

QCD and Hadron Physics

Gernot Eichmann
Lecture notes, 2021

Contents

1	Introduction	5
1.1	Standard Model	5
1.2	Questions in hadron physics	5
1.3	What is confinement?	5
1.4	QCD phase diagram	5
1.5	Theoretical methods	5
1.6	History of QCD	5
2	QCD	1
2.1	QCD Lagrangian	1
2.2	Quantization of QCD	11
2.2.1	Canonical quantization	11
2.2.2	Path-integral quantization	19
2.2.3	Gauge fixing in QCD	31
2.3	Renormalization	39
2.3.1	Feynman rules of QCD	39
2.3.2	Regularization and renormalization	45
2.3.3	β function and running coupling	55
3	Hadrons	65
3.1	Flavor symmetries and currents	65
3.1.1	Symmetries of the QCD Lagrangian	68
3.1.2	Symmetry relations at the quantum level	73
3.1.3	Extracting hadrons from QCD	83
3.2	Hadron spectrum	91
3.2.1	Mesons	93
3.2.2	Baryons	107
4	Low-energy QCD phenomenology	121
4.1	Quark potential models	121
4.2	Spontaneous chiral symmetry breaking	129
4.3	$U(1)_A$ anomaly	139
4.4	Chiral effective field theories	147
4.4.1	Sigma model	147
4.4.2	Chiral perturbation theory	161
4.5	Hadron matrix elements	163

4.5.1	Scattering amplitudes	163
4.5.2	Form factors	171
5	High-energy phenomenology	185
5.1	Deep inelastic scattering	185
A	SU(N)	199
A.1	Basic properties of $SU(N)$	199
A.2	SU(N) representations	201
A.3	Product representations	202
B	Poincaré group	211
B.1	Lorentz and Poincaré group	211
B.2	Representations of the Lorentz group	216
B.3	Poincaré invariance in field theories	220
C	Euclidean conventions	227

Chapter 1

Introduction

- 1.1 Standard Model
- 1.2 Questions in hadron physics
- 1.3 What is confinement?
- 1.4 QCD phase diagram
- 1.5 Theoretical methods
- 1.6 History of QCD

Chapter 2

QCD

Quantum Chromodynamics (QCD) is the theory of the strong interaction. It describes the ‘color force’ that binds quarks and gluons to colorless hadrons (protons, neutrons, pions, etc.) and hadrons to nuclei. At hadronic scales, the strong force is ~ 100 times stronger than the electromagnetic interaction, extremely short-ranged (the typical interaction range is the size of a hadron $\sim 1 \text{ fm} = 10^{-15} \text{ m}$), and its typical energy scale is the mass of the proton $\sim 1 \text{ GeV}$.

The strong interaction is described by a local, non-Abelian $SU(3)_C$ gauge theory with several peculiar features. While quarks and gluons are asymptotically free at short distances, they are confined at large distances: only colorless bound states (hadrons) can be detected in experiments, and no quark or gluon has ever been observed directly. Nevertheless, nature has given us an abundance of evidence that these constituents exist, and their theoretical description in terms of a non-Abelian gauge theory has evolved from being considered a mere mathematical trick to a quite fundamental framework. In this chapter we will recapitulate the properties of QCD and its fundamental degrees of freedom and postpone the discussion of hadrons to Chapter 3.

2.1 QCD Lagrangian

Field content. The definition of a quantum field theory starts with constructing its Lagrangian \mathcal{L} (or, equivalently, its action $S = \int d^4x \mathcal{L}$), based on the desired underlying symmetries. The symmetries of QCD are: Poincaré invariance, local color gauge invariance and various flavor symmetries, and the fields in the Lagrangian should transform under representations of these groups. The QCD Lagrangian contains quark and antiquark fields, and (as a consequence of color gauge invariance) gluon fields which mediate the strong interaction:

$$\psi_{\alpha,i,f}(x), \quad \bar{\psi}_{\alpha,i,f}(x), \quad A_a^\mu(x). \quad (2.1.1)$$

The quark fields are Dirac spinors (index α) and transform under the fundamental representation of $SU(3)_C$ (color index $i = 1, 2, 3$ or red, green blue). The additional index $f = 1 \dots N_f$ labels the flavor quantum number ($f = \text{up, down, strange, charm, bottom, top}$). The eight gluon fields $A_a^\mu(x)$ are Lorentz vectors; there is one field

for each generator \mathbf{t}_a of the group ($a = 1 \dots 8$). In the fundamental representation: $\mathbf{t}_a = \lambda_a/2$, where the λ_a are the eight Gell-Mann matrices; see Appendix A for a collection of basic $SU(N)$ relations. Gluons are flavor-blind and carry no flavor labels.

Gauge invariance. A free fermion Lagrangian $\bar{\psi}(i\cancel{D} - m)\psi$ constructed from the quark and antiquark fields (we leave the summation over Dirac, color and flavor indices implicit) is invariant under global $SU(3)_C$ transformations

$$\psi'(x) = U\psi(x), \quad \bar{\psi}'(x) = \bar{\psi}(x)U^\dagger \quad \text{with} \quad U = e^{i\varepsilon} = e^{i\sum_a \varepsilon_a \mathbf{t}_a}, \quad (2.1.2)$$

where $\varepsilon_a = \text{const.}$ and the U_{ij} act on the color indices of the quarks. This invariance is no longer satisfied if we impose a local $SU(3)_C$ gauge symmetry $\psi'(x) = U(x)\psi(x)$ with spacetime-dependent group parameters $\varepsilon_a(x)$. The mass term is still invariant, but the derivative in the kinetic term now also acts on the spacetime argument of $U(x)$, and invariance of the Lagrangian (or the action) cannot be satisfied with an ordinary partial derivative. To ensure local color gauge invariance, we introduce a **covariant derivative** and thus gluon fields:

$$D_\mu = \partial_\mu - igA_\mu, \quad (2.1.3)$$

where $A^\mu(x) = \sum A_a^\mu(x) \mathbf{t}_a$ is an element of the Lie algebra. From the new Lagrangian $\bar{\psi}(i\cancel{D} - m)\psi$ we see that $\cancel{D}\psi$ must transform in the same way as the quark field itself, which fixes the transformation properties of the gluon fields:

$$\bar{\psi}'\cancel{D}'\psi' \stackrel{!}{=} \bar{\psi}\cancel{D}\psi \quad \Rightarrow \quad D'_\mu\psi' = UD_\mu\psi = UD_\mu U^\dagger\psi' \quad (2.1.4)$$

$$\begin{aligned} &\Rightarrow (\partial_\mu - igA'_\mu)\psi' = U(\partial_\mu - igA_\mu)U^\dagger\psi' \\ &\Rightarrow A'_\mu = UA_\mu U^\dagger + \frac{i}{g}U(\partial_\mu U^\dagger). \end{aligned} \quad (2.1.5)$$

The second term in A'_μ is particular to local gauge transformations; for a global symmetry we don't need a covariant derivative and could simply set $A_\mu = 0$. Note also that we can generate gluon fields out of nothing ($A_\mu = 0$) by a local gauge transformation: such gauge fields $\sim U(\partial_\mu U^\dagger)$ are called pure gauge configurations.

Why do we actually impose local gauge invariance in the first place? In fact, only *global* symmetries are true 'symmetries' which lead to conserved charges and quantum numbers. A local gauge symmetry reflects a redundancy in the description, which can be seen if we turn the argument around and start from Eq. (2.1.5), for example in the Abelian case where $U(x) = e^{i\varepsilon(x)}$ is just a phase. The action of a free massless vector field contains redundant degrees of freedom which are related to each other by local gauge transformations $A'_\mu = A_\mu + \partial_\mu\varepsilon/g$. The standard way to eliminate them is to modify the Lagrangian and impose a gauge-fixing condition on the state space (cf. Sec. 2.2.3). As a consequence, longitudinal and timelike photons decouple from physical processes and S-matrix elements are transverse: $q_\mu \mathcal{M}^\mu = 0$. To preserve this feature when including interactions (e.g., when adding fermions), the interacting part of the action must couple to a conserved current corresponding to the global symmetry of the full action, $\delta S_{\text{int}}/\delta A^\mu = j^\mu$, which is equivalent to imposing *local* gauge invariance for the matter fields. Thus, Eq. (2.1.5) is tied to the invariance under $\psi'(x) = U(x)\psi(x)$, and even though we needed an underlying global symmetry in the fermion sector to begin with, the local gauge invariance is not truly a symmetry but rather a *consistency* constraint that generates dynamics. In QCD, it introduces a quark-gluon interaction of the form $g\bar{\psi}\cancel{A}\psi$.

Another way to motivate the covariant derivative is the following. We can write the ordinary derivative as

$$n^\mu \partial_\mu \psi(x) = \lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} [\psi(x + \epsilon n) - \psi(x)]. \quad (2.1.6)$$

For a local gauge transformation $\psi'(x) = U(x)\psi(x)$ the first term becomes $U(x + \epsilon n)\psi(x + \epsilon n)$ but the second $U(x)\psi(x)$, so we are comparing objects at different spacetime points. To remedy this, we define the **parallel transporter** or **link variable** $C(y, x)$ by

$$C'(y, x) = U(y) C(y, x) U^\dagger(x), \quad C(x, x) = 1, \quad (2.1.7)$$

because then the quantity $C(y, x)\psi(x)$ has a simple transformation behavior:

$$[C(y, x)\psi(x)]' = U(y) C(y, x) U^\dagger(x) U(x) \psi(x) = U(y) C(y, x) \psi(x). \quad (2.1.8)$$

Now, if we define the covariant derivative as

$$n^\mu D_\mu \psi(x) = \lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} [\psi(x + \epsilon n) - C(x + \epsilon n, x) \psi(x)] \quad (2.1.9)$$

and perform a gauge transformation, then $U(x + \epsilon n)$ can be pulled out so that $[D_\mu \psi(x)]' = U(x) D_\mu \psi(x)$, and thus $\bar{\psi} \not{D} \psi$ is invariant under the local symmetry.

Moreover, we can write down the Taylor expansion of the parallel transporter:

$$C(x + \epsilon n, x) = 1 + \epsilon n^\mu ig A_\mu(x) + \mathcal{O}(\epsilon^2), \quad (2.1.10)$$

where igA_μ is just a *name* for the coefficient of the linear term. Inserting this into (2.1.9) yields

$$n^\mu D_\mu \psi(x) = \lim_{\epsilon \rightarrow 0} \frac{1}{\epsilon} [\psi(x + \epsilon n) - \psi(x) - \epsilon n^\mu ig A_\mu \psi(x)] = n^\mu \partial_\mu \psi(x) - n^\mu ig A_\mu \psi(x) \quad (2.1.11)$$

and therefore $D_\mu = \partial_\mu - igA_\mu$. Similarly, the transformation of the gluon field follows from

$$\begin{aligned} C'(x + \epsilon n, x) &= U(x + \epsilon n) C(x + \epsilon n, x) U^\dagger(x) \\ &= [U(x) + \epsilon n^\mu \partial_\mu U(x) + \mathcal{O}(\epsilon^2)] [1 + \epsilon n^\nu ig A_\nu(x) + \mathcal{O}(\epsilon^2)] U^\dagger(x) \\ &= 1 + \epsilon n^\mu ig \left[U(x) A_\mu(x) U^\dagger(x) - \frac{i}{g} (\partial_\mu U(x)) U^\dagger(x) \right] + \mathcal{O}(\epsilon^2) \\ &\stackrel{!}{=} 1 + \epsilon n^\mu ig A'_\mu(x) + \mathcal{O}(\epsilon^2), \end{aligned} \quad (2.1.12)$$

which reproduces the result (2.1.5) since $\partial_\mu(UU^\dagger) = (\partial_\mu U) U^\dagger + U \partial_\mu U^\dagger = 0$.

Gluon dynamics. Next, we need a kinetic term that describes the dynamics of the gluons. To this end we define the **gluon field strength tensor** as the commutator of two covariant derivatives:

$$F_{\mu\nu}(x) = \frac{i}{g} [D_\mu, D_\nu] = \partial_\mu A_\nu - \partial_\nu A_\mu - ig [A_\mu, A_\nu]. \quad (2.1.13)$$

It is then also an element of the Lie algebra and we can write it as

$$F_{\mu\nu} = \sum_a F_{\mu\nu}^a \mathbf{t}_a, \quad (2.1.14)$$

where the \mathbf{t}_a are again taken in the fundamental representation because ∂_μ , D_μ and A_μ act on quark fields in the (three-dimensional) fundamental representation of $SU(3)_C$. $F_{\mu\nu}$ inherits the transformation properties from (2.1.4): $F'_{\mu\nu} = U F_{\mu\nu} U^\dagger$.

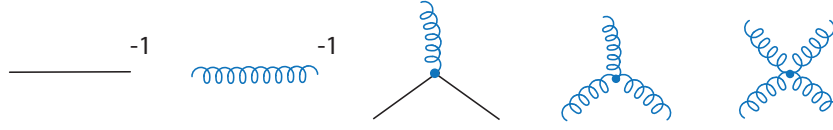


FIG. 2.1: Tree-level (inverse) propagators and interactions in the QCD action.

The contraction of two field-strength tensors is not gauge invariant; only its color trace is invariant due to the cyclic property of the trace:

$$\text{Tr} \{ F'_{\mu\nu} F'^{\mu\nu} \} = \text{Tr} \left\{ U F_{\mu\nu} U^\dagger U F^{\mu\nu} U^\dagger \right\} = \text{Tr} \{ F_{\mu\nu} F^{\mu\nu} \}. \quad (2.1.15)$$

Only the trace can therefore appear in the Lagrangian. We can write it as

$$\text{Tr} \{ F_{\mu\nu} F^{\mu\nu} \} = F_{\mu\nu}^a F_b^{\mu\nu} \text{Tr} \{ \mathbf{t}_a \mathbf{t}_b \} = T(R) F_{\mu\nu}^a F_a^{\mu\nu}, \quad (2.1.16)$$

where $T(R) = 1/2$ in the fundamental representation of $SU(N)$, cf. Appendix A. From Eq. (2.1.5) we also conclude that a gluon mass term $\sim m_g A_\mu A^\mu$ cannot appear in the Lagrangian because it would violate gauge invariance: gluons must be **massless**.

