).

(2) This term gives the gauge boson mass matrix $M^2_{AB} = g^2(iT^Av)_n(iT^Bv)_n$.

(3) This term gives masses of the scalar particles

$$M^2_{nm} = \left.\frac{\partial^2 V}{\partial\phi_n \partial\phi_m}\right|_{\boldsymbol{\phi}=\mathbf{v}}.$$

The identity $M^2_{nm}(T^A v)_m = 0$ holds (eq. 3.1) and the remaining eigenvalues correspond to the massive, non-Goldstone bosons.

The term 1 vanishes in unitary gauge, since the Goldstone fields are eliminated by the gauge condition $\tilde{\phi}_n T^A_{nm} v_m = 0$.

't Hooft gauge (R_ξ gauge)

Recall from eq. 1.8 that

$$\mathcal{L}_{\text{GF+ghost}} = -\frac{1}{2\xi} f^A f^A - \bar{c}^A \Delta^A, \quad \Delta^A = -g \int \mathrm{d}^4 y \left.\frac{\delta f^A[\mathbf{A}_\varepsilon(x)]}{\delta\varepsilon^B(y)}\right|_{\varepsilon=0} c^B(y).$$

The idea is to modify the covariant gauge fixing function such that the $\tilde{\phi}_n A^{A,\mu}$ term is cancelled:

$$f^A[A,\phi] = \partial^\mu A^A_\mu + \xi g \tilde{\phi}_n (iT^A v)_n. \tag{4.2}$$

Thus, the bilinear terms in $\mathcal{L}$, including $\mathcal{L}_{\text{GF}}$, become

$$\begin{aligned}\mathcal{L}_{\text{bilin}} = &-\frac{1}{4} G^A_{\mu\nu} G^{A,\mu\nu} + \frac{1}{2} M_{AB} A^A_\mu A^{M,\mu} - \frac{1}{2\xi}(\partial_\mu A^{A,\mu})(\partial_\nu A^{A,\nu}) + \frac{1}{2}\partial_\mu\tilde{\phi}_n \partial^\mu \tilde{\phi}_n + \\ &- \frac{1}{2}\underbrace{\xi g^2 (iT^A v)_n (iT^A v)_m}_{\equiv (M^{\text{G}})^2_{nm}} \tilde{\phi}_n \tilde{\phi}_m - \frac{1}{2} M^2_{mn} \tilde{\phi}_n \tilde{\phi}_m + \text{interactions}.\end{aligned}$$

We introduced the mass matrix for the Goldstone modes, M^{G}_{nm}. Let's have a look at the ghost Lagrangian. The infinitesimal gauge transformation acts on the scalar field with

$$\boldsymbol{\phi} \to e^{i\varepsilon^A T^A} \boldsymbol{\phi} = \boldsymbol{\phi} + i\varepsilon^A T^A (\mathbf{v} + \tilde{\boldsymbol{\phi}}).$$

Thus, we have

$$\frac{\delta f^A[\mathbf{A}_\varepsilon, \boldsymbol{\phi}_\varepsilon](x)}{\delta\varepsilon^B(y)} = \delta^{(4)}(x-y)\frac{1}{g}(\partial^2_{(x)}\delta^{AB} + g f^{ABC}\partial_\mu A^{C,\mu} + \xi g^2 (iT^A v)_n i T^B_{nk}(v+\tilde{\phi})_k).$$

The first two terms are the same as in eq. 1.7. The remaining term gives both the ghost mass matrix ξM^2_{AB} and the ghost interaction

$$\mathcal{L}_{\text{FP}} = \partial_\mu \bar{c}^A \partial^\mu c^A - g f^{ABC}(\partial_\mu \bar{c}^A) c^B A^{C,\mu} - \xi g^2 (iT^A v)_n i T^B_{nk} \bar{c}^A c^B (v_k + \tilde{\phi}_k).$$

In *'t Hooft gauge* (also called *R_ξ gauge*) Goldstone bosons and ghosts become massive with a guage-parameter dependent mass.

1. **Gauge bosons**: there are $\dim H$ linear combinations $T^A\mathbf{v} = 0$ and the corresponding number of zero eigenvalues. The remaining gauge bosons are massive. Assume that M^2_{AB} has been diagonalized with eigenvalues $m^2_A \geq 0$. The propagator is

$$\frac{i}{k^2 - m_A^2}\left(-g_{\mu\nu} + (1-\xi)\frac{k_\mu k_\nu}{k^2 - \xi m_A^2}\right) = \frac{i}{k^2 - m_A^2}\left(-g_{\mu\nu} + \frac{k_\mu k_\nu}{k^2}\right) - \frac{i k_\mu k_\nu}{k^2}\frac{\xi}{k^2 - \xi m_A^2}.$$

On the right-hand side of the equation, the first term corresponds to the three polarization states (two transverse and longitudinal) and the second part is a scalar polarization with mass $\sqrt{\xi}m_A$. For $k^2 \to \infty$ the expression behaves as $1/k^2$ as for massless gauge bosons. This implies renormalizability of dimension-4 interactions by power counting. However, there are unphysical degrees of freedom and as in unbroken gauge theory one needs to show that they do not contribute to the S-matrix on the physical Hilbert space. For $\xi \to \infty$ one recovers unitary gauge and the unphysical mode is absent (with infinite mass).

2. **Physical Higgs particles**: correspond to the non-vanishing eigenvalues (m_H^2) of M_{nm}^2. The propagator is $i/(k^2 - m_H^2)$.
3. **Goldstone modes**: the Goldstone mass matrix $(M^{\mathrm{G}})_{nm}^2 = \xi g^2 (iT^a v)_n (iT^A v)_m$ has $\dim G - \dim H$ non-vanishing eigenvalues (there are $\dim H$ linear combinations with $T^A v = 0$, see above). There are $\sqrt{\xi}\times$ the non-vanishing eigenvalues of M_{AB}^2, i.e. $m_{\mathrm{G}} = \sqrt{\xi} m_a$.

   Proof. Let $M_{AB}^2 \lambda^B = m^2 \lambda^A$. Then

   $$\begin{aligned}(M^{\mathrm{G}})_{nm}^2 (i\lambda^B T_{mk}^B v_k) &= \xi g^2 (iT^A v)_m (iT^A v)_m (iT^B v)_m \lambda^B \\ &= \xi (iT^A v)_m M_{AB}^2 \lambda_B = \xi m^2 (i\lambda^A T^A v)_k.\end{aligned}$$

   This means $i\lambda^A T^A \mathbf{v}$ is eigenvector of $(M^{\mathrm{G}})_{nm}^2$ to eigenvalue ξm^2. □
4. **Ghosts**: have masses $\sqrt{\xi} m_A$ after diagonalization and propagator $i/(k^2 - \xi m_A^2)$.

Hence, we have the following particle content in renormalizable gauge.

- Physical modes:
  - $\dim H$ massless gauge bosons (2 polarizations),
  - $\dim G - \dim h$ massive gauge bosons with mass m_A (3 polarizations),
  - $\dim H$ scalar, non-Goldstone bosons, generically massive.
- Unphysical modes:
  - $\dim H$ massless longitudinal and scalar gauge bosons and ghosts and anti-ghosts (also massless),
  - $\dim G - \dim H$ scalar gauge bosons, Goldstone bosons (also called *pseudo-Goldstone bosons*), ghosts and anti-ghosts, all with mass $\sqrt{\xi} m_A$.

The physical modes correspond to unitary gauge, where the unphysical modes become infinitely heavy and decouple. (However, the limit $\xi \to \infty$ does not commute with $k \to \infty$ in loops, which makes unitary gauge non-renormalizable.) The contributions of the unphysical modes cancel in the S-matrix, separately between the states of mass 0 and $\sqrt{\xi} m_A$.

The case of SU(2) × U(1)

We have a non-simple group and a complex Higgs field

$$\boldsymbol{\phi} = \begin{pmatrix} \phi^+ \\ \frac{1}{\sqrt{2}}(v + H + i\chi) \end{pmatrix}.$$

The Higgs doublet, expressed as a perturbation around the vacuum expectation value $v/\sqrt{2}$ consists of a complex scalar field ϕ^+ (for which it holds $(\phi^+)^\dagger = \phi^-$) and two real scalar fields

H, χ. The 't Hooft gauge for the SM means

$$\mathcal{L}_{\text{GF}} = -\frac{1}{2\xi}(\partial_\mu A^\mu)^2 - \frac{1}{2\xi_Z}(\partial_\mu Z^\mu - \xi_Z m_Z \chi)^2 - \frac{1}{\xi_W}(\partial_\mu W^{\mu-} + i\xi_W m_W \phi^-)(\partial^\mu W_\mu^+ - i\xi_W m_W \phi^+)$$

which eliminates the mixed gauge-boson-Higgs bilinear terms. The field content is

$$(W_\mu^\pm, \phi^\pm, c_W^\pm, \bar{c}_W^\pm), (Z_\mu, \chi, c_Z, \bar{c}_Z), (A_\mu, c_A, \bar{c}_A), H.$$

From those,

- the pseudo-Goldstone bosons $\phi^\pm$ and ghosts $c_W^\pm, \bar{c}_W^\pm$ have mass $\sqrt{\xi_W} m_W$ and electric charge ± 1,
- the pseudo-Goldstone boson χ and ghosts $c_Z, \bar{c}_Z$ have mass $\sqrt{\xi_Z}$ and are electrically neutral,
- all of $A_\mu, c_A, \bar{c}_A$ are massless, and
- the Higgs boson H has mass $m_H = v\sqrt{\lambda/2} = \sqrt{2}\mu$.

This results in many different vertices (see for instance the appendix of Böhm, Denver, Joss [citation needed]). This and the presence of different masses makes perturbative computations of electroweak quantum effects technically more complicated than those of the strong interaction. But they are usually also smaller and therefore less relevant, because the couplings are smaller: $g^2, g'^2 \ll g_s^2$.

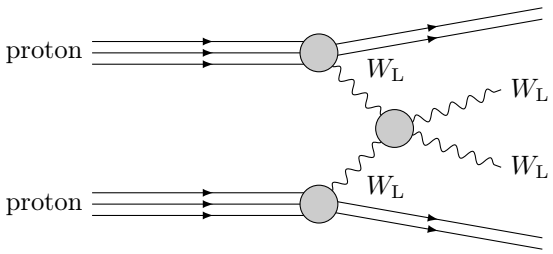
4.2 Scattering of longitudinal vector bosons

The subsequent considerations show that without the Higgs boson, the interaction of W-bosons at energies $E \gg m_W$ becomes strong (non-perturbative). Hence, either the Higgs boson must not be too heavy (as is now known to be true, if the scalar particle is the Higgs boson) or some other new phenomenon must happen to make the electroweak theory self-consistent. This was the reason why it was clear that *something* must be discovered at the LHC – as indeed it was.

Consider the scattering of *longitudinally polarized* W-bosons, $W_{\text{L}}^+ W_{\text{L}}^- \to W_{\text{L}}^+ W_{\text{L}}^-$. In the rest frame of W-boson, moving in z-direction the polarization vector is $\varepsilon_{\text{L}}^\mu = (0, 0, 0, 1)^\intercal$. After a Lorentz boost in z-direction to a frame, where the boson has momentum p^μ, we obtain

$$\varepsilon_{\text{L}}^\mu = \frac{1}{m_W} \begin{pmatrix} |\mathbf{p}| \\ 0 \\ 0 \\ E(\mathbf{p}) \end{pmatrix} \simeq \frac{p^\mu}{m_W}.$$

The last equality holds for $W \gg m_W$ where $E(\mathbf{p}) \approx |\mathbf{p}|$. The transverse polarization vectors remain $\mathcal{O}(1)$. For example, at the Large Hadron Collider one might have the following process.



Consider the amplitude for the process

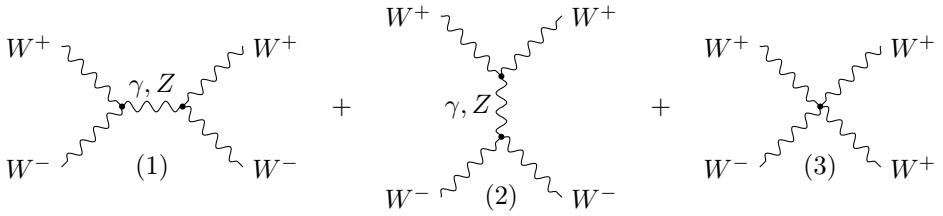
$$W_{\mathrm{L}}^{+}(p_1) + W_{\mathrm{L}}^{-}(p_2) \to W_{\mathrm{L}}^{+}(p_3) + W_{\mathrm{L}}^{-}(p_4).$$

In the centre-of-mass framer it holds

$$p_{1,2} = (E, 0, 0, \pm p), \quad E = \sqrt{m_W^2 + p^2}, \quad \varepsilon_{\mathrm{L}}(p_{1,2}) = \frac{1}{m_W}(p, 0, 0, \pm E),$$

$$p_{3,4} = (E, 0, \pm p\sin\theta, \pm p\cos\theta), \quad \varepsilon_{\mathrm{L}}(p_{3,4}) = \frac{1}{m_W}(p.0, \pm E\sin\theta, \pm E\cos\theta),$$

where θ is the scattering angle. The Feynman diagrams that contribute in the theory with pure gauge interactions (no Higgs field) are



Consider first the s-channel diagrams (1) and expand in $m_{W,Z}^2/E^2$:

$$\begin{aligned}\mathcal{A}^{s}(W_{\mathrm{L}}W_{\mathrm{L}} \to W_{\mathrm{L}}W_{\mathrm{L}}) &= \frac{-ie^2}{4}\cos\theta\left(\frac{s^2}{m_W^4} + \mathcal{O}(s^0)\right) - \frac{ig^2\cos\theta_{\mathrm{W}}}{4}\cos\theta\left(\frac{s^2}{m_W^4} + \frac{m_Z^2 s}{m_W^2} + \mathcal{O}(s^0)\right) \\ &= -\frac{ig^2}{4}\cos\theta\left(\frac{s^2}{m_W^4} + \frac{s}{m_W^2} + \mathcal{O}(s^0)\right).\end{aligned}$$

One should distinguish between the scattering angle θ and the Weinberg angle θ_{W}. The centre-of mass energy simply relates to the s variable $s = (p_1 + p_2)^2 = 4E^2$, and the electron charge relates to the SU(2) coupling via $e = g\sin\theta_{\mathrm{W}}$. Similar steps can be taken to calculate the t-channel diagram and the four-gauge-boson interaction:

$$\begin{aligned}\mathcal{A}^{s+t\text{-ch}}(W_{\mathrm{L}}W_{\mathrm{L}} \to W_{\mathrm{L}}W_{\mathrm{L}}) &= -ig^2\left(\frac{s^2}{m_W^4}(-3 + 6\cos\theta + \cos^2\theta) + \frac{s}{8m_W^2}(3 - 12\cos\theta) + \mathcal{O}(s^0)\right), \\ \mathcal{A}^{(3)}(W_{\mathrm{L}}W_{\mathrm{L}} \to W_{\mathrm{L}}W_{\mathrm{L}}) &= +ig^2\left(\frac{s^2}{m_W^4}(-3 + 6\cos\theta + \cos^2\theta) + \frac{2}{4m_W^2}(2 - 6\cos\theta) + \mathcal{O}(s^0)\right).\end{aligned}$$

The leading term for $E \gg m_W$ cancels and the total expression gives

$$\mathcal{A}^{(1)+(2)+(3)}(W_{\mathrm{L}}W_{\mathrm{L}} \to W_{\mathrm{L}}W_{\mathrm{L}}) = ig^2\frac{s}{4m_W^2}\frac{1 + \cos\theta}{2} = ig^2\frac{s+t}{4m_W^2} + \mathcal{O}(s^0).$$

This still grows with the scattering energy. Such behaviour *violates the unitarity of the scattering matrix*, more precisely perturbative unitarity. That is, unitarity is violated, or loop corrections must be as large as the tree diagrams, i.e. the interaction must become strong to restore unitarity.

4.2.1 Partial wave unitarity

Unitarity of the scattering operator $S = \mathbb{1} + iT$ requires that

$$\mathbb{1} = S^\dagger S = \mathbb{1} + i(T - T^\dagger) + T^\dagger T, \quad \Longrightarrow \quad T^\dagger T = i(T^\dagger - T). \tag{4.3}$$

This is called the *optical theorem* for the scattering matrix element T. For $2 \to 2$ scattering and neglecting the masses of external particles relative to s, t the differential cross-section reads

$$\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \frac{1}{64\pi^2 s}|T(s,t)|^2.$$

Applying the optical theorem gives

$$\sigma = \frac{1}{2s}T^\dagger T = \frac{1}{2s}(-i)(T - T^\dagger) = \frac{1}{s}\mathrm{Im}\,(T(\theta = 0)).$$

The expression $T^\dagger T$ represents the square of the amplitude over all intermediate states, and $T(\theta = 0)$ is the *forward* scattering amplitude since one takes $\langle i|T^\dagger T|i\rangle$ in the same state. Recall also the phase space integral $1/(8\pi)$ which explains the difference between $1/(2s)$ and $4\pi/(64\pi^2 s)$. In short, optical theorem relates the square of the scattering amplitude to the forward scattering amplitude.

Now perform a partial wave decomposition on the scattering matrix element:

$$T(s,t) = 16\pi \sum_{j=0}^{\infty}(2j+1)a_j(s)P_j(\cos\theta)$$

and compute σ in two different ways, from both sides of the optical theorem. Remember the orthogonality relation between the Legendre polynomials

$$\int_{-1}^{1} \mathrm{d}\,\cos\theta\, P_j(\cos\theta)P_k(\cos\theta) = \frac{2}{2j+1}\delta_{jk}.$$

The first calculation is via the square of the amplitude and the second calculation (in the last line) uses the forward scattering

$$\begin{aligned}
\sigma &= \int \mathrm{d}\Omega\, \frac{1}{64\pi^2 s}|T(s,t)|^2 \\
&= \frac{(16\pi)^2 2\pi}{64\pi^2 s}\int_{-1}^{1}\mathrm{d}\,\cos\theta \sum_{j,k=0}^{\infty}(2j+1)(2k+1)a_j(s)a_k^*(s)P_j(\cos\theta)P_k(\cos\theta) \\
&= \frac{16\pi}{s}\sum_{j=0}^{\infty}(2j+1)|a_j(s)|^2; \\
\sigma &\overset{4.3}{=} \frac{16\pi}{s}\sum_{j=0}^{\infty}(2j+1)\mathrm{Im}\,(a_j(s)).
\end{aligned}$$

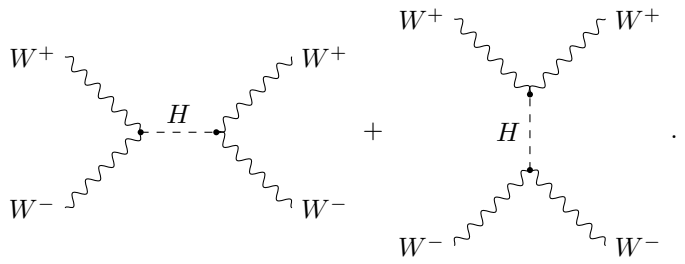
This gives an important condition $|a_j(s)|^2 = \mathrm{Im}\,(a_j(s))$, or alternatively,

$$(\mathrm{Re}\,(a_j(s)))^2 + \left(\mathrm{Im}\,(a_j(s)) - \frac{1}{2}\right)^2 = \frac{1}{4} \quad \Longrightarrow \quad |\mathrm{Re}\; a_j(s)| < \frac{1}{2} \quad \text{and} \quad |a_j(s)| < 1.$$

No partial wave can grow with energy indefinitely.

4.2.2 Unitarization by Higgs exchange

One needs a mechanism that unitarizes the $W_\mathrm{L}W_\mathrm{L}$ scattering amplitude. In the SM this comes through Higgs exchange:



We will discuss the vertices of the Higgs boson in detail below. The Higgs contribution to the scattering amplitude gives

$$\mathcal{A}^{\mathrm{Higgs}}(W_\mathrm{L}W_\mathrm{L}\to W_\mathrm{L}W_\mathrm{L}) = ig^2\left(-\frac{s+t}{4m_W^2}-\frac{m_H^2}{4m_W^2}\left(2+\frac{m_H^2}{s-m_H^2}+\frac{m_H^2}{t-m_H^2}\right)+\mathcal{O}(s^0m_H^0)\right).$$

The first term cancels precisely the growing term in the pure gauge-boson diagrams. Using $m_W^2 = g^2v^2/4$, the total amplitude can be expressed as

$$iT = \mathcal{A}^{W_\mathrm{L}W_\mathrm{L}\to W_\mathrm{L}W_\mathrm{L}}(s,t) = -i\frac{m_H^2}{v^2}\left(2+\frac{m_H^2}{s-m_H^2}+\frac{m_H^2}{t-m_H^2}\right). \tag{4.4}$$

The prefactor can be exchanged for $-i\lambda/2$ where λ is the Higgs self-coupling, i.e. gauge coupling completely drops out! For the s-wave amplitude ($j=0$)

$$a_0(s) = \frac{1}{16\pi}\frac{1}{2}\int_{-1}^{1}\mathrm{d}\,\cos\theta\,\frac{1}{i}\mathcal{A}^{W_\mathrm{L}W_\mathrm{L}\to W_\mathrm{L}W_\mathrm{L}}(s,t) \overset{s,t\gg m_h}{\approx} -\frac{m_H^2}{8\pi v^2}.$$

Hence, $|\mathrm{Re}\,a_0| < 1/2$ implies $m_h^2 < 4\pi v^2$, $m_H < 870\,\mathrm{GeV}$. Higgs exchange unitarizes longitudinal gauge boson scattering if its mass is not too large. With $m_H^2 = \lambda v^2/2$ the perturbative unitarity bound on the Higgs mass is equivalent to $\lambda < 8\pi$, i.e. for large m_H the Higgs self-interaction becomes non-perturbatively strong.

This is also seen in the Higgs decays. For $m_H \gtrsim 2m_W, 2m_Z$ the dominant decay modes are $H\to W^+W^-, Z^0Z^0$. The decay rates are

$$\Gamma(H\to W^+W^-) = \frac{g^2}{4\pi}\frac{m_H}{16}\frac{m_H^2}{m_W^2},\quad \Gamma(H\to Z^0Z^0) = \frac{g^2}{4\pi}\frac{m_H}{32}\frac{m_H^2}{m_Z^2}.$$

There is enhancement of $(m_H/m_{W,Z})^2$ from decay into longitudinal gauge bosons. The total decay width is approximately

$$\Gamma_H \approx \frac{g^2}{4\pi}\frac{3m_H^2}{32m_W^2} = \frac{3}{32\pi}\frac{m_H^3}{v^2}.$$

Hence, the total width is as large as the mass m_H itself if

$$m_H \gtrsim \sqrt{\frac{32\pi}{3}}v \approx 1.4\,\mathrm{TeV}.$$

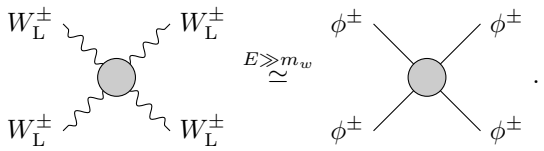
In this case one cannot speak of a Higgs particle or resonance any more. (A particle is a δ-function contribution to the spectral function, a resonance a narrow peak. When $\Gamma \gtrsim m$, there is no peak any more which one could associate with a long-lived particle.) Again, this occurs when the Higgs self-coupling is non-perturbatively strong.

4.2.3 Goldstone boson equivalence theorem

Both the unitarity constraint and the growth of the Higgs boson decay width are ultimately related to the Higgs self-interaction. This appears unexpected since neither $W_{\mathrm{L}}W_{\mathrm{L}} \to W_{\mathrm{L}}W_{\mathrm{L}}$ nor $H \to WW$ involve the Higgs self-coupling. The relation is clarified by the following.

Theorem 4.2.1 (*Goldstone boson equivalence theorem*). *In the high-energy limit $E/m_W \to \infty$ an external W_{L} can be replaced by the pseudo-Goldstone boson, which supplies the longitudinal gauge boson polarization state.*

In diagrammatic form, the theorem reads



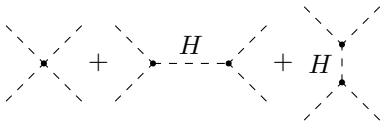
We set of the gauge fields to 0 and compute the scattering process in the linear Σ-model (eq. 4.1) with $\phi = (\phi^+, (v + H + i\phi^0)(\sqrt{2}))$.

Example 3. Instead of $W_{\mathrm{L}}^+W_{\mathrm{L}}^- \to W_{\mathrm{L}}^+W_{\mathrm{L}}^-$ one could calculate the amplitude for $\phi^+\phi^- \to \phi^+\phi^-$, where the interaction vertices follow from

$$\mu^2\phi^\dagger\phi - \frac{\lambda}{4}(\phi^\dagger\phi) \supset -\frac{m_H^2}{2v^2}(\phi^+\phi^-)^2 - \frac{m_H^2}{v}H\phi^+\phi^-.$$

The amplitude

$$\mathcal{A}(\phi^+\phi^- \to \phi^+\phi^-) =$$



$$= (-i)\frac{2m_H^2}{v^2} + (-i)\frac{m_H^2}{v}\left(\frac{i}{s-m_H^2} + \frac{i}{t-m_H^2}\right)(-i)\frac{m_H^2}{v}$$

is in agreement with the previous result (eq. 4.4). The only coupling here is the Higgs self-coupling.

4.3 Aspects of the Standard model Higgs sector

4.3.1 Higgs couplings in the SM (and beyond)

Recall that the SM Higgs sector assumes the simplest possible elementary field and SU(3) × SU(2) × U(1)$_Y$ representation required to break SU(2) × U(1)$_Y$ spontaneously. The representation of ϕ is $(\mathbf{1}, \mathbf{2}, 1/2)$, which means a singlet under SU(3), a doublet under SU(2) and with hypercharge 1/2. This determines the Lagrangian, which contains all terms compatible with the symmetry:

$$\begin{aligned}\mathcal{L}_{\mathrm{SM}} = (\mathrm{D}_\mu\phi)_n^\dagger(\mathrm{D}^\mu\phi)_n + \mu^2\phi_n^\dagger\phi_n - \frac{\lambda}{4}(\phi_n^\dagger\phi_n)^2 + \\ - \overline{L}_{in}\lambda_i^E\phi_n E_i - \overline{Q}_{in}\lambda_i^D\phi_n D_i - \overline{Q}_{in}\lambda_i^U\tilde{\phi}_n U_i + \mathrm{h.c.} + \mathcal{L}^{(5)} + \mathcal{L}^{(6)}.\end{aligned}$$

The dimension-5 and 6 operators can be added if we view the SM as an effective theory of some more fundamental theory up to a scale Λ. The index n represents the SU(2) component (1 for neutrino or up-type quark and 2 for electron-type lepton or down-type quark), i the generation (goes from 1 to 3, summation is implied over all occurrences of the index) and α the spinor component (1 or 2, because we are working with Weyl spinors). The index α is suppressed since it is evident from the structure of the terms that all spinor structures gives a Lorentz scalar.

- $L_{in\alpha}$ is the vector of left-handed lepton doublets with $Y = -1/2$:
$$L = \left(\begin{pmatrix} \nu_{e,\mathrm{L}} \\ e_{\mathrm{L}} \end{pmatrix}, \begin{pmatrix} \nu_{\mu,\mathrm{L}} \\ \mu_{\mathrm{L}} \end{pmatrix}, \begin{pmatrix} \nu_{\tau,\mathrm{L}} \\ \tau_{\mathrm{L}} \end{pmatrix} \right).$$
- $E_{i\alpha}$ is the vector of right-handed leptons with $Y = -1$: $E = (e_{\mathrm{R}}, \mu_{\mathrm{R}}, \tau_{\mathrm{R}})$.
- $Q_{in\alpha}$ is the vector of left-handed quarks with $Y = +1/6$:
$$Q = \left(\begin{pmatrix} u_{\mathrm{L}} \\ d_{\mathrm{L}} \end{pmatrix}, \begin{pmatrix} c_{\mathrm{L}} \\ s_{\mathrm{L}} \end{pmatrix}, \begin{pmatrix} t_{\mathrm{L}} \\ b_{\mathrm{L}} \end{pmatrix} \right).$$
- $D_{i\alpha}$ is the vector of right-handed down-like quarks with $Y = -1/3$: $E = (d_{\mathrm{R}}, s_{\mathrm{R}}, b_{\mathrm{R}})$.
- $U_{i\alpha}$ is the vector of right-handed leptons with $Y = +2/3$: $E = (u_{\mathrm{R}}, c_{\mathrm{R}}, t_{\mathrm{R}})$.
- λ_i^E, λ_i^D and λ_i^U are Yukava couplings to leptons and quarks.

Remark.

- There are only a few dimension-5 operators. A particularly interesting one is
$$\mathcal{L}^{(5)} \supset \sum_{i,j=1}^{3} \frac{f_{ij}}{\Lambda} (\overline{L}_{in} \tilde{\phi}_n)(\tilde{\phi}_n^{\mathrm{T}} L_{jn}^{\mathrm{c}}) + \mathrm{h.c.}$$
where L^{c} refers to charge conjugation and i, j are generation indices. Note that $\tilde{\phi}$ and L_i have the same $\mathrm{SU}(2) \times \mathrm{U}(1)_Y$ transformation. In the unitary gauge the operator becomes
$$\sum_{i,j=1}^{3} \frac{f_{ij} v^2}{2\Lambda} \overline{\nu}_{i\mathrm{L}} \nu_{j\mathrm{L}}^{\mathrm{c}} (1 + H/v)^2 + \mathrm{h.c.}$$
This operator, called *Weinberg operator*, generates naturally small Majorana masses ($m_\nu \sim v^2/\Lambda$) for the neutrinos and violates lepton number conservation.
- Possible dimension-6 terms which contain only ϕ are
$$\mathcal{L}_{\mathrm{Higgs}}^{(6)} = -\frac{f_1}{6\Lambda^2} (\phi^\dagger \phi)^3 + \frac{f_2}{2\Lambda^2} \partial_\mu (\phi^\dagger \phi) \partial^\mu (\phi^\dagger \phi) + \frac{f_3}{2\Lambda^2} (\phi^\dagger \mathrm{D}_\mu \phi)^* (\phi^\dagger \mathrm{D}^\mu \phi).$$
The second term has ordinary derivative, since $\phi^\dagger \phi$ is gauge-invariant. It must hold $f_1 > 0$ so that $V(\phi^\dagger \phi)$ is bounded from below.
- $\phi^\dagger \phi$ is the only gauge-invariant operator with mass dimension less than 4 which is a scalar. (The other gauge-invariant operator is $\overline{L}\tilde{\phi}$ with dimension 5/2, but it is a spinor.) Hence, the Higgs (and only the Higgs) can couple to yet unknown gauge-singlet fields at the renormalizable level:
$$\mathcal{L}_{\text{Higgs portal}} = \alpha \phi^\dagger \phi \mathcal{O}$$
where $\mathcal{O}$ is a gauge-invariant dimension-2 operator, e.g. $S^\dagger S$ where S is $(\mathbf{1}, \mathbf{1}, 0)$ under $\mathrm{SU}(3) \times \mathrm{SU}(2) \times \mathrm{U}(1)_Y$.

Couplings of the SM Higgs from renormalizable interactions

The Higgs boson couples to all particles proportional to their masses (or mass squared for bosons), even the self-interaction strength is $\propto m_H^2$. The proportionality factor involves Higgs vacuum expectation value v, which is known from the Fermi constant

$$\phi_0^\dagger\phi_0 = \frac{2\mu^2}{\lambda} = \frac{v^2}{2}, \quad m_H^2 = \frac{\lambda v^2}{2} = 2\mu^2, \quad \frac{G_\mathrm{F}}{\sqrt{2}} = \frac{g^2}{2m_W^2} = \frac{1}{2v^2}.$$

- Higgs self-interaction H^4 and H^3 is proportional to $\lambda = 2m_H^2/v^2$,
- Higgs gauge interaction HWW and $HHWW$ is proportional to $gv^2 = 4m_W^2/v$ (similar for Z),
- Higgs fermion interaction $H\overline{\Psi}\Psi$ is proportional to the Yukawa coupling $\lambda_f = \sqrt{2}m_f/v$.