We can work out the components of the field-strength tensor as

$$\begin{aligned} F_{\mu\nu} &= F_{\mu\nu}^a \mathbf{t}_a = \partial_\mu A_\nu^a \mathbf{t}_a - \partial_\nu A_\mu^a \mathbf{t}_a - ig A_\mu^a A_\nu^b [\mathbf{t}_a, \mathbf{t}_b] \\ &= (\partial_\mu A_\nu^a - \partial_\nu A_\mu^a + gf_{abc} A_\mu^b A_\nu^c) \mathbf{t}_a, \end{aligned} \quad (2.1.17)$$

where we used $[\mathbf{t}_a, \mathbf{t}_b] = if_{abc} \mathbf{t}_c$. Note that in an Abelian gauge theory such as QED this commutator would vanish, leaving only the linear terms in the gluon fields. The non-Abelian nature of $SU(3)_C$ introduces gluonic self-interactions which lead to significant complications. Inserting Eq. (2.1.17) into the term $F_{\mu\nu}^a F_a^{\mu\nu}$ and partial integration yields

$$\begin{aligned} -\frac{1}{4} F_{\mu\nu}^a F_a^{\mu\nu} &\cong \frac{1}{2} A_\mu^a (\square g^{\mu\nu} - \partial^\mu \partial^\nu) A_\nu^a \\ &\quad - \frac{g}{2} f_{abc} (\partial^\mu A_\nu^a - \partial^\nu A_\mu^a) A_\mu^b A_\nu^c - \frac{g^2}{4} f_{abe} f_{cde} A_\mu^a A_b^\mu A_\nu^c A_\nu^d, \end{aligned} \quad (2.1.18)$$

where \cong means ‘up to surface terms in the action’, e.g. $\partial_\mu A_\nu^a \partial^\mu A_\nu^a \cong -A_\nu^a \square A_\nu^a$ after partial integration. In contrast to the Abelian theory, where the F^2 term only produces a photon propagator, we can see that in the non-Abelian case we end up with the gluon propagator, a three-gluon interaction $\sim A^3$ and a four-gluon interaction $\sim A^4$.

Feynman rules. The terms $\bar{\psi} (i\not{D} - m) \psi$ and $-\frac{1}{4} F_{\mu\nu}^a F_a^{\mu\nu}$ in the Lagrangian allow us to read off the **Feynman rules** for the tree-level correlation functions of the QFT. In particular, the action contains the 1PI (one-particle irreducible, see Sec. 2.2.2) quantities, which means the vertices and *inverse* propagators that define the theory (Fig. 2.1). The procedure is as follows: symmetrize the respective term in the action (if necessary), transform it to momentum space, split off the integrals, fields and symmetry factors, and multiply with i to get the Feynman rule for the propagator or vertex.

For example, the inverse **quark propagator** corresponds to the term $\bar{\psi}(i\cancel{\partial} - m)\psi$. The Fourier transform of the fields is

$$\psi(x) = \int \frac{d^4p}{(2\pi)^4} e^{-ip \cdot x} \psi(p). \quad (2.1.19)$$

Abbreviating $\int_p = \int d^4p/(2\pi)^4$, the term in the action becomes

$$\int d^4x \bar{\psi}(i\cancel{\partial} - m)\psi = \iint_{p' p} \bar{\psi}(p') (\cancel{p} - m) \psi(p) \int d^4x e^{i(p' - p) \cdot x} = \int_p \bar{\psi}(p) (\cancel{p} - m) \psi(p)$$

and dividing by i , the inverse tree-level propagator is

$$S_0^{-1}(p) = -i(\cancel{p} - m) \quad \Leftrightarrow \quad S_0(p) = \frac{i(\cancel{p} + m)}{p^2 - m^2 + i\epsilon}. \quad (2.1.20)$$

Likewise, the inverse **gluon propagator** can be read off from Eq. (2.1.18). Replacing $\square \rightarrow -p^2$ and $\partial^\mu \partial^\nu \rightarrow -p^\mu p^\nu$, we find

$$(D_0^{-1})^{\mu\nu}(p) = ip^2 \left(g^{\mu\nu} - \frac{p^\mu p^\nu}{p^2} \right). \quad (2.1.21)$$

The symmetry factor 1/2 does not enter in the Feynman rule. Here we encounter, however, a difficulty: the inverse gluon propagator is proportional to a transverse projector, which is not invertible and thus the gluon propagator does not exist. We will cure the problem in Sec. 2.2.3 by the Faddeev-Popov method, where we follow analogous steps as in QED and add gauge-fixing terms to the action (which will also introduce ghost fields). Before we get there, keep in mind that the gluon propagator is not yet well-defined.

The **quark-gluon vertex** comes from the term $g\bar{\psi}\cancel{A}\psi$ induced by the covariant derivative. If we denote the incoming and outgoing quark momenta by p and p' and the incoming gluon momentum by q , we have

$$\int d^4x \bar{\psi} g\cancel{A}\psi = \sum_a \iint_{p' p q} (2\pi)^4 \delta^4(p' - p - q) \bar{\psi}(p') A_\mu^a(q) g\gamma^\mu \mathbf{t}_a \psi(p),$$

so the tree-level vertex is $ig\gamma^\mu \mathbf{t}_a$.

The **three-gluon vertex** must be fully symmetric under exchange of any two legs, but this symmetry is not yet manifest in the A^3 term of Eq. (2.1.18). To this end, we abbreviate $\partial^{\mu\nu\rho} = \partial^\mu g^{\nu\rho} - \partial^\nu g^{\mu\rho}$ and write

$$f_{abc} (\partial^\mu A_\alpha^\nu - \partial^\nu A_\alpha^\mu) A_\mu^b A_\nu^c = f_{abc} (\partial^{\mu\nu\rho} A_\rho^a) A_\mu^b A_\nu^c = f_{abc} A_\mu^a A_\nu^b (\partial^{\mu\nu\rho} A_\rho^c) = \dots$$

In the last step we renamed the color indices and used $f_{abc} = f_{bca} = f_{cab}$. For three Lorentz indices there are $3! = 6$ possible permutations; $\partial^{\mu\nu\rho}$ is already antisymmetric in $\mu \leftrightarrow \nu$ so we only need to add the two remaining cyclic permutations:

$$\begin{aligned} \dots &= \frac{1}{3} f_{abc} \left[A_\mu^a A_\nu^b (\partial^{\mu\nu\rho} A_\rho^c) + A_\nu^a A_\rho^b (\partial^{\nu\rho\mu} A_\mu^c) + A_\rho^a A_\mu^b (\partial^{\rho\mu\nu} A_\nu^c) \right] \\ &= \frac{1}{3} f_{abc} \left[A_\mu^a A_\nu^b (\partial^{\mu\nu\rho} A_\rho^c) + (\partial^{\nu\rho\mu} A_\mu^a) A_\nu^b A_\rho^c + A_\mu^a (\partial^{\rho\mu\nu} A_\nu^b) A_\rho^c \right]. \end{aligned} \quad (2.1.22)$$

In the first line we renamed the Lorentz indices and in the second line the color indices. Now we can pull out $A_\mu^a(p_1) A_\nu^b(p_2) A_\rho^c(p_3)$ in momentum space and the term in the action becomes

$$-\frac{ig}{6} f_{abc} \int \int \int_{p_1 p_2 p_3} (2\pi)^4 \delta^4(p_1 + p_2 + p_3) A_\mu^a(p_1) A_\nu^b(p_2) A_\rho^c(p_3) \times \left[(p_1 - p_2)^\rho g^{\mu\nu} + (p_2 - p_3)^\mu g^{\nu\rho} + (p_3 - p_1)^\nu g^{\rho\mu} \right], \quad (2.1.23)$$

from where we read off the Feynman rule for the vertex:

$$\Gamma_{3g,0}^{\mu\nu\rho} = gf_{abc} \left[(p_1 - p_2)^\rho g^{\mu\nu} + (p_2 - p_3)^\mu g^{\nu\rho} + (p_3 - p_1)^\nu g^{\rho\mu} \right]. \quad (2.1.24)$$

The symmetry factor $1/6$ does again not enter, and $p_1 + p_2 + p_3 = 0$. The resulting vertex is Bose-symmetric, i.e., symmetric under a combined exchange of any two momenta with corresponding Lorentz and color indices.

The same strategy applies to the **four-gluon vertex** from the A^4 term in (2.1.18), which is also not yet manifestly symmetric:

$$f_{abe} f_{cde} A_\mu^a A_\nu^b A_c^\mu A_d^\nu = f_{abe} f_{cde} g^{\mu\rho} g^{\nu\sigma} A_\mu^a A_\nu^b A_\rho^c A_\sigma^d = \frac{1}{2} f_{abe} f_{cde} (g^{\mu\rho} g^{\nu\sigma} - g^{\nu\rho} g^{\mu\sigma}) A_\mu^a A_\nu^b A_\rho^c A_\sigma^d. \quad (2.1.25)$$

Denoting $\Gamma^{\mu\nu\rho\sigma} = g^{\mu\rho} g^{\nu\sigma} - g^{\nu\rho} g^{\mu\sigma}$, then with four Lorentz indices there are $4! = 24$ possible permutations of $(\mu\nu\rho\sigma) \equiv (1234)$:

$$\begin{array}{cccccc} 1234 & 3412 & 2314 & 1423 & 3124 & 2431 \\ 1243 & 3421 & 2341 & 1432 & 3142 & 2413 \\ 2134 & 4312 & 3214 & 4123 & 1324 & 4231 \\ 2143 & 4321 & 3241 & 4132 & 1342 & 4213 \end{array}. \quad (2.1.26)$$

The permutations in the first two columns are already covered because $\Gamma^{\mu\nu\rho\sigma} = -\Gamma^{\mu\nu\sigma\rho}$, etc., so we only need to add (2314) and (3124):

$$\begin{aligned} f_{abe} f_{cde} A_\mu^a A_\nu^b A_c^\mu A_d^\nu &= \frac{1}{6} f_{abe} f_{cde} \left[(g^{\mu\rho} g^{\nu\sigma} - g^{\nu\rho} g^{\mu\sigma}) A_\mu^a A_\nu^b A_\rho^c A_\sigma^d \right. \\ &\quad + (g^{\nu\mu} g^{\rho\sigma} - g^{\rho\mu} g^{\nu\sigma}) A_\nu^a A_\rho^b A_\mu^c A_\sigma^d \\ &\quad \left. + (g^{\rho\nu} g^{\mu\sigma} - g^{\mu\nu} g^{\rho\sigma}) A_\rho^a A_\mu^b A_\nu^c A_\sigma^d \right] \\ &= \frac{1}{6} A_\mu^a A_\nu^b A_\rho^c A_\sigma^d \left[f_{abe} f_{cde} (g^{\mu\rho} g^{\nu\sigma} - g^{\nu\rho} g^{\mu\sigma}) \right. \\ &\quad + f_{bce} f_{ade} (g^{\nu\mu} g^{\rho\sigma} - g^{\rho\mu} g^{\nu\sigma}) \\ &\quad \left. + f_{cae} f_{bde} (g^{\rho\nu} g^{\mu\sigma} - g^{\mu\nu} g^{\rho\sigma}) \right]. \end{aligned} \quad (2.1.27)$$

Together with $-g^2/4$ from (2.1.18), the combined symmetry factor for the A^4 term is indeed $1/24$. The resulting four-gluon vertex is Bose-symmetric and given by

$$\begin{aligned} \Gamma_{4g,0}^{\mu\nu\rho\sigma} &= -ig^2 \left[f_{abe} f_{cde} (g^{\mu\rho} g^{\nu\sigma} - g^{\nu\rho} g^{\mu\sigma}) \right. \\ &\quad + f_{ace} f_{bde} (g^{\mu\nu} g^{\rho\sigma} - g^{\nu\rho} g^{\mu\sigma}) \\ &\quad \left. + f_{ade} f_{cbe} (g^{\mu\rho} g^{\nu\sigma} - g^{\mu\nu} g^{\rho\sigma}) \right]. \end{aligned} \quad (2.1.28)$$

QCD action. Putting everything together, the resulting QCD action constructed from the fields ψ , $\bar{\psi}$ and A_a^μ has the most general form that is invariant under Poincaré transformations, invariant under local gauge transformations, and renormalizable:

$$S_{QCD} = \int d^4x \mathcal{L}_{QCD}, \quad \mathcal{L}_{QCD} = \bar{\psi}(x) (i\not{D} - M) \psi(x) - \frac{1}{4} F_{\mu\nu}^a F_a^{\mu\nu}. \quad (2.1.29)$$

The summation over the Dirac, color and flavor indices of the quarks is again implicit, and we generalized the quark mass m to a quark mass matrix $M = \text{diag}(m_1 \dots m_{N_f})$. Some further remarks:

■ Eq. (2.1.29) also conserves **charge conjugation** and **parity**, where the charge conjugation operation is defined by

$$\psi'_\alpha = \bar{\psi}_\beta C_{\beta\alpha}, \quad \bar{\psi}'_\alpha = C_{\alpha\beta} \psi_\beta, \quad C = i\gamma^2 \gamma^0 \quad (2.1.30)$$

and the parity transformation by

$$\psi'(x') = \gamma^0 \psi(x), \quad \bar{\psi}'(x') = \bar{\psi}(x) \gamma^0, \quad x' = (t, -\mathbf{x}). \quad (2.1.31)$$

Since CPT is always conserved, this implies that the QCD action is also invariant under time reversal.

■ In principle, another gauge-invariant and renormalizable (but parity-violating) term could appear in the Lagrangian, namely a topological charge density:

$$\mathcal{Q}(x) = \frac{g^2}{8\pi^2} \text{Tr} \left\{ F_{\mu\nu} \tilde{F}^{\mu\nu} \right\} \quad \text{with} \quad \tilde{F}^{\mu\nu} = \frac{1}{2} \varepsilon^{\mu\nu\alpha\beta} F_{\alpha\beta}, \quad (2.1.32)$$

where $\tilde{F}^{\mu\nu}$ is the dual field strength tensor. The resulting ‘ θ term’ in the Lagrangian $\mathcal{L} + \theta \mathcal{Q}(x)$ violates parity and would give rise to an electric dipole moment of the neutron, whose experimental upper limit is however tiny ($\theta \leq 10^{-10}$). So it would seem that QCD *does* conserve parity; unfortunately, even if we started with $\theta = 0$ in QCD, the CP -violating weak interactions would renormalize it to $\theta \neq 0$. There are several possible scenarios how $\theta = 0$ could be enforced beyond the Standard Model, e.g. by promoting θ to a field (axions). Then again, CP *must* have been violated in the early universe, because otherwise the Big Bang would have created matter and antimatter in equal portions, which would have annihilated and resulted in a radiation universe without matter. This leads to the **strong CP problem**. On the other hand, since $\mathcal{Q} = \partial_\mu K^\mu$ can be written as the divergence of the Chern-Simons current K^μ , it only contributes a surface term to the action and in principle we could discard it (unless topological gauge field configurations play a role).

■ We could have defined the gluon fields so that they absorb the coupling constant g (i.e., by replacing $A \rightarrow A/g$ and $F \rightarrow F/g$). From Eqs. (2.1.13), (2.1.18) and (2.1.29) we see that the only place in the Lagrangian where the coupling would then appear is in front of the gluon kinetic term, as a prefactor $1/g^2$. This shows that the sign of g is physically irrelevant.