With $m_H = 125\,\mathrm{GeV}$ now (since 2012) known, all parameters of the renormalizable part of the SM are known. The observed mass is below the unitarity limit and the Higgs particle is a very narrow resonance. Measuring the interaction strength of the H particle is a strong test of the SM, since the presence of non-renormalizable operators changes the above relations.

As an example, consider the effect of $\mathcal{L}^{(6)}_\mathrm{Higgs}$ (in reality there are 59 more dimension-6 operators). Assume that $f_i v^2/\Lambda^2 \ll 1$, so that one can expand to first order. The $(\phi^\dagger\phi)^3$ changes the potential. To determine the new minimum, write

$$\frac{\partial V}{\partial(\phi^\dagger\phi)} = -\mu^2 + \frac{\lambda}{2}\phi^\dagger\phi + \frac{f_1}{2\Lambda^2}(\phi^\dagger\phi)^2 \stackrel{!}{=} 0.$$

If we denote with $v_0 = 2\mu/\sqrt{\lambda}$ the vacuum expectation value in absence of f_1, the corrected expectation becomes

$$2\phi_0^\dagger\phi_0 \equiv v^2 = v_0^2\left(1 - \frac{f_1 v_0^2}{2\lambda\Lambda^2} + \mathcal{O}(f_1^2)\right).$$

In unitary gauge, the Higgs field is $\phi = (0, v + \tilde{H})/\sqrt{2}$. Inserting this into the Lagrangian gives

$$\begin{aligned}\mathcal{L}^{(4)}_\mathrm{SMHiggs} + \mathcal{L}^{(6)}_\mathrm{Higgs} &= \frac{1}{2}\left(1 + \frac{(4f_2 + f_3)v_0^2}{4\Lambda^2} + \dots\right)\partial_\mu\tilde{H}\partial^\mu\tilde{H} - \frac{1}{2}\left(\frac{\lambda v_0^2}{2} - \frac{f_1 v_0^4}{8\Lambda^2} + \dots\right)\tilde{H}\tilde{H} + \\ &\quad + \#\tilde{H}^2 - \left(\frac{\lambda}{16} + \frac{5f_1 v_0^2}{16\Lambda^2}\right)\tilde{H}^4 + \dots \\ &= \frac{1}{2}\partial_\mu H\partial^\mu H - \frac{1}{2}\underbrace{\lambda v_0^2\left(1 - \frac{(f_1/\lambda - 4f_2 - f_3)v_0^2}{4\Lambda^2}\right)}_{m_H^2} HH + \dots\end{aligned}$$

In the last step we rescaled $\tilde{H}$ to H so that the kinetic term became canonically normalized. The coefficient of $\tilde{H}^4$ also changes and the relation $2m_H^2/v = \lambda$ is no longer valid.

Gauge boson masses are also modified by f_3. Since

$$(\mathrm{D}_\mu\phi)^\dagger(\mathrm{D}^\mu\phi) \supset \frac{g^2v^2}{4}W_\mu^- W^{+\mu} + \frac{g^2 + g'^2}{8}v^2 Z_\mu Z^\mu$$

we calculate

$$\begin{aligned}\phi^\dagger(\mathrm{D}_\mu\phi) \supset \frac{v}{\sqrt{2}}(0,1)\left[ig\frac{1}{2}\begin{pmatrix} W_\mu^3 & \sqrt{2}W_\mu^+ \\ \sqrt{2}W_\mu^- & -W_\mu^3\end{pmatrix} + ig'Y_\phi\begin{pmatrix} B_\mu & 0 \\ 0 & B_\mu\end{pmatrix}\right]\frac{v}{\sqrt{2}}\begin{pmatrix}0\\1\end{pmatrix} \\ = \frac{1}{2}v^2(0,1)\begin{pmatrix} igW_\mu^+/\sqrt{2} \\ -i\sqrt{g^2+g'^2}Z_\mu/2\end{pmatrix} = -i\frac{\sqrt{g^2+g'^2}v^2}{4}Z_\mu.\end{aligned}$$

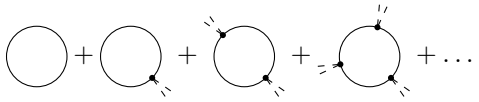
Thus, f_3 contributes to the Z mass, but not the $W^\pm$ mass:

$$m_W^2 = \frac{g^2v^2}{4}, \quad m_Z^2 = \frac{(g^2+g'^2)v^2}{4}\left(1+\frac{f_3v^2}{4\Lambda^2}\right), \quad \Longrightarrow \quad \rho = \frac{m_W^2}{m_Z^2\cos^2\theta_\mathrm{W}} = 1 - \frac{f_3v^2}{4\Lambda^2} + \dots$$

The ρ-parameter is measured to be 1 within few ‰ accuracy. Hence, $f_3v^2/(2\Lambda)^2$ must be very small. Note that the f_3-coupling violates custodial symmetry, contrary to f_1 and f_2. Custodial symmetry guarantees $\rho = 1$. Hence, either Λ is very large or the theory underlying the SM is also approximately custodially symmetric, so $f_3 \ll 1$.

4.3.2 Effective potential and vacuum stability

We want to determine the effective Higgs potential in the SM in the one-loop approximation. Diagrammatically, the effective potential would be calculated as a sum of vacuum bubbles with n insertions of the background field (dashed line) with zero momentum



Here, we use functional rather than diagrammatic methods to compute V_eff, using that the loop expansion corresponds to an expansion in $\hbar$ (page 56, item c)). Consider first a single real scalar field with $-\lambda\phi^4/4!$ interaction. The generating functional reads

$$Z[J] = e^{iW[j]} = N\int\mathcal{D}[\phi]\,\exp\left(iS[\phi] + i\int\mathrm{d}^4x\;J(x)\phi(x)\right).$$

Expand $S[\phi]$ around the field configuration $\varphi(x)$, i.e. $\phi(x) = \varphi(x) + \hat\phi(x)$ and

$$S[\phi] = S[\varphi] + \int\mathrm{d}^4x\,\frac{\delta S}{\delta\phi(x)}\bigg|_\varphi\hat\phi(x) + \frac{1}{2}\int\mathrm{d}^4x\,\mathrm{d}^4y\,\frac{\delta S}{\delta\phi(x)\delta\phi(y)}\bigg|_\varphi\hat\phi(x)\hat\phi(y) + \dots$$

With this, the expression for the effective action $\Gamma[\varphi]$ becomes

$$\begin{aligned}e^{i\Gamma(\varphi)} = N\int\mathcal{D}[\hat\phi]\,\exp\Bigg(iS[\varphi] + i\int\mathrm{d}^4x\,\left(\frac{\delta S}{\delta\phi(x)}\bigg|_\varphi + J(x)\right)\hat\phi(x) + \\ + \frac{1}{2}\int\mathrm{d}^4x\,\mathrm{d}^4y\,\frac{\delta^2S}{\delta\phi(x)\delta\phi(y)}\bigg|_\varphi\hat\phi(x)\hat\phi(y) + \dots\Bigg).\end{aligned}$$

The first term in the expansion simply gives $\Gamma[\varphi] = S[\varphi] + \mathcal{O}(\hbar)$. Since $J(x)$ is chosen to be $-\delta\Gamma[\varphi]/\delta\varphi$, the second term vanishes when one inserts the zeroth order result $\Gamma = S$. This guarantees that the one-point function vanishes, $\langle\Omega|\hat\varphi(x)|\Omega\rangle = 0$. Since we want ti compute $\Gamma(\varphi)$

only in the one-loop approximation, we can neglect this term. The third term is a Gaussian path integral which gives the functional determinant

$$\det\left(-i\frac{\delta^2 S}{\delta\phi(x)\delta\phi(y)}\bigg|_\varphi\right)^{-1/2}.$$

The higher order terms can be treated as perturbations which give diagrams such as  , where the propagators come from the inverse of the two field variations of the action and the vertices come from the three field variations of the action. However, this is a two-loop correction, which we won't consider. We have thus far derived the following expression for the effective action

$$\Gamma[\varphi] = S[\varphi] + i\frac{1}{2}\ln\det\left(-i\frac{\delta^2 S}{\delta\phi(x)\delta\phi(y)}\bigg|_\varphi\right) + \text{higher loops}.$$

For constant field configurations φ, the effective potential follows from the effective action simply by dividing by the volume of the space-time (VT)

$$V_{\text{eff}}^{\text{1-loop}} = V(\varphi) - \frac{i}{2}\frac{1}{VT}\operatorname{tr}\ln\left(-i\frac{\delta^2 S}{\delta\phi(x)\delta\phi(y)}\bigg|_\varphi\right),$$

where the classical effective potential follows directly from the Lagrangian by ignoring the field derivatives: $V(\varphi) = -\mathcal{L}(\varphi)|_{\partial_\mu\varphi=0}$. The determinant and the trace refer to the variables x and y which can be thought of as indices of ϕ_x and ϕ_y. For the ϕ^4 theory it holds

$$-\frac{\delta^2 S}{\delta\phi(x)\delta\phi(y)}\bigg|_\varphi = \delta^{(4)}(x-y)\left(\partial_\mu\partial^\mu + m^2 + \frac{\lambda}{2}\varphi(x)^2\right) \equiv \delta^{(4)}(x-y)\left(\partial^2 + M^2(\varphi(x))\right).$$

This is a diagonal "matrix" in variables x and y, so the trace gives VT. This is better seen in the momentum space representation, where

$$\ln\left(-\frac{\delta^2 S}{\delta\phi(x)\delta\phi(y)}\bigg|_\varphi\right) \xrightarrow{\text{Fourier transformation}} (2\pi)^4\delta^{(4)}(p-q)\ln(-p^2 + M^2(\varphi))$$

for constant φ. To check this, write

$$\begin{aligned}\int\frac{\mathrm{d}^4p}{(2\pi)^4}\frac{\mathrm{d}^4q}{(2\pi)^4}\,e^{ipx}e^{-iqx}(2\pi^4)\delta^{(4)}(p-q)\ln(-p^2+M^2(\varphi)) &= \int\frac{\mathrm{d}^4p}{(2\pi)^4}\,e^{ip(x-y)}\ln(-p^2+M^2(\varphi))\\ &= \delta^{(4)}(x-y)\ln(\partial^2+M^2(\varphi)).\end{aligned}$$

Taking the trace in the momentum space amounts to setting $p = q$ and then integrating over all p:

$$\operatorname{tr}\ln\left(-\frac{\delta^2 S}{\delta\phi(x)\delta\phi(y)}\bigg|_\varphi\right) = \underbrace{(2\pi)^4\delta^{(4)}(0)}_{VT}\int\frac{\mathrm{d}^4p}{(2\pi)^4}\ln(-p^2+M^2(\varphi)).$$

The factor of i in the argument of the logarithm was dropped since it gives an irrelevant constant. Taking the determinant expression into the potential calculation gives

$$V_{\text{eff}}^{\text{1-loop}}(\varphi) = V(\varphi) - \frac{i}{2}\int\frac{\mathrm{d}^4p}{(2\pi)^4}\,\ln(-p^2+M^2(\varphi)).$$

This integral is divergent, so we use dimensional regularization

$$\begin{aligned}\tilde{\mu}^{2\varepsilon}\int\frac{\mathrm{d}^dp}{(2\pi)^d}\,\ln(-p^2+M^2) &= -\frac{\partial}{\partial\alpha}\tilde{\mu}^{2\varepsilon}\int\frac{\mathrm{d}^dp}{(2\pi)^d}\frac{1}{(-p^2+M^2)^\alpha}\bigg|_{\alpha=0}\\ &= -\frac{\partial}{\partial\alpha}\left(\frac{i}{(4\pi)^{d/2}}\frac{\Gamma(\alpha-d/2)}{\Gamma(\alpha)}\frac{\tilde{\mu}^{2\varepsilon}}{(M^2)^{\alpha-d/2}}\right)\bigg|_{\alpha=0}\\ &= \frac{(-i)}{(4\pi)^{d/2}}\Gamma(-d/2)\frac{\tilde{\mu}^{2\varepsilon}}{(M^2)^{-d/2}}\\ &= \frac{i}{32\pi^2}M^4\left(-\frac{1}{\varepsilon}+\ln\frac{M^2}{\mu^2}-\frac{3}{2}\right)+\mathcal{O}(\varepsilon).\end{aligned}$$

When we insert the classical effective potential (Lagrangian at tree level), we should not forget about the counter-terms which cancel the $1/\varepsilon$ divergences introduced with dimensional regularization

$$\begin{aligned}V_{\text{eff}}^{\text{1-loop}} &= \frac{1}{2}m^2\varphi^2+\frac{\lambda}{4!}\varphi^4+\frac{1}{2}\delta m^2\varphi^2+\delta_\lambda\frac{\lambda}{4!}\varphi^4+\frac{(m^2+\lambda\varphi^2/2)^2}{64\pi^2}\left(-\frac{1}{\varepsilon}+\ln\frac{m^2+\lambda\varphi^2/2}{\mu^2}-\frac{3}{2}\right)\\ &= \text{const}+\frac{1}{2}m^2\varphi^2+\frac{\lambda}{4!}\varphi^4+\frac{(m^2+\lambda\varphi^2/2)^2}{64\pi^2}\left(\ln\frac{m^2+\lambda\varphi^2/2}{\mu^2}-\frac{3}{2}\right).\end{aligned}\tag{4.5}$$

We used $\delta m^2=\lambda m^2/(32\pi^2\varepsilon)$ and $\delta_\lambda=3\lambda/(32\pi^2\varepsilon)$. There is a divergent, field-independent zero-point energy proportional to m^4. The constant $3/2$ at the end depends on the renormalization scheme – we used $\overline{\text{MS}}$. The result is finite as it should be when expressed in terms of renormalized parameters (up to irrelevant vacuum energy). The result 4.5 is called *Coleman–Weinberg potential* (named after Eric Weinberg, not Steven).

Generalization to the Standard Model

There are a number of subtleties when calculating the effective potential of the standard model, since we have a gauge theory. However, we will be able to rewrite everything in terms of a scalar loop correction, so it is convenient to define

$$f(x^2)=\frac{x^4}{64\pi^2}\left(-\frac{1}{\varepsilon}+\ln\frac{x^2}{\bar{\mu}^2}-\frac{3}{2}\right).$$

The $\overline{\text{MS}}$ scale has here been denoted by $\bar{\mu}$ in order to avoid confusion with the quadratic Higgs coupling μ. Since the Higgs field is the only scalar in the SM, it is the only field that is allowed by Lorentz invariance to acquire a non-zero vacuum expectation value

$$\boldsymbol{\phi}=\frac{1}{\sqrt{2}}\begin{pmatrix}\phi_1+i\phi_2\\ \varphi+H+i\chi\end{pmatrix}\quad\Longrightarrow\quad\langle\Omega|\boldsymbol{\phi}|\Omega\rangle=\boldsymbol{\varphi}=\frac{1}{\sqrt{2}}\begin{pmatrix}0\\ \varphi\end{pmatrix}.$$

The background field is thus a real scalar φ which we will use from now on with replacement $\boldsymbol{\varphi}^\dagger\boldsymbol{\varphi}=\varphi^2/2$.

- There are four real quantum fields in the vector $\boldsymbol{\phi}$ that can run in the vacuum bubble. Three of them, the pseudo-Goldstone bosons χ,ϕ_1,ϕ_2 give the mass term $M_{\text{G}}(\varphi)=-\mu^2+\lambda\varphi^2/4$ and the Higgs boson gives mass $M_H^2(\varphi)=-\mu^2+3\lambda\varphi^2/4$. Notice that setting $\varphi=v$ gives $M_H^2(v)=m_H^2$ and $M_{\text{G}}^2(v)=0$. It is tempting to say that M_{G} is the Goldstone mass matrix, however after one-loop corrections, both v and M_G are changed, but Goldstones remain massless.

- The kinetic term of the Higgs field gives masses to gauge bosons

$$M_W^2(\varphi) = \frac{g^2\varphi^2}{4}, \quad M_Z^2(\varphi) = \frac{g^2+g'^2}{4}\varphi^2, \quad M_A^2(\varphi) = 0.$$

  The rotation from B_μ, W_μ^3 to Z_μ, A_μ is independent of the Higgs expectation value, so the photon remains massless.

- For the gauge fixing one needs to use a renormalizable gauge, because the limits $\xi \to \infty$ and $k \to \infty$ in loops do not commute. After introducing the gauge-fixing term that cancels ϕ-A_μ mixing (eq. 4.2) one obtains the following field content
  - $d-1$ modes for each massive gauge boson with mass $M_i^2(\varphi)$ as before (for $i = W, Z$),
  - one scalar polarization for each massive gauge boson with mass $\xi_i M_i^2(\varphi)$,
  - one would-be Goldstone boson for each massive gauge boson, with a corrected mass $M_G^2(\varphi) + \xi_i M_i^2(\varphi)$,
  - a ghost and anti-ghost for each massive gauge boson with mass $\xi_i M_i^2(\varphi)$ (with a minus sign because of the Grassmann determinant).

  In total, the contributions are

$$(d-1)f(M_i^2) + f(M_G^2 + \xi_i M_i^2) + f(\xi_i M_i^2) - 2f(\xi_i M_i^2)$$

  per broken generator. The effective potential is gauge-dependent, but its minimum value $V_{\text{eff}}(v)$ is a physical observable, so it should be gauge-independent. Thus, we choose the most convenient value for the gauge parameter, namely $\xi_i = 0$ (which is called the *Landau gauge*).

- All massive fermions of the Standard Model also contribute to the effective potential, but since the effect is proportional to m^4, we will only consider the most massive one – the top quark. There are some complications with calculating a Dirac loop, but in the end one obtains $M(\varphi) = m + y\varphi$ for a generically coupled fermion field $\mathcal{L} \supset -m\overline{\Psi}\Psi - y\overline{\Psi}\phi\Psi$. Fermions from the SM only have Yukawa coupling, so for the top quark we write $M_t^2(\varphi) = y_t^2\varphi^2/2$.

For each of the massive particles considered, we need to determine how many degrees of freedom are running in the loop:

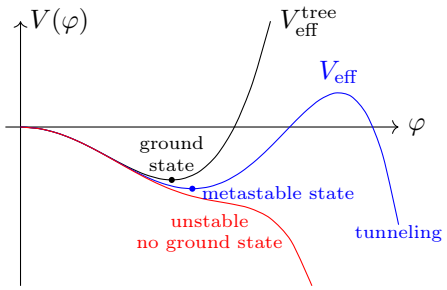
- Higgs boson: 1, would-be Goldstone bosons: 3,
- W boson: $2(3-2\varepsilon)$ (it's a complex field, only non-scalar modes in dimreg $d = 4-2\varepsilon$),
- Z boson: $3-2\varepsilon$ (real field),
- t quark: 12 (3 flavours, 2 spins, particle/antiparticle)

The factor $(3-2\varepsilon)$ in front of the massive gauge bosons changes the constant factor $-3/2$ to $-5/6$. All divergences are cancelled by the counter-terms. Since we are working in $\overline{\text{MS}}$-scheme, all $1/\varepsilon$ divergencies (and only these divergencies) are absorbed into the counter-terms. The

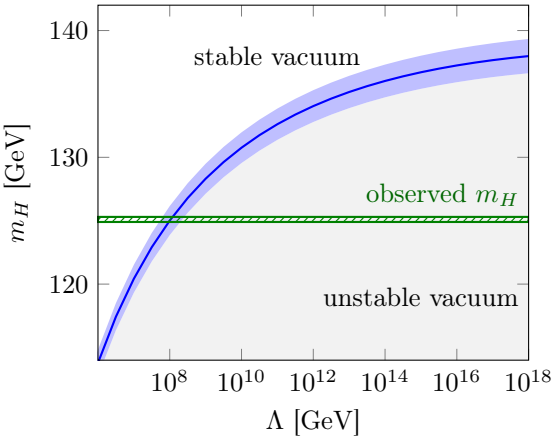
one-loop corrected effective potential for the Standard Model thus reads

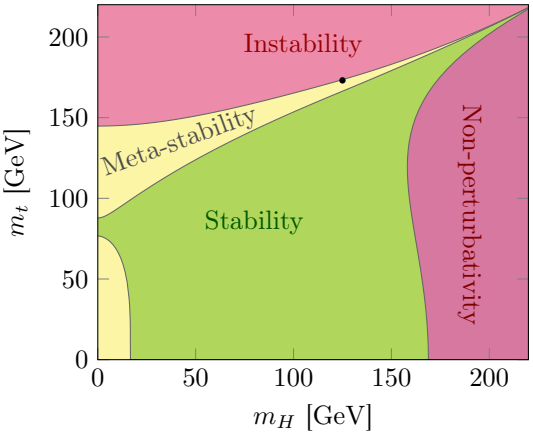
$$\begin{aligned}
V_{\text{eff,SM}}^{\text{1-loop}}(\varphi^2) = &-\frac{\mu^2}{2}\varphi^2 + \frac{\lambda}{16}\varphi^4 + \frac{(-\mu^2+3\lambda\varphi^2/4)^2}{64\pi^2}\left[\ln\left(\frac{-\mu^2+3\lambda\varphi^2/4}{\bar{\mu}^2}\right) - \frac{3}{2}\right] + \\
&+ 3\frac{(-\mu^2+\lambda\varphi^2/4)^2}{64\pi^2}\left[\ln\left(\frac{-\mu^2+\lambda\varphi^2/4}{\bar{\mu}^2}\right) - \frac{3}{2}\right] + 6\frac{g^4\varphi^4}{16\cdot 64\pi^2}\left[\ln\left(\frac{g^2\varphi^2}{4\bar{\mu}^2}\right) - \frac{5}{6}\right] + \\
&+ 3\frac{(g^2+g'^2)^2\varphi^4}{16\cdot 64\pi^2}\left[\ln\left(\frac{(g^2+g'^2)\varphi^2}{4\bar{\mu}^2}\right) - \frac{5}{6}\right] - 12\frac{y_t^4\varphi^4}{4\cdot 64\pi^2}\left[\ln\left(\frac{y_t^2\varphi^2}{2\bar{\mu}^2}\right) - \frac{3}{2}\right] \\
= &-\frac{\mu^2}{2}\varphi^2 + \frac{\lambda}{16}\varphi^4 + \sum_{i\in\{H,\text{G},W,Z,t\}} n_i \frac{M_i^4(\varphi)}{64\pi^2}\left[\ln\left(\frac{M_i^2(\varphi)}{\bar{\mu}^2}\right) - C_i\right].
\end{aligned}$$

Here, n_i is the number of degrees of freedom weighted by a possible minus sign from the fermion loop and C_i are renormalization-scheme dependent constants (in dimreg, 3/2 and 5/6 for bosons and fermions, respectively). The most important contributions are those from the top quark (the other fermions can be neglected; their masses are small) and from the Higgs loop. The top-quark makes a *negative* contribution for large φ and destabilizes the ground state unless its contribution is compensated by bosonic ones, in particular the one proportional to $\lambda^2 v^4 \propto m_H^4$. Qualitatively, this gives the following picture.



If one assumes that the SM is valid up to a scale Λ, then V_{eff} can be modified only for field values $|\varphi| > \Lambda$ by unknown physics. Hence, requiring that metastability or instability does not occur for $|\varphi| < \Lambda$, one obtains constraints for m_t and m_H. At one-loop order with RG improvement, the constraints are as shown on the plots below.





The blue band comes from the top mass uncertainty, $m_t = (173.1 \pm 0.7)\,\text{GeV}$. For a 2-loop

analysis see Degrassi et al. (ArXiv: 1205.6497) and updated analysis by Buttazzo et al. (ArXiv: 1307.3536). For a Higgs mass of 125 GeV the SM cannot be valid up to the Planck scale.

4.3.3 Triviality

Start with ϕ^4 theory with $\mathcal{L}_{\text{int}} = -\lambda\phi^4/4!$. The scale dependent coupling satisfies the renormalization group equation (eq.??)

$$\mu^2 \frac{\mathrm{d}}{\mathrm{d}\mu^2}\lambda = \beta(\lambda) = \frac{3\lambda^2}{32\pi^2} + \mathcal{O}(\lambda^3).$$

Note that λ and $\beta(\lambda)$ depend on the renormalization scheme but not $\beta_0 = 3/(32\pi^2)$ for $\mu \gg m$. The factor of β_0 comes from the divergent parts of the one-loop four-point diagrams  ,  and  . Then we integrate form energy scale μ_1 to μ_2

$$\ln\frac{\mu_2^2}{\mu_1^2} = \int_{\lambda(\mu_1)}^{\lambda(\mu_2)} \frac{\mathrm{d}\lambda}{\beta(\lambda)}. \tag{4.6}$$

At one-loop order, the running of the coupling is described by

$$\lambda(\mu_2) = \frac{\lambda(\mu_1)}{1 - \beta_0\lambda(\mu_1)\ln(\mu_2^2/\mu_1^2)}.$$

This expression leads to a singularity at energy Λ_{LP} called the *Landau pole*. If this singularity exists, the theory can be valid only at energy scales $E < \Lambda_{\text{LP}}$, despite being renormalizable. However, since $\lambda(\mu) \sim \mathcal{O}(1)$ near $\mu = \Lambda_{\text{LP}}$, this equation cannot be answered in perturbation theory. There are several options.

- There exists λ_* with $\beta(\lambda_*) = 0$, such that $\lambda(\mu_2) \to \lambda_*$. Because of 4.6 this requires $\ln(\mu_2^2/\mu_1^2)$ to grow to infinity, hence $\mu_2 \to \infty$. In other words, $\lim_{\mu\to\infty}\lambda(\mu) = \lambda_*$ and λ_* is a UV fixed point (p.??).
- The integral on the right-hand side of 4.6 diverges when $\lambda(\mu_2) \to \infty$, i.e., $\lim_{\mu\to\infty}\lambda(\mu) = \infty$, but remains finite for every finite μ.
- The integral converges for any $\lambda(\mu_2)$ including $\lambda(\mu_2) = \infty$. But in the limit $\mu_2 \to \infty$ the logarithm on the left-hand side also diverges and hence 4.6 has no solution for large μ_2. In this case the theory cannot be defined consistently for all scales.

Non-perturbative simulations of ϕ^4 theory on a space-time lattice suggest that the third option is realized for this theory. One also says that the theory is *trivial*, since the only solution which does not lead to an inconsistency is $\lambda(\mu) = 0$ for all μ, i.e. a free theory.

The Standard Model Higgs sector in the approximation that the coupling of the Higgs field to all other fields is neglected, has the interaction $-\lambda(\phi^\dagger\phi)^2/4$ and the RGE is

$$\mu^2\frac{\mathrm{d}\lambda}{\mathrm{d}\mu^2} = \frac{3}{16\pi^2}\lambda^2, \quad \Lambda_{\text{LP}} = v\exp\left(\frac{1}{2\beta_0\lambda(v)}\right) = v\exp\left(\frac{4\pi^2v^2}{3m_H^2(v)}\right).$$

Scale dependent $\lambda(\mu)$ defines a scale-dependent Higgs mass $m_H^2(\mu)$. The minimal requirement is that the SM should be consistent up to the scale v, i.e. there should be no Landau pole in the Higgs coupling for $\mu \lesssim v$. Above the scale v there should be no Landau pole either but if

the SM is replaced by a more fundamental theory, the new interactions will change the evolution equation:

$$\Lambda_{\mathrm{LP}} \gtrsim v \xrightarrow{\text{exponent}\gtrsim 1} m_H^2(v) \lesssim \frac{4\pi^2 v^2}{3} \sim (1\,\mathrm{TeV})^2.$$

If the SM should be valid to high scales, the limit on the Higgs mass decreases. Hence, as before ($W_{\mathrm{L}}W_{\mathrm{L}}$ scattering), the Higgs mass cannot be too large in order that the self-coupling does not become too large.

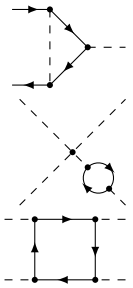
A more precise calculation includes the other couplings of the SM and considers the system of RGEs:

$$\mu^2 \frac{\mathrm{d}}{\mathrm{d}\mu^2} g_i = f_i(g_j).$$

In the present context, the most important other coupling is the Yukawa coupling of the top quark to the Higgs boson, $m_t = \lambda_t v/\sqrt{2}$. Since $m_t = 173\,\mathrm{GeV}$ and $v = 246\,\mathrm{GeV}$, the Yukava coupling is $y_t \approx 1$. The system of equations is

$$\mu^2 \frac{\mathrm{d}\lambda_t^2}{\mathrm{d}\mu^2} = \frac{9}{32\pi^2}\lambda_t^4,$$
$$\mu^2 \frac{\mathrm{d}\lambda}{\mathrm{d}\mu^2} = \frac{3}{16\pi^2}(\lambda^2 + 2\lambda\lambda_t^2 - 4\lambda_t^4).$$

The various contributions to the RGE equations come from the following diagrams, where all solid lines represent top quark and all dashed lines Higgs boson:

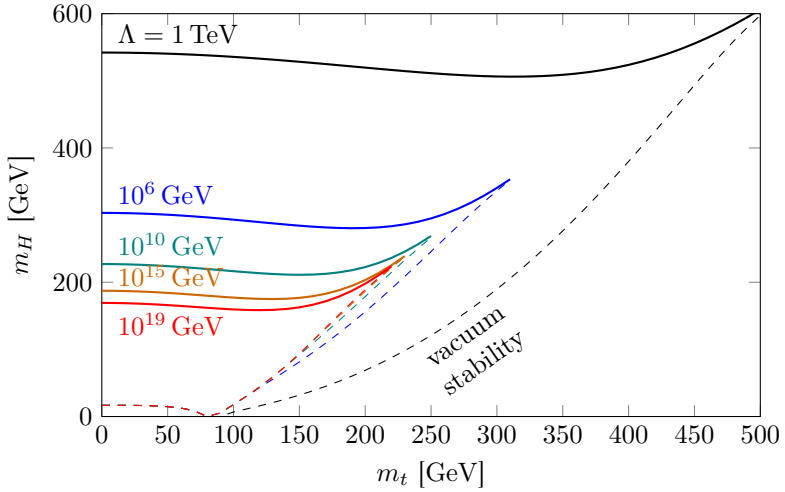


λ_t^4 term in the top quark RGE,

$\lambda\lambda_t^2$ term in the Higgs self-coupling RGE,

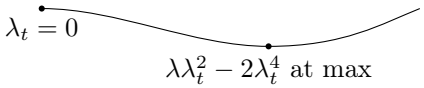
λ_t^4 term in the Higgs self-coupling RGE,

Requiring that no Landau pole appears for $\mu \leq \Lambda$, i.e. the SM is valid up to scale Λ one obtains the following limit on m_H as a function of m_t. (note that $\lambda, \lambda_t \leftrightarrow m_H, m_t$).



One usually does not consider $\Lambda > M_{\mathrm{Pl}}$ since one assumes that the SM should certainly be modified above the Planck scale. We can understand the shape of the triviality bound qualitatively.

- Start with small λ_t, increasing it. Then $\lambda\lambda_t^2 - 2\lambda_t^4 > 0$ and growing. Hence, λ grows faster, the bound becomes stronger, i.e. negative slope.
- At certain point for λ_t, $\lambda\lambda_t^2 - 2\lambda_t^4$ stops growing, then the $-\lambda_t^4$ becomes dominant and $\lambda\lambda_t^2 - 2\lambda_t^4 < 0$. The triviality bound weakens. Hence, the shape



- For even larger λ_t, $\lambda^2 + 2\lambda\lambda_t^2 - 4\lambda_t^4$ becomes negative and hence as μ increases $\lambda(\mu)$ can become negative. This is a manifestation of vacuum instability (since $\lambda(\mu)\phi^4$ is then unbounded from below) which has already been discussed (dashed line).