Quark masses and flavor structure. With regard to the flavor structure, we can simply ignore the gluons since they are flavor independent. The quark-gluon interaction is flavor-blind, and the distinction between different quarks only comes from their masses. If the masses of all quark flavors were equal, the Lagrangian would have an additional $SU(N_f)$ flavor symmetry. This is not realized in nature, where

$$m_u \sim m_d \sim 2 \dots 6 \text{ MeV}, \quad m_s \sim 100 \text{ MeV}, \quad \begin{array}{l} m_c \sim 1.3 \text{ GeV}, \\ m_b \sim 4.2 \text{ GeV}, \\ m_t \sim 173 \text{ GeV}. \end{array} \quad (2.1.33)$$

These **current-quark masses** have their origin in the Higgs sector and from the point of view of QCD they are external parameters that enter through the quark mass matrix $M = \text{diag}(m_1 \dots m_{N_f})$. Because M is diagonal in flavor space, the flavor pieces in the Lagrangian simply add up: $\bar{\psi} M \psi = \sum_f m_f \bar{\psi}_f \psi_f$. The flavor structure of the Lagrangian is crucial for the properties of hadrons and we will return to it in Chapter 3.

Infinitesimal gauge transformations. For later convenience it is useful to work out the infinitesimal transformations of the fields. The covariant derivative as defined in Eq. (2.1.3) acts on fields that transform under the fundamental representations of $SU(3)_C$, i.e., the group elements. When acting on elements of the algebra (those containing the matrix generators \mathfrak{t}_a , for example ε , A_μ or $F_{\mu\nu}$), we need an additional commutator in its definition: $D_\mu = \partial_\mu - ig[A_\mu, \cdot]$, or written in components:

$$\begin{aligned} (D_\mu \varepsilon)^a &= (\partial_\mu \varepsilon - ig[A_\mu, \varepsilon])^a = \partial_\mu \varepsilon^a - ig A_\mu^c \varepsilon^b i f_{cba} \\ &= (\partial_\mu \delta_{ab} - g f_{abc} A_\mu^c) \varepsilon^b = D_\mu^{ab} \varepsilon^b. \end{aligned} \quad (2.1.34)$$

In the fundamental representation, the group generators are the Gell-Mann matrices; in the adjoint representation they are given by $(\mathfrak{t}_c)_{ab} = -i f_{abc}$. Inserting this into Eq. (2.1.3), we see that D_μ^{ab} is the covariant derivative in the adjoint representation:

$$(D_\mu)_{ab} = (\partial_\mu - ig A_\mu)_{ab} = \partial_\mu \delta_{ab} - ig A_\mu^c (\mathfrak{t}_c)_{ab} = \partial_\mu \delta_{ab} - g f_{abc} A_\mu^c. \quad (2.1.35)$$

In an Abelian gauge theory such as QED, the commutator vanishes and $D_\mu^{ab} = \partial_\mu \delta_{ab}$ is the ordinary partial derivative.

With $U = 1 + i\varepsilon$, the infinitesimal gauge transformation of the fields is given by

$$\begin{aligned} \psi' &= U \psi \approx (1 + i\varepsilon) \psi, \\ \bar{\psi}' &= \bar{\psi} U^\dagger \approx \bar{\psi} (1 - i\varepsilon), \\ A'_\mu &= U A_\mu U^\dagger + \frac{i}{g} U (\partial_\mu U^\dagger) \approx A_\mu + i [\varepsilon, A_\mu] + \frac{1}{g} \partial_\mu \varepsilon = A_\mu + \frac{1}{g} D_\mu \varepsilon, \\ F'_{\mu\nu} &= U F_{\mu\nu} U^\dagger \approx F_{\mu\nu} + i [\varepsilon, F_{\mu\nu}], \end{aligned} \quad (2.1.36)$$

from where we obtain:

$$\delta\psi = i\varepsilon\psi, \quad \delta\bar{\psi} = -i\bar{\psi}\varepsilon, \quad \delta A_\mu = \frac{1}{g} D_\mu \varepsilon, \quad \delta F_{\mu\nu} = i [\varepsilon, F_{\mu\nu}]. \quad (2.1.37)$$

Classical equations of motion. The classical Euler-Lagrange equations of motion follow from the action principle:

$$\begin{aligned} S[\phi] = \int d^4x \mathcal{L}(\phi, \partial_\mu \phi) &\Rightarrow \delta S[\phi] = \int d^4x \left(\frac{\partial \mathcal{L}}{\partial \phi} \delta \phi(x) + \frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi)} \delta (\partial_\mu \phi) \right) \\ &= \int d^4x \left(\frac{\partial \mathcal{L}}{\partial \phi} - \partial_\mu \frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi)} \right) \delta \phi(x) = 0, \end{aligned}$$

which means that the functional derivative of the action vanishes:

$$\frac{\delta S[\phi]}{\delta \phi(x)} = \frac{\partial \mathcal{L}}{\partial \phi} - \partial_\mu \frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi)} = 0. \quad (2.1.38)$$

If the action contains several fields, there is one equation of motion for each component: $\delta S[\phi_1, \dots, \phi_n]/\delta \phi_i(x) = 0$.

As a reminder, the **functional derivative** $\delta F[\phi]/\delta \phi(x)$ of a functional $F[\phi]$ is defined as

$$F[\phi + \delta \phi] = F[\phi] + \delta F[\phi] = F[\phi] + \int_{-\infty}^{\infty} dx \frac{\delta F[\phi]}{\delta \phi(x)} \delta \phi(x), \quad (2.1.39)$$

where the last term is the continuum version of $\sum_i (\delta F/\delta \phi_i) \delta \phi_i$ written for one spacetime dimension. Here are some examples:

$F[\phi]$	$F[\phi + \delta \phi]$	$\frac{\delta F[\phi]}{\delta \phi(x)}$
$\int dx \phi(x) J(x)$	$\int dx (\phi + \delta \phi) J = F[\phi] + \int dx J \delta \phi$	$J(x)$
$\int dx f(\phi(x)) J(x)$	$\int dx [f(\phi) + f'(\phi) \delta \phi] J$	$f'(\phi(x)) J(x)$
$\int dx f(\phi(x), \phi'(x))$	$\int dx \left[f(\phi, \phi') + \frac{\partial f}{\partial \phi} \delta \phi + \frac{\partial f}{\partial \phi'} \delta \phi' \right]$ $= F[\phi] + \int dx \left[\frac{\partial f}{\partial \phi} - \frac{d}{dx} \frac{\partial f}{\partial \phi'} \right] \delta \phi$	$\frac{\partial f}{\partial \phi(x)} - \frac{d}{dx} \frac{\partial f}{\partial \phi'(x)}$
$\int_0^\infty dx \phi(x) = \int_{-\infty}^\infty dx \phi(x) \Theta(x)$	$F[\phi] + \int dx \Theta(x) \delta \phi(x)$	$\Theta(x)$
$\exp \left[i \int dx \phi(x) J(x) \right]$	$\exp \left[i \int dx (\phi(x) + \delta \phi(x)) J(x) \right]$ $= F[\phi] (1 + i \int dx J(x) \delta \phi(x))$	$iJ(x) \exp \left[i \int dy \phi(y) J(y) \right]$
$\phi(z) = \int dx \phi(x) \delta(x - z)$	$F[\phi] + \int dx \delta \phi(x) \delta(x - z)$	$\delta(x - z)$
$f(\phi(z))$	$f(\phi(z)) + f'(\phi(z)) \delta \phi(z)$ $= F[\phi] + \int dx f'(\phi(x)) \delta \phi(x) \delta(x - z)$	$f'(\phi(x)) \delta(x - z)$
$\phi'(z) = \int dx \phi'(x) \delta(x - z)$ $= - \int dx \phi(x) \delta'(x - z)$	$F[\phi] - \int dx \delta \phi(x) \delta'(x - z)$	$-\delta'(x - z)$

Let us work out the classical equations of motion of QCD defined by the Lagrangian $\mathcal{L} = \bar{\psi} (i\not{D} + g\not{A} - M) \psi - \frac{1}{4} F_{\mu\nu}^a F_a^{\mu\nu}$. Although they are not directly relevant for our purposes, they will later enter in the quantum equations of motion (Sec. 2.2.2) and conservation laws (Sec. 3.1). If we take the derivatives of \mathcal{L} with respect to ψ and $\bar{\psi}$,

$$\frac{\partial \mathcal{L}}{\partial \bar{\psi}} = (i\not{D} - M) \psi, \quad \frac{\partial \mathcal{L}}{\partial(\partial_\mu \bar{\psi})} = 0, \quad \frac{\partial \mathcal{L}}{\partial \psi} = \bar{\psi} (g\not{A} - M), \quad \frac{\partial \mathcal{L}}{\partial(\partial_\mu \psi)} = \bar{\psi} i\gamma^\mu \quad (2.1.40)$$

we obtain the **Dirac equations** for the quark and antiquark fields:

$$\frac{\delta S}{\delta \bar{\psi}} = (i\not{D} - M) \psi = 0, \quad \frac{\delta S}{\delta \psi} = \bar{\psi} (-i\overleftarrow{\not{D}} + g\not{A} - M) = 0. \quad (2.1.41)$$

For the gluons, we first work out the derivatives of the field-strength tensor:

$$\frac{\partial F_{\mu\nu}^a}{\partial A_\rho^c} = gf_{abc} (A_\mu^b \delta_\nu^\rho - A_\nu^b \delta_\mu^\rho), \quad \frac{\partial F_{\mu\nu}^a}{\partial(\partial_\rho A_\sigma^c)} = (\delta_\mu^\rho \delta_\nu^\sigma - \delta_\nu^\rho \delta_\mu^\sigma) \delta_{ac}. \quad (2.1.42)$$

With the product rule we then obtain

$$\frac{\partial \mathcal{L}}{\partial A_\mu^a} = g \bar{\psi} \gamma^\mu \mathbf{t}_a \psi + gf_{abc} A_\nu^c F_b^{\mu\nu}, \quad \frac{\partial \mathcal{L}}{\partial(\partial_\nu A_\mu^a)} = F_a^{\mu\nu} \quad (2.1.43)$$

and finally

$$\frac{\delta S}{\delta A_\nu^a} = gf_{abc} A_\nu^c F_b^{\mu\nu} - \partial_\nu F_a^{\mu\nu} + g \bar{\psi} \gamma^\mu \mathbf{t}_a \psi = 0. \quad (2.1.44)$$

The first two terms on the r.h.s. can be combined to

$$-(\partial_\nu \delta_{ab} - gf_{abc} A_\nu^c) F_b^{\mu\nu} = -D_\nu^{ab} F_b^{\mu\nu} = -(D_\nu F^{\mu\nu})^a, \quad (2.1.45)$$

whereas the last term is the vector current corresponding to the global $SU(3)_C$ transformation: $J_a^\mu = \bar{\psi} \gamma^\mu \mathbf{t}_a \psi$. Then the quantity $J^\mu = \sum_a J_a^\mu \mathbf{t}_a$ lives in the Lie algebra and Eq. (2.1.44) becomes

$$D_\nu F^{\mu\nu} = g J^\mu. \quad (2.1.46)$$

These are the **classical Yang-Mills equations** for the gluon field, i.e., the Maxwell equations in the non-Abelian theory. They are the direct generalization from electrodynamics, where the covariant derivative in the adjoint representation would reduce to the ordinary derivative.

It is not too much of a stretch to ask whether there is also a generalization of the Maxwell equation for the dual field strength tensor $\tilde{F}^{\mu\nu}$. Indeed we find

$$\begin{aligned} D_\nu \tilde{F}^{\mu\nu} &= \frac{1}{2} \varepsilon^{\mu\nu\alpha\beta} D_\nu F_{\alpha\beta} \\ &= \frac{1}{6} \left(\varepsilon^{\mu\nu\alpha\beta} + \varepsilon^{\mu\alpha\beta\nu} + \varepsilon^{\mu\beta\nu\alpha} \right) D_\nu F_{\alpha\beta} \\ &= \frac{1}{6} \varepsilon^{\mu\nu\alpha\beta} (D_\nu F_{\alpha\beta} + D_\alpha F_{\beta\nu} + D_\beta F_{\nu\alpha}) = 0, \end{aligned} \quad (2.1.47)$$

where the parenthesis vanishes due to the **Bianchi identity**, which follows from the Jacobi identity for the generators, Eq. (A.1.3). Similarly, one can establish covariant current conservation $D_\mu J^\mu = 0$ for the solutions of the equations of motion.

2.2 Quantization of QCD

So far we have discussed the Lagrangian and action of QCD, Eq. (2.1.29), which define the classical field theory. What are the consequences of going to the quantum field theory (QFT)? There are two standard methods to transform the classical action into a QFT. One is the canonical formalism, where the fields are treated as operators on a Fock space and canonical (anti-)commutation relations are imposed. The other is the path-integral formalism where an integral over all fields is performed, which provides an intuitive picture of how quantum corrections augment the classical field configurations. Both methods are equivalent and we will use them in combination, depending on what better suits our needs.

In the following we briefly recall some basic concepts of QFT. For illustration, we work in a generic theory with one species of fields, $\phi(x)$, defined by the classical action $S[\phi] = \int d^4x \mathcal{L}(\phi, \partial_\mu \phi)$; in QCD, ϕ would stand for the fields $\phi \in \{\psi, \bar{\psi}, A_a^\mu\}$.

2.2.1 Canonical quantization

Assuming that you have heard a standard QFT course and went through all the machinery, how would you summarize the basic ideas in a few words? Fortunately, the formalism behind a QFT requires only a small number of axioms:

Relativity and unitarity. In a QFT, the field $\phi(x)$ is interpreted as an operator on a Fock space with elements $|p_1 \dots p_n\rangle$, $n = 0, 1, 2, \dots$. This includes the vacuum $|0\rangle$, single-particle states $|p\rangle$ with four-momentum p , as well as multi-particle states. These states transform under *unitary* representations $U(\Lambda, a)$ of the Poincaré group, which provide a probability interpretation for S -matrix elements $\langle p_1 \dots p_n | q_1 \dots q_r \rangle$.

More details on the Poincaré group can be found in Appendix B. The group consists of translations (with group element a^μ), rotations and boosts (which define the Lorentz group with group element Λ). Because of the boosts, the Lorentz group is not compact and therefore its finite-dimensional representations are not unitary. However, the classical fields transform just under such finite-dimensional representations: a scalar field $\phi'(x') = \phi(x)$ is invariant, a Dirac field $\psi(x)$ transforms under a four-dimensional spinor representation $\psi'(x') = D(\Lambda)\psi(x)$, a vector field $A^\mu(x)$ under a four-dimensional vector representation $A'^\mu(x') = \Lambda^{\mu\nu}A_\nu(x)$, etc. Because these representations are not unitary, there is no hope for extracting probability amplitudes from the classical fields. *Unitary* representations are *infinite-dimensional*, which is why we need an infinite-dimensional Fock space. In other words, implementing relativity (which enters through the boosts) *while* maintaining unitarity takes us from quantum mechanics to QFT, where we interpret the fields as operators on the Fock space — instead of fields like in the classical field theory or wave functions like in quantum mechanics.

Causality. The second basic requirement is that two measurements with a spacelike distance cannot affect each other:

$$[\mathcal{O}_1(x), \mathcal{O}_2(y)] = 0 \quad \text{for} \quad (x - y)^2 < 0. \quad (2.2.1)$$

Here, \mathcal{O}_1 and \mathcal{O}_2 are observables such as quark bilinears $\bar{\psi} \Gamma \psi$, where Γ is some Dirac matrix. For bosonic fields, this requirement is implemented by imposing spacelike commutation relations $[\phi(x), \phi(y)] = 0$ for $(x - y)^2 < 0$, whereas for fermionic fields we need anticommutation relations $\{\psi_\alpha(x), \bar{\psi}_\beta(y)\} = 0$. The equal-time commutation or anticommutation relations are special cases of those.

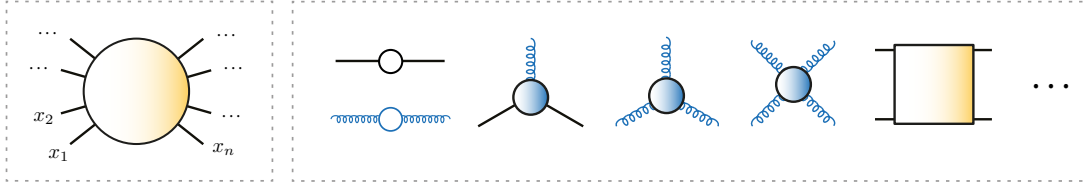


FIG. 2.2: Generic correlation function (left) and some of the elementary correlation functions in QCD (right)

These are some of the pillars on which a QFT is built and they summarize the first part of a typical QFT course: one develops the formalism for free scalar, spinor and vector fields, based on the representations of the Lorentz group, and studies the implementation of symmetries through the Noether theorem. Below we will discuss another pillar, the spectral condition, and depending on what we are after we could also add renormalizability or (for gauge theories) local gauge invariance to the list.

Interacting QFT. Unfortunately, after switching on interactions, one quickly runs into trouble at the operator level (cf. Haag’s theorem: a free field remains always free). To avoid these problems, the hard statements that can be made in an interacting QFT are those for *matrix elements* of operators, which therefore become the central objects of interest. The **correlation functions**, also called n -point functions or Green functions, are the vacuum expectation values of time-ordered products of fields:

$$G(x_1, \dots, x_n) := \langle 0 | T \phi(x_1) \dots \phi(x_n) | 0 \rangle. \quad (2.2.2)$$

Pictorially, these are blobs with n legs, one for each spacetime point $x_1 \dots x_n$, which contain all possible interactions between the particles that can happen inside (Fig. 2.2). The simplest example is a two-point function, a **propagator**, which contains the (self-) interactions of a single particle and which for free fields becomes the usual Feynman propagator. For theories with different types of fields there can be interactions between different particles, and Fig. 2.2 shows some of the correlation functions in QCD: the quark and gluon propagators and some of their three- and four-point functions.

Why are these correlation functions relevant? For one, the LSZ formula tells us that they are related to S -matrix elements:

$$\begin{aligned} & \text{FT} [G(x_1 \dots x_n, y_1 \dots y_r)] \\ &= \left(\prod_{i=1}^n \frac{i\sqrt{Z}}{p_i^2 - m_i^2 + i\epsilon} \right) \left(\prod_{j=1}^r \frac{i\sqrt{Z}}{q_j^2 - m_j^2 + i\epsilon} \right) \langle p_1 \dots p_n, \text{out} | q_1 \dots q_r, \text{in} \rangle_{\text{conn.}}, \end{aligned} \quad (2.2.3)$$

That is, we can extract the connected S -matrix element for an $n \rightarrow r$ scattering process (the invariant amplitude \mathcal{M}) if we calculate the respective correlation function in momentum space (‘FT’ stands for Fourier transform), go to the kinematic limit where the external propagators are onshell, and amputate those external propagators. If the particles carry spinor or vector quantum numbers, we must also contract with onshell spinors or polarization vectors. Once we know \mathcal{M} , we can compute the cross section from $|\mathcal{M}|^2$ and compare it to experiment.

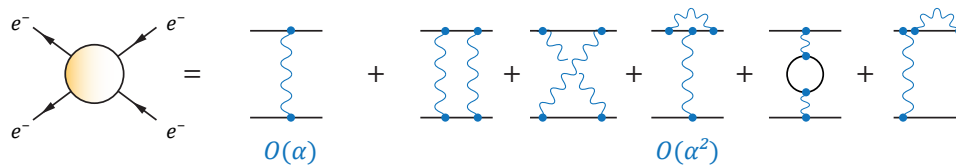


FIG. 2.3: Lowest-order perturbative diagrams that contribute to Møller scattering in QED. Topologies like those in the last diagram do not survive the amputation.

This is the standard recipe for a typical QFT calculation. The practical problem is of course that we first need to know *how* to calculate the correlation functions in question. If the coupling constant of the theory is small at the momentum scales of interest, we can use **perturbation theory** and expand them into Feynman diagrams, where we neglect diagrams above a certain loop order that are suppressed by the small coupling constant in front. (The other practical issue is renormalization but this is well understood, see Sec. 2.3.)

This recipe has turned out to be extremely successful, with the prime example given by **QED**. In that case, the dimensionless coupling $\alpha_{\text{QED}} = e^2/(4\pi)$ is indeed small; inserting dimensions, we have¹

$$\alpha_{\text{QED}} = \frac{e^2}{4\pi} \frac{1}{\hbar c \epsilon_0} \approx \frac{1}{137}. \quad (2.2.4)$$

If we perform a loop expansion for a given correlation function, then higher loop diagrams come with higher powers of the coupling and in practice it may even be sufficient to stick with the lowest (tree-level) order. For example, the electron four-point function in Fig. 2.3 describes both the Møller ($e^-e^- \rightarrow e^-e^-$) and Bhabha ($e^+e^- \rightarrow e^+e^-$) scattering processes. The leading contribution to Møller scattering is the one-photon exchange diagram, which leads to the Mott cross section plus spin terms (more on that later). The smallness of α_{QED} has contributed to the successes of QED, where many observables can be calculated quite precisely by going to higher orders in perturbation theory. This has led to a variety of precision measurements of α_{QED} , from the anomalous magnetic moment ($g-2$) of electrons and muons, measurements of the Rydberg constant, the energy level splittings in atoms, etc.

Unfortunately, when we try to apply the same principles to QCD we are confronted with two challenges that complicate matters enormously. The first difficulty is that the coupling α_{QCD} becomes large at low momenta (see Section 2.3) and invalidates a perturbative expansion. Unfortunately this is just the region that is relevant for hadron physics, so we must look for **nonperturbative methods** to calculate these correlation functions at low momenta. The second difficulty is more fundamental: confinement entails that it is pointless to calculate invariant scattering amplitudes of quarks and gluons because we can never measure such processes. What we can measure are reactions between hadrons (e.g. NN or $N\pi$ scattering), or hadrons that interact with leptons through the electroweak interaction (e^+e^- annihilation, eN scattering etc.). We will return to this point in Section 3.1.3, where we show how one can still extract measurable information from QCD's elementary correlation functions.

¹In natural units $\hbar = 1$ sets the units of action [ML^2/T], $c = 1$ sets the units of velocity [L/T] and $\hbar c \epsilon_0 = 1$ the units of charge [C^2].

In any case, what still stands is that the correlation functions in Eq. (2.2.2) encode the full content of the QFT, namely through the quantum effective action (more on this in Sec. 2.2.2). Thus, if we knew all of them — and there are infinitely many — it would be equivalent to having *solved* the QFT. For this reason, QCD's correlation functions will be the central quantities of interest throughout this course.

Spectral representation. Returning to the remaining basic property of QFT that we passed over before, let us assume that each Fock state, which we generically denote by $|\lambda\rangle$, is an eigenstate of the Hamiltonian with definite energy and momentum. The vacuum $|0\rangle$ has vanishing energy and momentum. A one-particle state $|\mathbf{p}\rangle$ has momentum \mathbf{p} and energy $E_p = \sqrt{\mathbf{p}^2 + m^2}$. Multiparticle states, which are characterized by a center-of-mass momentum \mathbf{p} and relative momenta among the particles, form a continuum: For example, the lowest possible energy of a two-particle state in its rest frame ($\mathbf{p} = 0$) is $2m$, but since the two particles can have relative momentum, which contributes to their total energy, the state can have *any* energy $E_p(\lambda) = (\mathbf{p}^2 + m_\lambda^2)^{1/2}$, where $m_\lambda \geq 2m$ is the invariant mass of the state (its energy in the rest frame). The eigenvalue spectrum of H then has the form shown in Fig. 2.4: a one-particle state sits on its mass shell, a two-particle state forms a continuum with $m_\lambda \geq 2m$ and so on.

Based on this, we write the completeness relation of the Fock space as

$$\mathbb{1} = \sum_\lambda \frac{1}{(2\pi)^3} \int d^4p \Theta(p^0) \delta(p^2 - m_\lambda^2) |\lambda\rangle\langle\lambda| = \sum_\lambda \frac{1}{(2\pi)^3} \int \frac{d^3p}{2E_p} |\lambda\rangle\langle\lambda|, \quad (2.2.5)$$

where the sum over λ is formal and includes integrals over relative momenta. The Lorentz-invariant integral measure ensures that we only integrate over $p^2 = m_\lambda^2 \Leftrightarrow p_0^2 = E_p^2$, and we used

$$\delta(p^2 - m_\lambda^2) = \delta(p_0^2 - E_p^2) = \frac{\delta(p_0 - E_p) + \delta(p_0 + E_p)}{2E_p}, \quad (2.2.6)$$

where only the first term survives because the step function enforces $p_0 > 0$.

It is easy to show that if we insert the completeness relation into the propagator $G(x, y) = \langle 0 | T \phi(x) \phi(y) | 0 \rangle$, we arrive at the **spectral representation** for the propagator in momentum space:

$$G(p^2) = \sum_\lambda \frac{iR_\lambda}{p^2 - m_\lambda^2 + i\epsilon} = \int_0^\infty \frac{ds}{2\pi} \frac{i\rho(s)}{p^2 - s + i\epsilon}. \quad (2.2.7)$$

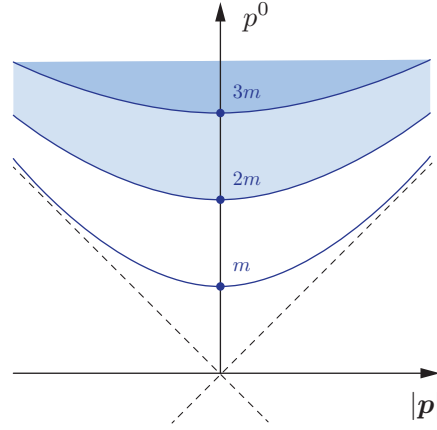


FIG. 2.4: Eigenvalue spectrum of the Hamiltonian in terms of one-particle states with mass m and multiparticle states with invariant mass $m_\lambda \geq 2m$.

Here we defined the **spectral function**

$$\rho(s) = \sum_{\lambda} 2\pi\delta(s - m_{\lambda}^2) R_{\lambda}, \quad (2.2.8)$$

where the sum over λ is again formal and denotes relative-momentum integrations.

The proof goes as follows: By inserting the completeness relation, we obtain

$$\begin{aligned} G(x, y) &= \langle 0 | \mathbb{T} \phi(x) \phi(y) | 0 \rangle \\ &= \Theta(x^0 - y^0) \langle 0 | \phi(x) \phi(y) | 0 \rangle + \Theta(y^0 - x^0) \langle 0 | \phi(y) \phi(x) | 0 \rangle \\ &= \sum_{\lambda} \frac{1}{(2\pi)^3} \int \frac{d^3 p}{2E_p} \left\{ \begin{array}{l} \Theta(x^0 - y^0) \langle 0 | \phi(x) | \lambda \rangle \langle \lambda | \phi(y) | 0 \rangle \\ + \Theta(y^0 - x^0) \langle 0 | \phi(y) | \lambda \rangle \langle \lambda | \phi(x) | 0 \rangle \end{array} \right\}. \end{aligned} \quad (2.2.9)$$

Now we use the transformation properties of the operator $\phi(x)$ under Poincaré transformations $U(\Lambda, a)$ and in particular translations $U(1, a)$:

$$U(\Lambda, a) \phi(x) U(\Lambda, a)^{-1} = \phi(\Lambda x + a) \quad \Rightarrow \quad U(1, a) \phi(0) U(1, a)^{-1} = \phi(a). \quad (2.2.10)$$

On the other hand, applying a translation to a Fock state only produces a phase, whereas the vacuum remains invariant:

$$U(1, a) |\lambda(p)\rangle = e^{ip \cdot a} |\lambda(p)\rangle, \quad U(1, a) |0\rangle = |0\rangle. \quad (2.2.11)$$

In combination, we have

$$\begin{aligned} \langle 0 | \phi(x) | \lambda \rangle &= \langle 0 | U(1, x) \phi(0) U(1, x)^{-1} | \lambda \rangle = \langle 0 | \phi(0) | \lambda \rangle e^{-ip \cdot x}, \\ \langle \lambda | \phi(x) | 0 \rangle &= \langle 0 | \phi(0) | \lambda \rangle^* e^{ip \cdot x}. \end{aligned} \quad (2.2.12)$$

The remaining matrix element $\langle 0 | \phi(0) | \lambda(p) \rangle$ is Lorentz-invariant, so for a single-particle state it can only depend on $p^2 = m^2$ which is just a number (for multiparticle states it still depends on the relative momenta). Denoting $|\langle 0 | \phi(0) | \lambda \rangle|^2 = R_{\lambda}$ and $z = x - y$, we arrive at

$$G(z) = \sum_{\lambda} R_{\lambda} \int \frac{d^3 p}{2E_p} \left[\frac{\Theta(z^0) e^{-ip \cdot z} + \Theta(-z^0) e^{ip \cdot z}}{(2\pi)^3} \right]_{p^0 = E_p(\lambda)} = \sum_{\lambda} R_{\lambda} D_F(z, m_{\lambda}). \quad (2.2.13)$$

The integral is nothing but the Feynman propagator

$$D_F(z, m_{\lambda}) = \int \frac{d^4 p}{(2\pi)^4} e^{-ip \cdot z} \frac{i}{p^2 - m_{\lambda}^2 + i\epsilon}, \quad (2.2.14)$$

and we arrive at the result

$$G(z) = \int \frac{d^4 p}{(2\pi)^4} e^{-ip \cdot z} \sum_{\lambda} \frac{iR_{\lambda}}{p^2 - m_{\lambda}^2 + i\epsilon} \quad (2.2.15)$$

from where we can read off the propagator $G(p^2)$ in momentum space.