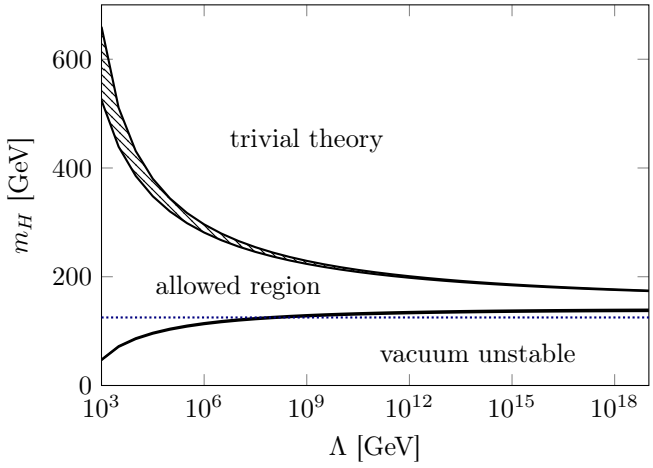
The simple system of equations gives correct qualitative behaviour. For more accurate results, one needs to consider the running of all involved couplings $(g, g', g_{\mathrm{s}}, \lambda, y_t)$ and the correction to the relation between y_t and m_t

$$\mu^2\frac{\mathrm{d}g'}{\mathrm{d}\mu^2} = \frac{1}{32\pi^2}\frac{41}{6}g'^3, \qquad \mu^2\frac{\mathrm{d}g}{\mathrm{d}\mu^2} = -\frac{1}{32\pi^2}\frac{19}{6}g^3, \qquad \mu^2\frac{\mathrm{d}g_{\mathrm{s}}}{\mathrm{d}\mu^2} = \frac{7}{32\pi^2}g_{\mathrm{s}}^3,$$

$$\mu^2\frac{\mathrm{d}y_t}{\mathrm{d}\mu^2} = \frac{1}{32\pi^2}y_t\left(\frac{9}{2}y_t^2 - 8g_{\mathrm{s}}^2 - \frac{9}{4}g^2 - \frac{17}{12}g'^2\right), \qquad y_t = \frac{\sqrt{2}m_t}{v}\left(1 - \frac{4\alpha_{\mathrm{s}}}{3\pi}\right)$$

$$\mu^2\frac{\mathrm{d}\lambda}{\mathrm{d}\mu^2} = \frac{3}{32\pi^2}\left(2\lambda^2 - 8y_t^4 + g^4 + \frac{3}{2}(g^2 + g'^2)^2 - \lambda(3g^2 + g'^2 - 4y_t^2)\right).$$

All the plots in this section were made using this set of equations. For $m_t = (173.1 \pm 0.7)\,\mathrm{GeV}$, this gives the following picture.



For given Λ, m_H is constrained from above and below. At $m_H \sim (130\text{–}175)\,\mathrm{GeV}$ the SM could have been a valid theory up to the Planck scale. For $m_H = 125\,\mathrm{GeV}$, triviality is not an issue, the self-coupling λ is small enough. But vacuum stability constrains Λ to be below the Planck scale.

Chapter 5

Infrared divergences and the strong interaction in high-energy processes

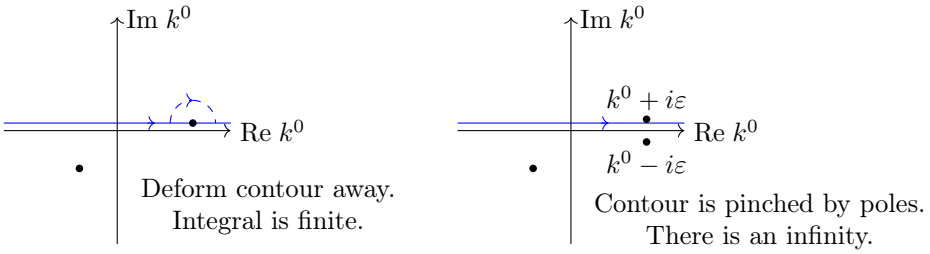
Renormalization deals with the behaviour of quantum fluctuations in the ultraviolet region, $k^\mu \to \infty$. The essence of renormalization is that observables are insensitive to the UV region when expanded in terms of suitably defined renormalized parameters. In this chapter we assume that this has been done. Instead, we consider the regions when some integration variables become small. This generically leads to *infrared (IR) divergences* – the aim is to understand their physical origin. This is crucial to define IR-finite observables in QED and QCD.

A generic Feynman amplitude has the structure

$$I(p_i, m_j) = \int \prod_{\ell=1}^{L} \frac{\mathrm{d}^4 k_\ell}{(2\pi)^4} N(p_i, m_j, k_\ell) \prod_{n=1}^{I} \frac{i}{q_n^2 - m_n^2 + i\varepsilon},$$

where p_i are external momenta, m_j are masses, $N(p_i, m_j, k_\ell)$ is some numerator, the last term is line integral propagator with $g_n = q_n(p_i, k_\ell)$ and the loop integrals include phase space integrals if we allow N to also contain $2\pi\theta(k_\ell^0)\delta(k_\ell^2 - m_\ell^2)$.

Since we are not interested here in UV behaviour $k_\ell \to \infty$, singularities of I can only occur if the integration over k_ℓ goes through a pole of one of the propagators, i.e. $q_n^2(p_i, k_\ell) = m_n^2$ for some n. Because of the $+i\varepsilon$ the integrals over k_ℓ^0 and $\mathbf{k}_\ell$ should be viewed as complex integrations of an analytic function. Hence, the pole is harmless if the contour of integration can be deformed away from the pole. Singularities precisely occur when this is not possible.

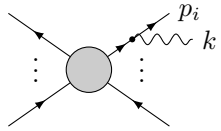


This can be pursued further mathematically with Landau equations. We will not do this here but proceed by examples that highlight the physical interpretation.

5.1 Soft and collinear divergences in QED

5.1.1 Soft photon emission amplitude

Consider a generic non-radiative scattering amplitude for $i \to f$ with only external electron and positron lines. The symbol ◯ will symbolize the non-radiative process (no photons emitted). Now consider adding a photon emission (with momentum k) from line i:



$$= \mathcal{A}(i \to f + \gamma) = \overline{u}(p_i) ie Q_i \gamma^\mu \frac{i(\not{p}_i + \not{k} + m)}{(p_i + k)^2 - m^2 + i\varepsilon} \mathcal{A}(p_i + k, \dots)$$

where Q_i is the electric charge of line i and $\mathcal{A}(p_i + k, \dots)$ is the non-radiative process. Now assume the photon is soft, i.e. $k \to 0$ or $k \ll p_i$. Then using Dirac equation $\overline{u}(-\not{p}_i + m) = 0$ and $p_i^2 = m^2$ gives

$$\overline{u}(p_i) \frac{\gamma^\mu(\not{p}_i + \not{k} + m)}{(p_i + k^2) - m^2 + i\varepsilon} = \overline{u}(p_i) \frac{2p_i^\mu + \gamma^\mu \not{k}}{2p_i k + k^2 + i\varepsilon} \approx \overline{u}(p_i) \frac{p_i^\mu}{p_i \cdot k + i\varepsilon}.$$

Adding the contributions from all external legs results in the following expression

$$\mathcal{A}^{i \to f+\gamma}(p_i, k) \overset{k \to 0}{=} \mathcal{A}^{i \to f}(p_i) \sum_j \frac{(-eQ_j) p_j^\mu}{\eta_j p_j \cdot k + i\varepsilon} \tag{5.1}$$

were $\eta_j = +1$ for an outgoing line (propagator $p_j + k$) and -1 for an ingoing line (propagator $p_j - k$). Now compute the cross-section for the radiative process $i \to f + \gamma$ with the additional requirement that the photon energy is smaller than certain detection threshold, $k^0 = |\mathbf{k}| < \Delta E_\gamma$:

$$\begin{aligned} \mathrm{d}\sigma_{\Delta E_\gamma}^{i \to f+\gamma}(p_i) &= (\text{flux}) \prod_{i \text{ out}} \frac{\mathrm{d}^3\mathbf{p}_i}{(2\pi)^3} \frac{1}{2E_i} \int \frac{\mathrm{d}^3\mathbf{k}}{(2\pi)^3} \frac{1}{2|\mathbf{k}|} (2\pi)^4 \delta^{(4)} \left(\sum_{\text{out}} p_i + k - \sum_{\text{in}} p_i \right) |A^{i \to f+\gamma} \varepsilon_\mu(k)|^2 \\ &= \mathrm{d}\sigma^{i \to f}(p_i) \sum_{i,j} e^2 Q_i Q_j \underbrace{\int_{|\mathbf{k}| < \Delta E_\gamma} \frac{\mathrm{d}^3\mathbf{k}}{(2\pi)^3} \frac{1}{2|\mathbf{k}|} \frac{(-1)(p_i \cdot p_j)}{\eta_i (p_i \cdot k) \eta_j (p_j \cdot k)}}_{I_{ij}}. \end{aligned}$$

The negative sign in the numerator comes from the photon polarization sum, which gives $-g_{\mu\nu}$.

Note. This can be generalized to an arbitrary number of emissions of photons with the remarkable result of exponentiation

$$\mathrm{d}\sigma^{i \to f+n\gamma}(p_i) = \mathrm{d}\sigma^{i \to f}(p_i) \exp\left(\sum_{i,j} e^2 Q_i Q_j I_{ij} \right).$$

Let's analyse the integral I_{ij}

$$\begin{aligned} I_{ij} &= -\eta_i \eta_j \int_{|\mathbf{k}| < \Delta E_\gamma} \frac{\mathrm{d}^3\mathbf{k}}{(2\pi)^3} \frac{1}{2|\mathbf{k}|} \frac{p_i \cdot p_j}{(p_i^0 |\mathbf{k}| - \mathbf{p}_i \cdot \mathbf{k})(p_j^0 |\mathbf{k}| - \mathbf{p}_j \cdot \mathbf{k})} \\ &= -\frac{\eta_i \eta_j}{16\pi^3} \int_0^{\Delta E_\gamma} \mathrm{d}|\mathbf{k}| \frac{|\mathbf{k}|^2}{|\mathbf{k}|^3} \underbrace{\int \mathrm{d}\Omega_{\mathbf{n}} \frac{p_i \cdot p_j}{(p_i^0 - \mathbf{p}_i \cdot \mathbf{n})(p_j^0 - \mathbf{p}_j \cdot \mathbf{n})}}_{J_{ij}}. \end{aligned}$$

We have used $\mathbf{k} = |\mathbf{k}|\mathbf{n}$ with $\mathbf{n}^2 = 1$.

- The first integral over the photon energy $\mathbf{k}$ is (infrared)divergent from $k \to 0$. This is the *soft divergence*. The radiative cross-section is infinite! After temporarily introducing an IR cut-off $\Lambda_{\mathrm{IR}} < |\mathbf{k}|$ the divergence becomes logarithmic

$$\int_{\Lambda_{\mathrm{IR}}}^{\Delta E_\gamma} \frac{\mathrm{d}|\mathbf{k}|}{|\mathbf{k}|} = \ln \frac{\Delta E_\gamma}{\Lambda_{\mathrm{IR}}}.$$

- Since $p_{i,j}^2 = \sqrt{m^2 + \mathbf{p}_{i,j}^2}$ the angular integral is finite for $m^2 > 0$ because $p_{i,j}^0 > |\mathbf{p}_{i,j} \cdot \mathbf{n}|$. Consider however the case $m = 0$, and let θ_i be the angle between $\mathbf{p}_i$ and $\mathbf{n}$. Then

$$J_{ij} = \int_{-1}^{1} \mathrm{d}\,\cos\theta_i \int_0^{2\pi} \mathrm{d}\varphi \, \frac{p_i \cdot p_j}{(1 - \cos\theta_i)|\mathbf{p}_i|(p_j^0 - \mathbf{p}_j \cdot \mathbf{n})} = \infty.$$

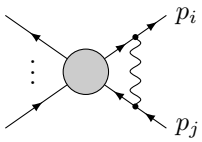
  The divergence arises from $\cos\theta_i = 1$, i.e. $\mathbf{k}$ being parallel to $\mathbf{p}_i$. Similarly holds for $\mathbf{k} \parallel \mathbf{p}_j$. This is the *collinear divergence*. It is also logarithmic.

Note. Although we derived this in the soft photon approximation, the collinear divergence arises also when $|\mathbf{k}|$ is not small (since $k^2 = 0$ in the denominator gives the same). The only requirement is that $\mathbf{k}$ is collinear to $|\mathbf{p}|$. Although there is no literal divergence when $m \neq 0$, the collinear singularity still leads to a large enhancement of the amplitude for scattering energies $p_i^0 \gg m$. The explicit evaluation gives

$$J_{ij} = \frac{2\pi}{\beta_{ij}} \ln \frac{1 + \beta_{ij}}{1 - \beta_{ij}} \overset{m^2 \ll p_i \cdot p_j}{\approx} 2\pi \ln \frac{(p_i \cdot p_j)^2}{m^4}, \quad \beta_{ij} = \sqrt{1 - \frac{m^4}{(p_i \cdot p_j)^2}}.$$

5.1.2 Virtual corrections and the Bloch–Nordsieck cancellation

To understand the interpretation of the above result, consider the virtual (loop) correction to the non-radiative process.



We can construct this from the soft radiative amplitude 5.1 by tying two emissions together with a photon propagator $-ig_{\mu\nu}/(k^2 + i\varepsilon)$:

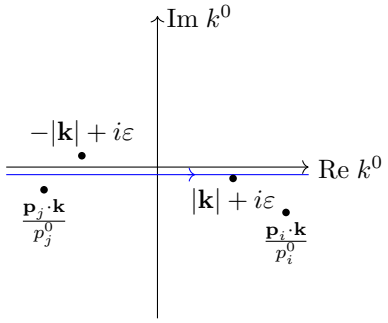
$$\mathrm{d}\sigma_{\text{1-loop}}^{i \to f}(p_i) = \mathrm{d}\sigma^{i \to f}(p_i) \sum_{i,j} e^2 Q_i Q_j \frac{1}{2}(V_{ij} + V_{ij}^*).$$

The factor $1/2$ comes because we sum over i and j and this includes the same term twice. The factor V_{ij} is complex, because the firs one-loop correction to the square of the amplitude is $\mathcal{A}^{\text{1-loop}}\mathcal{A}^{\text{tree}} + \text{c.c.}$ and it reads

$$V_{ij} = (-i) \int \frac{\mathrm{d}^4 k}{(2\pi)^4} \frac{p_i \cdot p_j}{(k^2 + i\varepsilon)(\eta_i p_i \cdot k + i\varepsilon)(-\eta_j p_j \cdot k + i\varepsilon)}.$$

The minus sign in front and in the second propagator are there because k is incoming at the p_j line while it was always outgoing in the convention of 5.1. The loop momentum goes over all k, but the expression is valid only for small k. We therefore assume $|\mathbf{k}| < \Lambda_{\text{UV}}$. This is fine, since we are interested only in the behaviour for small k. We compute V_{ij} by applying Cauchy's theorem to the integration over k^0. The poles in k^0 can arise from $k^2 = 0$, $p_i \cdot k = 0$ or $p_j \cdot k = 0$:

$$k^0 = |\mathbf{k}| - i\varepsilon, \quad k^0 = -|\mathbf{k}| + i\varepsilon, \quad k^0 = \frac{\mathbf{p}_i \cdot \mathbf{k} - \eta_i i\varepsilon}{p_i^0}, \quad k^0 = \frac{\mathbf{p}_j \cdot \mathbf{k} + \eta_j i\varepsilon}{p_j^0}.$$



This diagram holds if p_i is outgoing and p_j incoming. If both are outgoing or both incoming the particle propagator poles are on different sides of the real axis.

Now close the contour in the lower half-plane which sets $k^0 \to |\mathbf{k}|$ and gives residue factor $-2\pi i/(2|\mathbf{k}|)$ (negative sign from the counter-clockwise contour) from the photon pole. Similar contributions also come from the particle poles. The integral becomes

$$\begin{aligned} V_{ij} =& \eta_i \eta_j \int_{|\mathbf{k}|<\Lambda_{\text{UV}}} \frac{\mathrm{d}^3\mathbf{k}}{(2\pi)^3} \frac{1}{2|\mathbf{k}|} \frac{p_i \cdot p_j}{(p_i^0|\mathbf{k}| - \mathbf{p}_i \cdot \mathbf{k})(p_j^0|\mathbf{k}| - \mathbf{p}_j \cdot \mathbf{k})} \\ &+ \theta(\eta_i \eta_j)(\text{contribution from particle pole}). \end{aligned}$$

The second row vanishes if η_i, η_j have opposite sign. If not, one can calculate the particle pole contribution and finds it is purely imaginary. But we only need the real part of V_{ij}, so we drop these terms. The integral in the first line is exactly the same as I_{ij} except for an opposite sign and now $|\mathbf{k}| < \Lambda_{\text{UV}}$ rather than $|\mathbf{k}| < \Delta E_\gamma$. Hence, the infrared divergence for $|\mathbf{k}| \to 0$ (soft divergence) cancels when the virtual correction to the non-radiative process and the radiative process are added together.

Why should one do this? Experimentally, a final state of electrons and a final state of the same electrons and one (or many) very low energy photons cannot be distinguished. The observable quantity is always of the form

$$\sigma^{i\to f} + \sigma^{i\to f+\gamma} + \sigma^{i\to f+2\gamma} + \dots$$

where energy of each photon is smaller than the resolution ΔE_γ that can be achieved for soft photons. Any observable quantity does not have the soft divergence. We showed this above for the emission of one soft photon, but because of exponentiation it is not difficult to show this to all loop orders and any number of emitted soft photons. This is called the *Bloch–Nordsieck cancellation*.

A similar reasoning applies to energetic (i.e. not soft) but collinear photons. One cannot distinguish a photon from an electron unless they are at least an angle θ_{min} apart. Hence, the

singularity at $\theta \to 0$ will be removed by the angular resolution that can be achieved experimentally. Then the calculation of the QED quantum corrections to a process $i \to f$ at NLO (*next-to-leading order*) will have the form

$$\mathrm{d}\sigma_{\mathrm{NLO}}^{i\to f(+\gamma)}(p_i) = \mathrm{d}\sigma_{\mathrm{LO}}^{i\to f}(p_i)\left(1 + \frac{\alpha_{\mathrm{EM}}}{\pi}\left[\mathrm{const} \times \ln\frac{E^2}{\Delta E_\gamma^2} \times \ln\frac{E^2}{m_e^2} + \text{less relevant terms}\right] + \dots\right). \tag{5.2}$$

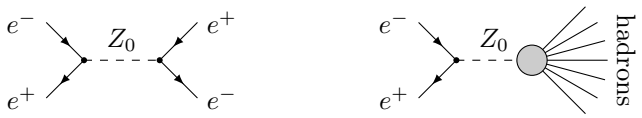
For a collinear divergence E^2/m_e^2 is replaced by $1/\theta_{\min}^2$. A double logarithm is very typical for such IR-divergences, it is called *Sudakov double logarithm*. If $\Delta E_\gamma \ll E$ or $\theta_{\min} \ll 1$ or $m_e \ll E$ these corrections are large and may have to be summed to *all orders* in perturbation theory, even though α_{EM} is very small.

5.2 IR divergences in QCD and jets

One can generalize the steps from the last section from QED to QCD. It is more complicated because T^A stands at the quark-gluon vertex and there is the gluon self-interaction after radiation. But there is a conceptual problem: no quarks and gluons are allowed in a final/initial state (confinement). What are the observables?

5.2.1 Hadron production at high energies

We will consider the process $e^+e^- \to Z^0 \to$ hadrons and compare it to one where Z^0 decays back into electron-positron pair.



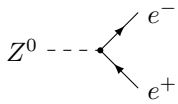
We only consider hadrons as the final states of the second process, no leptons, scalars or gauge bosons involved. Since the mass of the Z boson $M_Z = 91\,\mathrm{GeV}$ which determines the energy scale of the process is much larger than QCD confinement scale Λ_{QCD}, we can safely talk about high-energy process. In fact, it is enough to look at the decay of the Z^0 boson. Its coupling to fermions is

$$\mathcal{L}_{\mathrm{int}} = \frac{e}{\sin 2\theta_{\mathrm{W}}}\sum_f \overline{\Psi}_f\gamma^\mu(v_f - a_f\gamma_5)\Psi_f, \quad v_f = T_f^3 - 2Q_f\sin\theta_{\mathrm{W}}, \quad a_f = T_f^3, \quad T_f^3 = \pm\frac{1}{2}.$$

The sum runs over all fermions in the SM, $f \in \{e, \mu, \tau, \nu_e, \nu_\mu, \nu_\tau, u, d, s, c, b, t\}$. The decay into electrons is calculated with

$$\Gamma(Z^0 \to e^+e^-) = \frac{1}{2m_Z}\frac{1}{3}\int\frac{\mathrm{d}^3\mathbf{p}_1}{(2\pi)^3}\frac{1}{2p_1^0}\frac{\mathrm{d}^3\mathbf{p}_2}{(2\pi)^3}\frac{2}{2r_2^0}(2\pi)^4\delta^{(4)}(p_Z - p_1 - p_2)\sum_{\mathrm{pols}}|\mathcal{M}|^2.$$

The factor 1/3 comes from the average over three polarization states of Z^0 and the amplitude reads

$\mathcal{M} =$ 

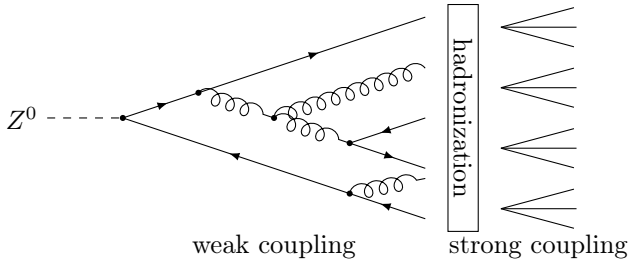
$$= \frac{ie}{\sin(2\theta_{\mathrm{W}})}\overline{u}(p_1)\gamma^\mu(v_e - a_e\gamma_5)v(p_2)\varepsilon_\mu(p_Z, \lambda).$$

Plugging all together and neglecting electron mass gives the decay width

$$\Gamma(Z^0 \to e^+e^-) = (v_e^2 + a_e^2)\frac{e^2}{\sin^2(2\theta_{\mathrm{W}})}\frac{m_Z}{12\pi} = 81.1\,\mathrm{MeV}.$$

The experimental result is 83.8 MeV which is consistent with our calculation since the corrections to our result are of order $\mathcal{O}(\alpha/\sin^2\theta_{\mathrm{W}})$.

Now consider the decay into hadrons. The strong coupling at the mass of Z^0 boson is $\alpha_{\mathrm{s}}(m_Z) \approx 0.12$ which indicates perturbative expansion. In every loop order the final state consists of quarks and gluons, not hadrons?! How do we paint this picture physically.



The decay cascades into energetic quarks and gluons (collectively called *partons*), which can be calculated in perturbation theory. After the partons *hadronize* the coupling becomes strong, and the states cannot be computed in detail. The final state consists of *jets*, i.e. bundles of energetic hadrons. Roughly speaking, an energetic parton turns into a jet.

As long as one does not ask for specific properties of the hadronic final state (the number of protons, etc.) but only whether hadroproduction takes place or how many jets there are, one can compute certain properties of hadroproduction at high energy (so that $\alpha_{\mathrm{s}}(Q) \ll 1$) in perturbation theory. Devising suitable observables is still an active area of research and important to diagnose new particles in a hadronic final state. The validity of the above statements cannot be rigorously derived in QCD up to now. But they are empirically in quantitatively confirmed. Hence,

$$\Gamma_{\mathrm{had}} = \Gamma(Z^0 \to \text{hadrons}) = \frac{1}{2m_Z}\frac{1}{3}\sum\!\!\!\!\!\!\int_X (2\pi)^4\delta^{(4)}(p_Z - p_X)|\mathcal{M}|^2$$

where X are all hadronic final states X including phase space integral. The scattering amplitude reads

$$\mathcal{M} = \frac{ie}{\sin 2\theta_{\mathrm{W}}}\,\langle X|j^\mu_{\mathrm{had}}(0)|\Omega\rangle\,\varepsilon_\mu(p_Z,\lambda).$$

Here, Ω represents the QCD vacuum and j^μ_{had} is the hadronic current

$$j^\mu_{\mathrm{had}} = \sum_f \overline{q}_f\gamma^\mu(v_f - a_f\gamma_5)q_f.$$

The sum goes over all quarks except top which is too heavy to be produced in Z^0 decay. Plugging all together give the decay width

$$\Gamma_{\mathrm{had}} = \frac{1}{6m_Z}\frac{e^2}{\sin^2(2\theta_{\mathrm{W}})}\left(-g_{\mu\nu} + \frac{p_{Z\mu}p_{Z\nu}}{m_Z^2}\right)\sum\!\!\!\!\!\!\int_X (2\pi)^4\delta^{(4)}(p_Z - p_X)\,\langle\Omega|j^\mu_{\mathrm{had}}(0)|X\rangle\,\langle X|j^\nu_{\mathrm{had}}(0)|\Omega\rangle\,. \tag{5.3}$$

. To continue the calculation, we assume *parton-hadron duality*, which states that the sum over hadrons (what is observed) equals the sum over states in perturbation theory (quarks, gluons, what can be computed):

$$\sum_{X=\text{hadrons}} |X\rangle \langle X| = \sum_{X=q,g} |X\rangle \langle X| .$$

In leading order in α_s, the partonic final state X is only $q\bar{q}$. Radiation of a gluon or a loop correction costs a power of α_s. Hence, to leading order

$$\frac{\Gamma_\text{had}}{\Gamma_{e^+e^-}} = \frac{N_\text{c}}{v_e^2 + a_e^2} \sum_{f=u,c,s,c,b} (v_f^2 + a_f^2) = 20.2 \left(1 + \mathcal{O}(\alpha_\text{s}, m_b^2/m_Z^2)\right) .$$

Here, $N_\text{c} = 3$ for gauge group $\text{SU}(3)_\text{c}$ (three colour states for every quark flavour). The experimental result for this ratio is 20.8, hence this simple calculation already gives a very good approximation.

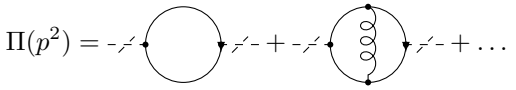
We can bring Eq. 5.3 into a form where the explicit sum over X does no longer appear:

$$\begin{aligned}
&\sum\!\!\!\!\!\!\int_X (2\pi)^4 \delta^{(4)}(p_Z - p_X) \langle\Omega|j^\mu_\text{had}(0)|X\rangle \langle X|j^\nu_\text{had}(0)|\Omega\rangle = \\
&\qquad = \int \mathrm{d}^4x\, e^{i(p_Z - p_X)x} \sum\!\!\!\!\!\!\int_X \langle\Omega|j^\mu_\text{had}(0)|X\rangle \langle X|j^\nu_\text{had}(0)|\Omega\rangle \\
&\qquad = \int \mathrm{d}^4x\, e^{ip_Z x} \sum\!\!\!\!\!\!\int_X \langle\Omega|e^{iPx} j^\mu_\text{had}(0) e^{-iPx}|X\rangle \langle X|j^\nu_\text{had}(0)|\Omega\rangle \\
&\qquad = \int \mathrm{d}^4x\, e^{ip_Z x} \sum\!\!\!\!\!\!\int_X \langle\Omega|j^\mu_\text{had}(x)|X\rangle \langle X|j^\nu_\text{had}(0)|\Omega\rangle \\
&\qquad = \int \mathrm{d}^4x\, e^{ip_Z x} \langle\Omega|j^\mu_\text{had}(x) j^\nu_\text{had}(0)|\Omega\rangle \\
&\qquad = 2\text{Im}\left(i \int \mathrm{d}^4x\; e^{ip_Z x} \langle\Omega|T(j^\mu_\text{had}(x) j^\nu_\text{had}(0))|\Omega\rangle\right) \\
&\qquad = 2\text{Im}\left(\Pi_{\mu\nu}(p_Z)\right) = 2(p_{Z\mu} p_{z_\nu} - p_Z^2 g_{\mu\nu})\text{Im}\;\Pi(p_Z^2).
\end{aligned}$$

In the second step we used momentum operator P, then the completeness relation $|X\rangle \langle X| = \mathbb{1}$ and the optical theorem (see p.??). The factorization of the two-point function is a consequence of Lorentz invariance and current conservation $\partial_\mu j^\mu_\text{had} = 0$ which implies $p_Z^\mu \Pi_{\mu\nu}(p_Z) = 0$. Finally, we write the decay width as

$$\Gamma_\text{had} = \frac{1}{6 m_Z} \frac{e^2}{\sin^2(2\theta_\text{W})} 6\text{Im}\;\Pi(m_Z^2). \tag{5.4}$$

Remember the diagrammatic expansion for the two-point function

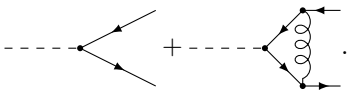
$\Pi(p^2) =$  $+ \ldots$

That is, the inclusive hadronic decay width of the Z^0 can be computed from the discontinuity of the two-point function. Except for $|\Omega\rangle$, hadronic particle states no longer appear explicitly. This does not mean that Γ_had can be computed in perturbation theory to arbitrary accuracy. The QCD vacuum is itself a complicated, strongly interacting state, different from the Fock space vacuum of quarks and gluons. For less inclusive observables there is no representation similar to 5.4 and one has to consider explicitly scattering amplitudes of quarks and gluons.

Loop corrections and final states with a gluon

We want to compute the $\mathcal{O}(\alpha_s)$ correction to $R_{\text{had}} \equiv \Gamma_{\text{had}}/\Gamma_{e^+e^-}$. For that we could use 5.4 and compute 2-loop diagrams. However, for later discussion, we will go back to 5.3. There are two effects.

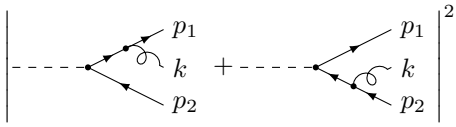
- Loop correction to $X = q\bar{q}$ (also called *virtual correction*)



  This equals $\langle q\bar{q}|j^\mu_{\text{had}}|\Omega\rangle$ in the 1-loop approximation. We also need the quark field renormalization in the on-shell scheme (from ).

- *Real emission correction* for a final state $X = q\bar{q}g$:

$$|\mathcal{M}|^2 =$$



Note the similarities to the discussion within QED in section Soft and collinear divergences in QED. We first compute the real correction. Introduce the dimensionless variables

$$x_{1,2} = \frac{2p_{1,2}\cdot p_Z}{p_Z^2} = \frac{2E_{q,\bar{q}}}{m_Z}, \quad x_3 = \frac{2k\cdot p_Z}{p_Z^2} = \frac{2E_g}{m_Z}.$$

The expressions involving m_Z only hold in the rest frame of Z^0. The energy conservation implies $x_1 + x_2 + x_3 = 2$. Let's calculate the real emission correction to the Z^0 decay at first order:

$$\begin{aligned}\Gamma^{(1,\text{real})}_{\text{had}} =& \frac{1}{6m_Z}\frac{e^2}{\sin^2(2\theta_W)}\left(-g_{\mu\nu} + \frac{p_{Z\mu}p_{Z\nu}}{m_Z^2}\right)(ig_s)(-ig_s)\,\text{tr}(T^AT^A)\sum_f (v_f^2 + a_f^2)(-g_{\rho\sigma})\cdot \\ &\cdot \int \frac{d^3\mathbf{p}_1}{(2\pi)^3 2p_1^0}\frac{d^3\mathbf{p}_2}{(2\pi)^3 2p_2^0}\frac{d^3\mathbf{k}}{(2\pi)^3 2k^0}(2\pi)^4(p_Z - p_1 - p_2 - k)\cdot \\ &\cdot \left(\frac{\text{tr}(\gamma^\mu(\not{p}_1 + \not{k})\gamma^\rho \not{p}_1 \gamma^\sigma(\not{p}_1 + \not{k})\gamma_\mu \not{p}_2)}{(p_1+k)^4} + 3 \text{ more terms}\right).\end{aligned}$$

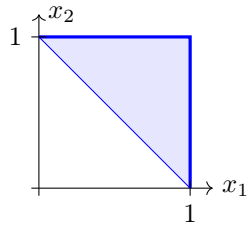
The gluon polarization sum gives δ^{AB} and hence $\text{tr}(T^AT^A) = C_F N_c$. The gluon propagator was used in Feynman gauge $\xi = 1$ and there is still no interference between γ_μ and $\gamma_\mu\gamma_5$ terms in the $Zq\bar{q}$ vertex, hence $v_f^2 + a_f^2$. The phase space integral together with energy conservation can be expressed in terms of x_1 and x_2

$$\int \frac{d^3\mathbf{p}_1}{(2\pi)^3 2p_1^0}\frac{d^3\mathbf{p}_2}{(2\pi)^3 2p_2^0}\frac{d^3\mathbf{k}}{(2\pi)^3 2k^0}(2\pi)^4(p_Z - p_1 - p_2 - k) = \frac{m_Z^2}{128\pi^3}\int_0^1 dx_1 \int_0^1 dx_2\,\theta(x_1 + x_2 - 1).$$

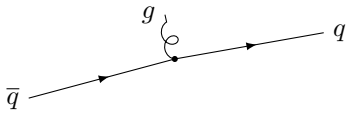
Expressing also the propagator terms with x variables and collecting the leading order decay rate in a single symbol $\Gamma^{(0)}_{\text{had}}$ gives

$$\Gamma^{(1,\text{real})}_{\text{had}} = \Gamma^{(0)}_{\text{had}}\frac{\alpha_s C_F}{2\pi}\int\limits_{\substack{0\le x_1,x_2\le 1\\ x_1+x_2\ge 1}} dx_1\,dx_2\,\frac{x_1^2 + x_2^2}{(1-x_1)(1-x_2)}.$$

This is a divergent integral! But it shouldn't come as a surprise – we expect the soft and collinear divergence of real emission. On the figure below, the thick blue edges denote the diverging regions in the integration space.



a) When $x_1, x_2 \to 1$, that is $x_3 \to 0$, the gluon energy becomes very small and we get the soft divergence.



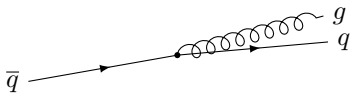
b) When $x_1 \to 1$ and x_2 is fixed, then

$$k \cdot p_2 = |\mathbf{k}||\mathbf{p}_2|(1 - \cos\theta(\mathbf{k}, \mathbf{p}_2)) = \frac{1}{2}(k + p_2)^2 = \frac{1}{2}(p_Z - p_1)^2 = \frac{m_Z^2}{2}(1 - x_1).$$

Hence, $\cos\theta(\mathbf{k}, \mathbf{p}_2) \to 1$ and the gluon becomes parallel to the antiquark. This is collinear divergence.



c) For $x_2 \to 1$ with x_1 fixed, the gluon is parallel to the quark.



These are also the regimes where propagators $1/((p+k)^2 + i\varepsilon)$ can become singular. From QED discussion we know that we must split the $q\overline{q}g$ final state into a piece where the energy of the gluon is sufficiently large, and its direction is separated by a minimal angle from q or $\overline{q}$. This naturally defines a final state with three jets of hadrons and has no IR singularities. The IR singular reminder must be combined with the virtual correction. The final state looks like a two-jet event.

final state	divergence	number of jets	IR finite
$X = q\overline{q}$	1-loop virtual	two-jet like	sum is IR finite
$X = q\overline{q}g$	soft or collinear	two-jet like	
	hard g	three-jet like	IR finite

This works at $\mathcal{O}(\alpha_\mathrm{s})$. In the next order, one must divide the $q\bar{q}gg$ final state into a two-, three- and four-jet regions, the one loop correction to the $q\bar{q}g$ final state into two- and three-jet events and so on.

To carry this out in practice, one needs a precise definition of a "jet". We come back to this in the next section. The present purpose is to compute the $\mathcal{O}(\alpha_\mathrm{s})$ corrections to the *inclusive* rate Γ_had. We therefore compute the real and virtual correction by applying an *infrared regularization*. We can remove the regulator at the end of the calculation of IR-finite observables. Possible regulators are:

- give the gluon a small mass λ (this would work here, but causes difficulties (technical and gauge invariance violation) at $\mathcal{O}(\alpha_\mathrm{s}^2)$)
- dimensional regularization – the method of choice.

Dimensional regularization means that

- we compute $\Gamma^{(0)}_\mathrm{had}$ in $d = 4 - 2\varepsilon$ dimensions,
- $|\mathcal{M}|^2$ in d dimensions for real and virtual processes,
- use d dimensions also for the phase-space integration.

Repeating the steps from before gives

$$\Gamma^{(0)}_\mathrm{had} = \underbrace{\frac{\Gamma(2-\varepsilon)}{\Gamma(2-2\varepsilon)}\left(\frac{m_Z^2}{4\pi\tilde{\mu}^2}\right)^{-\varepsilon}}_{1+\mathcal{O}(\varepsilon)} N_\mathrm{c}\sum_f (v_f^2+a_f^2)\frac{e^2}{\sin^2(2\theta_\mathrm{W})}\frac{m_Z}{12\pi}$$

for the leading order decay rate,

$$|\mathcal{M}|^2 = \frac{(1-\varepsilon)^2(x_1^2+x_2^2) - 2\varepsilon(1-\varepsilon)(2-2x_1-2x_2+x_1x_2)}{(1-x_1)(1-x_2)}$$

for the $X = q\bar{q}g$ amplitude and

$$\begin{aligned}\tilde{\mu}^{4\varepsilon}\int \frac{\mathrm{d}^{d-1}\mathbf{p}_1}{(2\pi)^3}\frac{1}{2p_1^0}\frac{\mathrm{d}^{d-1}\mathbf{p}_2}{(2\pi)^3}\frac{1}{2p_2^0}\frac{\mathrm{d}^{d-1}\mathbf{k}}{(2\pi)^3}\frac{1}{2k^0}(2\pi)^d\delta^{(d)}(p_Z-p_1-p_2-k) = \\ = \frac{m_Z^2}{128\pi^3}\left(\frac{m_Z^2}{4\pi\tilde{\mu}^2}\right)^{-2\varepsilon}\frac{1}{\Gamma(d-2)}\int\limits_{\substack{0\le x_1,x_2\le 1\\ x_1+x_2\ge 1}} \mathrm{d}x_1\,\mathrm{d}x_2\,(x_1x_2)^{-2\varepsilon}\left(\frac{1-v_\star^2}{4}\right)^{-\varepsilon}\end{aligned}$$

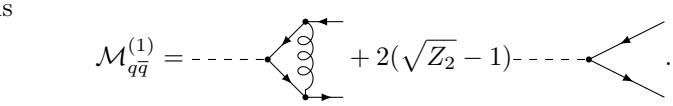
for the phase space integral. Note the exponent 4ε on $\tilde{\mu}$ – this is because the leading order calculation used 2 instead of 4. In the singular regions (soft, collinear)

$$v_\star \equiv 1 + 2\frac{1-x_1-x_2}{x_1x_2}$$

converges to 1 and the factor $(1-v_\star^2)^{-\varepsilon}$ regulates the previously divergent integral. Now we integrate over x_1 and x_2 and expand in ε:

$$\Gamma^{(1,\mathrm{real})}_\mathrm{had} = \Gamma^{(0)}_\mathrm{had}\frac{\alpha_\mathrm{s}C_\mathrm{F}}{2\pi}\left(\frac{m_Z^2}{4\pi\tilde{\mu}^2}\right)^{-\varepsilon}\frac{\Gamma(1-\varepsilon)}{\Gamma(1-3\varepsilon)}\left(\frac{2}{\varepsilon^2}+\frac{3}{\varepsilon}+\frac{19}{2}+\mathcal{O}(\varepsilon)\right).$$

Notice the double pole, with contributions from both soft and collinear divergence. The virtual correction reads

$\mathcal{M}^{(1)}_{q\bar{q}} =$  $+ 2(\sqrt{Z_2} - 1)$.

Both the loop and the renormalization coefficient contain UV and IR divergences, all regulated dimensionally. The UV divergences cancel within the amplitude, IR don't. One then needs

$$|\mathcal{M}^{(0)}_{q\bar{q}} + \mathcal{M}^{(1)}_{q\bar{q}}|^2 = |\mathcal{M}^{(0)}_{q\bar{q}}|^2 + 2\mathrm{Re}\,(\mathcal{M}^{(0)*}_{q\bar{q}}\mathcal{M}^{(1)}_{q\bar{q}}) + \mathcal{O}(\alpha_\mathrm{s}^2)$$

integrated over the d-dimensional two-particle $q\bar{q}$ phase space:

$$\Gamma^{(1,\mathrm{virt})}_{\mathrm{had}} = \Gamma^{(0)}_{\mathrm{had}} \frac{\alpha_\mathrm{s} C_\mathrm{F}}{2\pi} \left(\frac{m_Z^2}{4\pi\tilde{\mu}^2}\right)^{-\varepsilon} \frac{\Gamma(1-\varepsilon)}{\Gamma(1-3\varepsilon)} \left(-\frac{2}{\varepsilon^1} - \frac{3}{\varepsilon} - 8 + \mathcal{O}(\varepsilon)\right).$$

Then

$$\Gamma^{(1)}_{\mathrm{had}} = \lim_{\varepsilon\to 0}(\Gamma^{(1,\mathrm{virt})}_{\mathrm{had}} + \Gamma^{(1,\mathrm{real})}_{\mathrm{had}}) = \Gamma^{(0)}_{\mathrm{had}} \frac{3\alpha_\mathrm{s} C_\mathrm{F}}{4\pi}$$

which is IR finite. The hadron-to-electron decay ratio is

$$\frac{\Gamma_{\mathrm{had}}}{\Gamma_{e^+e^-}} = \frac{\Gamma^{(0)}_{\mathrm{had}}}{\Gamma_{e^+e^-}} \left(1 + \frac{\alpha_\mathrm{s}}{\pi} + \mathcal{O}(\alpha_\mathrm{s}^2, m_b^2/m_Z^2)\right) = 21.0.$$

The process involves the scale m_Z as the only scale variable, so naturally choose $\alpha_\mathrm{s}^{\overline{\mathrm{MS}}}(m_Z)/\pi \approx 0.038$. This is in a very good agreement with experimental value of 20.8. The corrections have now been calculated to $\mathcal{O}(\alpha_\mathrm{s}^4)$ and Γ_{had} is used to *determine* $\alpha_\mathrm{s}(m_Z)$.

5.2.2 Calculation of jet production

To define a *jet* one needs an algorithm that determines how the hadrons in a real event and the partons in a theoretical calculation are assigned to jets. There are several algorithms. Here, we illustrate the procedure with the *JADE algorithm*.

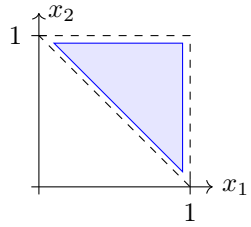
The IR divergences in the $q\bar{q}g$ cross-section arise from $(p_1 + k)^2 \to 0$ and/or $(p_2 + k)^2 \to 0$, i.e. the qg or $\bar{q}g$ system has small variation in mass. Hence, define $p_3 = k$ and

$$\begin{aligned} q\bar{q} \text{ or } q\bar{q}g \text{ with } (p_i + p_j)^2 < m_Z^2 \quad &\Longrightarrow \quad \text{two jets,} \\ q\bar{q}g \text{ with } (p_i + p_j)^2 < m_Z^2 \text{ for all } i,j \quad &\Longrightarrow \quad \text{three jets.} \end{aligned}$$

Here, y is a jet resolution parameter. The smaller y, the narrower the bundles of hadrons in a single jet. Concretely at $\mathcal{O}(\alpha_\mathrm{s})$ in $Z^0 \to q\bar{q}g$ the three-jet condition is

$$\begin{aligned} (p_1 + k)^2 &= m_Z^2(1 - x_2) \geq m_Z^2 y, \\ (p_2 + k)^2 &= m_Z^2(1 - x_1) \geq m_Z^2 y, \\ (p_1 + p_2)^2 &= M_Z^2(x_1 + x_2 - 1) \geq m_Z^2 y. \end{aligned}$$

This corresponds to the following picture.



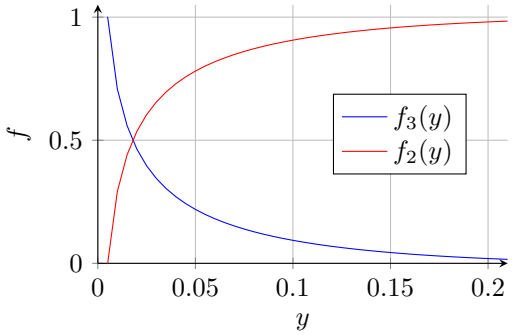
Integration region now excludes the singular regions. In terms of the integral expression, we write

$$\begin{aligned}\frac{\Gamma_{\text{had}}^{(3\text{-jet})}}{\Gamma_{\text{had}}^{(0)}} &= \frac{\alpha_{\text{s}}C_{\text{F}}}{2\pi}\int\limits_{\substack{0\le x_1,x_2\le 1-y\\ x_1+x_2\ge 1+y}} \mathrm{d}x_1\,\mathrm{d}x_2\,\frac{x_1^2+x_2^2}{(1-x_1)(1-x_2)} \equiv f_3(y)\\ &= \frac{\alpha_{\text{s}}C_{\text{F}}}{2\pi}\int_{2y}^{1-y}\mathrm{d}x_2\int_{1+x-x_1}^{1-x}\mathrm{d}x_1\,\frac{x_1^2+x_2^2}{(1-x_1)(1-x_2)}\\ &= \frac{\alpha_{\text{s}}C_{\text{F}}}{2\pi}\left(2\ln^2\left(\frac{y}{1-y}\right)+3(1-2y)\ln\left(\frac{y}{1-2y}\right)+4\mathrm{Li}_2\left(\frac{y}{1-y}\right)+\frac{5}{2}-6y-\frac{9y^2}{2}-\frac{\pi^2}{3}\right),\end{aligned}$$

where

$$\mathrm{Li}_2(x) \equiv -\int_0^x \mathrm{d}\xi\,\frac{\ln(1-\xi)}{\xi}$$

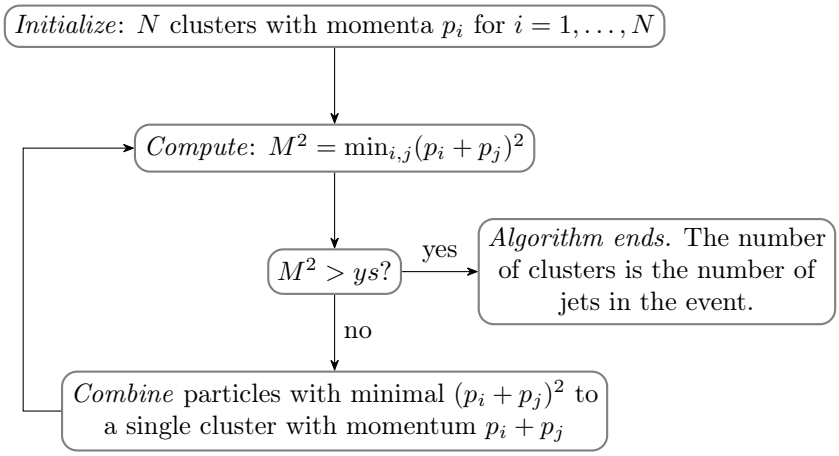
is called *dilogarithm*. The index 3 with the function f_3 indicates the three-jet amplitude. At $\mathcal{O}(\alpha_{\text{s}})$ the two-jet function is simply $f_2(y) = 1 - f_3(y)$ so adding the two together give the decay rate independent of the jet resolution. The four-jet function $f_4(y)$ also exists, but is non-zero only at $\mathcal{O}(\alpha_{\text{s}}^2)$. For smaller y one resolves more events into three jets (see picture below, where the value for α_{s} has been chosen at scale m_Z).



One cannot make y arbitrarily small, since then one would resolve the soft and collinear configurations where perturbation theory breaks down. The calculation shows this as for $y \to 0$

$$f_3(y) \approx \frac{\alpha_{\text{s}}C_{\text{F}}}{2\pi}\left(2\ln^2 y + 3\ln y + \frac{5}{2} - \frac{\pi^2}{3}\right) \to \infty$$

i.e. a Sudakov double logarithmic divergence similar to the QED result (for $\Delta E_\gamma \to 0$ and $\theta_{\text{min}} \to 0$ in Eq. 111). For a final state with many hadrons (easily $\mathcal{O}(100)$ in a real high-energy event) we define the algorithm as follows. Let $p_1, p_2, \ldots, p_N$ be the four-momenta of an event with N particles in a scattering with invariant mass s (this would be m_Z^2 in Z^0 decay).



This is the specific use of JADE (Joint Approximation Diagonalization of Eigen-matrices) algorithm for determining the number of jets, however JADE can in general be more complicated and have vastly different uses.

Infrared finiteness

In any application of QCD perturbation theory one must ensure that the observables are IR-finite, i.e. not sensitive to configurations of soft and collinear partons. One needs to

- construct IR-finite observable and
- factorize IR-sensitive contributions into non-perturbative objects and then apply perturbation theory only to the reminder.

We will discuss an important example of the second case in the following section. Here are some final remarks on the first case, again for $Z^0 \to q, g$. Nor an N-parton final state, define the observable by a function $S^{(N)}(p_1, \ldots, p_N)$ of the momenta. Then

$$S = \sum_N \mathrm{dPS}[p_1, \ldots, p_n] \, |\mathcal{M}_{Z^0 \to N \text{ partons}}|^2 S^{(N)}(p_1, \ldots, p_N)$$

is the value of the observable S. For such an observable to be IR finite, the following must hold

(1) it has to be insensitive to soft partons

$$S^{(N)}(p_1, \ldots, p_i, \ldots, p_N) \xrightarrow{p_i \to 0, \text{ any } i} S^{(N-1)}(p_1, \ldots, \hat{p}_i, \ldots, p_N),$$

(2) it has to be insensitive to combining two partons with collinear momenta into a single parton with the sum of the momenta (insensitive to collinear partons),

$$S^{(N)}(p_1, \ldots, p_i, \ldots, p_j, \ldots, p_N) \xrightarrow{\mathbf{p}_i \| \mathbf{p}_j, \text{ any pair } i,j} S^{(N-1)}(p_1, \ldots, p_i + p_j, \ldots, \hat{p}_j, \ldots, p_N).$$

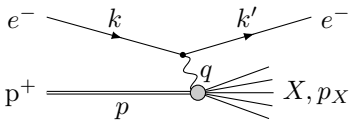
Observable	$S^{(2)}(p_1, p_2)$	$S^{(3)}(p_1, p_2, p_3)$	IR-finite
Γ_{had}	1	1	Yes
$\Gamma_{Z \to q\bar{q}}$	1	0	No
two-jet rate in the JADE algorithm	1	1 if $(p_i + p_j)^2 < y m_Z^2$ 0 otherwise	Yes

5.3 Deep inelastic lepton–nucleon scattering (DIS) and the concept of the parton distribution

Imagine we collide electron (or positron or neutrino) to a proton (or pion) at large electron energy in the rest frame of the proton:

$$e^- + \mathrm{p}^+ \to e^- + X.$$

This can produce many complex hadronic states X. Denote the momentum of the proton with p and the momentum of the exchange particle with q. For small enough q, the exchange particle can only be photon, but for $Q^2 \equiv -q^2 \gg m_Z^2$ it can also be Z^0 (and in the case of neutrino scattering also $W^\pm$).



For electron momenta k^μ and k'^μ we have $q^\mu = k^\mu - k'^\mu$, and we also define $q^2 = -Q^2$ and $p^2 = M_\mathrm{p}^2$ (the proton mass). If the momentum transfer is large, $Q^2 \gg M_\mathrm{p}^2$, then we are talking about *deep inelastic scattering* (DIS). The proton typically disintegrates into many hadrons. In the proton's rest frame define

$$v \equiv p \cdot q = M_\mathrm{p}(E - E'), \quad y \equiv \frac{p \cdot q}{p \cdot k} = \frac{E - E'}{E}, \quad x \equiv \frac{Q^2}{2p \cdot q}.$$

The latter, x, is called *Bjorken's scaling variable*. Historically (SLAC in 1967) DIS provided firs evidence that the proton is made of point-like constituents, which were then identified with the quarks in Gell-Mann classification scheme.

In the following we will assume $Q^2 \ll m_{W,Z}^2$ so that only photon exchange needs to be considered. Also assume that the proton is unpolarized. An average of initial-state polarizations is performed, but not written explicitly. We also neglect the electron mass, $m_e = 0$. The scattering matrix element can be written as

$$iT_{\mathrm{ep}\to eX} = -ie\overline{u}(k')\gamma^\alpha u(k)\frac{-i}{q^2}ie\,\langle X|j_\alpha(0)|p\rangle\,.$$

The part of the photon propagator that is proportional to $1 - \xi$ drops out by Dirac equation $\overline{u}(k')\not{q}u(k) = 0$. We used the QCD matrix element of the electromagnetic current

$$j_\alpha = \sum_{f\in\{u,d,s,c,b\}} Q_f \overline{q}_f \gamma_\alpha q_f.$$

Differential scattering cross-section in the momentum of the scattered electron has the form

$$(2\pi)^3 2k'^0 \frac{\mathrm{d}\sigma}{\mathrm{d}^2\mathbf{k}'} = \frac{1}{2}\frac{1}{2s}\,\mathrm{tr}(\not{k}'\gamma^\alpha \not{k}\gamma^\beta)\frac{4\pi e^4}{Q^4}W_{\alpha\beta}(p,q)$$

where $1/2$ comes from the e^- spin average. At DIS energies, the Mandelstam variable s can be simplified to $s = (p+k)^2 \approx 2p\cdot k = Q^2/(xy)$ and the trace gives $4(k^\alpha k'^\beta + k^\beta k'^\alpha k \cdot k\prime\, g^{\alpha\beta})$. The *hadronic tensor* $W_{\alpha\beta}$ contains all information about the strong interaction and the hadronic

final state X. In inclusive DIS one does not measure any specific property of X and therefore sums over all X:

$$\begin{aligned} W_{\alpha\beta}(p,q) &= \frac{1}{4\pi}\sum\!\!\!\!\!\!\int_X (2\pi)^4\delta^{(4)}(p+q-p_X)\,\langle \mathrm{p}|j_\beta^\dagger(0)|X\rangle\,\langle X|j_\alpha(0)|\mathrm{p}\rangle \\ &= \frac{1}{4\pi}\int \mathrm{d}^4x\; e^{iqx}\sum\!\!\!\!\!\!\int_X \langle \mathrm{p}|j_\beta^\dagger(x)|X\rangle\,\langle X|j_\alpha(0)|\mathrm{p}\rangle \\ &= \frac{1}{4\pi}\int \mathrm{d}^4x\; e^{iqx}\,\langle \mathrm{p}|j_\beta^\dagger(x)j_\alpha(0)|\mathrm{p}\rangle\,. \end{aligned}$$

The spin average of $|p\rangle$ is understood and all explicit reference to the hadronic final state has disappeared. Compare this with $Z^0 \to$ hadrons: we have the same operator, but now we need the proton rather than vacuum matrix element.

Structure function

Use Lorentz covariance to express the tensor $W_{\alpha\beta}$ in terms of scalar function. The available objects are $g_{\alpha\beta}$, p_α, q_α and $\varepsilon_{\alpha\beta\gamma\delta}$ (importantly, no quantities related to polarization, since the proton is unpolarized). Instead of dependence on p and q we will rather use the dependence on x and Q^2, so each W_i will be implicitly a function $W_i = W_i(x,Q^2)$. The tensor decomposition gives

$$W_{\alpha\beta} = W_a g_{\alpha\beta} + W_b p_\alpha p_\beta + W_c q_\alpha q_\beta + W_d p_\alpha q_\beta W_e q_\alpha p_\beta + W_f i\varepsilon_{\alpha\beta\gamma\delta}p^\gamma q^\delta.$$

- Parity conservation requires $W_f = 0$ (however only up to $Q^2 \sim m_Z^2$, when the parity-violating weak interaction starts playing a role).
- The tensor has to be symmetric in α and β, so $W_e = W_d$.
- The electromagnetic charge conservation (gauge invariance, Ward identities) give two further conditions
$$q^\alpha W_{\alpha\beta} = q^\beta W_{\alpha\beta} = 0.$$

These conditions help reduce the matrix to merely two *structure functions*, which contain all information about the strong interaction

$$W_{\alpha\beta}(p,q) = \left(-g_{\alpha\beta} + \frac{q_\alpha q_\beta}{q^2}\right)F_1(x,Q^2) + \left(p_\alpha - \frac{p\cdot q}{q^2}q_\alpha\right)\left(p_\beta - \frac{p\cdot q}{q^2}q_\beta\right)\frac{F_2(x,Q^2)}{p\cdot q}.$$

It will be useful to write the integration measure in terms of Bjorken scaling parameter x and the momentum transfer Q^2

$$\frac{\mathrm{d}^3\mathbf{k}'}{(2\pi)^3}\frac{1}{2k'^0} = \frac{E'\,\mathrm{d}E'\,\mathrm{d}\cos\theta'}{8\pi^2} = \frac{M_pEy}{8\pi^2}\,\mathrm{d}x\,\mathrm{d}y = \frac{sy^2}{(4\pi)^2Q^2}\,\mathrm{d}x\,\mathrm{d}Q^2.$$

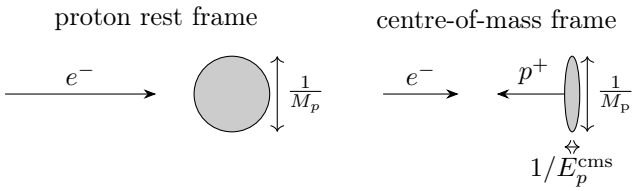
Consequently, a quantity measured in an experiment is

$$\frac{\mathrm{d}\sigma}{\mathrm{d}x\,\mathrm{d}Q^2} = \frac{4\pi\alpha^2}{Q^4}\left((1+(1-y)^2)F_1 + \frac{1-y}{x}\underbrace{(F_2-2xF_1)}_{F_\mathrm{L}}\right). \tag{5.5}$$

Structure functions F_1 and F_2 can only partially be predicted from QCD calculations. In the following, we will try to describe DIS in terms of photon scattering off quarks and gluons in the proton. Note there are two very different scales involved: Q^2 and $M_p^2 \sim \Lambda^2_\mathrm{QCD}$ – the scale of "hard" quark/gluon-photon scattering and the scale of the strong interaction bound state.

The parton model

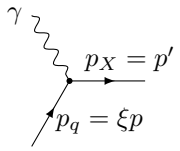
Proton is a bound state of size $1/M_\mathrm{p}$. The typical momenta of the constituents in the proton rest frame are $\mathcal{O}(M_p)$. Consider DIS in the centre-of-mass frame. The proton moves relativistically in the z-direction, so $E_p^\mathrm{cm} \gg M_\mathrm{p}$.



The transverse momenta are unaffected by the boost and $\mathcal{O}(M_\mathrm{p})$. Neglect them and consider scattering off constituents with longitudinal momenta $\xi p^\mu \gg M_\mathrm{p}$. Because of Lorentz contraction in the z-direction, the time which the electron needs to traverse the proton is small compared to $1/M_\mathrm{p}$, the timescale on which the distribution of constituents in the proton changes. The electron "sees" a "frozen" *parton distribution function* $f_i(\xi)$ of constituents of type i with momentum fraction ξ. Hence, in the parton model we can replace the proton matrix element with the sum over all constituents i (quarks and gluons) and integral over all momentum fractions ξ

$$\begin{aligned} W_{\alpha\beta} &= \frac{1}{4\pi} \sum\!\!\!\!\!\!\int_X (2\pi)^4 \delta^{(4)}(p+q-p_X) \, \langle \mathrm{p}|j_\beta^\dagger|X\rangle \, \langle X|j_\alpha|\mathrm{p}\rangle \\ &= \sum_i \int \mathrm{d}\xi \, f_i(\xi) \frac{1}{4\pi} \frac{s}{\xi s} \sum\!\!\!\!\!\!\int_X (2\pi)^4 \delta^{(4)}(p+q-p_X) \, \langle i|j_\beta^\dagger|X\rangle \, \langle X|j_\alpha|j\rangle \, . \end{aligned}$$

The lowest order process is scattering off quarks with flavour $i = u, d, s, c, b$



To calculate the hadronic tensor, we use

$$\int \frac{\mathrm{d}^3\mathbf{p}'}{(2\pi)^3} \frac{1}{2p'^0} = \int \frac{\mathrm{d}^4 p'}{(2\pi)^4} 2\pi\theta(p'^0)\delta(p'^2)$$

where we neglected quark masses. Now we write

$$\begin{aligned} W^{\alpha\beta} &= \sum_i \int \mathrm{d}\xi \, f_i(\xi) \frac{1}{4\pi\xi} \int \frac{\mathrm{d}^3\mathbf{p}'}{(2\pi)^3 2p'^0} (2\pi)^4 \delta^{(4)}(p+q-p_X) \cdot \frac{Q_i^2}{2N_\mathrm{c}} \sum_{\substack{\mathrm{pols} \\ \mathrm{colours}}} \overline{u}(p')\gamma^\beta u(\xi p)\overline{u}(\xi p)\gamma^\alpha u(p') \\ &= \sum_i \int \mathrm{d}\xi \, f_i(\xi) \frac{1}{4\pi\xi} 2\pi\delta_+((\xi p+q)^2) Q_i^2 \frac{1}{2N_\mathrm{c}} N_\mathrm{c} \, \mathrm{tr}(\not{p}'\gamma^\beta \xi \not{p} \gamma^\alpha). \end{aligned}$$

The factor $1/(2N_\mathrm{c})$ comes from the final spin and colour average of the initial quark. The expression in the δ-function can be simplified $(\xi p+q)^2 \approx 2p\cdot q(\xi - x)$. Hence, the remaining δ-function sets $\xi = x$, i.e. the momentum fraction is Bjorken's x. From that we obtain

$$\frac{\mathrm{d}\sigma}{\mathrm{d}x\mathrm{d}Q^2} = \frac{4\pi\alpha^2}{Q^4} \sum_i Q_i^2 f_i(x) \frac{1+(1-y)^2}{2}.$$

Comparing this equation with Eq. 5.5, we see $F_L(x, Q^2) = 0$ and

$$F_2(x, Q^2) = 2xF_1(x, Q^2) = \sum_i Q_i^2 c f_i(x). \tag{5.6}$$

This is called *Callan–Gross relation*.

- Structure functions are independent of Q^2 – this is referred to as *Bjorken scaling*, a consequence of the point-like nature of the constituents.
- Structure functions contain information about the electric charge (Q_i) and spin of the constituents. $F_L = 0$ is related to spin-1/2.
- The parton distribution functions fulfil sum rules. The electric charge of the proton is one (neutron's charge is zero):

  $$\int_0^1 \mathrm{d}x \sum_{i=q,\overline{q}} Q_i f_i(x) = 1 \text{ or } 0$$

  and the baryon number of the proton (and neutron) is one

  $$\int_0^1 \mathrm{d}x \sum_{i=q} \frac{1}{3}(d_i(x) - f_{\overline{i}}(x)) = 1.$$

  The antiquark distribution function is denoted by $f_{\overline{i}}(x)$.