Note that Eq. (2.2.7) holds for the full ('dressed') propagator of the theory, i.e., knowledge of the spectral function is equivalent to the knowledge of the full two-point function. The spectral function is related to the analytic structure of the propagator. For a typical theory it is positive and has the form of Fig. 2.5: The one-particle states lead to an isolated δ -function peak at $s = m^2$, which defines the mass m of the particle in the full theory (which is not the mass parameter in the Lagrangian!) from the lowest-lying pole location of the propagator in momentum space. The continuum of n -particle states begins at $s \geq (2m)^2$, which leads to a branch cut in the propagator starting at $p^2 = 4m^2$.

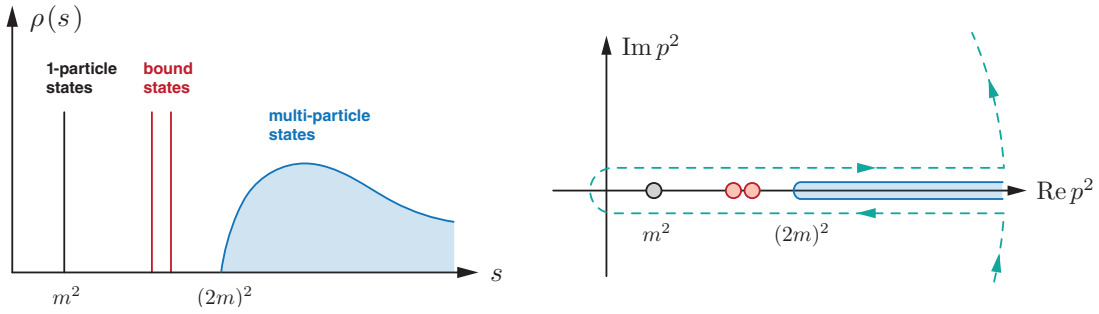


FIG. 2.5: Left: spectral function of a typical field theory, with a single-particle peak at $s = m^2$ and a multiparticle continuum for $s \geq 4m^2$, together with possible bound states. Right: Analytic structure of the corresponding propagator with a single-particle pole, bound-state poles and a branch cut above $p^2 = 4m^2$. The dashed line is the contour integration path that encloses the domain of analyticity.

To see this, define $G_E(w) = iG(w)$ with $w = u + iv \in \mathbb{C}$ and $u, v \in \mathbb{R}$. This function (which is the *Euclidean* propagator) satisfies the Schwartz reflection principle $G_E^*(w) = G_E(w^*)$ for analytic functions. We can use the Cauchy formula to determine $G_E(w)$ for w inside its domain of analyticity:

$$G_E(w) = \frac{1}{2\pi i} \oint ds \frac{G_E(s)}{s - w}. \quad (2.2.16)$$

If $G_E(w)$ has only singularities on the positive real axis, we can choose the integration path in Fig. 2.5, and if it falls off sufficiently fast for $|w| \rightarrow \infty$, what remains is the integral

$$\begin{aligned} G_E(w) &= \frac{1}{2\pi i} \lim_{\epsilon \rightarrow 0} \int_0^\infty ds \left[\frac{G_E(s + i\epsilon)}{s - w + i\epsilon} - \frac{G_E(s - i\epsilon)}{s - w - i\epsilon} \right] \\ &= \frac{1}{2\pi i} \lim_{\epsilon \rightarrow 0} \int_0^\infty ds \left[\frac{G_E(s + i\epsilon)}{s - u - i(v - \epsilon)} - \frac{G_E^*(s + i\epsilon)}{s - u - i(v + \epsilon)} \right]. \end{aligned} \quad (2.2.17)$$

For $|v| > \epsilon$, w always lies in the domain of analyticity and we can take the limit $\epsilon \rightarrow 0$ to obtain

$$G_E(w) = \frac{1}{\pi} \int_0^\infty ds \frac{\text{Im } G_E(s)}{s - u - iv}. \quad (2.2.18)$$

This is the same formula as Eq. (2.2.7) if we identify $u = p^2$ and $v = \epsilon$, and therefore

$$\rho(s) = 2 \text{Im } G_E(s). \quad (2.2.19)$$

Thus, if the spectral function is non-zero for $s > 4m^2$, the imaginary part of the propagator is discontinuous above that threshold — it has a branch cut.

In this way, single-particle asymptotic states produce poles in the propagator and multi-particle states produce cuts. In fact, every new multi-particle production threshold opens up another branch cut and thus another Riemann sheet, so the propagator becomes a multi-valued function. Vice versa, the existence of a spectral representation implies that the propagator has only singularities on the timelike axis $p^2 > 0$ but not in the complex plane of the first Riemann sheet: had there been singularities inside the integration contour in Eq. (2.2.16), their residues would produce further terms in (2.2.18) and the identification with Eq. (2.2.7) would no longer go through.

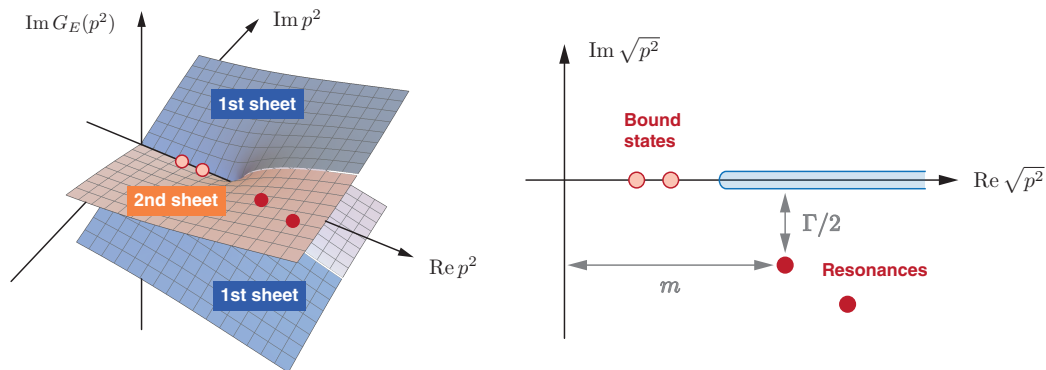


FIG. 2.6: Bound states have poles on the positive real axis below the threshold, whereas resonances have poles on the second Riemann sheet. (The left figure is only meant for illustration, since what is shown is really just the square-root function.)

At this point there is no consensus to what extent a spectral representation can be formulated for gauge theories and/or theories with confinement. For the scalar example the situation is clear: the operator $\phi(x)$ creates single- or multiparticle states $|\lambda\rangle$ with definite mass m , momentum \mathbf{p} and energy E_p . But what about QCD? Certainly, we can count pions $|\pi\rangle$, nucleons $|N\rangle$ and other stable states as asymptotic states in QCD — but these are all *bound states*. The corresponding multiparticle states are $|\pi\pi\rangle$, $|\pi\pi\pi\rangle$, $|NN\rangle$, $|N\bar{N}\rangle$ etc. Indeed, the sensible matrix elements in QCD are all of the form

$$\langle 0|\dots|0\rangle, \quad \langle 0|\dots|\pi\rangle, \quad \langle N|\dots|N\rangle, \quad \langle 0|\dots|\pi\pi\rangle, \quad \text{etc.}, \quad (2.2.20)$$

where the dots denote time-ordered products of quark and gluon field operators, and the initial and final states contain either the vacuum or stable particles. But what about quarks and gluons? What is the ‘mass’ or ‘energy’ of a quark or gluon if we cannot measure it? Should we even count them as Fock states in the sense of Eq. (2.2.5)? If we didn’t, we would break the link between the ‘masses’ of quarks and gluons and the poles in their propagators, and there would be no reason why QCD’s elementary correlation functions *should* have singularities on the positive real axis only — they could also lie in the complex plane, as long as their effects cancel out in observable scattering amplitudes like those in Eq. (2.2.20). (It is often said that *causality* alone restricts the singularities to the positive real axis, but the proof of this statement also assumes a spectral condition.) It may still be possible to formulate *generalized* spectral representations also for quarks and gluons, which restrict their singularities again to the positive real axis. In any case, in practice from now on we assume that $|\lambda\rangle$ only refers to asymptotic states with a well-defined mass and energy, which in QCD are bound states.

Resonances. There is another way how poles can move away from the real axis, even in the presence of a spectral representation: A bound-state pole can pass a multiparticle production threshold and become a resonance, i.e., an unstable state. By doing so, it acquires a width, which means an imaginary part. This does not contradict our earlier statements, since the spectral representation still implies that the first (‘physical’) sheet must be free of singularities. As a consequence, resonances can only appear as poles on ‘unphysical’ higher Riemann sheets.

The situation is sketched in Fig. 2.6. The resonance pole location is usually written as $p^2 = (m - i\Gamma/2)^2$, so that in the complex $\sqrt{p^2}$ plane the real part is quoted as the mass of the resonance and the imaginary part as half the width. The width is related to the inverse lifetime of the particle: a bound state has an infinite lifetime whereas very short-lived resonances have poles far in the complex plane.



FIG. 2.7: Integration paths according to the $i\epsilon$ prescription.

Where does $i\epsilon$ come from? As a final remark, what is the origin of the ‘ $i\epsilon$ prescription’ that shows up in formulas like (2.2.3) and (2.2.14)? You probably first encountered it when taking the Fourier transform of the propagator $i/(p^2 - m^2)$ in the free field theory. After splitting the d^4p integral into d^3p and dp_0 , this function has poles at $p_0 = \pm E_p = \pm\sqrt{\mathbf{p}^2 + m^2}$, which depend on $|\mathbf{p}|$ as indicated in the left of Fig. 2.7. To reproduce the Feynman propagator $\langle 0|\mathbb{T}\phi(x)\phi(y)|0\rangle$, one must integrate slightly below the p_0 axis for $\text{Re } p_0 < 0$ and slightly above for $\text{Re } p_0 > 0$.

In general, the $i\epsilon$ prescription follows from the imaginary-time boundary conditions when projecting correlation functions onto the interacting vacuum $|0\rangle$:

$$\sum_{n=0}^{\infty} e^{-iE_n T} |n_f\rangle \langle n_f|0\rangle \xrightarrow{T \rightarrow \infty(1-i\epsilon)} e^{-iE_0 T} |0_f\rangle \langle 0_f|0\rangle, \quad (2.2.21)$$

which removes the higher-energy contributions of the free-particle states $|n_f\rangle$. This is equivalent to imposing boundary conditions for every d^4x integral (such as the one in the action of the theory) and every d^4p integral,

$$\int d^4x = \int d^3x \int_{-\infty(1-i\epsilon)}^{\infty(1-i\epsilon)} dx_0 \quad \Leftrightarrow \quad \int d^4p = \int d^3p \int_{-\infty(1+i\epsilon)}^{\infty(1+i\epsilon)} dp_0, \quad (2.2.22)$$

or, alternatively, shifting the propagator poles in the p_0 variable by $i\epsilon$.

In practice it is convenient to employ a **Euclidean metric** by defining $x_4 = ix_0$ with $x_4 \in \mathbb{R}$ and formulate the theory directly in Euclidean spacetime. In that case boundary conditions become irrelevant and the weight factor $e^{-S_E[\phi]}$ in the path integral defines a probability measure. This is what is usually done in practical calculations using path integrals (here we will stick to the Minkowski metric, though). For momentum-space integrals it corresponds to a **Wick rotation**, where the integration path in Fig. 2.7 does not proceed from left to right but from bottom to top.

Strictly speaking, for any finite ϵ , Eq. (2.2.22) implies to integrate not *slightly* below and above the p_0 axis, but to start at the bottom left corner and integrate to the top right. This does not make any difference for the Fourier transform of the propagator, but it is useful to remember when calculating loop integrals, whose pole locations can line up like in the right of Fig. 2.7 for certain kinematical situations. If one integrates over d^3p first, the poles produce branch cuts in the complex p_0 plane, and to avoid them one has to deform the integration contour accordingly — both with or without a Wick rotation. If one integrates over p_0 first, one must pick the correct residues. If this is done properly, results in Minkowski and Euclidean space are the same.

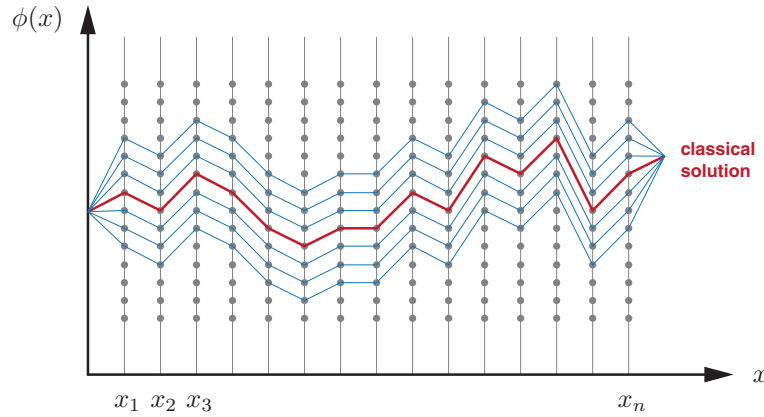


FIG. 2.8: Visualization of a path integral. The integration over all values of $\phi(x_i)$ at each spacetime point x_i is equivalent to an integration over all possible field configuration $\phi(x)$.

2.2.2 Path-integral quantization

Path integrals in QFT. Another way to quantize a field theory is the path-integral formalism. The central object is the path integral

$$\int \mathcal{D}\phi e^{iS[\phi]}(\dots) = \lim_{n \rightarrow \infty} \int d\phi(x_1) \dots d\phi(x_n) e^{iS[\phi]}(\dots), \quad (2.2.23)$$

which can be pictorially understood as in Fig. 2.8: at each spacetime point x_i we integrate over all possible values $\phi(x_i)$, which amounts to an integration over all *paths* in the field space, i.e., all possible field configurations. If we had different types of fields with additional group representation labels or Lorentz-Dirac indices, the product would go over all of them as well. With $e^{iS[\phi]}$ as probability measure (for this one needs the $i\epsilon$ prescription or go to Euclidean space), the integral picks up the quantum corrections to the classical path defined by the classical equations of motion $\delta S[\phi]/\delta\phi = 0$. As long as the system is sufficiently ‘classical’, the classical solutions will dominate the result; otherwise many trajectories will contribute.

One can show that the correlation functions (2.2.2) in the path-integral formalism are given by

$$G(x_1, \dots, x_n) = \langle 0 | \mathsf{T} \phi(x_1) \dots \phi(x_n) | 0 \rangle = \frac{\int \mathcal{D}\phi e^{iS[\phi]} \phi(x_1) \dots \phi(x_n)}{\int \mathcal{D}\phi e^{iS[\phi]}}. \quad (2.2.24)$$

Even though our notation does not distinguish it, one should remember that the $\phi(x_i)$ in the vacuum expectation value are *operators* acting on the state space, whereas the $\phi(x_i)$ in the path integral are merely *functions* but not operators. For fermionic fields, they are anticommuting Grassmann fields to implement the anticommutativity (see Sec. 2.2.3). To avoid the cumbersome distinction between operators and functions, and to reflect the statistical nature of the path integral as a quantum expectation value, it is common to denote the correlation functions by $G(x_1, \dots, x_n) = \langle \phi(x_1) \dots \phi(x_n) \rangle$.