Quantum field theoretical (operator) definition of the parton distribution

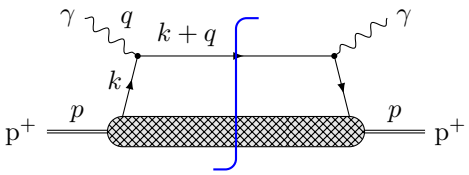
As we have seen earlier in this chapter, the hadron matrix is written as

$$W_{\alpha\beta}(p, q) = \frac{1}{4\pi} \int \mathrm{d}^4 z \, e^{iqz} \langle \mathrm{p}|j_\beta^\dagger(z) j_\alpha(0)|\mathrm{p}\rangle$$

with

$$\langle \mathrm{p}|j_\beta^\dagger(z) j_\alpha(0)|\mathrm{p}\rangle = \sum_{i,j} Q_i Q_j \langle \mathrm{p}|\overline{q}_i(z)\gamma_\beta q_i(z)\overline{q}_j(0)\gamma_\alpha q_j(0)|\mathrm{p}\rangle \,.$$

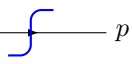
The expression can be represented pictorially.



The grey crosshatched area represents a collinear set of particles – a proton remnant that is one quark off. This quark carries momentum k (represented by vertical line) and is then scattered by a photon of momentum q. The scattered quark with large momentum $k+q$ is the horizontal fermion line. The blue line essentially represents the cut – square the amplitude and integrate over phase space. Since the process corresponds to high-energy (hard) scattering, the contraction $q_i(z)\overline{q}_j(0)$ can be calculated perturbatively. But what even is $q_i(z)\overline{q}_j(0)$? It is not a quark propagator,

since there is no time ordering. Rather, it is a "cut propagator":

$$\begin{aligned}\langle 0|q_{ia}(z)\overline{q}_{jb}|0\rangle &= \langle 0|\int \frac{\mathrm{d}^3\mathbf{p}_1}{(2\pi)^3 2p_1^0}\frac{\mathrm{d}^3\mathbf{p}_2}{(2\pi)^3 2p_2^0}\sum_{s_1,s_2} e^{-ip_1 z}u(p_1,s_1)\overline{u}(p_2,s_2)a_{ia}(p_1,s_1)a^\dagger_{jb}(p_2,s_2)|0\rangle \\ &= \int \frac{\mathrm{d}^3\mathbf{p}_1}{(2\pi)^3 2p_1^0}e^{-ip_1 z}\delta_{ij}\delta_{ab}\not{p}_1 \\ &= \delta_{ij}\delta_{ab}\int \frac{\mathrm{d}^4 p}{(2\pi)^4}\,2\pi\delta_+(p^2)e^{-ipz}\not{p} = \end{aligned}$$



The blue line is the symbol for a cut propagator. In the derivation we used the Fock vacuum $|0\rangle$ like for the free propagator. Now we write

$$W_{\alpha\beta}(p,q) = \frac{1}{4\pi}\sum_i Q_i^2\int\frac{\mathrm{d}^4 p'}{(4\pi)^4}\,2\pi\delta_+(p'^2)\int \mathrm{d}^4 z\, e^{i(q-p')z}\,\langle \mathrm{p}|\overline{q}_i(z)\gamma_\beta \not{p}'\gamma_\alpha q_i(0)|\mathrm{p}\rangle\,.$$

This indeed looks similar to the parton model expression. To extract the expression for $f_i(\xi)$, parametrize the momenta as follows:

$$p^\mu = (p,0,0,p),\quad n^\mu \equiv (\frac{1}{2p},0,0,\frac{1}{2p}),\quad n^2 = 0,\ n\cdot p = 1.$$

The momentum q decomposes as

$$q^\mu = (p\cdot q)n^\mu + q_\perp^\mu,\quad q_\perp\cdot p = q_\perp\cdot n = 0.$$

This is called the *Sudakov decomposition*. Since $F_L = 0$, we project on $F_2 = (p\cdot q)n^\alpha n^\beta W_{\alpha\beta}(p,q)$. Using $\not{n}\not{p}'\not{n} = 2(n\cdot p')\not{n}$ and $p' = q + k$ we obtain

$$F_2(x,Q^2) = \frac{p\cdot q}{4\pi}\sum_i Q_i^2\int\frac{\mathrm{d}^4 k}{(2\pi)^4}\,2\pi\delta_+((q+k)^2)2n\cdot(q+k)\times\int \mathrm{d}^4 z\, e^{-ikz}\,\langle \mathrm{p}|\overline{q}_{ia}(z)\not{n}q_{ia}(0)|\mathrm{p}\rangle\,. \tag{5.7}$$

Since $n\cdot q = 0$, the expression $2n\cdot(q+k)$ simplifies to $2n\cdot k$. Similarly, as we earlier decomposed q we can also parametrize k^μ in terms of ξ, k^2 and $k_\perp^\mu$:

$$k^\mu = \xi p^\mu + \frac{k^2 + k_\perp^2}{2\xi}n^\mu + k_\perp^\mu.$$

In this parametrization, ξp^μ is large and all the rest is small by nature of the proton state (the momentum of the proton is much greater than its mass, so the internal movement of partons inside the proton doesn't contribute much to their total momentum). Now, using $\xi = (n\cdot k)/(n\cdot p) = n\cdot k$ and $x = 2p\cdot q/Q^2$, we write

$$\delta((q+k)^2) \approx \delta(2\xi p\cdot q - Q^2)0\frac{1}{2p\cdot q}\delta(n\cdot k - x) = \frac{1}{2p\cdot q}\int\frac{\mathrm{d}\lambda}{2\pi}e^{-i\lambda(x-n\cdot k)}$$

to get

$$\begin{aligned}F_2(x,Q^2) &= \frac{p\cdot q}{4\pi}\sum_i Q_i^2 2\pi\, 2x\frac{1}{p\cdot q}\int\frac{\mathrm{d}\lambda}{4\pi}\,e^{-i\lambda x}\int\mathrm{d}^4 z\int\frac{\mathrm{d}^4 k}{(2\pi)^4}e^{-i(z-\lambda n)k}\,\langle \mathrm{p}|\overline{q}_{ia}(z)\not{n}q_{ia}(0)|\mathrm{p}\rangle \\ &= \sum_i Q_i^2 x\int\frac{\mathrm{d}\lambda}{4\pi}e^{-i\lambda x}\,\langle \mathrm{p}|\overline{q}_{ia}(\lambda n)\not{n}q_{ia}(0)|\mathrm{p}\rangle\end{aligned} \tag{5.8}$$

By comparing to eq. 5.6 we can identify the parton distribution function as

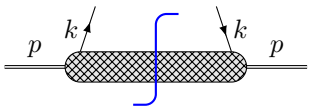
$$f_i(x) = \int \frac{\mathrm{d}\lambda}{4\pi} e^{-i\lambda x} \langle \mathrm{p}|\overline{q}_{ia}(\lambda n)\not{n} q_{ia}(0)|\mathrm{p}\rangle \,.$$

There is an alternative derivation of 5.8 from 5.7. Start again with

$$F_2(x,Q^2) = \frac{p\cdot q}{4\pi}\sum_i Q_i^2 \int \frac{\mathrm{d}^4k}{(2\pi)^4}\, 2\pi\delta_+(p'^2) 2n\cdot p' \times \int \mathrm{d}^4z\, e^{-i(q-p')z} \langle \mathrm{p}|\overline{q}_{ia}(z)\not{n} q_{ia}(0)|\mathrm{p}\rangle \,.$$

Define

$$B_i(p.k) \equiv \int \mathrm{d}^4z'\, e^{-ikz'} \langle \mathrm{p}|\overline{q}_{ia}(z)\not{n} q_{ia}(0)|\mathrm{p}\rangle =$$



Thus the structure function is expressed as

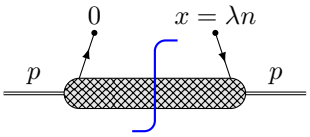
$$\begin{aligned} F_2(x,Q^2) &= p\cdot q \sum_i Q_i^2 \int \frac{\mathrm{d}^4p'}{(2\pi)^4}\frac{\mathrm{d}^4k}{(2\pi)^4}\delta_+(p'^2)\, n\cdot p'\, B_i(p,k) \int \mathrm{d}^4z\, e^{i(q+k-p')z} \\ &= p\cdot q \sum_i Q_i^2 \int \frac{\mathrm{d}^4k}{(2\pi)^4}\delta_+((k+q)^2)\, n\cdot(q+k)\, B_i(p,k). \end{aligned}$$

Now we use the same decomposition of k^μ as in the previous derivation and write $\delta((q+k)^2) = \delta(\xi - x)/(2p\cdot q)$. So one obtains

$$F_2(x,Q^2) = \sum_i Q_i^2 \frac{1}{2}\int \frac{\mathrm{d}^4k}{(2\pi)^4}\delta(x - n\cdot k) n\cdot k B_i(p,k) = \sum_i Q_i^2 x f_i(x).$$

We can interpret this as $f_i(\xi)$ being $B_i(p,k)$ integrated over $k_\perp$ and k^2:

$$f_i(\xi) = \frac{1}{4\pi}\int \mathrm{d}\lambda\, e^{-i\lambda\xi} \langle \mathrm{p}|\overline{q}_{ia}(\lambda n)\not{n} q_{ia}(0)|\mathrm{p}\rangle \,\Big|_{\substack{\text{spin}\\ \text{averaged}}} = \frac{1}{4\pi}\int \mathrm{d}\lambda\, e^{-i\lambda\xi}$$



This is the operator definition of the parton distribution (with an amendment to be made later). Note the fields $q_i(0)$ and $\overline{q}_i(\lambda n)$ are separated by a *light-like distance* $x^2 = \lambda^2 n^2 = 0$.

Operator renormalization

After introducing field renormalization $\phi_0(x) = \sqrt{Z_\phi}\phi(x)$ the Green functions and matrix elements of renormalized fields $\phi(x_1)\ldots\phi(x_n)$ are ultraviolet finite if all x_i are different. Products of fields at the same point require an additional renormalization of the composite local operator, since the limit $x_i \to x_j$ generates additional UV divergences (p. ??), e.g. $\phi_0^2(x) \equiv Z_{\phi^2}\phi^2(x)$ but $Z_{\phi^2} \neq \sqrt{Z_\phi}^2$. This is also the reason why mass counter-terms require an additional renormalization which in this case is associated with the mass counterterm.

Moreover, additional divergences in a field product $\phi(0)\phi(0)$ appear not only for $x \to 0$ but already for $x^2 \to 0$ (which is not the same in Minkowski space). These are called *light-cone divergencies*. In momentum space they correspond to the situation that some components

$k^\mu \to \infty$ while k^2 remains small or constant, as opposed to $k^\mu \to \infty$ for all components. Light-cone divergencies are associated with collinear divergencies (because we could also consider the limit where k^μ are fixed and $k^2 \to 0$). We renormalize a composite operator as follows

$$[\overline{q}(\lambda n)\not{n} q(0)]_0 = \int \mathrm{d}\lambda'\, Z(\lambda,\lambda')[\overline{q}(\lambda' n)\not{n} q(0)](\mu)$$

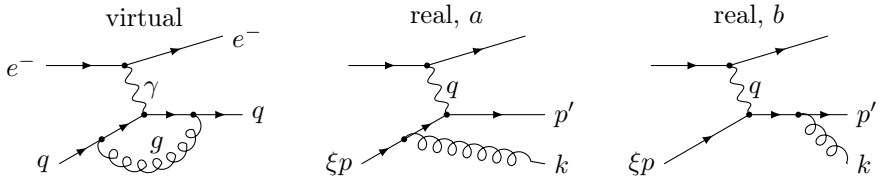
where the term on the left-hand side is the bare light-cone operator (or rather a family of operators with parameter λ), $Z(\lambda,\lambda')$ is a "matrix" of renormalization constants and the term on the right is renormalized operator that now depends on μ. From this we conclude that the parton distribution does not depend on Q but on the arbitrary renormalization scale μ. The formula cannot yet be complete

$$F_2(x,Q^2) = \sum_i Q_i^2 x f_i(x,\mu).$$

We naturally set $\mu = Q$ since Q is the only physical scale. However, if there are any scaling violations, they have to be calculated from the operator renormalization.

5.3.1 Factorization and DGLAP evolution

In this section we will try to compute quantum corrections to the parton model.



Apart from corrections listed above there is also the field renormalization (Z-factor) which counts towards virtual divergencies. The scattering matrix element for a is

$$\langle qg|\overline{q}_i\gamma^\mu q_i|q\rangle = i g_\mathrm{s}\overline{u}(p')\gamma^\mu \frac{i}{\xi\not{p} - \not{k}}\gamma^\rho T^A u(\xi p)\varepsilon_\rho(k,\lambda).$$

We choose *axial gauge* fixing with a light-like vector n^μ. Recall the ghost-fixing Lagrangian in axial gauge

$$\mathcal{L}_\mathrm{GF} = -\frac{1}{2}(n^\mu A_\mu^A)^2, \quad n^2 = 0$$

and the gluon propagator

$$\frac{i}{k^2 + i\varepsilon}\left(-g_{\mu\nu} + \frac{k_\mu n_\nu + k_\nu n_\mu}{n\cdot k}\right) = \frac{i}{k^2+i\varepsilon}\sum_{\lambda=1,2}\varepsilon_\mu(k,\lambda)\varepsilon_\nu^*(k,\lambda).$$

There are no ghosts and only two transverse polarization states in the propagator. The square of the matrix element summed over final polarizations and averaged over initial polarizations reads

$$\overline{|\mathcal{M}|^2}_{\mu\nu} = \frac{1}{2N_\mathrm{c}} g_\mathrm{s}^2\,\mathrm{tr}(T^A T^B)\frac{\mathrm{tr}(\gamma^\nu \not{p}'\gamma^\mu(\xi\not{p} - \not{k})\gamma^\rho \xi\not{p}\gamma^\sigma(\xi\not{p}-\not{k}))}{(\xi p - k)^4}\left(-g_{\rho\sigma} + \frac{k_\rho n_\sigma + k_\sigma n_\rho}{n\cdot k}\right).$$

This is the square of the part a alone. The full result would be $|a+b|^2$. We will come to this later. This time, we use a different parametrization for k:

$$k^\mu = \xi(1-\xi')p^\mu + \frac{k_\perp^2}{2\xi(1-\xi')}n^\mu + k_\perp^\mu \quad \text{such that} \quad (\xi p - k)^\mu = \xi\xi' p^\mu - \frac{k_\perp^2}{2\xi(1-\xi')}n^\mu - k_\perp^\mu$$

and $k^2 = 0$ as it must be. After radiating the gluon the quark carries momentum fraction $\xi\xi'$. After some elementary Dirac algebra one finds

$$n^\mu n^\mu \overline{|\mathcal{M}|^2_{\mu\nu}} = g_\mathrm{s}^2 C_\mathrm{F} \frac{8\xi'(1+\xi'^2)}{k_\perp^2}\xi^2.$$

For the phase space, we write

$$\begin{aligned}\mathrm{dPS} &= \frac{\mathrm{d}^3\mathbf{p}'}{(2\pi)^3 2p'^0}\,\frac{\mathrm{d}^3\mathbf{k}}{(2\pi)^3 2k^0}(2\pi)^4\delta^{(4)}(\xi p + q - p' - k)\\ &= \frac{\mathrm{d}^3\mathbf{k}}{(2\pi)^3 2k^0}2\pi\theta(k^0 - \xi p^0 - q^0)\delta((\xi p + q - k)^2).\end{aligned}$$

To convert the measure in k to measure in ξ' choose $k_\perp^\mu = (0, k^1, k^2, 0)$ and $\theta = \angle(k_\perp, q_\perp)$. Then $2\mathrm{d}^3\mathbf{k} = \mathrm{d}k_\perp^2\,\mathrm{d}k^3\mathrm{d}\theta$ and

$$\frac{\mathrm{d}k^3}{k^0} = \frac{\mathrm{d}(\xi(1-\xi')p^0 - \frac{1}{2p^0}\frac{k_\perp^2}{2\xi(1-\xi')})}{\xi(1-\xi')p^0 + \frac{1}{2p^0}\frac{k_\perp^2}{2\xi(1-\xi')}} = \frac{\mathrm{d}\xi'}{1-\xi'}.$$

It also holds

$$(q+\xi p - k)^2 = -Q^2 + 2\xi p\cdot q - 2k\cdot q - 2\xi p\cdot k = -Q^2\left(1 - \frac{\xi\xi'}{x}\right) + 2k_\perp\cdot q_\perp - \frac{k_\perp^2}{1-\xi}.$$

Putting everything together in the parton model formula gives

$$\begin{aligned}F_2^a(x,Q^2) &= \sum_i Q^2\, p\cdot q\frac{g_\mathrm{s}^2 C_\mathrm{f}}{4\pi}\int_0^1 \mathrm{d}\xi\, f_i^{(0)}(\xi)\frac{1}{16\pi^2\xi}\int\frac{\mathrm{d}\xi'}{1-\xi'}\,\mathrm{d}k_\perp^2\,\mathrm{d}\theta\cdot\\ &\qquad\cdot\delta\Big(-Q^2(1-\xi\xi'/x) + 2\sqrt{k_\perp^2}Q\cos\theta - k_\perp^2/(1-\xi')\Big)\,\xi^2\frac{8\xi'(1+\xi'^2)}{k_\perp^2}\\ &= \sum_i Q_i^2\alpha_\mathrm{s}C_\mathrm{F}\int_0^1\mathrm{d}\xi\, f_i(0)_i(\xi)\frac{1}{32\pi^2}\int\mathrm{d}k_\perp^2\,\mathrm{d}\theta\,\frac{\mathrm{d}\xi'}{1-\xi'}\cdot\\ &\qquad\cdot\delta\Big(\xi' - \frac{x}{\xi} - \frac{x}{\xi Q^2}\Big(\frac{k_\perp^2}{1-\xi'} - 2\sqrt{k_\perp^2}Q\cos\theta\Big)\Big)\,\frac{8\xi'(1+\xi'^2)}{k_\perp^2}.\end{aligned}$$

We used $k_\perp\cdot q_\perp = \sqrt{k_\perp^2}Q\cos\theta$ and $Q = \sqrt{Q^2}$. The superscript on $f_i^{(0)}(\xi)$ will be explained later. Now we split the integral over $k_\perp^2$ into intervals $(0,\mu^2)$ and (μ^2,∞) where $\mu^2 \ll Q^2$. In the integral over small $k_\perp^2$ one can approximate the δ-function to $\delta(\xi'-x/\xi)$. Then the ξ' integral and θ integral can be done trivially. Including also the leading-order contribution that we calculated before gives

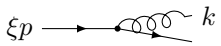
$$F_2(x,Q^2) = \sum_i Q_i^2\int_x^1\mathrm{d}u\,\frac{x}{u}f_i^{(0)}\Big(\frac{x}{u}\Big)\left(\delta(1-u) + \frac{\alpha_\mathrm{s}}{2\pi}\hat{P}_{qq}(u)\int_0^{\mu^2}\frac{\mathrm{d}k_\perp^2}{k_\perp^2} + \frac{\alpha_\mathrm{s}}{2\pi}\left[\hat{P}_{qq}(u)\ln\frac{Q^2}{\mu^2} + R(u)\right]\right). \tag{5.9}$$

Since $\xi' \leq 1$ and $\xi \geq x$, we defined u by $\xi = x/u$. A new function was introduced, called the quark-quark *splitting function*

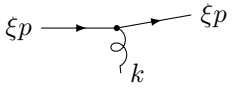
$$\hat{P}_{qq}(u) = C_{\mathrm{F}} \frac{1+u^2}{1-u}.$$

More on it later. The splitting of the integral over two intervals is still visible in the expression: the part that is left as an integral over $k_\perp^2$ corresponds to the interval $(0, \mu^2)$, the remaining part is in the square brackets – it is not explicitly calculated and is parametrized by $R(u)$ which could be obtained with some extra work. However, we know that the μ-dependence must cancel in the sum of the small $k_\perp$ and large $k_\perp$ contribution, so we know it must be $\hat{P}_{qq} \ln \mu^2$ and then Q^2 is the only dimensionful quantity that can render the argument of the logarithm dimensionless. The expression 5.9 is divergent in two different ways

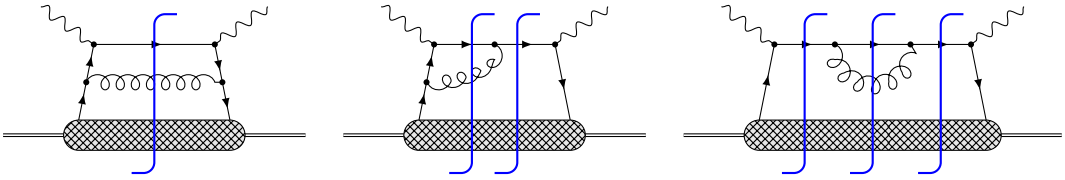
- $k_\perp \to 0$ gives collinear divergence, $\mathbf{k} \parallel \xi\mathbf{p}$,



- $u \to 1$ or $x \to \xi$ or $\xi' \to 1$ gives soft divergence (no component in k that scales with the large momentum p).



Some of these divergences remain when all other contributions are added. The latter have the following form.



$|a|^2$, calculated above ab^*, virtual correction $|b|^2$, final quark renormalization

- The first diagram represents a gluon emitted from the incoming quark (and the amplitude is squared), $|a|^2$.
- The left cut on the middle diagram represents a product of two different amplitudes – one where initial quark radiates a gluon and one where final quark radiates a gluon, ab^*. The right cut represents a product of tree level amplitude and an amplitude with virtual vertex correction. There is another such diagram with two cuts and gluon on the right side – the two cuts would represent the complex conjugates of the expressions in the middle drawn diagram.
- The right diagram has three possible cuts and three contributions to the total amplitude. The middle cut is the product of two amplitudes with gluon emitted from the final quark, $|b|^2$. The remaining two cuts represent a product of a leading order process and a process with final quark renormalization correction.

It can be shown, however, that none of the cuts in the middle and right diagrams give collinear divergence in axial gauge. The only diagram we haven't considered yet is the one loop field renormalization for the incoming quark.



These diagrams only contribute for $u = 1$ or $\xi = x$, since the emitted gluon is reabsorbed, so no matter its momentum, the incoming quark will have momentum $\xi\xi' p = \xi p$. This modifies the splitting function a bit, namely

$$\hat{P}_{qq} = C_{\mathrm{F}} \left(\frac{1+u^2}{[1-u]_+} + \frac{3}{2}\delta(1-u) \right).$$

Here, $[]_+$ denotes the "plus distributions" defined on a test function g as

$$\int_0^1 \mathrm{d}z \, [f(z)]_+ g(z) \equiv \int_0^1 \mathrm{d}z f(z)(g(z) - g(1)).$$

The soft divergence as $u \to 1$ therefore cancels in the sum of real and virtual corrections, but the *collinear divergence remains*. The scattering of a photon from a quark with definite momentum indeed violates the criterion for (collinear) IR safety, since it is sensitive to collinear splitting: the quark momentum changes, but the observable identifies it. The solution is a reinterpretation of the parton distribution. Physically the situation looks as follows.



$k_\perp$ large ($\sim Q$) $k_\perp$ small ($\ll Q$, $\sim \Lambda_{\mathrm{QCD}}$)

In the case of large virtuality $(\xi p - k)^2 = \mathcal{O}(Q^2)$ (figure on the left) the scattering $\gamma q \to qg$ is hard and should be calculable in perturbation theory. For small virtuality (figure on the right), the only hard process is $\gamma + q \to q$, and $q \to q + g$ only changes the quark momentum distribution. Hence, define the "renormalized" parton distribution

$$f_i(\xi) = f_i^{(0)}(\xi) + \delta f_i(x), \quad \text{with} \quad \delta f_i(\xi) \equiv \frac{\alpha_{\mathrm{s}}}{2\pi} \int_\xi^1 \mathrm{d}v \, f_i^{(0)}(v) \hat{P}_{qq}(\xi/v) \int_0^{\mu^2} \frac{\mathrm{d}k_\perp^2}{k_\perp^2}.$$

The correction is formally divergent, so it cannot be computed in perturbation theory. Hence, $f_i(\xi)$ must be determined from experiment and is finite. Exactly as for renormalized parameters, where also the bare parameter and counter-terms are divergent. See also the discussion of the operator definition of $f_i(\xi)$ and the operator renormalization above (eq. 5.9)

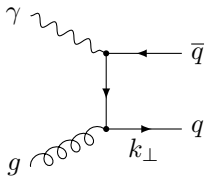
$$F_2(x, Q^2) = \sum_i Q_i^2 \int_x^1 \mathrm{d}u \, \frac{x}{u} f_i(x/u, \mu) \left(\delta(1-u) + \frac{\alpha_{\mathrm{s}}}{2\pi} \left[\hat{P}_{qq}(u) \ln \frac{Q^2}{\mu^2} + R(u) \right] \right). \tag{5.10}$$

This can be generalized to all orders in expansion in α_{s}:

$$F_2(x, Q^2) = \sum_i Q_i \int_x^1 \mathrm{d}u \, \frac{x}{u} f_i(x/u, \mu) C_i(\alpha_{\mathrm{s}}(\mu), \ln(Q^2/\mu^2); u).$$

It is the prototype of a "factorization formula": C_i carries the short-distance information at scale Q and expands in $\alpha_s \ll 1$, while f_i is non-perturbative long-distance information at scale $\Lambda_{\rm QCD}$. The principle is similar to effective field theory. A factorization formula refers to a specific process, an effective Lagrangian to the entire system. Factorization formulae are crucial to use in perturbation theory of quarks and gluons to describe physical processes, the parton distributions are the objects that mediate between hadrons and the quark-gluon description – up to corrections $\mathcal{O}(\Lambda_{\rm QCD}/Q^2)$.

The formula 5.10 is not yet complete. The right figure above shows that the $q \to qg$ process leads to a gluon distribution $f_g(\xi)$ in the proton. The following process occurs.



For large $k_\perp$ the scattering $\gamma g \to q\bar{q}$ is hard. If $k_\perp$ is small, $g \to q\bar{q}$ is collinear and the only hard part is $\gamma\bar{q} \to \bar{q}$. Apart from the gluon distribution, one also needs the anti-quark distribution in the proton. Combining all possible constituents i of a hadron H, we can write the differential cross-section for the process $e + H \to e + H$ as

$$\frac{\mathrm{d}\sigma}{\mathrm{d}x\mathrm{d}Q^2} = \sum_{i=q_f,\bar{q}_f,g} \int_x^1 \mathrm{d}u\, f_{i/H}(x/u,\mu)\, \hat{\sigma}_{e+i\to e+X}(\alpha_s(\mu), \ln(Q^2/\mu^2); u)$$

where $\hat{\sigma}$ denotes scattering cross-section in which collinear singularities are subtracted into $f_{i/H}$ (the momentum distribution function of parton i in hadron H).

- In particular calculations one does not separate the $k_\perp^2$ integral into two regions explicitly. Instead, we use dimensional regularization. Then

$$\int_0^{Q^2} \frac{\mathrm{d}k_\perp^2}{k_\perp^2} \to \tilde{\mu}^{2\varepsilon} \int_0^{Q^2} \frac{\mathrm{d}^{2-2\varepsilon}k_\perp}{k_\perp^2} = \pi\left(-\frac{1}{\varepsilon} + \ln\frac{Q^2}{\mu^2}\right).$$

  The divergent part $1/\varepsilon$ is absorbed in δf_i and the logarithm goes to $\hat{\sigma}$. Minimal subtraction of the $1/\varepsilon$ pole defines $f_{i/H}(\xi,\mu)$ in the $\overline{\rm MS}$scheme.

- C_i and $\hat{\sigma}$ respectively contain $\alpha_s(\mu)\ln(Q^2/\mu^2)$. To avoid the breakdown of perturbation theory one should use $\mu \sim Q$. This sums $\left(\alpha_s(\mu)\ln(Q^2/\mu^2)\right)^n$ to all orders (these are called *collinear logarithms*). One then needs to know how $f_{i/H}(\xi, u = Q)$ changes when Q changes.

DGLAP evolution

Even though $f_{i/H}(\xi,\mu)$ is non-perturbative, once its value is measured at certain μ_0 its μ-dependence can be computed using renormalization group flow. The renormalization group evolutions in this case are called *DGLAP evolution equations* (named after Dokshitzer, Gribov, Lipatov, Altarelli and Parisi). To find the equation use the fact that the physical cross-section must be independent of renormalization scale

$$0 = \mu^2 \frac{\mathrm{d}}{\mathrm{d}\mu^2}\frac{\mathrm{d}\sigma}{\mathrm{d}x\mathrm{d}Q^2} = \sum_{i=q_f,\bar{q}_f,g} \int_x^1 \mathrm{d}u \left(\hat{\sigma}\mu^2\frac{\mathrm{d}}{\mathrm{d}\mu^2} f_{i/H} + f_{i/H}\mu^2\frac{\mathrm{d}}{\mathrm{d}\mu^2}\hat{\sigma}\right).$$

Since the running of $\hat{\sigma}$ is known from computation of the parton process, we can extract the running of $f_{i/H}$. For example, from eq. 5.10 follows

$$0 = \int_x^1 \mathrm{d}u \left(\mu^2 \frac{\mathrm{d}}{\mathrm{d}\mu^2} \left(\frac{x}{u} f_i(x/u,\mu)\delta(1-u) \right) + \frac{x}{u} f_i(x/u) \frac{\alpha_\mathrm{s}}{2\pi} (-\hat{P}_{qq}(u)) + \mathcal{O}(\alpha_\mathrm{s}^2) \right).$$

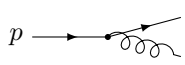
Next, divide by x and define $y \equiv x/u$ to obtain an integro-differential equation

$$\mu^2 \frac{\mathrm{d}}{\mathrm{d}\mu^2} f_i(x,\mu) = \frac{\alpha_\mathrm{s}}{2\pi} \int_x^1 \frac{\mathrm{d}y}{y} \hat{P}_{qq}(x/y) f_i(y,\mu).$$

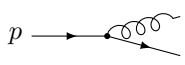
This is one of the DGLAP evolution equations. The general result, including higher-order α_s^n corrections and all partons has the form

$$\mu^2 \frac{\mathrm{d}}{\mathrm{d}\mu^2} f_{i/H}(x,\mu) = \sum_{j=q,\overline{q},g} \int_x^1 \frac{\mathrm{d}y}{y} P_{ij}(x/y) f_{j/H}(y,\mu)$$

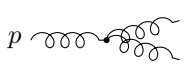
where $P_{ij} = \alpha_\mathrm{s} \hat{P}_{ij}/(2\pi) + \mathcal{O}(\alpha_\mathrm{s}^2)$. The splitting functions are related to the Z-factors of the light-cone operators and can be interpreted as their "anomalous dimension". The relevant cases are:



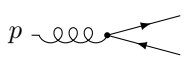
$$\hat{P}_{q\leftarrow q}(x) = C_\mathrm{F} \left(\frac{1+x^2}{[1-x]_+} + \frac{3}{2}\delta(1-x) \right),$$



$$\hat{P}_{g\leftarrow q}(x) = C_\mathrm{F} \frac{1+(1-x)^2}{x},$$



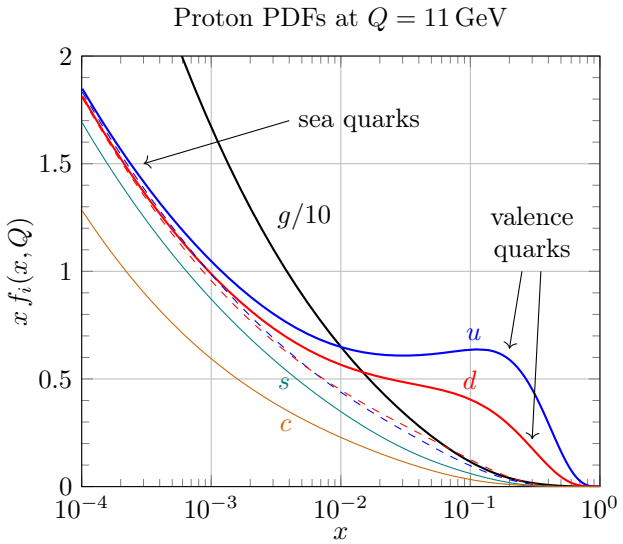
$$\hat{P}_{g\leftarrow g}(x) = C_\mathrm{F} \left(\frac{x}{[1-x]_+} + \frac{1-x}{x} + x(1-x) \right) + \frac{11C_\mathrm{A} - 4n_f T_\mathrm{F}}{6} \delta(1-x),$$



$$\hat{P}_{q\leftarrow g}(x) = n_f T_\mathrm{F} (x^2 + (1-x)^2).$$

Here, n_f is the number of relevant quark flavours and $T_\mathrm{F} = 1/2$. Identical equations hold for anti-quarks.

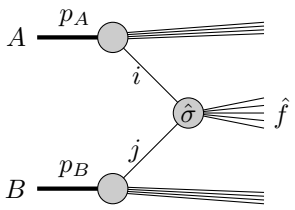
- $P_{i\leftarrow j}$ means that j splits into i with momentum fraction x.
- The $1/x$ factors in $P_{g\leftarrow q}$ and $P_{g\leftarrow g}$ imply that the gluon distribution function in the proton (and then because of $\hat{P}_{q\leftarrow g}$ also for q and $\overline{q}$) increases very rapidly for small x and large Q^2. Hence, for high-energy scattering, the proton is much more a bundle of gluons with relatively small momentum fraction than of three quarks with momentum fraction 1/3 each.



At $x \approx 1$ we see that approximately one third of momentum is carried by u quarks and one third by d quarks (these are called the *valence quarks*). At lower x however the interaction $g \to q\overline{q}$ produces $u, \overline{u}, d, \overline{d}$ in almost the same proportions – these are called the *sea quarks* and their mixture together with gluons is called the *sea*.

5.3.2 Proton-proton collisions

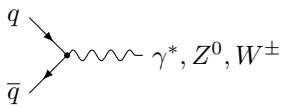
The above description can be generalized to hadron-hadron collisions $A+B \to f$. The important point, which makes this possible, is that the parton distribution functions are universal; $f_{i/H}(\xi, \mu)$ depend only on hadron H and are independent of the scattering process whenever f is such that a factorization formula can be derived. Consider the following process.



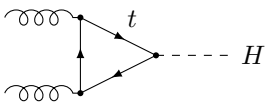
The final state f is the sum $\hat{f}$ and the fragments of hadrons A and B. The total cross-section is given by the formula

$$\mathrm{d}\sigma_{A+B\to f} = \sum_{i,j=q,\overline{q},g} \int \mathrm{d}\xi_1\,\mathrm{d}\xi_2\, f_{i/A}(\xi_1) f_{j/B}(\xi_2) \cdot \mathrm{d}\hat{\sigma}_{i+j\to\hat{f}}.$$

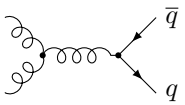
- One computes the partonic processes $i + j \to \hat{f}$. Three examples are shown below.



Drell–Yan process
$\hat{f} = Z^0, \gamma^*, W^\pm$



Higgs production
$\hat{f} = H$



Jet production
$\hat{f} = j + j$

The quantum corrections to these processes are IR divergent.

- For *suitable* final states the soft divergences and the collinear divergences associated with the final state cancel. The collinear divergences associated with the two initial partons, whose momentum is resolved, are factored into the parton distributions. This is a non-trivial statement on the structure of divergences.

Often $\mathrm{d}\hat{\sigma}_{i+j\to\hat{f}}$ depends on ξ_1 and ξ_2 only through the partonic centre-of-mass energy $\hat{s} = (p_i + p_j)^2 = (\xi_1 p_A + \xi_2 p_B)^2 = \xi_1\xi_2 s$. Then

$$\mathrm{d}\sigma_{A+B\to f} = \sum_{i,j} \int_{\tau_{\min}}^{1} \mathrm{d}\tau\, \mathscr{L}_{ij}(\tau) \mathrm{d}\hat{\sigma}_{i+j\to\hat{f}}$$

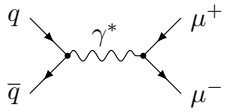
with *parton luminosity* defined as

$$\mathscr{L}_{ij}(\tau) = \int \mathrm{d}\xi_1\, \mathrm{d}\xi_2\, \delta(\tau - \xi_1\xi_2) f_{i/A}(\xi_1) f_{j/B}(\xi_2)$$

and $\hat{s}_{\min} = \tau_{\min} s$ the minimal partonic cms energy required to produce $\hat{f}$. Take for example the Higgs production, $\hat{s}_{\min} = m_H^2 = (125\,\mathrm{GeV})^2$. For $\sqrt{s} = 7\,\mathrm{TeV}$ where the Higgs was discovered in 2012 it holds $\xi_1\xi_2 \geq m_H^2/s \approx 3.2 \cdot 10^{-4}$. For $gg \to H$ one expects $\xi_1 \approx \xi_2$, so the typical momentum fraction of the gluon is around 0.02.

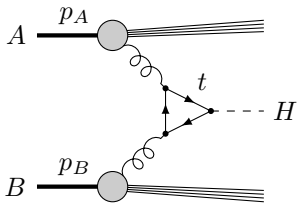
Drell–Yan process and Higgs production

The *Drell–Yan process* is a production of a colourless, single particle at leading order. Historically, the process first studied was $q\overline{q} \to \gamma^*$ with mass (virtuality) Q and then $\gamma^* \to \mu^+\mu^-$.



Alternatively, instead of photon there could be Z^0 or $W^\pm$ bosons. The cross-section is a simple expression $\hat{\sigma}_{q\overline{q}\to W} = A\delta(\hat{s} - m_W^2)$ for some factor of A.

The Higgs production at the LHC is predominantly also of this form.

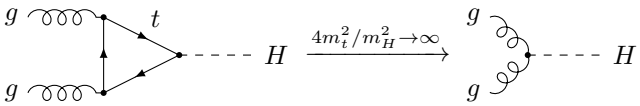


The top quark loop dominates because $\lambda_t \approx 1$. The differential cross-section for this process can be written as

$$\hat{\sigma}_{gg\to H} = \frac{\alpha_s^2 m_H^2}{576\pi v^2} |f(y)|^2 \delta(\hat{s} - m_H^2).$$

The function $f(y)$ can be calculated from the top quark loop, however for large arguments $y = 4m_t^2/m_H^2$ the function limits to one. Even though $y \approx 7.8$ the limit $y \to \infty$ still provides a

good approximation. This corresponds to replacing the top loop by a local Higgs-gluon coupling.



The effective vertex has the form (any other structure would be prohibited by gauge invariance)

$$\mathcal{L}_{\text{eff}} = \frac{\alpha_{\text{s}}}{12\pi v} H G^A_{\mu\nu} G^{A,\mu\nu}.$$

The Higgs production cross-section is then

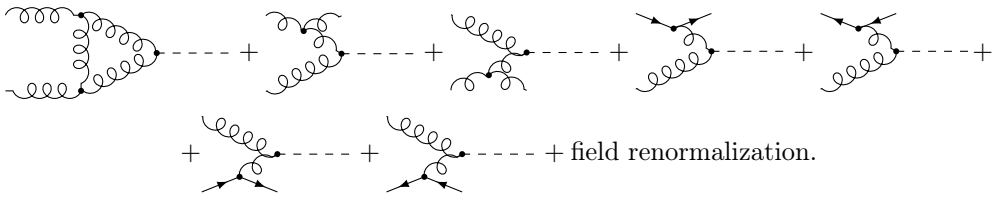
$$\sigma_{pp\to H} = \int_{m_H^2/s}^1 \mathrm{d}\tau \mathscr{L}_{gg}(\tau)\hat{\sigma}_{gg\to H}(\tau s) = \frac{\alpha_{\text{s}}^2 m_H^2}{576\pi v^2}\frac{1}{s}\mathscr{L}_{gg}(m_H^2/s).$$

Compare this to the W production

$$\sigma_{pp\to W} = \frac{\pi m_W^2}{6v^2}\frac{1}{s}\sum_{q\bar{q}',\bar{q}q'} |V_{qq'}|^2 \mathscr{L}_{q\bar{q}'}(m_W^2/s).$$

Quantum corrections to DY production

Let's use Higgs production as an example. The next-to leading processes are shown below.



All these processes are IR-divergent. In order for the factorization formula to be consistent, the IR singularities must have the following form

$$\sigma_{i+j\to H+X} = \sum_{i,j}\int \mathrm{d}\xi_1\,\mathrm{d}\xi_2\, f_{i/g}(\xi_1) f_{j/g}(\xi_2)\hat{\sigma}_{i+j\to H+X}.$$

We have replaced A and B with two gluons, so the parton distribution functions change, for example $f_{i/g}(\xi_1) = \delta(1-\xi_1)\delta_{ig} + \delta f^{(1)}_{i/g}(\xi_1) + \dots$. Importantly, however, $\hat{\sigma}_{i+j\to H+X}$ must remain IR-finite. Now we compare the amplitude order-by-order. The zeroth order gives

$$\sigma^{(0)}_{g+g\to H+X} = \hat{\sigma}^{(0)}_{g+g\to H+X} = \frac{\alpha_{\text{s}}^2 m_H^2}{576\pi v^2}\delta(\hat{s}-m_H^2) \equiv A\delta(\hat{s}-m_H^2).$$

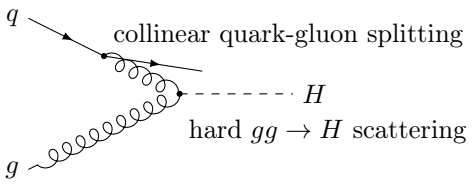
At the first order it holds

$$\begin{aligned}
\sigma^{(1)}_{gg\to Hx} &= \hat{\sigma}^{(1)}_{gg\to HX} + \sum_i \int \mathrm{d}\xi_1\, \delta f^{(1)}_{i/g}(\xi_1)\hat{\sigma}^{(0)}_{ig\to HX}(\xi_1 s) + \sum_j \int \mathrm{d}\xi_2\, \delta f^{(1)}_{j/g}(\xi_2)\hat{\sigma}_{gj\to HX}(\xi_2 s) \\
&= \hat{\sigma}^{(1)}_{gg\to HX} + 2A\int \mathrm{d}\xi\, \delta(\xi s - m_H^2)\delta f^{(1)}_{g/g}(\xi) \\
&= \hat{\sigma}^{(1)}_{gg\to HX} + 2A\int \mathrm{d}\xi\, \delta(\xi s - m_H^2)\frac{\alpha_{\text{s}}}{2\pi}\int_\xi^1 \mathrm{d}v\, \underbrace{f^{(0)}_{g/g}(v)}_{\delta(1-v)}\hat{P}_{gg}(\xi/v)\underbrace{\int_0^{\mu^2}\frac{\mathrm{d}k_\perp^2}{k_\perp^2}}_{-1/\varepsilon} \\
&= \hat{\sigma}^{(1)}_{gg\to HX} - \frac{2A}{s}\frac{\alpha_{\text{s}}}{2\pi\varepsilon}\hat{P}_{gg}(m_H^2/s).
\end{aligned}$$

If the factorization formula is correct, then the term starting with $2A/s$ must equal the divergent part of the partonic process $g + g \to H + X$, since otherwise $\hat{\sigma}^{(1)}_{g+g\to H+X}$ would not be finite. Similar holds for $q + g \to H + X$. In this case $f_{i/q} f_{j/g}$ gets contributions from $i = j = g$ since $g + g \to H + X$ is non-vanishing in zeroth order, so the divergence is

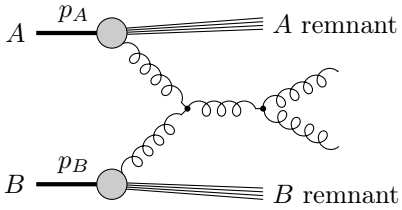
$$-\frac{A}{s}\frac{\alpha_s}{2\pi\varepsilon}\hat{P}_{gq}(m_H^2/s).$$

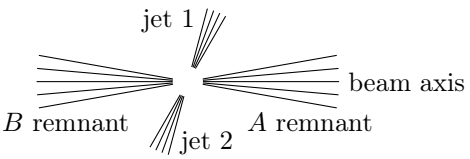
Exactly the same holds for antiquark with replacement $q \to \overline{q}$. The divergence corresponds to the following diagram.



Jet production in hadron collisions

Two jets might be produced in the following example.





The sketch on the right-hand side show the two jet final state in the frame of the detector. For this process to be IR-safe the two jets must have large transverse momentum relative to the beam axis in order to be separated from the beam remnants. Jet algorithms for hadron collisions must cluster two beam jets plus any number of further jets.

Chapter 6

Anomalies

A symmetry has an *anomaly* if the classical action is invariant under the symmetry, $\delta S = 0$, but the generating l or the effective action is not, $\delta\Gamma, \delta Z \neq 0$. That is, the symmetry is in fact not present in the quantum theory. The most important anomalies are:

- *conformal anomaly*: a classically scale-invariant theory (no dimensional parameter in $\mathcal{L}$, for example massless ϕ^4 theory or gauge theories with massless fermions) is not scale-invariant after quantization. The conformal anomaly is the rule rather than exception, since renormalization always introduces some kind of scale and all parameters in the renormalized theory are scale-dependent;
- *chiral anomaly*: in this chapter we focus on this anomaly. It is of great importance for
  - chiral symmetries in QCD, the so-called *strong CP problem* and the interaction of Goldstone bosons with gauge fields,
  - the consistency of chiral gauge theories, i.e. gauge theories where the gauge fields couple differently to left- and right-handed fermions (as in the SM),
  - the generation of matter-antimatter asymmetry in the early universe.

6.1 $\pi^0 \to \gamma\gamma$

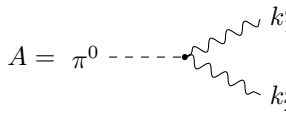
The interpretation and explanation of the decay rate of this process has been instrumental for the discovery and understanding of anomalies. As in Chapter 3 we consider the Goldstone boson Lagrangian for the spontaneously broken chiral symmetry of QCD coupled to external fields. The simplest term that gives $\pi^0 \to \gamma\gamma$ is

$$\mathcal{L} \supset g\pi^0 \varepsilon^{\mu\nu\rho\sigma} F_{\mu\nu} F_{\rho\sigma}$$

which is gauge invariant (π^0 is neutral and does not transform) and parity-conserving (QED and QCD conserve parity, π^0 has odd parity, hence the need for ε tensor). Any derivative terms and quark mass terms are suppressed.

Computation of the decay rate

Denoting the photon momenta as k_1^α and k_2^β, the term in the Lagrangian $4g\pi^2\varepsilon^{\mu\nu\rho\sigma}\partial_\mu A_\nu \partial_\rho A_\sigma$ gives the amplitude

$A =$  $= 8g\varepsilon^{\mu\alpha\rho\beta} i k_{1\mu} i k_{2\rho}.$

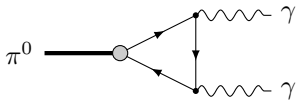
Summing over photon polarizations yields

$$\overline{|A|^2} = 64g^2 k_{1\mu} k_{2\rho} k_{1\mu'} k_{2\rho'} \varepsilon^{\mu\alpha\rho\beta}\varepsilon^{\mu'\alpha'\rho'\beta'}(-g_{\alpha\alpha'})(-g_{\beta\beta'}) = 128g^2(k_1\cdot k_2)^2 = 32g^2 m_\pi^4.$$

For the decay rate, we use the known expression for the phase space and divide by two for two identical particles in the final state

$$\Gamma(\pi^0\to\gamma\gamma) = \frac{1}{2m_\pi}\frac{1}{2}\frac{1}{8\pi}32g^2m_\pi^4 = \frac{g^2m_\pi^3}{\pi}.$$

The coupling g has mass dimension -1 and contains two electromagnetic couplings. At quark level we can estimate this as

 $g \approx \dfrac{e^2}{16\pi^2 f_\pi^2} = \dfrac{\alpha_{\rm em}}{4\pi f_\pi}.$

The factor $16\pi^2$ comes from a loop and $f_\pi = 92.4\,\text{MeV}$ is a charged pion decay constant. The estimate above gives $\Gamma(\pi^0\to\gamma\gamma)\approx 4.4\cdot 10^{16}\,\text{s}^{-1}$ to be compared with the measurement $1.2\cdot 10^{16}\,\text{s}^{-1}$. In view of the approximations, this is a reasonable agreement. The problem with this result is that the assumed interaction $\pi^0 F_{\mu\nu}\tilde{F}^{\mu\nu}$ cannot appear in the chiral Lagrangian, since all interactions with π^0 must involve a derivative of π^0 or be proportional to a light quark mass m_u or m_d. This suppresses the amplitude by $(4\pi f_\pi/m_\pi)^2\approx 10^2$ and the decay rate by a factor 10^4 which is then in conflict with observations. This is known as the *Sutherland–Veltman paradox*. Apparently, the understanding of chiral symmetry breaking and the couplings of pions to photons is not yet complete.

6.2 Anomalous symmetries in the path integral formalism

Consider a generating functional

$$Z[J_n] = |N|^2\int\mathcal{D}[\phi_n]\,\exp\!\left(iS[\phi_n] + i\int \mathrm{d}^4x\;\, J_n(x)\phi_n(x)\right).$$

In (Ch. ??) we considered infinitesimal transformations of the integration variables $\phi_n(x)\to\phi_n(x)+\varepsilon\theta^a(x)F_n^a[\phi_{n'}(x)]$. Since an integral does not change under a change of variables, one obtains

$$\int\mathcal{D}[\phi_n]\,\exp\!\left(iS[\phi_n] + i\int \mathrm{d}^4x\;\, J_n(x)\phi_n(x)\right)\cdot\int\mathrm{d}^4x\;\,\theta^a(x)\left[\sum_n \frac{\delta F_n^a[\phi_m](x)}{\delta\phi_n(x)} + i\left(\frac{\delta S}{\delta\phi_n(x)} + J_n(x)\right)F_n^a[\phi_m(x)]\right] = 0$$

The term with functional derivative of F_n^a arises from the transformation of the path integral measure

$$\left|\det \frac{\delta\phi_n'(y)}{\delta\phi_n(x)}\right| = \exp \operatorname{tr} \ln\left(1+\varepsilon\frac{\delta(\theta^a(x)F_{n'}^a[\phi_m](y))}{\delta\phi_n(x)}\right)$$
$$\approx \exp\left(\varepsilon \operatorname{tr}\theta^a(y)\frac{\delta F_{n'}^a[\phi_m(y)]}{\delta\phi_n(x)}\right) \approx 1+\varepsilon\int \mathrm{d}^4x\ \theta^a(x)\sum_n \frac{\delta F_n^a[\phi_m(x)]}{\delta\phi_n(x)}.$$

Note that the trace is in x, y and n, n'. If the change of variables is also a symmetry of the classical action for constant θ^a, then we can replace

$$\frac{\delta S}{\delta\phi_n(x)}F_n^a[\phi_m(x)] \to \partial_\mu^a(x)$$

where $j_\mu^a(x)$ is the Noether current of this symmetry (p.??). Hence, we obtain

$$\int \mathcal{D}[\phi_n]\ \exp\left(iS[\phi_n]+i\int \mathrm{d}^4x\ J_n(x)\phi_n(x)\right)\cdot\int \mathrm{d}^4x\ \theta^a(x)\Bigg[\\ \partial^\mu j_\mu^a[\phi_m(x)] - J_n(x)F_n^a[\phi_m(x)] + i\sum_n \frac{\delta F_n^a[\phi_m](x)}{\delta\phi_n(x)}\Bigg] = 0 \tag{6.1}$$

Differentiating with respect to $J(z_i)$ gives the Ward identities of the symmetry. The symmetry is anomalous if the last term, $\operatorname{tr}(\delta F_{n'}^a(x)/\delta\phi_n(y))$, does not vanish, that is, if the path integral measure is *not* invariant.

6.2.1 Computation of the chiral symmetry

The following method is due to *Fujikawa* (1979). Consider QED with a massless electron (a Dirac spinor)

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \overline{\Psi}i\not{D}\Psi.$$

The theory is invariant under $\Psi(x) \to e^{i\theta\gamma_5}\Psi(x)$ and $\overline{\Psi}(x) \to \overline{\Psi}(x)e^{i\theta\gamma_5}$. The Noether current for this symmetry is $j_\mu = -\overline{\Psi}\gamma_\mu\gamma_5\Psi$. The negative sign is important and comes from $\gamma_5\gamma_0 = -\gamma_0\gamma_5$. The first attempt to calculate the variation of the integral measure gives

$$\operatorname{tr}\frac{\delta F_{n'}^a[\phi_m(x)]}{\delta\phi_n(y)} \to (-1)\left(\operatorname{tr}\frac{\delta F_\Psi[\Psi,\overline{\Psi}](x)}{\delta\Psi(x)} + \operatorname{tr}\frac{\delta_{\overline{\Psi}}[\Psi,\overline{\Psi}](x)}{\delta\overline{\Psi}(y)}\right) = -2i\operatorname{tr}(\gamma_5)\delta^{(4)}(x-x)\int \mathrm{d}^4x.$$

- The traces run in x, y and the Dirac index.
- The negative sign comes because for Grassmann variables one gets the inverse of the Jacobi determinant (p.??), so in 6.1 replace i with $-i$.
- The factor 2 comes because Ψ and $\overline{\Psi}$ give same contributions.
- The variation of the field is $F_\Psi[\Psi,\overline{\Psi}] = i\gamma_5\Psi$ and $F_{\overline{\Psi}}[\Psi,\overline{\Psi}] = i\gamma_5\overline{\Psi}$.
- The trace of γ_5 is zero.
- The delta function $\delta^{(4)}(x-x)$ is a short-distance singularity (and is not the space-time volume as for $\delta^{(4)}(0)$ in momentum space).

- For non-chiral phase rotations $e^{i\theta}\Psi$, Ψ and $\overline{\Psi}$ contribute with opposite signs and the variation is manifestly zero.

The above expression is ill-defined. This is not surprising, since we have not been very precise about the definition of the path integral measure up to now. We now define the measure for the fermion field by expanding $\Psi, \overline{\Psi}$ in a complete set of eigenfunctions of the Dirac operator $i\not{\mathrm{D}}$ (for given gauge field configuration $A_\mu(x)$). The existence of such a set is guaranteed for a hermitian operator. This is the case for the Dirac operator in Euclidean space only. In the following, we will omit this step of rotating to Euclidean space-time ($t = i\tau$) and back. Assume $\phi_n, \overline{\phi}_n$ are left and right eigenfunctions with

$$i\not{\mathrm{D}}\phi_n = \lambda_n\phi_n, \quad \overline{\phi}_n i\overleftarrow{\not{\mathrm{D}}} = \lambda_n\overline{\phi}_n.$$

Then any field configuration $\Psi(x), \overline{\Psi}(x)$ can be written as

$$\Psi(x) = \sum_n a_n\phi_n(x), \quad \overline{\Psi}(x) = \sum_n \overline{a}_n\overline{\phi}_n(x),$$

and we define the measure As

$$\mathcal{D}[\Psi, \overline{\Psi}] = \prod_n \mathrm{d}a_n\, \mathrm{d}\overline{a}_n.$$

The change of variables amounts to

$$\Psi'(x) = \sum_n a'_n\phi_n(x) = (1 + i\theta(x)\gamma_5)\sum_n a_n\phi_n, \quad a'_m = a_m + i\sum_n \int \mathrm{d}^4x\ \theta(x)\overline{\phi}_m(x)\gamma_5\phi_n(x)a_n.$$

Similar holds for $\overline{a}_m$. The determinant calculation gives

$$\begin{aligned}\left|\det\frac{\delta(\Psi', \overline{\Psi}')}{\delta(\Psi, \overline{\Psi})}\right|^{-1} &= \exp\left(-\operatorname{tr}\ln\frac{\partial(a_{n'}, \overline{a}_{n'})}{\partial(a_n, \overline{a}_n)}\right) = \exp\left(-2i\sum_n\int \mathrm{d}^4x\ \theta(x)\overline{\phi}_n(x)\gamma_5\phi_n(x)\right)\\ &= \lim_{\Lambda\to\infty}\exp\left(-2i\sum_n\int \mathrm{d}^4x\ \theta(x)\overline{\phi}_n(x)\gamma_5\phi_n(x)e^{\lambda_n^2/\Lambda^2}\right).\end{aligned}$$

To regulate the sum over n (which corresponds to the trace in x, y which gave the ill-defined $\delta^{(4)}(x-x)$) we introduced the gauge-invariant regulator $e^{\lambda_n^2/\Lambda^2}$ (gauge-invariance follows from gauge-invariance of eigenvalues λ_n) called the *heat-kernel regulator*. Note that after the rotation to Euclidean space $\lambda_n^2 < 0$ (this will be seen below). Then

$$\begin{aligned}\lim_{\Lambda\to\infty}\sum_n \overline{\phi}_n(x)\gamma_5\phi_n(x)e^{\lambda_n^2/\Lambda^2} &= \lim_{\lambda\to\infty}\sum_n \overline{\phi}_n\gamma_5 e^{(i\not{\mathrm{D}})^2/\Lambda^2}\phi_n(x)\\ &= \lim_{\Lambda\to\infty}\int\frac{\mathrm{d}^4k}{(2\pi)^4}e^{ikx}\operatorname{tr}(\gamma_5 e^{(i\not{\mathrm{D}})^2/\Lambda^2})e^{-ikx}.\end{aligned} \tag{6.2}$$

In the last equality the trace was evaluated in the momentum basis rather than in the basis of eigenfunctions of $i\not{\mathrm{D}}$. Thus, the sum was replaced by an integral, but the trace over Dirac indices remained. It holds

$$(i\not{\mathrm{D}})^2 = \gamma_\mu\gamma_\nu i\mathrm{D}_\mu i\mathrm{D}_\nu = -\mathrm{D}^2 + \frac{1}{i}\sigma_{\mu\nu}i\mathrm{D}_\mu i\mathrm{D}_\nu = -\mathrm{D}^2 + \frac{1}{2i}\sigma_{\mu\nu}[i\mathrm{D}_\mu, i\mathrm{D}_\nu] = -\mathrm{D}^2 + \frac{e}{2}\sigma_{\mu\nu}F^{\mu\nu}.$$

Remember $\sigma_{\mu\nu} = i[\gamma_\mu, \gamma_\nu]/2$. Now we expand $\operatorname{tr}(\gamma_5 e^{(i\not{\partial})^2/\Lambda^2})$ in powers of the gauge field $A_\mu(x)$. To get a non-vanishing trace, we need at least two factors of $\sigma_{\mu\nu}F^{\mu\nu}$. This brings Λ^{-4} factor. Any further non-vanishing term will be of order Λ^{-6} or smaller, so the exponent becomes

$$\begin{aligned}
6.2 &= \lim_{\Lambda\to\infty} \int \frac{\mathrm{d}^4k}{(2\pi)^4}\, e^{ikx} \operatorname{tr}\left(\gamma_5 \frac{1}{2!}\left(\frac{e}{2}\sigma_{\mu\nu}\frac{F^{\mu\nu}}{\Lambda^2}\right)^2\right) e^{-\partial^2/\Lambda^2} e^{-ikx} + \mathcal{O}(\Lambda^{-6}) \\
&= \lim_{\Lambda\to\infty} \int \frac{\mathrm{d}^4k}{(2\pi)^4}\, e^{k^2/\Lambda^2} \frac{e^2}{8\Lambda^4} F^{\mu\nu}F^{\mu'\nu'} \operatorname{tr}(\gamma_5\sigma_{\mu\nu}\sigma_{\mu'\nu'}) \\
&= \lim_{\Lambda\to\infty} \frac{i}{16\pi^2}\Lambda^4 \frac{e^2}{8\Lambda^4} i^2 \operatorname{tr}(\gamma_5\gamma_\mu\gamma_\nu\gamma_{\mu'}\gamma_{\nu'}) F^{\mu\nu}F^{\mu'\nu'} \\
&= -\frac{e^2}{32\pi^2}\varepsilon^{\mu\nu\mu'\nu'} F_{\mu\nu}F_{\mu'\nu'}.
\end{aligned}$$

Hence, the path integral measure is not invariant

$$\left|\det \frac{\delta(\Psi', \overline{\Psi}')}{\delta(\Psi, \overline{\Psi})}\right|^{-1} = \exp\left(i\int \mathrm{d}^4x\, \theta(x) \frac{e^2}{16\pi^2}\varepsilon^{\mu\nu\rho\sigma} F_{\mu\nu}(x)F_{\rho\sigma}(x)\right)$$

and the symmetry is anomalous. Thus, in 6.1 we do the following replacement and reinterpretation of the anomalous term

$$i\sum_n \frac{\delta F_n^a[\phi_m(x)]}{\delta\phi_n(x)} \to -\frac{e^2}{16\pi^2}\varepsilon^{\mu\nu\rho\sigma} F_{\mu\nu}F_{\rho\sigma}, \quad \partial^\mu j_\mu = \frac{e^2}{16\pi^2}\varepsilon^{\mu\nu\rho\sigma} F_{\mu\nu}F_{\rho\sigma}$$

instead of $\partial^\mu j_\mu = 0$. Since the symmetry is not a symmetry of the quantum theory, the Noether current and the corresponding charge are not conserved. The path integral changes as follows under a chiral transformation of the fermion field variables

$$\int \mathcal{D}[\Psi', \overline{\Psi}']\, e^{iS[\Psi', \overline{\Psi}']} = \int \mathcal{D}[\Psi, \overline{\Psi}] \exp\left(iS[\Psi, \overline{\Psi}] + i\int \mathrm{d}^4x\, \theta(x)\left(-\partial^\mu j_\mu(x) + \frac{e^2}{16\pi^2}F_{\mu\nu}\tilde{F}^{\mu\nu}\right)\right)$$

i.e. the effect is equivalent to a change in the Lagrangian by

$$\delta_\theta \mathcal{L} = \theta(x)\frac{e^2}{16\pi^2}\varepsilon^{\mu\nu\rho\sigma} F_{\mu\nu}(x)F_{\rho\sigma}(x) \tag{6.3}$$

6.2.2 Chiral anomaly and $\pi^0 \to \gamma\gamma$

Consider the strong interaction with up and down quarks and the $\mathrm{SU(2)_A}$ axial symmetry operating on $(u, t)^\intercal$. For the π^0 we are interested in the transformation with the third generator T^3 (see the Goldstone matrix in Chapter 3 and below). The quarks couple to the electromagnetic field through $\overline{\Psi} i\not{D}\Psi$ with

$$i\not{D} = i\not{\partial} + g_\mathrm{s}\not{A}_\mathrm{gluon} + Qe\not{A}_\mathrm{photon}, \quad Q = \begin{pmatrix} 2/3 \\ -1/3 \end{pmatrix}.$$

Hence, the trace that appeared in the previous section now also needs to sum over the colour and flavour

$$\operatorname{tr}(\gamma_5 T^3 e^{(i\not{D})^2/\Lambda^2}) \xrightarrow{\text{extra factor}} \operatorname{tr}_\mathrm{flavour}(T^3Q^2)\operatorname{tr}_\mathrm{colour}\mathbb{1} = \operatorname{tr}\begin{pmatrix} 1/2 & 0 \\ 0 & -1/2 \end{pmatrix}\begin{pmatrix} 4/9 & 0 \\ 0 & 1/9 \end{pmatrix} N_\mathrm{c} = \frac{N_\mathrm{c}}{6}$$

Hence, eq. 6.3 becomes

$$\delta_\theta \mathcal{L}_{\text{QCD}} = \theta(x) \frac{N_\text{c}}{6} \frac{e^2}{16\pi^2} \varepsilon^{\mu\nu\rho\sigma} F_{\mu\nu}(x) F_{\rho\sigma}(x) \tag{6.4}$$

The effective charge of the QCD Lagrangian in the presence of an electromagnetic field must be reproduced by the Goldstone effective Lagrangian. This has no chiral fermions, the measure $\mathcal{D}[\mathcal{U}]$ is invariant under a chiral transformation. Hence, $\mathcal{L}^{(2)}$ itself must transform as eq. 6.4. To find the required additional term to $\mathcal{L}^{(2)}$, recall

$$\mathcal{U}(\pi) \xrightarrow{\varepsilon_\text{R} = -\varepsilon_\text{L} = \varepsilon_\text{A}} = e^{-i\theta T^3} \mathcal{U}(\pi) e^{-i\theta T^3}, \quad \mathcal{U}(\pi) = \exp \frac{i\sqrt{2}}{f_\pi} \begin{pmatrix} \pi^0/\sqrt{2} & \pi^+ \\ \pi^- & -\pi^0/\sqrt{2} \end{pmatrix}$$

which implies $\pi^0 \to \pi^0 - \theta f_\pi$ under the T^3 SU(2)_A transformation. Hence, we obtain the right-hand side of eq. 6.4 for $\delta_\theta \mathcal{L}^{(2)}$ if we add

$$\mathcal{L}^{(2)} \supset -\frac{N_\text{c}}{6} \frac{e^2}{16\pi^2 f_\pi} \pi^0 \varepsilon^{\mu\nu\rho\sigma} F_{\mu\nu} F_{\rho\sigma}$$

This is exactly the term that was used in the beginning, if one identifies

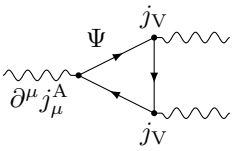
$$g = -\frac{N_\text{c}}{6} \frac{\alpha_\text{em}}{4\pi f_\pi}, \quad \Gamma_{\pi^0 \to \gamma\gamma} = \left(\frac{N_\text{c}}{6} \right)^2 \frac{\alpha_\text{em}^2}{64\pi^3} \frac{m_\pi^3}{f_\pi^2}$$

Since $N_\text{c} = 3$, g is by factor 2 smaller than the estimate on p. 137. The decay rate is now in very good agreement with experiment. Note that $\Gamma \propto N_\text{c}^2$ knows about the internal colour degree of freedom. With m_π and f_π we can now determine N_c.