The apparent drawback of the path-integral approach is that it is quite *hard* to actually calculate path integrals. With the exception of a few simple cases, such calculations usually have to be performed numerically. This is done in **lattice QFT**, where spacetime is discretized and path integrals are calculated by statistical Monte-Carlo sampling. This is the most direct way to compute correlation functions and the hadronic observables they encode from QCD, and over the last decades lattice QCD has made spectacular progress in that arena.

Functional derivatives. There is another way of making the path-integral approach useful without *actually* calculating path integrals. Namely, one can generate the correlation functions from the **partition function** $Z[J]$ by adding a source term with an external source $J(x)$:

$$Z[J] = \int \mathcal{D}\phi e^{i(S[\phi] - \int d^4x \phi(x)J(x))}. \quad (2.2.25)$$

If we take a functional derivative of the source term, we obtain

$$\frac{i\delta}{\delta J(x_1)} \left[-i \int d^4x \phi(x)J(x) \right] = \phi(x_1). \quad (2.2.26)$$

Taking the functional derivative of the partition function then yields

$$\begin{aligned} \frac{i\delta}{\delta J(x_1)} Z[J] &= \int \mathcal{D}\phi e^{i(\dots)} \phi(x_1), \\ \frac{i\delta}{\delta J(x_1)} \frac{i\delta}{\delta J(x_2)} Z[J] &= \int \mathcal{D}\phi e^{i(\dots)} \phi(x_1) \phi(x_2) \end{aligned} \quad (2.2.27)$$

etc., which can be generalized to an arbitrary polynomial function(al) of the fields:

$$f \left[\frac{i\delta}{\delta J} \right] Z[J] = \int \mathcal{D}\phi e^{i(\dots)} f[\phi]. \quad (2.2.28)$$

Comparing this with Eq. (2.2.24), we see that we can generate the correlation functions by an n -fold derivative of $Z[J]$, dividing by $Z[0]$ and finally letting $J \rightarrow 0$:

$$G(x_1 \dots x_n) = \frac{i\delta}{\delta J(x_1)} \cdots \frac{i\delta}{\delta J(x_n)} \Big|_{J=0} \frac{Z[J]}{Z[0]}. \quad (2.2.29)$$

The two-fold functional derivative of $Z[J]$ is the two-point function, the three-fold derivative the three-point function, etc. Vice versa, the partition function can be written as

$$\begin{aligned} Z[J] &= \int \mathcal{D}\phi e^{iS[\phi]} \sum_{n=0}^{\infty} \frac{(-i)^n}{n!} \left(\int d^4x J(x) \phi(x) \right)^n \\ &= \sum_{n=0}^{\infty} \frac{(-i)^n}{n!} \int d^4x_1 \dots d^4x_n J(x_1) \dots J(x_n) \underbrace{\int \mathcal{D}\phi e^{iS[\phi]} \phi(x_1) \dots \phi(x_n)}_{Z[0] G(x_1, \dots, x_n)}. \end{aligned} \quad (2.2.30)$$

Thus, we can generate all correlation functions from $Z[J]$, and $Z[J]$ can be reconstructed from the knowledge of all correlation functions. They contain the same information as the partition function, which defines the QFT.

In the following it will be convenient to leave the J -dependence intact, at least for intermediate steps in calculations, so that the ‘physics’ is recovered in the end when setting $J = 0$. Through Eq. (2.2.28), we define correlation functions in the presence of the source J as

$$\langle f[\phi] \rangle_J := \frac{\int \mathcal{D}\phi e^{i(S[\phi] - \int d^4x \phi(x)J(x))} f[\phi]}{\int \mathcal{D}\phi e^{i(S[\phi] - \int d^4x \phi(x)J(x))}} = \frac{1}{Z[J]} f \left[\frac{i\delta}{\delta J} \right] Z[J]. \quad (2.2.31)$$

Perturbation theory. The path-integral approach is very useful for doing calculations in perturbation theory. To do so, we split the action into a non-interacting and an interacting part:

$$S[\phi] = S_0[\phi] + g S_I[\phi]. \quad (2.2.32)$$

Then we can write

$$e^{ig S_I[\phi]} e^{-i \int d^4x \phi(x)J(x)} = e^{ig S_I \left[\frac{i\delta}{\delta J} \right]} e^{-i \int d^4x \phi(x)J(x)}, \quad (2.2.33)$$

since for a small coupling constant g we can expand $e^{ig S_I[\phi]}$ in powers of g and each term consists of polynomials of the fields. In this way we can pull out the interacting part of the exponential from the path integral,

$$Z[J] = e^{ig S_I \left[\frac{i\delta}{\delta J} \right]} \int \mathcal{D}\phi e^{i(S_0[\phi] - \int d^4x \phi(x)J(x))} = \sum_n \frac{(ig)^n}{n!} \left(S_I \left[\frac{i\delta}{\delta J} \right] \right)^n Z_0[J], \quad (2.2.34)$$

where the remaining path integral $Z_0[J]$ is calculable in a closed form since it only depends on the free action. For example, in the scalar theory one obtains

$$Z_0[J] = \int \mathcal{D}\phi e^{i(S_0[\phi] - \int d^4x \phi(x)J(x))} = Z_0[0] e^{-\frac{1}{2} \int d^4x J(x) D_F(x,y) J(y)}, \quad (2.2.35)$$

where the constant $Z_0[0]$ absorbs the remaining path integral and $D_F(x, y)$ is the Feynman propagator. In this way, $Z[J]$ can be computed order by order in perturbation theory: The result for $n = 0$ is the free theory, $n = 1$ gives the $\mathcal{O}(g)$ correction, and so on. The correlation functions (2.2.29) follow from functional derivatives, where the constant $Z_0[0]$ drops out.

Quantum equations of motion. While the machinery of perturbation theory is equally straightforward to set up using canonical quantization, the power of the path-integral approach lies in its non-perturbative applications. This leads to the so-called **functional methods**, where relations between the correlation functions of the theory are derived in the form of integral or differential equations (or both).

The **Dyson-Schwinger equations (DSEs)** are the quantum equations of motion of a QFT and can be derived with almost no assumptions. The partition function $Z[J]$ is invariant under a shift $\phi'(x) = \phi(x) + \epsilon(x)$, since this is just a relabeling of the fields under the integral, so we can write

$$\begin{aligned} Z[J] &= \int \mathcal{D}\phi' e^{i(S[\phi'] - \int d^4x \phi'(x)J(x))} \\ &= \int \mathcal{D}\phi e^{i(S[\phi] - \int d^4x \phi(x)J(x)) + i \int d^4x \epsilon(x) \left(\frac{\delta S[\phi]}{\delta \phi(x)} - J(x) \right)} \\ &= Z[J] \left\langle e^{i \int d^4x \epsilon(x) \left(\frac{\delta S[\phi]}{\delta \phi(x)} - J(x) \right)} \right\rangle_J. \end{aligned} \quad (2.2.36)$$

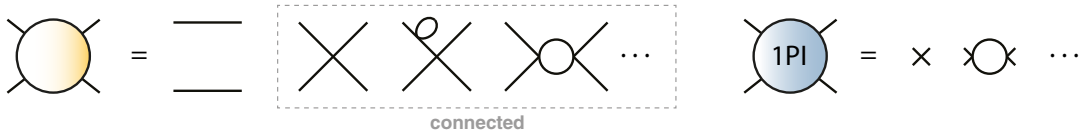


FIG. 2.9: Full, connected and 1PI diagrams in the four-point function of ϕ^4 theory (permutations not shown).

In the second line we assumed that the path integral measure $\mathcal{D}\phi$ is also invariant, and in the third line we inserted the definition (2.2.31). Since $\epsilon(x)$ is arbitrary, by expanding the exponential we find

$$\left\langle \frac{\delta S[\phi]}{\delta \phi(x)} \right\rangle_J = J(x), \quad (2.2.37)$$

which for $J \rightarrow 0$ is just the quantum average of the classical equations of motion.

But Eq. (2.2.37) is more useful than that. If we leave the dependence on the source J intact and exploit (2.2.31) again, it becomes

$$\frac{1}{Z[J]} \frac{\delta S}{\delta \phi} \left[\frac{i\delta}{\delta J} \right] Z[J] = J(x). \quad (2.2.38)$$

This should be read in the sense that we replace the functional dependence of $\delta S/\delta \phi$ on ϕ by a dependence on $i\delta/\delta J$ and apply it to the partition function. In this way, the equation serves as a ‘generating DSE’, because upon taking further functional derivatives $\delta^n/\delta J^n$ and setting $J \rightarrow 0$ in the end, we successively generate relations between the correlation functions (2.2.29) of the theory — the tower of DSEs. Note that the path integral no longer appears in these equations explicitly; instead, we calculate the correlation functions *from each other*.

As an example, we take another derivative $i\delta/\delta J$ of (2.2.31) to obtain

$$\begin{aligned} \frac{i\delta}{\delta J(y)} \langle f(\phi) \rangle_J &= \frac{1}{Z[J]} f \left[\frac{i\delta}{\delta J} \right] \frac{i\delta}{\delta J(y)} Z[J] - \frac{1}{Z[J]^2} \frac{i\delta Z[J]}{\delta J(y)} f \left[\frac{i\delta}{\delta J} \right] Z[J] \\ &= \langle f(\phi) \phi(y) \rangle_J - \langle f(\phi) \rangle_J \langle \phi(y) \rangle_J. \end{aligned} \quad (2.2.39)$$

Applied to Eq. (2.2.37) and setting the sources to zero, this yields

$$\left\langle \frac{\delta S[\phi]}{\delta \phi(x)} \phi(y) \right\rangle = i\delta^4(x-y). \quad (2.2.40)$$

From here one can derive the DSEs for the n -point functions with a tree-level counterpart in the action (which enter with $\propto \phi^n$ terms). For example, in a free scalar theory the propagator DSE becomes

$$\frac{\delta S[\phi]}{\delta \phi(x)} = -(\square + m^2) \phi(x) \quad \Rightarrow \quad -(\square_x + m^2) \langle \phi(x) \phi(y) \rangle = i\delta^4(x-y), \quad (2.2.41)$$

which returns the Feynman propagator $i/(p^2 - m^2 + i\epsilon)$ in momentum space.

Similar types of equations for the correlation functions can be derived from symmetry relations following from the Noether theorem. These are the **Ward-Takahashi identities (WTIs)** and **Slavnov-Taylor identities (STIs)**, which we will briefly touch upon in Sec. 2.2.3 and discuss in more detail in Sec. 3.1.2.

Quantum effective action. The relations above are rather formal and generic. What do they look like in practice? At this point it is useful to distinguish between

- the full correlation functions, generated by the partition function $Z[J]$,
- **connected** correlation functions, which enter in S-matrix elements and are thus of physical interest,
- and **one-particle-irreducible (1PI)** correlation functions, where external propagators are amputated and we keep only those diagrams that do not fall apart by cutting one line (see Fig. 2.9). In this way, they do away with the redundancy and describe the irreducible content of an n -point interaction vertex. For example, renormalizability can be determined from the 1PI vertices alone; hence they are the prime quantities of theoretical interest.

We also define two new generating functionals, which generate these new types of correlation functions by functional derivatives, namely $W[J] = -i \ln Z[J]$ and the **quantum effective action**

$$\Gamma[\varphi] = W[J] + \int d^4x \varphi(x) J(x). \quad (2.2.42)$$

$W[J]$ depends on the source field $J(x)$, whereas the source that appears in $\Gamma[\varphi]$ is the **averaged field** $\varphi(x)$. From Eq. (2.2.42) one can see that $W[J]$ and $\Gamma[\varphi]$ are Legendre transforms of each other. The meaning of these quantities becomes more clear if we write the partition function as

$$Z[J] = e^{iW[J]} = \int \mathcal{D}\phi e^{i(S[\phi] - \int d^4x \phi(x) J(x))} = e^{i(\Gamma[\varphi] - \int d^4x \varphi(x) J(x))}. \quad (2.2.43)$$

In this sense, $\Gamma[\varphi]$ and $\varphi(x)$ can be seen as the quantum versions of the classical action $S[\phi]$ and the classical field $\phi(x)$, integrated over quantum fluctuations and with the path integral exponential as a weight factor. More precisely, from Eq. (2.2.42) we have

$$\varphi(x) = -\frac{\delta W[J]}{\delta J(x)} = \frac{1}{Z[J]} \frac{i\delta}{\delta J(x)} Z[J] \stackrel{(2.2.31)}{=} \langle \phi(x) \rangle_J. \quad (2.2.44)$$

Thus, $\varphi(x)$ is the vacuum expectation value of the field $\phi(x)$ in the presence of the source J , which vanishes² in the limit $J = 0$. Vice versa, $J(x)$ satisfies

$$J(x) = \frac{\delta \Gamma[\varphi]}{\delta \varphi(x)}. \quad (2.2.45)$$

Like in thermodynamics, $J(x)$ and $\varphi(x)$ are conjugate variables and the generating functionals $W[J]$ and $\Gamma[\varphi]$ are the corresponding potentials. J is the ‘intensive’ and φ the ‘extensive’ variable, and differentiation with respect to one gives the other.