6.3 Diagrammatic derivation of the anomaly

Since the divergence $\partial^\mu j_\mu$ of an anomalous Noether current is non-vanishing, the matrix element $\langle \gamma\gamma | \partial^\mu j_\mu | \Omega \rangle \neq 0$. This matrix element is related to the three-point function

$$\langle \Omega | j_\text{V}^\alpha \, j_\text{V}^\beta \, \partial_\mu j_\text{A}^\nu | \Omega \rangle =$$



$$+ \text{NLO} + \ldots$$

We denoted the vector current $\overline{\Psi} \gamma^\alpha \Psi$ with j_V^α. Here, we will investigate a more general case of currents

$$j_\mu^A(x) = \overline{\Psi}(x) \gamma_\mu \frac{1 - \gamma_5}{2} T^A \Psi(x), \quad \delta\Psi = i\varepsilon^A T^A \frac{1 - \gamma_5}{2} \Psi$$

where Ψ is a vector of left-handed fermions (due to $P_{\mathrm{L}} = (1-\gamma)/2$) and the symmetry transformation can be global or local. The three-point function of this current is

$$\Gamma^{ABC}_{\mu\nu\rho}(x,y,z) \equiv \langle\Omega|T(j^A_\mu(x)j^B_\nu(y)j^C_\rho(z))|\Omega\rangle$$

$$= \quad + \quad + \text{more loops}_{\text{(neglected)}}$$

$$= \underbrace{\int \frac{\mathrm{d}^4k_1}{(2\pi)^4}\frac{\mathrm{d}^4k_2}{(2\pi)^4} e^{-ik_1x} e^{e^{-ik_2z}e^{i(k_1+k_2)x}}}_{\text{denote by FT}} \int \frac{\mathrm{d}^4\ell}{(2\pi)^4} i^3(-1)\Big\{$$

$$\mathrm{tr}\left(\gamma_\mu P_{\mathrm{L}} T^A \frac{1}{\not p - \not k_1 - \not a}\gamma_\nu P_{\mathrm{L}} T^B \frac{1}{\not l + \not a}\gamma_\rho P_{\mathrm{L}} T^C \frac{1}{\not l + \not k_2 + \not a}\right) +$$

$$+ \mathrm{tr}\left(\gamma_\mu P_{\mathrm{L}} T^A \frac{1}{\not p - \not k_1 - \not b}\gamma_\rho P_{\mathrm{L}} T^C \frac{1}{\not l + \not b}\gamma_\nu P_{\mathrm{L}} T^B \frac{1}{\not l + \not k_2 + \not b}\right)\Big\}.$$

- The factors P_{L} can be merged into a single P_{L} since γ_5 anticommutes with γ_μ and $P_{\mathrm{L}}^2 = P_{\mathrm{L}}$. We assumed the fermions are massless.
- We introduced two arbitrary vectors a^μ and b^μ. This allows us to parametrize different momentum routings. For instance, we could choose $a = 0$ or $a = k_2 - k_1$ or another choice. Naively, the integral does not depend on a and b at all, because we could change the loop momentum $l \to l - a$ and $l \to l - b$. However, one must be careful because the loop integral is *superficially linearly divergent* and needs to be regularized. If, for instance, one introduces a cut-off $|l^\mu| < \Lambda$, the integral is not invariant under a momentum shift.

Take the derivative with respect to x^μ on the three-point function. The Ward identity without the possible anomalous term is (p.??)

$$\partial^\mu_{(x)} \langle\Omega|T(j^A_\mu(x)j^B_\nu(y)j^C_\rho(z))|\Omega\rangle = i\delta^{(4)}(x-y)\,\langle\Omega|T(\delta^A j^B_\nu(y)j^C_\rho(z))|\Omega\rangle + \\ + i\delta^{(4)}(x-z)\,\langle\Omega|T(j^B_\nu(y)\delta^A j^C_\rho(z))|\Omega\rangle \tag{6.5}$$

where $\delta^A j^B_\nu = i\overline{\Psi}[T^A, T^B]\gamma_\nu P_{\mathrm{L}}\Psi$ is the variation of j^B_ν under the symmetry transformation δ^A with parameter ε^A. Note the prefactor $+i$ instead of $-i$ on (p.??) because the Noether current is $-j^A_\mu(x)$. We will now show that this equation receives an extra, anomalous contribution. From the explicit expression for $\Gamma^{ABC}_{\mu\nu\rho}(x,y,z)$ we see that $\partial^\mu_{(x)}$ gives a factor of $i(k_1+k_2)^\mu$. Then use

$$\not k_1 + \not k_2 = (\not l + \not k_2 + \not a) - (\not l - \not k_1 - \not b) = (\not l + \not k_1 + \not b) - (\not l - \not k_2 - \not a)$$

to cancel the propagators. This produces

$$\partial^\mu_{(x)}\Gamma^{ABC}_{\mu\nu\rho}(x,y,z) = (-i)i^3\,\mathrm{FT}\int\frac{\mathrm{d}^4l}{(2\pi)^4}\cdot\Big\{$$

$$\mathrm{tr}(T^AT^BT^C)\left(\mathrm{tr}\,\frac{1}{\not l - \not k_1 + \not a}\gamma_\nu\frac{1}{\not l + \not a}\gamma_\rho P_{\mathrm{L}} - \mathrm{tr}\,\frac{1}{\not l + \not a}\gamma_\rho\frac{1}{\not l + \not k_2 + \not a}\gamma_\nu P_{\mathrm{L}}\right) + \tag{6.6}$$

$$+\mathrm{tr}(T^AT^CT^B)\left(\mathrm{tr}\,\frac{1}{\not l - \not k_2 + \not b}\gamma_\rho\frac{1}{\not l + \not b}\gamma_\nu P_{\mathrm{L}} - \mathrm{tr}\,\frac{1}{\not l + \not b}\gamma_\nu\frac{1}{\not l + \not k_1 + \not b}\gamma_\rho P_{\mathrm{L}}\right)\Big\}$$

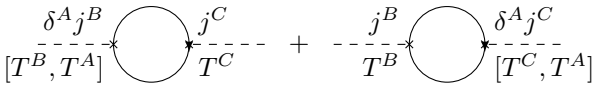
with

$$\operatorname{tr}(T^A T^B T^C) = \frac{1}{2}\operatorname{tr}(\{T^A, T^B\}T^C) + \frac{1}{2}\operatorname{tr}([T^A, T^B]T^C) = d^{ABC} + \frac{i}{4} f^{ABC}.$$

The anomaly coefficients d^{ABC} are totally symmetric, which is already evident for AB by definition and for the other combination we can write

$$d^{ACB} = \frac{1}{2}\operatorname{tr}(T^A T^C T^B + T^C T^A T^B) = \frac{1}{2}\operatorname{tr}(T^B T^A T^C + T^A T^B T^C) = \frac{1}{2}\operatorname{tr}(\{T^A, T^B\}T^C) = d^{ABC}.$$

More generally, in some representation R it holds $\operatorname{tr}(T_R^A\{T_R^B, T_R^C\}) = A(R)d^{abc}/2$ where $A(R)$ is called *anomaly coefficient* and equals 1 for fundamental representation. The right-hand side of the (non-anomalous) Ward identity 6.1 contains



One can verify that this corresponds exactly to the f^{ABC} terms of the generator traces in 6.6 above. Hence, the d^{ABC} part is (if non-zero) the anomalous contribution

$$\begin{aligned}\partial^\mu_{(x)}\Gamma^{ABC}_{\mu\nu\rho}(x,y,z)|_{\text{anomalous}} = (-1)d^{ABC}\,\mathrm{FT}\int\frac{\mathrm{d}^4 l}{(2\pi^4)}\Bigg\{ \\ \operatorname{tr}\frac{1}{\not l - \not k_1 + \not a}\gamma_\nu\frac{1}{\not l + \not a}\gamma_\rho P_\mathrm{L} - \operatorname{tr}\frac{1}{\not l + \not a}\gamma_\rho\frac{1}{\not l + \not k_2 + \not a}\gamma_\nu P_\mathrm{L} + \\ + \operatorname{tr}\frac{1}{\not l - \not k_2 + \not b}\gamma_\rho\frac{1}{\not l + \not b}\gamma_\nu P_\mathrm{L} - \operatorname{tr}\frac{1}{\not l + \not b}\gamma_\nu\frac{1}{\not l + \not k_1 + \not b}\gamma_\rho P_\mathrm{L}\Bigg\}\end{aligned}$$

As said above it would be wrong to conclude that this integral vanishes, because we can shift momenta $l \to l + k_1 - a + b$ in the first and $l \to l + k_2 - b + a$ in the third term so that the four terms cancel pairwise. After regularization, such shifts may not be allowed. The special feature of integrals that would be equal after momentum shifts is, however, that the difference is independent of the regularization. Therefore, we do not perform momentum shifts, but write

$$\int\frac{\mathrm{d}^4 l}{(2\pi)^4}\{\dots\} = \operatorname{tr}(\gamma^\mu\gamma^\nu\gamma^\sigma\gamma^\rho P_\mathrm{L})I_{\mu\sigma}(a-b-k_1, b, b+k_1) + \operatorname{tr}(\gamma^\mu\gamma^\rho\gamma^\sigma\gamma^\nu P_\mathrm{L})I_{\mu\sigma}(b-a-k_2, a, a+k_2)$$

where

$$I_{\mu\nu}(u,v,w)0\int\frac{\mathrm{d}^4 l}{(2\pi)^4}\left(f_{\mu\sigma}(l+u,v,w) - f_{\mu\sigma}(b-a-k_2, a, a+k_2)\right) \tag{6.7}$$

and

$$f_{\mu\sigma}(l,v,w) = \frac{(l+v)_\mu (l+w)_\sigma}{(l+v)^2(l+w)^2}.$$

Each integral over this function is infinite, but the difference is finite. Since I would be 0 if the integrals separately were not divergent, we may assume that any non-vanishing result arises from loop momenta larger than k_1, k_2, a, b and therefore can Taylor expand the integral in the difference u:

$$I_{\mu\sigma}(u,v,w) = \sum_{n=1}^{\infty}\frac{1}{n!}u^{\lambda_1}u^{\lambda_2}\dots u^{\lambda_n}\int\frac{\mathrm{d}^4 l}{(2\pi)^4}\frac{\partial^n f_{\mu\nu}(l,v,w)}{\partial l^{\lambda_1}\partial l^{\lambda_2}\dots\partial l^{\lambda_n}}.$$

The $n = 0$ cancels between the two terms. The integral is a total divergence. If the loop momenta are restricted to a four-dimensional solid ball of radius R, the surface term is of order R^{2-n} (3

powers from the surface, -2 from f and $1-n$ for $n-1$ derivatives). Hence, as $R\to\infty$, only $n=1,2$ in the sum survive. We will treat them separately.

$n=1$:

$$\begin{aligned}\int\frac{\mathrm{d}^4l}{(2\pi)^4}\frac{\partial}{\partial l^\lambda}f_{\mu\sigma}(l,v,w) &= \frac{i}{(2\pi)^4}\lim_{R\to\infty}\int\mathrm{d}\Omega^{(3)}\,R^2R_\lambda f_{\mu\sigma}(R,v,w)\\ &= \frac{i}{(2\pi)^4}\lim_{R\to\infty}\int\mathrm{d}\Omega^{(3)}\,R^2R_\lambda\left(\frac{R_\mu R_\sigma+R_\mu w_\sigma+v_\mu R_\sigma}{R^4}+\frac{R_\mu R_\sigma}{R^6}(2v\cdot R-2w\cdot R)+\mathcal{O}(R^{-4})\right)\\ &= \frac{i}{8\pi^2}\left(\frac{1}{4}g_{\lambda\mu}w_\sigma+\frac{1}{4}g_{\lambda\sigma}v_\mu-\frac{1}{12}(g_{\lambda\mu}g_{\sigma\rho}+g_{\lambda\sigma}g_{\mu\rho}+g_{\lambda\rho}g_{\mu\sigma})(v+w)^2\right).\end{aligned}$$

The factor of i comes from the Wick rotation and the angular integral gives $2\pi^2$. Also remember that integrals with odd number of R^α vanish by rotational invariance.

$n=2$:

$$\begin{aligned}\int\frac{\mathrm{d}^4l}{(2\pi)^4}\frac{\partial^2}{\partial l^{\lambda_1}\partial l^{\lambda_2}}f_{\mu\sigma}(l,v,w) &= \\ &= \frac{i}{(2\pi)^4}\lim_{R\to\infty}\int\mathrm{d}\Omega^{(3)}\,R^2R_{\lambda_1}\left(\frac{g_{\mu\lambda_2}R_\sigma+g_{\sigma\lambda_2}R_\mu}{R^4}-\frac{4R_\mu R_\sigma R_{\lambda_2}}{R^6}+\mathcal{O}(R^{-4})\right)\\ &= \frac{i}{8\pi^2}\left(\frac{1}{4}g_{\mu\lambda_2}g_{\sigma\lambda_1}+\frac{1}{4}g_{\mu\lambda_1}g_{\sigma\lambda_2}-\frac{1}{6}g_{\mu\lambda_2}g_{\sigma\lambda_1}-\frac{1}{6}g_{\mu\lambda_1}g_{\sigma\lambda_2}-\frac{1}{6}g_{\mu\sigma}g_{\lambda_1\lambda_2}\right).\end{aligned}$$

The $n=1$ and $n=2$ contributions together give

$$I_{\mu\sigma}=\frac{i}{32\pi^2}\left(u_\mu w_\sigma+u_\sigma w_\mu-\frac{u_\mu(v+w)_\sigma+u_\sigma(v+w)_\mu+g_{\mu\sigma}u\cdot(v+w)}{3}+\frac{u_\mu u_\sigma-g_{\mu\sigma}u^2}{3}\right).$$

Together with

$$\mathrm{tr}(\gamma^\mu\gamma^\nu\gamma^\sigma\gamma^\rho P_\mathrm{L})=2(g^{\mu\nu}g^{\sigma\rho}-g^{\mu\sigma}g^{\nu\rho}+g^{\mu\rho}g^{\nu\sigma})+2i\varepsilon^{\mu\nu\sigma\rho}$$

one obtains

$$\begin{aligned}\int\frac{\mathrm{d}^4l}{(2\pi)^4}\{\dots\} =& \frac{i}{16\pi^2}\left(-\frac{(a+b)^\nu(k_1+k_2)^\rho+(a+b)^\rho(k_1+k_2)^\nu+(a+b)\cdot(k_1+k_2)g^{\rho\nu}}{3}+\right.\\ &\left.-i\varepsilon^{\alpha\beta\nu\sigma}(a-b)_\alpha(k_1+k_2)_\beta\right).\end{aligned}$$

Now we choose $b=-a$ so that the term that arises from the trace without γ_5 vanishes. The reason for this is that this would be the (only) term if we replaced j^A_μ by vector currents $\overline{\Psi}\gamma_\mu\Psi$. However, the vector current is the Noether current of the QED gauge theory. Hence, this current *must* be conserved (i.e. the Ward identity for the three-point function must not have an anomalous term), since without the coupling to a constant current, one cannot obtain a consistent theory of a massless vector field – the unphysical gauge field polarizations would not decouple from the scattering matrix. Now, we can finally write the derivative of the three-point function

$$\partial^\mu_{(x)}\Gamma^{ABC}_{\mu\nu\rho}(x,y,z)|_\mathrm{anomalous}=\mathrm{FT}\left(-\frac{\mathrm{d}^{ABC}}{8\pi^2}\varepsilon_{\alpha\beta\nu\rho}a^\alpha(k_1+k_2)^\beta\right). \tag{6.8}$$

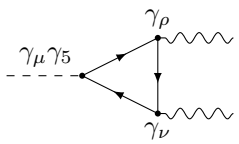
It appears that this can also be made to vanish by choosing $s=0$ or $a=k_1+k_2$. However, there are also two further equations for $\partial^\nu_{(y)}\Gamma$ and $\partial^\rho_{(z)}\Gamma$. The corresponding results are obtained from the right-hand side of 6.8 as follows

$$\partial^\nu_{(y)}:\ {}^{\nu\to\mu}_{k_1+k_2\to k_1},\quad {}^{a\to a+k_2}_{b\to b-k_2},\qquad \partial^\rho_{(z)}:\ {}^{\nu\to\mu}_{\rho\to\nu}{}^{}_{k_1+k_2\to k_2},\quad {}^{a\to a-k_1}_{b\to b+k_1}.$$

In both cases $a+b \to a+b$, that is, the choice $b=-a$ makes the vector current anomaly 0 for all three derivatives. For the axial vector current anomaly to vanish one would need $a \propto k_1+k_2$, $a+k_2 \propto k_1$ and $a-k_1 \propto k_2$ which cannot be satisfied simultaneously. Thus, if $d_{ABC} \neq 0$ at least one of the currents j^A_μ, j^B_ν, j^C_ρ has an anomaly.

6.3.1 The Abelian anomaly

This is the case we considered in Sec. 6.2 (also called *Adler–Bell–Jackiw anomaly*). Since $T^A = T^B = T^C = 1$, the anomaly coefficient becomes $d^{ABC} = 1$. Consider the following diagram.



Since the currents with indices ν, ρ are vector currents and the Noether currents of the gauge symmetry, the anomalous term must vanish for $\partial^\nu_{(y)}\Gamma$ and $\partial^\rho_{(z)}\Gamma$. Hence, we must choose $b=-a$ and $a=k_1-k_2$:

$$\begin{aligned}\partial^\mu_{(x)}\Gamma^{\text{anomal}}_{\mu\nu\rho}(x,y,z) &= \int \frac{\mathrm{d}^4k_1}{(2\pi)^4}\frac{\mathrm{d}^4k_2}{(2\pi)^4}\, e^{-ik_1y}e^{-ik_2z}e^{i(k_1+k_2)x}(+2)\frac{-1}{8\pi^2}\varepsilon_{\alpha\beta\nu\rho}(k_1-k_2)^\alpha(k_1+k_2)^\beta \\ &= \frac{1}{2\pi^2}\varepsilon_{\alpha\beta\nu\rho}\partial^\alpha_{(y)}\delta^{(x)}(y-x)\partial^\beta_{(z)}\delta^{(4)}(z-x).\end{aligned}$$

The factor 2 came from the fact that the Abelian axial Noether current is $-\gamma_\mu\gamma_5$ instead of $\gamma_\mu(1-\gamma_5)/2$. The term $(k_1-k_2)^\alpha(k_1+k_2)^\alpha$ simplifies to $2k_1^\alpha k_2^\alpha$ which gives $-2\partial/\partial y_\alpha\,\partial/\partial z_\beta$ before the integral. We claim that the result is equal to the previously found divergence of the axial Noether current $\partial^\mu j_\mu = e^2/(16\pi^2)\varepsilon^{\mu\nu\rho\sigma}F_{\mu\nu}F_{\rho\sigma}$. To this end we compute the matrix element

$$\langle\gamma(k_1,\lambda_1)\gamma(k_2,\lambda_2)|j_\mu(x)|\Omega\rangle = (ie)^2\int \mathrm{d}^4y\,\mathrm{d}^4z\, e^{ik_1y+ik_2z}\varepsilon_1^{*\nu}\varepsilon_2^{*\rho}\,\langle\Omega|T(j_\mu(x)j^{\mathrm{V}}_\nu(y)j^{\mathrm{V}}_\rho(z))|\Omega\rangle\,.$$

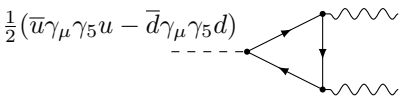
The three-point function is $\Gamma_{\mu\nu\rho}(x,y,z)$ from above. Hence,

$$\begin{aligned}\langle\gamma\gamma|\partial^\mu j_\mu(x)|\Omega\rangle &= (ie)^2\int \mathrm{d}^4x\,\mathrm{d}^4y\, e^{ik_1y+ik_2z}\varepsilon_1^{*\nu}\varepsilon_2^{*\rho}\frac{1}{2\pi^2}\varepsilon_{\alpha\beta\nu\rho}\partial^\alpha_{(y)}\delta^{(x)}(y-x)\partial^\beta_{(z)}\delta^{(4)}(z-x) \\ &= (ie)^2\frac{1}{2\pi^2}e^{i(k_1+k_2)x}(ik_1)^\alpha(ik_2)^\beta\varepsilon_{\alpha\beta\nu\rho}\varepsilon^{*\nu}(k_1,\lambda_1)\varepsilon^{*\rho}(k_2,\lambda_2) \\ &\overset{2\times\text{IBP}}{=} (ie)^2\frac{1}{2}\frac{1}{2\pi^2}\varepsilon_{\alpha\beta\nu\rho}\,\langle\gamma\gamma|\partial^\alpha A^\nu(x)\partial^\beta A^\rho(x)|\Omega\rangle \\ &= \frac{e^2}{16\pi^2}\varepsilon^{\alpha\beta\rho\sigma}\,\langle\gamma\gamma|F_{\mu\nu}(x)F_{\rho\sigma}(x)|\Omega\rangle\,.\end{aligned}$$

The "normal part" of the Ward identity 6.1 vanishes here, since the variation δj^{V} of the vector current under the axial transformation vanishes

$$j^{\mathrm{V}} \to \overline{\Psi}e^{i\theta\gamma_5}\gamma^\mu e^{i\theta\gamma_5}\Psi = j^{\mathrm{V}}.$$

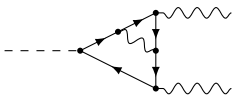
The above result demonstrates the identity, however, strictly speaking, one should show this for any matrix element, so that the divergence equation for $\partial^\mu j_\mu$ holds as an operator identity. For $\pi^0 \to \gamma\gamma$ the diagrammatic treatment corresponds to the following picture.



The axial vertex is the current for T^3 of $\mathrm{SU}(2)_\mathrm{A}$ which gives the factor $N_\mathrm{c}(Q_u^2 - Q_d^2)/2 = N_\mathrm{c}/6$ as already calculated earlier. The Abelian anomaly is a $\mathrm{U}(1)_\mathrm{A}$ anomaly. The above treatment raises two questions.

- Are there other contributions than from triangle diagrams?
- Are there contributions from multi-loop triangle diagrams?

The agreement with the path integral derivation suggests that this is not the case. One can also show this diagrammatically (this is called *Adler–Bardeen theorem* from year 1969). Heuristically, for a diagram such as



the fermion loop has five propagators and the fermion-loop subgraph is no longer divergent. Hence, a momentum shift is allowed and the anomalous term vanishes.

6.3.2 Chiral anomaly in dimensional regularization

In dimensional regularization, loop integrals are defined by analytic continuation to $d = 4 - 2\varepsilon$ dimensions. For the integral 6.7

$$I_{\mu\sigma}(u,v,w) = \int \frac{\mathrm{d}^d l}{(2\pi)^d}(f_{\mu\sigma}(l+u,v,w) - f_{\mu\sigma}(l,v,w)) = 0$$

the results is zero because the shift $l \to l' = l + u$ is allowed, since the integral is well-defined. Since there are no integration boundaries, the first term equal second. It thus appears there is no anomaly. But that would be wrong!

The subtlety arises from the relation $\{\gamma_5, \gamma_\mu\} = 0$, which was used to commute γ_5 to a position such that $I_{\mu\sigma}$ appears as above. The standard definition of γ_5 is

$$\gamma_5 = i\gamma^0\gamma^1\gamma^2\gamma^3 = -\frac{i}{4!}\varepsilon_{\mu\nu\rho\sigma}\gamma^\mu\gamma^\nu\gamma^\rho\gamma^\sigma.$$

In four dimensions, this implies

$$\gamma_5^2 = 1, \quad \{\gamma_5, \gamma^\mu\} = 0, \quad \mathrm{tr}\,\gamma_5 = 0, \quad \mathrm{tr}(\gamma_5\gamma_\mu\gamma_\nu) = 0, \quad \mathrm{tr}(\gamma_5\gamma_\mu\gamma_\nu\gamma_\rho\gamma_\sigma) = -4i\varepsilon_{\mu\nu\rho\sigma}.$$

In d dimensions, the basic relation $\{\gamma^\mu, \gamma^\nu\} = 2g^{\mu\nu}$ still continues to be valid, which implies $\gamma^\mu\gamma_\mu = d\mathbb{1}$, but the definition of γ_5 is intrinsically four-dimensional (explicit reference to $0, 1, 2, 3$ or to ε-tensor). If one insists on $\{\gamma_5, \gamma^\mu\} = 0$, it follows

$$(d-4)\,\mathrm{tr}(\gamma_5\gamma_\mu\gamma_\nu\gamma_\rho\gamma_\sigma) = 0$$

hence the trace must vanish whenever $d \neq 4$. An analytic continuation to the four-dimensional

result $-4i\varepsilon_{\mu\nu\rho\sigma}$ is not possible. To show this, write

$$\begin{aligned} d\,\mathrm{tr}(\gamma_5\gamma_\mu\gamma_\nu\gamma_\rho\gamma_\sigma) &= \mathrm{tr}(\gamma_5\gamma_\mu\gamma_\nu\gamma_\rho\gamma_\sigma\gamma_\lambda\gamma^\lambda) \\ &= -\,\mathrm{tr}(\gamma_5\gamma^\lambda\gamma_\mu\gamma_\nu\gamma_\rho\gamma_\sigma\gamma_\lambda) \\ &= -2\,\mathrm{tr}(\gamma_5\gamma_\mu\gamma_\rho\gamma_\sigma\gamma_\mu) + 2\,\mathrm{tr}(\gamma_5\gamma_\mu\gamma_\rho\gamma_\sigma\gamma_\nu) - 2\,\mathrm{tr}(\gamma_5\gamma_\mu\gamma_\nu\gamma_\sigma\gamma_\rho) + \\ &\qquad + 2\,\mathrm{tr}(\gamma_5\gamma_\mu\gamma_\nu\gamma_\rho\gamma_\sigma) - d\,\mathrm{tr}(\gamma_5\gamma_\mu\gamma_\nu\gamma_\rho\gamma_\sigma) \\ &= 2\,\mathrm{tr}(\gamma_5\{\gamma_\mu,\gamma_\nu\}\gamma_\rho\gamma_\sigma) + 4\,\mathrm{tr}(\gamma_5\gamma_\rho\gamma_\sigma)g_{\mu\nu} - 2\,\mathrm{tr}(\gamma_5\gamma_\mu\gamma_\nu\{\gamma_\sigma\gamma_\rho\}) + \\ &\qquad + 4\,\mathrm{tr}(\gamma_5\gamma_\mu\gamma_\nu)g_{\rho\sigma} - d\,\mathrm{tr}(\gamma_5\gamma_\mu\gamma_\nu\gamma_\rho\gamma_\sigma) \\ &= (8-d)\,\mathrm{tr}(\gamma_5\gamma_\mu\gamma_\nu\gamma_\rho\gamma_\sigma). \end{aligned}$$

Hence, $d = 4$ and no analytic continuation is allowed. One possible definition of γ_5 in d dimensions is the *'t Hooft–Veltman scheme*, which keeps $\gamma_5 \equiv i\gamma^0\gamma^1\gamma^2\gamma^3$ and $\{\gamma_5,\gamma^\mu\} = 0$ for $\mu = 0,1,2,3$ but uses $[\gamma_5,\gamma^\mu] = 0$ for all other μ (since $\{\gamma^\mu,\gamma^\nu\} = 0$ for $\nu = 0,1,2,3$ and $\mu \neq 0,1,2,3$). With this definition, we go back to the Abelian anomaly of the axial and two vector currents. The corresponding expression for $\Gamma_{\mu\nu\rho}(x,y,z)$ read

$$\Gamma_{\mu\nu\rho}(x,y,z) = \mathrm{FT}\,(-1)i^3\int\frac{\mathrm{d}^4l}{(2\pi)^d}\left\{\mathrm{tr}\left(\gamma^\mu\gamma_5\frac{1}{\not l-\not k_1}\gamma^\nu\frac{1}{\not l}\gamma^\rho\frac{1}{\not l-\not k_2}\right) + \right. \\ \left. + \mathrm{tr}\left(\gamma^\mu\gamma_5\frac{1}{\not l-\not k_2}\gamma^\rho\frac{1}{\not l}\gamma^\nu\frac{1}{\not l-\not k_1}\right)\right\}.$$

We don't need vectors a^μ and b^μ any more, so we set them to zero. The loop momentum l^μ is split into 4-dimensional component l_4^μ and the $(d-4)$-dimensional component $\hat l^\mu$ while the external momenta remain 4-dimensional ($\hat k_1^\mu = \hat k_2^\mu = 0$). The derivative in x gives additionally

$$(\not k_1+\not k_2)\gamma_5 = (\not l+\not k_2-(\not l-\not k_1))\gamma_5 = (\not l+\not k_2)\gamma_5 + \gamma_5(\not l-\not k_1) - 2\gamma_5\hat{\not l}.$$

The term $2\gamma_5\hat{\not l}$ is of order $\mathcal{O}(\varepsilon)$. Now, the anomaly diagram gives

$$\partial^\mu_{(x)}\Gamma_{\mu\nu\rho}(x,y,z) = \mathrm{FT}\,(-1)\int\frac{\mathrm{d}^4l}{(2\pi)^d}\left\{\underbrace{\mathrm{tr}\left(\gamma_5\frac{1}{\not l-\not k_1}\gamma_\nu\frac{1}{\not l}\gamma_\rho\right)}_{1}\right. \\ + \underbrace{\mathrm{tr}\left(\gamma_5\gamma_\nu\frac{1}{\not l}\gamma_\rho\frac{1}{\not l+\not k_2}\gamma_\rho\right)}_{2} + \underbrace{\mathrm{tr}\left(\gamma_5\frac{1}{\not l-\not k_2}\gamma_\rho\frac{1}{\not l}\gamma_\nu\right)}_{3} + \underbrace{\mathrm{tr}\left(\gamma_5\gamma_\rho\frac{1}{\not l}\gamma_\rho\frac{1}{\not l+\not k_1}\gamma_\nu\right)}_{4} + \\ \left. - \mathrm{tr}\left(2\gamma_5\hat{\not l}\frac{1}{\not l-\not k_1}\gamma_\nu\frac{1}{\not l}\gamma_\rho\frac{1}{\not l+\not k_2}\right) - \mathrm{tr}\left(2\gamma_5\hat{\not l}\frac{1}{\not l-\not k_2}\gamma_\rho\frac{1}{\not l}\gamma_\nu\frac{1}{\not l+\not k_1}\right)\right\}.$$

Perform the change of variables $l \to l + k_1$ in term 1 and use $\gamma_5\gamma_\nu = -\gamma_\nu\gamma_5$ in term 2 (allowed since ν and ρ can be chosen as 4-dimensional vectors). Similarly, shift $l \to l + k_2$ in term 3 and $\gamma_5\gamma_\rho = -\gamma_\rho\gamma_5$ in term 4. Then 2+3 and 1+4 terms cancel. The remaining two integrals in the last line can be computed with standard methods. The result is the same as before for the choice $b = -a$ and $a = k_1 - k_2$, i.e. dimensional regularization automatically leads to a vanishing of the anomaly for the vector current. In dimensional regularization the anomaly arises technically from $\mathcal{O}(\varepsilon)$ in the trace multiplied by $1/\varepsilon$ from the divergent loop integral, which gives a finite result that is a polynomial in external momenta.