While the classical action $S[\phi]$ contains the content of the classical field theory, either of the functionals $Z[J]$, $W[J]$ or $\Gamma[\varphi]$ determines the QFT completely since all

²Unless in the case of spontaneous symmetry breaking, but even then one can redefine the field so that its vacuum expectation value vanishes.

correlation functions can be derived from them: the connected ones are derivatives of $W[J]$ with respect to J , and 1PI vertices are derivatives of the effective action $\Gamma[\varphi]$ with respect to φ . For example, the 1PI two- and three-point functions are given by

$$\frac{\delta^2 \Gamma[\varphi]}{\delta\varphi(x_1) \delta\varphi(x_2)} \Big|_{\varphi=0}, \quad \frac{\delta^3 \Gamma[\varphi]}{\delta\varphi(x_1) \delta\varphi(x_2) \delta\varphi(x_3)} \Big|_{\varphi=0}. \quad (2.2.46)$$

Relations between full, connected and 1PI correlation functions. To relate the full correlation functions with their connected and 1PI counterparts, we must reformulate Eq. (2.2.31) in terms of $W[J]$ instead of $Z[J]$. To do so, we exploit the useful relation

$$e^{-A} f(\partial) e^A = f(\partial + \partial A), \quad (2.2.47)$$

where ∂ stands for a generic derivative acting on A . The r.h.s. acts on 1, so all unsaturated derivatives vanish. It is straightforward to verify this for polynomial functions, e.g. $f(\partial) = \partial^2$:

$$\begin{aligned} e^{-A} \partial^2 e^A &= e^{-A} \partial (A' e^A) = A'' + A'^2, \\ (\partial + \partial A)^2 &= (\partial + \partial A) \partial A = A'' + A'^2. \end{aligned} \quad (2.2.48)$$

Applied to Eq. (2.2.31), with $A = iW[J]$ and $\partial = i\delta/\delta J$, we arrive at

$$\langle f[\phi] \rangle_J = f \left[\frac{i\delta}{\delta J} - \frac{\delta W[J]}{\delta J} \right]. \quad (2.2.49)$$

Here the dependence of f on each field value $\varphi(x_i)$ has to be replaced by a dependence on the bracket above, with $\delta/\delta J(x_i)$, and all unsaturated derivatives vanish since it acts on 1. In this way we have expressed the vacuum expectation value of $f(\phi)$ through functional derivatives of $W[J]$, which are the connected n -point functions. For example, if we abbreviate

$$W''_{xy}[J] = \frac{\delta^2 W[J]}{\delta J(x) \delta J(y)}, \quad \Gamma''_{xy}[\varphi] = \frac{\delta^2 \Gamma[\varphi]}{\delta\varphi(x) \delta\varphi(y)}, \quad \text{etc.} \quad (2.2.50)$$

we obtain for the two-point function:

$$\begin{aligned} \langle \phi(x) \phi(y) \rangle_J &= \left(\frac{i\delta}{\delta J(x)} - W'_x[J] \right) \left(\frac{i\delta}{\delta J(y)} - W'_y[J] \right) \\ &= W'_x[J] W'_y[J] - i W''_{xy}[J] \\ &= \varphi(x) \varphi(y) - i W''_{xy}[J]. \end{aligned} \quad (2.2.51)$$

The unsaturated derivative in the first line vanishes, and in the third line we used Eq. (2.2.44). In the limit $J = 0$ the vacuum expectation value $\varphi(x)$ is zero, and we find that the ‘connected’ propagator $W''_{xy}[0]$ is the same as the usual one, $\langle \phi(x) \phi(y) \rangle$, modulo a factor i . The same exercise can be repeated for higher n -point functions to express them in terms of connected and disconnected parts.

Now let us reformulate Eq. (2.2.49) in terms of $\Gamma[\varphi]$ instead of $W[J]$. From Eqs. (2.2.44–2.2.45) we have $W'_x[J] = -\varphi(x)$ and $\Gamma'_x[\varphi] = J(x)$ and therefore

$$\int d^4y W''_{xy}[J] \Gamma''_{yz}[\varphi] = - \int d^4y \frac{\delta\varphi(x)}{\delta J(y)} \frac{\delta J(y)}{\delta\varphi(z)} = -\frac{\delta\varphi(x)}{\delta\varphi(z)} = -\delta^4(x-z). \quad (2.2.52)$$

In other words, the 1PI two-point function $\Gamma''_{xy}[0]$ is the inverse of $W''_{xy}[0]$. This explains why the *inverse* tree-level propagators appear in the classical action together with the tree-level vertices: the classical action contains the *1PI correlation functions* at tree level.

Let us denote the dressed propagator in the presence of the external source φ by $\Delta_{xy}[\varphi] = \Gamma''_{xy}[\varphi]^{-1}$. The product rule entails that the derivative of the propagator with respect to φ is

$$\frac{\delta}{\delta\varphi(z)} \Delta_{xy}[\varphi] = - \int d^4a \int d^4b \Delta_{xa}[\varphi] \Gamma'''_{azb}[\varphi] \Delta_{by}[\varphi]. \quad (2.2.53)$$

The derivative with respect to J becomes

$$\frac{\delta}{\delta J(x)} = \int d^4y \frac{\delta\varphi(y)}{\delta J(x)} \frac{\delta}{\delta\varphi(y)} = - \int d^4y W''_{xy}[J] \frac{\delta}{\delta\varphi(y)} = \int d^4y \Delta_{xy}[\varphi] \frac{\delta}{\delta\varphi(y)}, \quad (2.2.54)$$

so we can express Eq. (2.2.49) in terms of the effective action and its derivatives:

$$\langle f[\phi] \rangle_J = f \left[\varphi(x) + \int d^4y \Delta_{xy}[\varphi] \frac{i\delta}{\delta\varphi(y)} \right]. \quad (2.2.55)$$

This is the identity that we were after, because it allows us to relate the full n -point functions with their 1PI counterpart. Evaluating it for the two-point function, we recover our previous result (2.2.51):

$$\langle \phi(x) \phi(y) \rangle_J = \left(\varphi(x) + \int d^4a \Delta_{xa}[\varphi] \frac{i\delta}{\delta\varphi(a)} \right) \varphi(y) = \varphi(x) \varphi(y) + i\Delta_{xy}[\varphi]. \quad (2.2.56)$$

For the three-point function we obtain

$$\begin{aligned} \langle \phi(x) \phi(y) \phi(z) \rangle_J &= \left(\varphi(x) + \int d^4a \Delta_{xa}[\varphi] \frac{i\delta}{\delta\varphi(a)} \right) \left(\varphi(y) + \int d^4b \Delta_{yb}[\varphi] \frac{i\delta}{\delta\varphi(b)} \right) \varphi(z) \\ &= \left(\varphi(x) + \int d^4a \Delta_{xa}[\varphi] \frac{i\delta}{\delta\varphi(a)} \right) (\varphi(y) \varphi(z) + i\Delta_{yz}[\varphi]) \\ &= \varphi(x) \varphi(y) \varphi(z) + i\Delta_{xy}[\varphi] \varphi(z) + i\Delta_{xz}[\varphi] \varphi(y) + i\Delta_{yz}[\varphi] \varphi(x) \\ &\quad + \int d^4a \int d^4b \int d^4c \Delta_{xa}[\varphi] \Delta_{yb}[\varphi] \Gamma'''_{bac}[\varphi] \Delta_{cz}[\varphi], \end{aligned} \quad (2.2.57)$$

where we used Eq. (2.2.53) in the final step. After setting the sources to zero, only the last line survives and we find that the full three-point function is just the 1PI three-point vertex with external propagator legs attached. Going further to the four-point function, the relation would become more complicated since the four-point function has 1-particle-reducible topologies that fall apart by cutting one line (see below).

Dyson-Schwinger equations for 1PI correlation functions. Eq. (2.2.55) is not only useful for relating full and 1PI correlation functions, but also for working out the quantum versions of classical equations $f(\phi) = 0$ between the fields: we just need to replace the (usually non-linear) dependence on ϕ by the bracket on the r.h.s. – which generates further derivatives and derivatives of propagators – and set all fields to zero in the end, together with all unsaturated derivatives. If the classical action contains more than one field, then the functional dependence holds for each $\varphi_i(x)$, and the integral over y also goes over all intermediate (mixed!) propagators, which drop out in the end when setting the sources to zero.

With this we can express the generating DSE (2.2.37) in terms of the 1PI correlation functions:

$$\frac{\delta\Gamma[\varphi]}{\delta\varphi(x)} = \frac{\delta S}{\delta\phi} \left[\varphi(x) + \int d^4y \Delta_{xy}[\varphi] \frac{i\delta}{\delta\varphi(y)} \right]. \quad (2.2.58)$$

Now we can better see what the equation actually means: On the r.h.s. we have the classical equations of motion, but with the ϕ dependence replaced by the content of the bracket. Like in Eqs. (2.2.56) and (2.2.57), this will reproduce the classical equations of motion for $\varphi(x)$ plus further terms which describe quantum corrections. The l.h.s. is the derivative of the quantum effective action. If we take further derivatives, we generate the 1PI two-, three-, ... n -point functions, which therefore always contain a classical part together with quantum contributions.

In practice Eq. (2.2.55) amounts to repeated applications of the bracket with the derivative, which suggests to use a **diagrammatic language**. We need graphical expressions for the source field $\varphi(x)$, the propagator Δ_{xy} , the higher n -point functions Γ'''_{xyz} , Γ''''_{xyzw} , ... and the effect of the functional derivative $\delta/\delta\varphi(x)$ on these quantities:

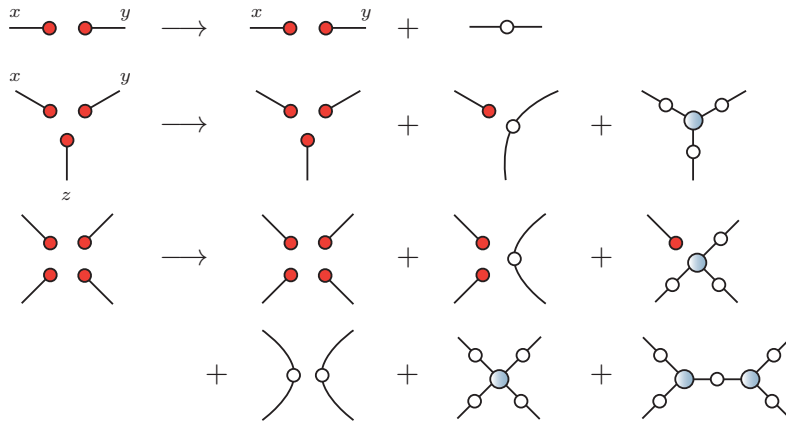
$$\begin{array}{lcl} \varphi(x) & = & \text{---} \bullet \\ \Delta_{xy} & = & \text{---} \circ \text{---} \\ \Gamma'''_{xyz} & = & \begin{array}{c} | \\ \circ \\ / \quad \backslash \end{array} \\ \Gamma''''_{xyzw} & = & \begin{array}{c} \diagup \quad \diagdown \\ \circ \\ \diagdown \quad \diagup \end{array} \end{array} \quad \begin{array}{lcl} \frac{\delta}{\delta\varphi(x)} \text{---} \bullet & = & \begin{array}{c} \text{---} \bullet \\ | \\ \text{---} \end{array} \\ \frac{\delta}{\delta\varphi(x)} \text{---} \circ \text{---} & = & \begin{array}{c} | \\ \circ \\ / \quad \backslash \end{array} \\ \frac{\delta}{\delta\varphi(x)} \begin{array}{c} | \\ \circ \\ / \quad \backslash \end{array} & = & \begin{array}{c} \diagup \quad \diagdown \\ \circ \\ \diagdown \quad \diagup \end{array} \\ \frac{\delta}{\delta\varphi(x)} \begin{array}{c} \diagup \quad \diagdown \\ \circ \\ \diagdown \quad \diagup \end{array} & = & \begin{array}{c} \diagup \quad \diagdown \\ \circ \\ \diagdown \quad \diagup \end{array} \end{array}$$

and so on. In the graphical notation we no longer distinguish between a correlation function that depends on $\varphi(x)$ as opposed to one where the field is set to zero, and for simplicity we also suppress all minus signs, i factors and multiplicities that arise from derivatives.

What we still need is a graphical analogue for Eq. (2.2.55), where $\phi(x)$ is replaced by $\varphi(x) + \int d^4y \Delta_{xy} i\delta/\delta\varphi(y)$. If we work this out for products of fields like in Eqs. (2.2.56) and (2.2.57) we arrive at:

$$\varphi(x) \dots \varphi(z) \rightarrow \left(\varphi(x) + \int \dots \right) \dots \left(\varphi(z) + \int \dots \right), \quad (2.2.59)$$

or graphically:



For simplicity we also absorbed the different symmetrizations with the same topology into one diagram. If we set $\varphi = 0$, these graphs tell us how ordinary n -point functions are related with their 1PI counterparts: the three-point functions are identical except for the external propagator legs, whereas the full four-point function is the sum of disconnected parts, a 1PI term, and 1-particle-reducible diagrams that contain 1PI three-point functions.

On the other hand, we can interpret these diagrams also differently: without the extra integral term in Eq. (2.2.55) we would return to the classical quantity $f(\varphi)$ expressed in terms of φ instead of ϕ . In the graphical notation we can then also drop the distinction between $\phi(x)$ and its quantum expectation value $\varphi(x)$ and use the same symbols for the fields that appear in the Lagrangian. In that way we can transform equations for the classical fields (equations of motion, symmetry relations etc.) into quantum identities for the 1PI correlation functions. Going from ‘classical’ to ‘quantum’ in the picture above then entails to connect the legs in all possible ways and equip them with dressed propagators.

ϕ^4 theory. We illustrate this by considering the simplest scalar field theory, ϕ^4 theory. The classical action and its functional derivative are given by

$$S = \int d^4x \left[\frac{1}{2} (\partial^\mu \phi \partial_\mu \phi - m^2 \phi^2) - \frac{g}{4!} \phi^4 \right], \quad \frac{\delta S}{\delta \phi} = -(\square + m^2)\phi - \frac{g}{3!} \phi^3. \quad (2.2.60)$$

Setting $\delta S/\delta \phi = 0$ yields the classical equations of motion. Diagrammatically, this amounts to

$$S = \text{---}^{-1}\text{---} + \text{---} \text{---} \text{---} \text{---}$$

$$\frac{\delta S}{\delta \phi} = \text{---}^{-1}\text{---} + \text{---} \text{---} \text{---} \text{---} = 0$$

The line with ‘-1’ is the inverse tree-level propagator. In the classical action it is contracted with the field ϕ ; the functional derivative removes one instance of ϕ . As before we ignore all prefactors and multiplicities.

Now ‘connect the dots’ in all possible ways to obtain the quantum eq. of motion:

$$\frac{\delta \Gamma}{\delta \varphi} = \text{---} \overset{-1}{\bullet} + \begin{array}{c} \bullet \\ | \\ \text{---} \bullet \\ | \\ \bullet \end{array} + \begin{array}{c} \circ \\ \diagup \quad \diagdown \\ \text{---} \bullet \end{array} + \begin{array}{c} \circ \\ \diagup \quad \diagdown \\ \text{---} \bullet \\ \diagdown \quad \diagup \\ \circ \end{array}$$

We arrive at the same result in formulas if we replace ϕ in Eq. (2.2.60) with the bracket in (2.2.58) and let it act on 1:

$$\frac{\delta \Gamma}{\delta \varphi(x)} = -(\square + m^2) \varphi(x) - \frac{g}{3!} \left[\varphi(x)^3 + 3i \varphi(x) \Delta_{xx} + \iiint_{z z' z''} \Delta_{xz} \Delta_{xz'} \Delta_{xz''} \Gamma'''_{zz'z''} \right], \quad (2.2.61)$$

where we suppressed the arguments in $\Delta[\tilde{\varphi}]$ and $\Gamma'''[\tilde{\varphi}]$ and abbreviated the integrals $\int d^4z$ by \int_z . Applying another derivative and setting $\varphi = 0$ yields the DSE for the inverse scalar propagator:

$$\left. \frac{\delta^2 \Gamma}{\delta \varphi^2} \right|_{\varphi=0} = \text{---} \circ \overset{-1}{=} = \text{---} \overset{-1}{=} + \begin{array}{c} \circ \\ \diagup \quad \diagdown \\ \text{---} \end{array} + \begin{array}{c} \circ \\ \diagup \quad \diagdown \\ \text{---} \bullet \\ \diagdown \quad \diagup \\ \circ \end{array}$$

$$\Delta_{xy}^{-1} = -(\square + m^2) \delta^4(x - y) - \frac{g}{3!} \left[3i \delta^4(x - y) \Delta_{xx} + \iiint_{z z' z''} \Delta_{xz} \Delta_{xz'} \Delta_{xz''} \Gamma'''_{zz'z''y} \right]. \quad (2.2.62)$$

In principle, the derivatives of the propagators in the two-loop diagram produce further terms including three-point vertices; however, as a consequence of the Lagrangian’s Z_2 symmetry under $\varphi \rightarrow -\varphi$, there are no odd n -point functions in ϕ^4 theory and therefore $\Gamma'''_{xyz}[0] = 0$.