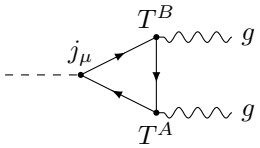
6.3.3 $\mathrm{U}(1)_\mathrm{A}$ and flavour anomaly in QCD

Assume an $\mathrm{SU}(N_\mathrm{c})$ gauge theory with N_f massless quarks. It has the global symmetry

$$\mathrm{SU}(N_\mathrm{f})_\mathrm{L} \times \mathrm{SU}(N_\mathrm{f})_\mathrm{R} \times \mathrm{U}(1)_\mathrm{B} \times \mathrm{U}(1)_\mathrm{A}$$

at the level of classical action.

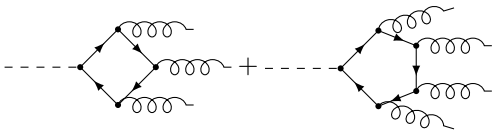
- $\mathrm{U}(1)_\mathrm{A}$: in Chapter. 3 we did not discuss the axial U(1) transformation as a symmetry, because it is anomalous. We prove this now. Noether current reads $j_\mu = \overline{q}\gamma_\mu\gamma_5 q$ which implicitly contains the sum over flavour f and colour a.



  The vertex factor comes from the $\overline{q}T^A\gamma_\mu q$ quark-gluon vertex. Compared to the Abelian triangle diagram we get an additional flavour and colour trace $\mathrm{tr}\,\mathbb{1}_{N_\mathrm{f}}\,\mathrm{tr}\,T^AT^B = N_\mathrm{f}\delta^{AB}/2$:

$$\partial^\mu j_\mu = -\frac{N_\mathrm{f} g_\mathrm{s}^2}{32\pi^2} 4\varepsilon^{\mu\nu\rho\sigma}\partial_\mu A^A_\nu \partial_\rho A^A_\sigma + \cdots = -\frac{N_\mathrm{f} g_\mathrm{s}^2}{32\pi^2}\varepsilon^{\mu\nu\rho\sigma} G^A_{\mu\nu} G^A_{\rho\sigma}.$$

  The sign is different from the Abelian anomaly because now the currents are $+\gamma_\mu\gamma_5$. The additional diagrams such as



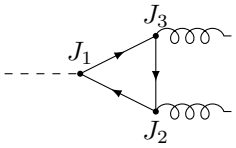
  must contribute due to gauge invariance so that one obtains the gluon field strength.

- $\mathrm{SU}(N_\mathrm{f})_\mathrm{A}$: if we denote t^a the generator of the $\mathrm{SU}(N_\mathrm{f})_\mathrm{A}$ flavour symmetry, the Noether current reads $j^a_\mu = \overline{q}t^a\gamma_\mu\gamma_5 q$. There is no anomaly in QCD, since $\mathrm{tr}\,t^a = 0$. The $\mathrm{SU}(N_\mathrm{f})_\mathrm{A}$ is non-anomalous, but spontaneously broken in QCD (and explicitly by quark masses – see Ch. 3). However, $\mathrm{SU}(N_\mathrm{f})_\mathrm{A}$ is anomalous in the presence of photons (external electromagnetic fields and also $W^\pm, Z^0$) because then the flavour trace is not zero. For instance $\mathrm{tr}(t^3Q^2) \neq 0$ which produces the anomalous $\pi^0 \to \gamma\gamma$ decay.

Note the following.

- The $\mathrm{U}(1)_\mathrm{A}$ anomaly involves gluons which are themselves strongly interacting. Hence, the anomaly has an $\mathcal{O}(1)$ effect on the hadronic spectrum. There is no trace of a Goldstone boson, because the symmetry is non-existent at the quantum level.
- The anomaly of the $\mathrm{SU}(3)_\mathrm{A}$ current is a weak (electromagnetic) effect. Hence, we can treat $\mathrm{SU}(3)_\mathrm{A}$ as a spontaneously broken symmetry of QCD and there are Goldstone bosons in the spectrum. The anomaly affects weak decays and can be implemented as a small perturbation term in the Goldstone boson effective Lagrangian (see remarks below).

Consider more generally which triangle diagrams can lead to anomalies:



where J_i are the currents from $\mathrm{U}(1)_{\mathrm{B,A}}$ and $\mathrm{SU}(3)_{\mathrm{V,A}}$. The anomalous contributions are

$$\mathrm{U}(1)_{\mathrm{A}}\,\mathrm{U}(1)_{\mathrm{B}}^2,\quad \mathrm{U}(1)_{\mathrm{A}}^3,\quad \mathrm{SU}(3)_{\mathrm{A}}\,\mathrm{SU}(3)_{\mathrm{V}}^2,\quad \mathrm{SU}(3)_{\mathrm{A}}^3,\quad \mathrm{U}(1)_{\mathrm{A}}\,\mathrm{SU}(3)_{\mathrm{V}}^2,\quad \mathrm{U}(1)_{\mathrm{A}}\,\mathrm{SU}(3)_{\mathrm{A}}^2.$$

For $\mathrm{SU}(3)_{\mathrm{A}}\,\mathrm{U}(1)_{\mathrm{B,A}}^2$ the flavour trace vanishes. Similarly, $\mathrm{U}(1)_{\mathrm{A}}^2\,\mathrm{U}(1)_{\mathrm{B}}$ doesn't contribute because two γ_5 matrices give $\mathbb{1}$ and no γ_5 is left (same happens for $\mathrm{SU}(3)_{\mathrm{A}}^2\,\mathrm{SU}(3)_{\mathrm{V}}$).

To compute $\mathrm{SU}(3)_{\mathrm{A}}$ anomalies, couple the $\mathrm{SU}(3)_{\mathrm{A}}$ Noether currents to an external vector and axial vector via

$$\mathcal{L}_{\mathrm{QCD}} \to \mathcal{L}_{\mathrm{QCD}} + j^{\mu}_{\mathrm{SU}(3)_{\mathrm{V}}} V_\mu + j^{\mu}_{\mathrm{SU}(3)_{\mathrm{A}}} A_\mu$$

and introduce covariant derivative such that $Z[J,V,A]$ is locally invariant under the $\mathrm{SU}(3)_{\mathrm{V}}$ and $\mathrm{SU}(3)_{\mathrm{A}}$ symmetries (see Ch. 3). Then compute the triangle graphs and obtain

$$\mathrm{D}^\mu j_\mu = \text{expression in terms of "gauge fields" } V, A.$$

This is called the *Bardeen anomaly*. One then has to modify the action of the effective chiral Lagrangian, such that its generating functional has the same anomalous transformation as the one of the QCD coupled to V and A. The additional terms are called *Wess–Zumino–Witten terms* and appear first at $\mathcal{O}(p^4)$. The WZW terms contain the $\pi^0 F\tilde{F}$ term responsible for fast $\pi^0 \to \gamma\gamma$ decay and other terms.

6.4 Consistency of chiral gauge theories

Consider a Lagrangian with a U(1) or non-Abelian gauge field and left-handed fermions transforming in a representation R of the gauge group with generators T_R^A. The projection to left-handed components is $\Psi_{\mathrm{L}} = P_{\mathrm{L}}\Psi$ for a Dirac spinor Ψ. The Lagrangian reads

$$\mathcal{L} = -\frac{1}{4} G^A_{\mu\nu} G^{A,\mu\nu} + \overline{\Psi}_{\mathrm{L}} i \not{\mathrm{D}} \Psi_{\mathrm{L}}.$$

The sum over species of fermions is implicit. In two component notation, the fermion term can also be written as $\Psi_{\mathrm{L}}^\dagger i\overline{\sigma}^\mu \mathrm{D}_\mu \Psi_{\mathrm{L}}$.

Remark. Without loss of generality one can assume all fermion fields are left-handed. If the theory contained a right-handed fermion Ψ_{R} transforming with T_R^A, one can instead use the charge-conjugated field $\Psi_{\mathrm{L}} = i\sigma^2 \Psi_{\mathrm{R}}^*$ which transforms as a left-handed spinor (p.??) and in the complex conjugate representation of the gauge group with generator $T_{\overline{R}}^A \equiv -T_R^{A*} = -(T_R^A)^\intercal$. To explicitly check this, write

$$\begin{aligned}
\Psi_{\mathrm{R}}^\dagger i\sigma^\mu(\partial_\mu - igA_\mu^A T_R^A)\Psi_{\mathrm{R}} &= \Psi_{\mathrm{L}}^\intercal i\sigma^2\sigma^\mu\sigma^2(\partial_\mu - igA_\mu^A T^A)\Psi_{\mathrm{L}}^* \\
&= (-1)\Psi_{\mathrm{L}}^\dagger i(\overleftarrow{\partial}_\mu - igA_\mu^A (T_R^A)^\intercal)\overline{\sigma}^{\mu *}\Psi_{\mathrm{L}} = \Psi_{\mathrm{L}}^\dagger i\overline{\sigma}^\mu(\partial_\mu - igA_\mu^A(-T_R^A)^\intercal)\Psi_{\mathrm{L}} \\
&= \Psi_{\mathrm{L}}^\dagger i\overline{\sigma}^\mu(\partial_\mu - igA_\mu^A T_{\overline{R}}^A)\Psi_{\mathrm{L}}.
\end{aligned}$$

The gauge currents are now all the form $\overline{\Psi}_{\mathrm{L}}\gamma^{\mu}T^{A}\Psi_{\mathrm{L}}$, which corresponds to the case considered at the beginning of Sec. 6.3. These currents must be non-anomalous. Otherwise, the gauge symmetry would be absent at the quantum level and all considerations in Ch. 1 about decoupling of physical states, the physical Hilbert space and the unitarity of the S-matrix would be invalidated. The theory could not be consistently defined. The necessary condition for anomaly freedom is

$$d^{ABC} = \frac{1}{2}\operatorname{tr}(\{T^{A}, T^{B}\}T^{C}) = 0 \tag{6.9}$$

where the trace runs over all (left-handed by convention) fermion species in the theory. If a species does not couple to a specific gauge current, it transforms as a singlet, and the corresponding $T^{A}_{\mathrm{singlet}} = 0$, so there is no contribution to 6.9.

6.4.1 General results

(A) Non-chiral gauge theories (such as QED, QCD) are anomaly-free. Non-chiral theory means fermions come in pairs $\overline{\Psi}_{\mathrm{L}}i\not{D}\Psi_{\mathrm{L}} + \overline{\Psi}_{\mathrm{R}}i\not{D}\Psi_{\mathrm{R}} = \overline{\Psi}i\not{D}\Psi$ (Ψ being a Dirac spinor), i.e. pairs of left-handed fermions in complex-conjugate representations to each other. Then

$$d^{ABC} = \frac{1}{2}\operatorname{tr}(\{T^{A}_{R}, T^{B}_{R}\}T^{C}_{R}) + \frac{1}{2}\operatorname{tr}(\{T^{A}_{\overline{R}}, T^{B}_{\overline{R}}\}T^{C}_{\overline{R}}) = 0$$

since

$$\begin{aligned}\operatorname{tr}(\{T^{A}_{\overline{R}}, T^{B}_{\overline{R}}\}T^{C}_{\overline{R}}) &= -\operatorname{tr}(\{(T^{A}_{R})^{\intercal}, (T^{B}_{R})^{\intercal}\}(T^{C}_{R})^{\intercal}) = -\operatorname{tr}(T^{C}_{R}T^{B}_{R}T^{A}_{R} + T^{C}_{R}T^{A}_{R}T^{B}_{R}) \\ &= -\operatorname{tr}(\{T^{A}_{R}, T^{B}_{R}\}T^{C}_{R}),\end{aligned}$$

i.e. right and left-handed fermions transforming in the same R contribute an opposite amount to anomaly, which cancels.

(B) Fermions which transform in a real or pseudo-real representation do not contribute to the anomaly. For real generators it holds $(iT^{A})^{*} = iT^{A}$ and for pseudo-real there exists a unitary matrix S such that $(iT^{A})^{*} = SiT^{A}S^{\dagger}$. Proof of the statement above goes as follows:

$$\begin{aligned}\operatorname{tr}(\{T^{A}, T^{B}\}T^{C}) &= \operatorname{tr}(\{T^{A}, T^{B}\}T^{C})^{\intercal} = \operatorname{tr}((T^{C})^{\intercal}\{(T^{A})^{\intercal}, (T^{B})^{\intercal}\}) = \\ &= \operatorname{tr}(ST^{C}S^{\dagger}\{ST^{B}S^{\dagger}, ST^{A}S^{\dagger}\}) = -\operatorname{tr}(T^{C}\{T^{B}, T^{A}\}) \\ &= -\operatorname{tr}(\{T^{A}, T^{B}\}T^{C}).\end{aligned}$$

Hence, it must hold $\operatorname{tr}(\{T^{A}, T^{B}\}T^{C}) = 0$.

(C) To see if a chiral gauge theory is anomaly free, consider a general gauge group

$$G = G_1 \times G_2 \times \cdots \times G_m \times \mathrm{U}(1)^{n}, \quad G_i \in \{\mathrm{SU}(N), \mathrm{SO}(N), \mathrm{Sp}(2N), G_2, F_4, E_6, E_7, E_8\}.$$

The listed are compact simple Lie groups. The following groups have *only* real or pseudo-real representations and are therefore anomaly-free:

$$\mathrm{SU}(2) \cong \mathrm{SO}(3), \quad \mathrm{SO}(N)\ (\text{except}\ \mathrm{SO}(4N+2)), \quad \mathrm{Sp}(2N), \quad , G_2, F_4, E_7, E_8.$$

Of the remaining groups some are such that d^{ABC} always vanishes, even for the complex representations. These are

$$\mathrm{SO}(4N+2)\ (\text{except}\,\mathrm{SO}(2) \cong \mathrm{U}(1)\text{and}\,\mathrm{SO}(6) \cong \mathrm{SU}(4)), \quad E_6.$$

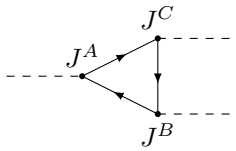
Hence, only $\mathrm{SU}(N)$ for $N \geq 3$ (which includes SO(6)) and U(1) (which includes SO(2)) can be anomalous.

6.4.2 Anomaly freedom of the Standard Model

The fermion content and their representations in the Standard Model are summarized in the following table.

Fermion	$\mathrm{SU}(3)_\mathrm{c}$	$\mathrm{SU}(2)_\mathrm{L}$	$\mathrm{U}(1)_Y$
$Q = \begin{pmatrix} u_\mathrm{L} \\ d_\mathrm{L} \end{pmatrix}$	$\mathbf{3}$	$\mathbf{2}$	$1/6$
$U = u_\mathrm{L}^\mathrm{c} \simeq u_\mathrm{R}^*$	$\overline{\mathbf{3}}$	$\mathbf{1}$	$-2/3$
$D = d_\mathrm{L}^\mathrm{c} \simeq d_\mathrm{R}^*$	$\overline{\mathbf{3}}$	$\mathbf{1}$	$+1/3$
$L = \begin{pmatrix} \nu_\ell \\ e_\ell \end{pmatrix}$	$\mathbf{1}$	$\mathbf{2}$	$-1/2$
$E = e_\mathrm{L}^\mathrm{c}$	$\mathbf{1}$	$\mathbf{1}$	1

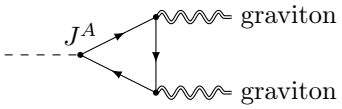
Note that the content is made exclusively of left-handed fields. The $\mathrm{U}(1)_Y$ charges of $\mathrm{SU}(2)_\mathrm{L}$ singlets are opposite of those for the right-handed particles (E has hypercharge $+1$ while e_R has hypercharge -1). Often, one will see $\square$ instead of $\mathbf{3}$ or $\mathbf{2}$, because one box depicts the fundamental representation as a Young tableau. Consider a triangle diagram



where the J's can be combination of the $\mathrm{SU}(3)_\mathrm{C}$, $\mathrm{SU}(2)_\mathrm{L}$ and $\mathrm{U}(1)_Y$ currents. One has a trace over the SU(3) and over the SU(2) indices. Hence, we need 0, 2 or 3 SU(3) currents, otherwise $\operatorname{tr} T^A_{\mathrm{SU}(3)} = 0$. Similar holds for SU(2). Hence, we need to consider the following contributions

$$\mathrm{SU}(3)^3, \quad \mathrm{SU}(2)^3 \quad \mathrm{SU}(3)^2\,\mathrm{U}(1)_Y, \quad \mathrm{SU}(2)^2\,\mathrm{U}(1)_Y, \quad \mathrm{U}(1)_Y^3.$$

From those, we can ignore $\mathrm{SU}(2)^3$ as it only has pseudo-real representations. In addition, all SM particles experience gravitational interactions, hence there is also the following diagram.



This does not require a quantum theory of gravity, only the SM particles appear in the loop. Now we check all possible cases.

- $\mathrm{SU}(3)^3$: only quarks can run in the loop, two from Q and one for U and D

$$d^{ABC} = \operatorname{tr} \mathbb{1}_{\mathrm{SU}(2)} \operatorname{tr}(\{T_3^A, T_3^B\}, T_3^C) + 2 \operatorname{tr}(\{T_{\bar{3}}^A, T_{\bar{3}}^B\}, T_{\bar{3}}^C) = 0$$

  because there are as many triplets as anti-triplets. This explains why for one SU(2) doublet of quarks one must have two SU(2) singlets. (The same does not hold for the SU(2) gauge group, because there is no $\mathrm{SU}(2)^3$ anomaly.)

- $\mathrm{SU}(3)^2\,\mathrm{U}(1)_Y$: again, only quarks and anti-quarks can run in the loop, so

$$d^{ABC} = \operatorname{tr} \mathbb{1}_{\mathrm{SU}(2)} \operatorname{tr}\{T_3^A, T_b^3\} Y_Q + \operatorname{tr}\{T_{\bar{3}}^A, T_{\bar{3}}^B\}(Y_U + Y_D) \propto 2Y_Q + Y_U + Y_D.$$

- $\mathrm{SU}(2)^2\,\mathrm{U}(1)_Y$: particles running in the loop can be left-handed set of quarks Q and left-handed doublet of leptons L:

$$d^{ABC} = \mathrm{tr}\,\mathbb{1}_{\mathrm{SU}(3)}\,\mathrm{tr}\{T_2^A, T_2^B\}Y_Q + \mathrm{tr}\{T_2^A, T_2^2\}Y_L \propto 3Y_Q + Y_L.$$

- $\mathrm{U}(1)_Y^3$: all fermions can run in the loop:

$$d^{ABC} \propto 6y_Q^3 + 3(y_U^3 + Y_D^3) + 2Y_L^3 + Y_E^3.$$

- Graviton-graviton-$\mathrm{U}(1)_Y$:

$$d^{ABC} \propto 6Y_Q + 3(Y_U + Y_D) + 2Y_L + Y_E.$$

There is no gravitational anomaly with SU(2) or SU(3) since $\mathrm{tr}(T^A_{\mathrm{SU}(2)}) = \mathrm{tr}(T^A_{\mathrm{SU}(3)}) = 0$. The cancellation of anomalies requires

$$Y_U + Y_D = -2Y_Q, \quad Y_L = -3Y_Q, \quad Y_E = 6Y_Q, \quad Y_U^3 + Y_D^3 = -56Y_Q^3$$

with solutions

1) $y_Q = 0$ and $Y_E = Y_L = 0$ and $Y_U = -Y_D$ which implies electric charges $Q_\Psi = T_{3\Psi} + Y_\Psi$

U_{L}	D_{L}	$U^{\mathrm{c}}_{\mathrm{L}}$	$D^{\mathrm{C}}_{\mathrm{L}}$	ν_{L}	e_{L}	$e^{\mathrm{c}}_{\mathrm{L}}$
$1/2$	$-1/2$	Y_U	$-Y_U$	$1/2$	$-1/2$	0

   which is not realized in Nature.

2) $Y_Q \neq 0$ which implies

$$Y_U = -4Y_Q, \quad Y_D = 2Y_Q, \quad Y_L = -3Y_Q, \quad Y_E = 6Y_Q.$$

   This corresponds to the SM for $Y_Q = 1/6$. Note that the normalization of the hypercharge is arbitrary and corresponds to a charge in g'. Hence, the choice of y_Q is arbitrary.

The consequences of the discussion above are:

- the SM gauge symmetry is anomaly free,
- anomaly freedom requires the seemingly arbitrary hypercharge assignments of the quarks and leptons to be related to each other,
- the number of colours of the quarks is interwoven in the equations.

A natural connection of quarks and leptons, of colours, isospin, and hypercharge arises, if one assumes that the SM gauge group is embedded into a larger simple group

$$\mathrm{SU}(3)_{\mathrm{c}} \times \mathrm{SU}(2)_{\mathrm{L}} \times \mathrm{U}(1)_Y \subset G$$

where the full symmetry G is spontaneously broken at a yet inaccessible high-energy scale. Anomaly freedom is then automatic if G is a group which cannot have anomalies. The simplest possibility that can accommodate all SM fermions and is automatically anomaly-free is SO(10).

6.4.3 Violation of baryon and lepton number conservation

Before we discussed how external electromagnetic and SU(2) fields can make the $\mathrm{SU}(3)_\mathrm{A}$ flavour symmetry of QCD anomalous. We did not discuss the $\mathrm{U}(1)_\mathrm{B}$ symmetry so far. We assumed it was exact. The entire SM has chiral, global U(1) symmetries

- $\mathrm{U}(1)_\mathrm{B}$: the *baryon number*

$$j^\mu_\mathrm{B} = \frac{1}{3}\sum_{\Psi=Q,U,D}\overline{\Psi}\gamma^\mu\Psi = \frac{1}{3}(\overline{Q}\gamma^\mu Q - \overline{u}^\mathrm{c}_\mathrm{L}\gamma^\mu u^\mathrm{c}_\mathrm{L} - \overline{d}^\mathrm{c}_\mathrm{L}\gamma^\mu d^\mathrm{c}_\mathrm{L})$$

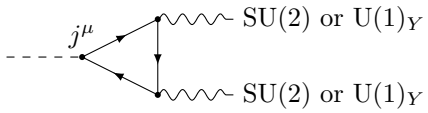
  where the sum over generations $I = 1, 2, 3$ is implicit.

- $\mathrm{U}(1)_{L_I}$: the *lepton flavour number*:

$$j^\mu_{L_I} = \sum_{\Psi=L_I,E_I}\overline{\Psi}\gamma^\mu\Psi = \overline{L}_I\gamma^\mu L_I - \overline{e}^\mathrm{c}_\mathrm{L}\gamma^\mu e^\mathrm{c}_\mathrm{L}.$$

  There is no sum over I, there is a separate current for every I.

$\mathrm{U}(1)_\mathrm{B}$ is an exact symmetry of QCD, but all four above currents are anomalous in the SM.



$$\begin{aligned}
\partial^\mu j_{\mathrm{B}\mu} &= \frac{1}{3}N_\mathrm{c}n_\mathrm{G}\frac{1}{2}\frac{g^2}{32\pi^2}F_{\mu\nu}\tilde{F}^{\mu\nu}|_{\mathrm{SU}(2)} + \frac{1}{3}N_\mathrm{c}n_\mathrm{G}\left(2\left(\frac{1}{6}\right)^2 - \left(\frac{2}{3}\right)^2 - \left(\frac{1}{3}\right)^2\right)\frac{g'^2}{32\pi^2}F_{\mu\nu}\tilde{F}^{\mu\nu}|_{\mathrm{U}(1)_Y} \\
&= \frac{n_\mathrm{G}}{2}\frac{g^2}{32\pi^2}F_{\mu\nu}\tilde{F}^{\mu\nu}|_{\mathrm{SU}(2)} - \frac{n_\mathrm{G}}{2}\frac{g'^2}{32\pi^2}F_{\mu\nu}\tilde{F}^{\mu\nu}|_{\mathrm{U}(1)_Y}, \\
\partial^\mu j_{L_I\mu} &= \frac{1}{2}\frac{g^2}{32\pi^2}F_{\mu\nu}\tilde{F}^{\mu\nu}|_{\mathrm{SU}(2)} + \left(2\left(-\frac{1}{2}\right)^2 - 1^2\right)\frac{g'^2}{32\pi^2}F_{\mu\nu}\tilde{F}^{\mu\nu}|_{\mathrm{U}(1)_Y} \\
&= \frac{1}{2}\frac{g^2}{32\pi^2}F_{\mu\nu}\tilde{F}^{\mu\nu}|_{\mathrm{SU}(2)} - \frac{1}{2}\frac{g'^2}{32\pi^2}F_{\mu\nu}\tilde{F}^{\mu\nu}|_{\mathrm{U}(1)_Y}.
\end{aligned}$$

Symbol n_G denotes the number of generations, factor $1/2$ in the SU(2) part comes from $\mathrm{tr}(T_2^A T_2^B)$ and the numerator of the anomaly $32\pi^2$ differs by factor 2 from 6.4 because the current is $(1-\gamma_5)/2$. B(aryon number) and L_I for $I = 1, 2, 3$ both have an $\mathrm{SU}(2)\,\mathrm{U}(1)_Y$ anomaly. But B_L (where $L = L_1 + L_2 + L_3$ is the total lepton number) and $L_I - L_J$ are non-anomalous and conserved. Hence, processes as $\mu \to e\gamma$ ($L_\mu - L_e$ changes) or proton decay ($B - L$ changes, at the beginning $B = 1$ and $L = 0$, at the end $B = 0$ so L should be -1, which is not possible) are not possible in the SM. On the other hand

$$^3\mathrm{He}_2{}^+ \longrightarrow e^+ + \mu^+ + \overline{\nu}_\tau \tag{6.10}$$

is possible.

Note. The existence of neutrino oscillations shows that $L_I - L_J$ is not conserved in Nature. The considerations here are based on the SM with renormalizable interactions. The dimension-5 operator discussed in Sec. 4.3.1, which generates neutrino masses, breaks lepton-flavour number conservation and even lepton number conservation explicitly. This effect is small, because the neutrino masses are so small.

Nevertheless, processes such as 6.10 which violate baryon number – more precisely $B + L$ since $B - L$ is conserved – are not observed in Nature. The reason is that the integral of the anomaly term in the current conservation equation

$$\int \mathrm{d}^4x \, F_{\mu\nu}\tilde{F}^{\mu\nu}$$

can be written as a surface integral, because $F_{\mu\nu}\tilde{F}^{\mu\nu}$ can in fact be expressed as a total divergence. Normally we would therefore conclude that the anomaly has no physical effect at all! This is not the case here, because there exist non-trivial ("topological") gauge configurations which do not vanish for $x^\mu \to \infty$ (after Wick rotation to Euclidean space) while nevertheless $F_{\mu\nu} = 0$. These are the so-called *instantons*. Unfortunately, there is no time to discuss this subject in this course (it is one of the main highlights of the QFT3 course). One can show that instantons are related to tunnelling transitions between topologically different vacuum states. The tunnelling is accompanied by $B + L$ violating process, however, the tunnelling amplitude is exponentially small, $\propto e^{-8\pi^2/g^2}$, hence, in practice $B + L$ violating processes such as 6.10 do not happen at any observable rate.

This is different in the very early Universe, where the energy barrier may be overcome by thermal fluctuations of the hot plasma. It is possible that the baryon number excess (matter-antimatter asymmetry) is generated from an initial lepton number asymmetry through $B + L$ violation in the early Universe. The lepton number asymmetry can be naturally generated by whatever gives rise to the L-violating dimension-5 operator in the SM, which also causes the neutrino masses.

Index

Symbols
ρ parameter 88

A
adjoint representation 9
Adler–Bardeen theorem 146
Adler–Bell–Jackiw anomaly 145
anomalous dimension 30
anomaly 136
anomaly coefficient 143
anti-ghost field 18
asymptotic freedom 29
axial gauge 126

B
Bardeen anomaly 149
baryon 75
baryon number 153
Bjorken scaling 123
Bjorken's scaling variable 120
Bloch–Nordsieck cancellation 110
BRST transformation 20

C
Callan–Gross relation 123
chiral anomaly 136
chiral logarithms 84
chiral perturbation theory 83
Christoffel symbols 43
cohomology 36
Coleman–Weinberg potential 101
collinear divergence 109
collinear logarithms 130
colour 74
colour indices 11
confinement 74
conformal anomaly 136
covariant derivative 4, 6
covariant gauge 18
custodial symmetry 86

D
de Donder gauge 46
deep inelastic scattering 120
DGLAP evolution equations 130
dilogarithm 118
dimensional transmutation 30
Drell–Yan process 133
Dynkin index 9

E
effective action 55
effective potential 55
Einstein field equations 44
Einstein–Hilbert action 44
electric charge current 4
electroweak symmetry 85
Euler–Lagrange equations 4

F
Faddeev–Popov method 46
Feynman gauge 27
Fierz–Pauli Lagrangian 41
Fujikawa method 138
fundamental representation 9

G
gauge symmetry 4
gauge-fixing term 18
Gell-Mann matrices 10
Gell-Mann–Oakes–Renner relation 77
ghost field 18
ghost Lagrangian 18
gluon 74

Goldstone boson 61
Goldstone boson equivalence theorem ... 96
Goldstone's theorem 59
gradient expansion 55
graviton 46
group
special orthogonal 8
special unitary 8
symplectic 9

H
hadron 74
hadronic tensor 120
hadronization 112
heat-kernel regularization 139
Higgs sector 85
hypercharge 85

I
infrared divergence 107
infrared regularization 116
instanton 154
irreducible representation 9

J
Jacobi identity 9
JADE algorithm 117
jet 112, 117

K
kaon 78

L
Landau gauge 102
Landau pole 104
lepton flavour number 153
light-cone divergence 125
linear σ-model 64
Lippmann–Schwinger equation 35

M
meson 75
minimal coupling 48

N
Nakanishi–Lautrup field 20
next-to-leading order 111
non-linear representation 71

O
optical theorem 94
order parameter 52

P
parton 112
parton distribution function 122
parton luminosity 133
parton-hadron duality 113
Pauli matrices 10, 65
pion 77
pseudo-Goldstone bosons 91
pure Yang–Mills theory 19

Q
quantum chromodynamics 74
quantum electrodynamics 4
quark 74

R
real emission correction 114
Ricci scalar 43
Ricci tensor 43
Riemann curvature tensor 43
R_ξ gauge 90

S
sea 132
sea quarks 132
Slavnov–Taylor formula 22
soft divergence 109
spectral density function 61
splitting function 128
strong CP problem 136
strong interaction 74
structure constants 5
structure functions 121
Sudakov decomposition 124
Sudakov double logarithm 111
superficial degree of divergence 24
Sutherland–Veltman paradox 137

T
't Hooft gauge 90
't Hooft–Veltman scheme 147
triviality 104

U
unitary gauge 85

V
valence quarks 132
vierbein formalism 48
virtual correction 114

W
Ward identity 59
Weinberg angle 88
Weinberg operator 97
Wess–Zumino–Witten term 149
Wess–Zumino–Witten terms 84