Eq. (2.2.62) says that the inverse full (‘dressed’) propagator is the sum of the inverse classical (tree-level) propagator plus quantum loop corrections, which define the self-energy. The equation is exact but depends on the four-point vertex, which satisfies its own DSE. Still, if we happened to *know* the exact four-point vertex, the equation would give us the exact two-point function. This goes back to the comment below Eq. (2.2.38): instead of working out the path integral explicitly, one can calculate the correlation functions from each other. In practice the equations are solved in momentum space, where each loop becomes a four-momentum integration, so they have the usual structure of Feynman diagrams.

Since the DSEs are exact, they also reproduce **perturbation theory** if the coupling is small. To get a DSE for the propagator instead of its inverse, multiply Eq. (2.2.62) with the tree-level propagator (Δ_0) from the left and the full propagator (Δ) from the right (or vice versa):

$$\Delta^{-1} = \Delta_0^{-1} - \Sigma \quad \Rightarrow \quad \Delta = \Delta_0 + \Delta_0 \Sigma \Delta, \quad (2.2.63)$$

where Σ is the sum of the self-energy terms. Reinsert the equation again for each instance of the dressed propagator on the right,

$$\Delta = \Delta_0 + \Delta_0 \Sigma \Delta = \Delta_0 + \Delta_0 \Sigma \Delta_0 + \Delta_0 \Sigma \Delta = \dots, \quad (2.2.64)$$

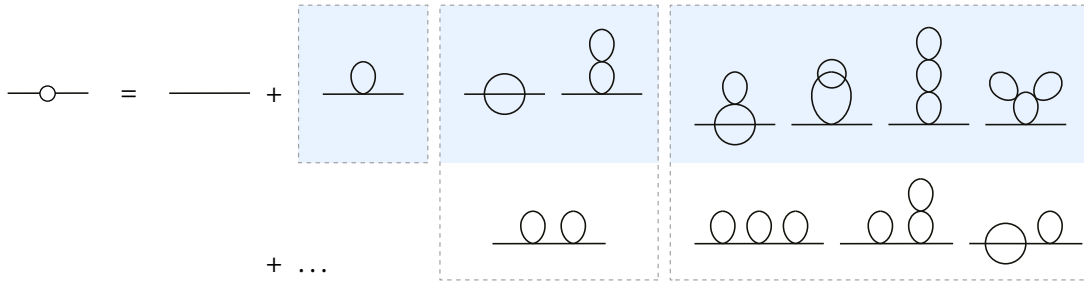


FIG. 2.10: Perturbative expansion of the propagator in ϕ^4 theory. The colored boxes highlight the 1PI diagrams which appear in the self-energy.

and do the same for every instance of the propagator and the vertex inside the self-energy Σ . Because the self-energy comes with a factor g and the four-point vertex includes another factor g , every additional loop carries a factor g ; and as long as $|g| < 1$, higher loop diagrams are suppressed. In this way we successively generate the perturbative series for the propagator (Fig. 2.10).

Even if the perturbation series does not converge, it is useful to remember that there are non-perturbative, exact DSEs behind it, which can be derived for any QFT. A simple analogue is the geometric series: The solution of the equation

$$f(x) = 1 + xf(x) = 1 + x + x^2 f(x) = 1 + x + x^2 + x^3 f(x) = \dots, \quad (2.2.65)$$

is $f(x) = 1/(1 - x)$. The geometric series $f(x) = \sum_{n=0}^{\infty} x^n$ converges to that result only if $|x| < 1$, whereas each step in Eq. (2.2.65) is non-perturbative and exact. The difference comes from the last term, where $f(x)$ appears again and pulls the result back even if x becomes explosively large.

Dyson-Schwinger equations in QCD. Let us be bold and try to apply the procedure to QCD right away. There are still issues we have not yet dealt with (see Sec. 2.2.3), which will produce additional ghost diagrams, but otherwise the diagrammatic derivation goes through as before. The classical action of QCD takes the form:

$$S = \text{---} \overset{-1}{\bullet} \text{---} \bullet + \text{---} \overset{\text{gluon}}{\curvearrowright} \bullet + \text{---} \overset{-1}{\text{gluon}} \text{---} \bullet + \text{---} \overset{\text{gluon}}{\text{gluon}} \text{---} \bullet + \text{---} \overset{\text{gluon}}{\text{gluon}} \text{---} \bullet$$

We use the convention that the left circles in the fermion terms represent the antiquark fields $\bar{\psi}$ and the right circles the quark fields ψ . Taking a functional derivative with respect to $\bar{\psi}$ yields the classical equation of motion for the quark, the Dirac equation:

$$\frac{\delta S}{\delta \bar{\psi}} = \text{---} \overset{-1}{\bullet} + \text{---} \overset{\text{gluon}}{\curvearrowright} \bullet = 0$$

Consequently, the quantum equation of motion becomes (connect the dots)

$$\frac{\delta \Gamma}{\delta \bar{\psi}} = \text{---} \overset{-1}{\bullet} + \text{---} \overset{\text{gluon}}{\curvearrowright} \bullet + \text{---} \overset{\text{gluon}}{\text{gluon}} \text{---} \bullet$$

Note that this involves a mixed, field-dependent ‘quark-gluon propagator’ on the right. Taking another functional derivative with respect to ψ and setting all fields to zero gives us the **quark DSE** for the inverse quark propagator:

$$\frac{\delta^2 \Gamma}{\delta \bar{\psi} \delta \psi} \Big|_{A, \psi, \bar{\psi}=0} = \text{---} \circ \text{---}^{-1} = \text{---}^{-1} + \text{---} \circ \text{---} \text{---}$$

The quark DSE tells us that the inverse full propagator is the sum of the inverse tree-level propagator plus quantum loop corrections, which are contained in the quark self-energy. In practice it enables us to compute the quark propagator if the gluon propagator and quark-gluon vertex are known.

We can repeat the same steps also for the gluon propagator. The classical equation of motion for the gluon (the Maxwell equation) reads

$$\frac{\delta S}{\delta A} = \text{---} \text{---} + \text{---} \text{---}^{-1} + \text{---} \text{---} \text{---} + \text{---} \text{---} \text{---} = 0$$

and the corresponding quantum equation of motion:

$$\frac{\delta \Gamma}{\delta A} = \frac{\delta S}{\delta A} + \text{---} \text{---} \text{---} + \text{---} \text{---} \text{---} + \text{---} \text{---} \text{---} + \text{---} \text{---} \text{---}$$

After taking another derivative with respect to A and setting all sources to zero, we obtain the **gluon DSE**:

$$\frac{\delta^2 \Gamma}{\delta A^2} \Big|_{A, \psi, \bar{\psi}=0} = \text{---} \text{---} \text{---}^{-1} = \text{---} \text{---} \text{---}^{-1} + \text{---} \text{---} \text{---} + \text{---} \text{---} \text{---} + \text{---} \text{---} \text{---} + \text{---} \text{---} \text{---} + \text{---} \text{---} \text{---} + \text{---} \text{---} \text{---}$$

The inverse dressed gluon propagator is the sum of the inverse tree-level propagator, a quark loop, a gluon loop, a tadpole diagram, a ‘sunset’ diagram and a ‘squint’ diagram. The gauge fixing procedure in the next section will also produce a ghost loop, which has the same form as the quark loop. The inputs of the equation are then the quark and ghost propagators and the quark-gluon, ghost-gluon, three-gluon and four-gluon vertices. If we solved the quark, gluon and ghost DSEs in combination (i.e., all two-point functions), the remaining inputs would be the vertices.

Actually we have ignored another subtlety: if the action contains several fields, one has to sum over them when taking the derivative of the propagators in Eq. (2.2.53), which also leads to mixed vertices. The general rule is that, after taking functional derivatives, for each internal ‘half-propagator’ connected to a dressed vertex one has to sum over all types of fields. This does not modify the quark and gluon DSEs but it will produce additional diagrams, for example, in the quark-gluon vertex DSE.

2.2.3 Gauge fixing in QCD

Path integrals in QCD. We have discussed most of the things so far at the level of a generic QFT, so we should get a bit more specific. The straightforward generalization of the partition function in Eq. (2.2.25) to QCD with quark (ψ), antiquark ($\bar{\psi}$) and gluon fields (A^μ) is:

$$Z[J, \bar{\eta}, \eta] = \int \mathcal{D}A \mathcal{D}\psi \mathcal{D}\bar{\psi} e^{i(S[A, \psi, \bar{\psi}] - \int d^4x (J_\mu A^\mu + \bar{\eta} \psi + \eta \bar{\psi}))}, \quad (2.2.66)$$

where we have added a vector source J^μ for the gluon field and spinor sources $\bar{\eta}$ and η for the quark and antiquark fields, respectively. The fermion fields in the path integral are anticommuting Grassmann numbers, whereas their corresponding field operators in the canonical approach satisfy equal-time anticommutation relations:

$$\begin{aligned} \{\psi_{\alpha,i}(x), \psi_{\beta,j}^\dagger(y)\}_{x^0=y^0} &= \delta^3(\mathbf{x} - \mathbf{y}) \delta_{\alpha\beta} \delta_{ij}, \\ \{\psi_{\alpha,i}(x), \psi_{\beta,j}(y)\}_{x^0=y^0} &= \{\psi_{\alpha,i}^\dagger(x), \psi_{\beta,j}^\dagger(y)\}_{x^0=y^0} = 0. \end{aligned} \quad (2.2.67)$$

For example, the **quark propagator** is given by

$$S_{\alpha\beta}(x_1, x_2) = \langle 0 | \mathbf{T} \psi_\alpha(x_1) \bar{\psi}_\beta(x_2) | 0 \rangle = \frac{i^2 \delta^2}{\delta \bar{\eta}_\alpha(x_1) \delta \eta_\beta(x_2)} \Big|_{J, \eta, \bar{\eta}=0} \frac{Z[J, \eta, \bar{\eta}]}{Z[0, 0, 0]}. \quad (2.2.68)$$

We recollect some relations for **Grassmann variables**. A Grassmann algebra is an algebra generated by a basis $\{\theta_1 \dots \theta_n\}$ which satisfies anticommutation relations $\{\theta_i, \theta_j\} = 0$. In particular, this implies $\theta_i^2 = 0$. As a consequence, a general element of the algebra is at most linear in each θ_i ,

$$f(\theta) = c_0 + c_1 \theta, \quad f(\theta_1, \theta_2) = c_0 + c_1 \theta_1 + c_2 \theta_2 + c_{12} \theta_1 \theta_2, \quad \text{etc.} \quad (2.2.69)$$

which also means that a Taylor expansion stops: $e^{a\theta} = 1 + a\theta$. A derivative can then be defined by replacing $\theta_i \rightarrow 1$, but the θ_i must first be permuted to the derivative operator:

$$\frac{\partial f(\theta)}{\partial \theta} = c_1, \quad \frac{\partial f(\theta_1, \theta_2)}{\partial \theta_1} = c_1 + c_{12} \theta_2, \quad \frac{\partial f(\theta_1, \theta_2)}{\partial \theta_2} = c_2 - c_{12} \theta_1, \quad \text{etc.} \quad (2.2.70)$$

The integration can be defined by

$$\int d\theta 1 = 0, \quad \int d\theta \theta = 1, \quad (2.2.71)$$

where the first relation guarantees translation invariance, $\int d\theta f(\theta + \eta) - \int d\theta f(\theta) = \int d\theta (c_1 \eta) \stackrel{!}{=} 0$, and the second is a normalization. Also here the integration variable must be permuted to the integral measure. As a consequence, the integration and derivative are the same:

$$\int d\theta f(\theta) = \int d\theta (c_0 + c_1 \theta) = c_1 = \frac{df(\theta)}{d\theta}. \quad (2.2.72)$$

An integral over n Grassmann variables $d^n \theta = d\theta_n \dots d\theta_1$ becomes

$$\int d^n \theta \theta_1 \dots \theta_n = 1, \quad \int d^n \theta \theta_{i_1} \dots \theta_{i_n} = \varepsilon_{i_1 \dots i_n}, \quad (2.2.73)$$

where $\varepsilon_{i_1 \dots i_n}$ is the totally antisymmetric tensor with $\varepsilon_{1 \dots n} = 1$. In the integral over $f(\theta) = f(\theta_1, \dots, \theta_n)$ only the last term in Eq. (2.2.69) survives, since every $d\theta_i$ must be saturated by θ_i :

$$\int d^n \theta f(\theta) = \int d^n \theta c_{1 \dots n} \theta_1 \dots \theta_n = c_{1 \dots n} \quad (2.2.74)$$

To obtain the Jacobian for the transformation $\theta'_i = A_{ij} \theta_j$, we write

$$\int d^n \theta f(\theta') = \int d^n \theta c_{1\dots n} A_{1i_1} \dots A_{ni_n} \theta_{i_1} \dots \theta_{i_n} = \underbrace{\varepsilon_{1\dots n} A_{1i_1} \dots A_{ni_n}}_{=\det A} c_{1\dots n}, \quad (2.2.75)$$

and since $c_{1\dots n} = \int d^n \theta f(\theta) = \int d^n \theta' f(\theta')$ we find

$$d^n \theta' = \frac{1}{\det A} d^n \theta. \quad (2.2.76)$$

This is in contrast to the bosonic case, where $x'_i = A_{ij} x_j$ leads to $d^n x' = |\det A| d^n x$.

We define complex Grassmann variables by $(\theta_i \theta_j)^* = \theta_j^* \theta_i^*$, so that

$$d^n \theta d^n \theta^* = d\theta_n \dots d\theta_1 d\theta_1^* \dots d\theta_n^* = d\theta_1 d\theta_1^* \dots d\theta_n d\theta_n^* = d\theta_n d\theta_n^* \dots d\theta_1 d\theta_1^*. \quad (2.2.77)$$

Pairs like $d\theta_i d\theta_i^*$ are bosonic and can be permuted through, whereas $d^n \theta d^n \theta^* = (-1)^n d^n \theta^* d^n \theta$. This yields $\int d^n \theta d^n \theta^* \theta_{i_1}^* \theta_{j_1} \dots \theta_{i_n}^* \theta_{j_n} = \varepsilon_{i_1 \dots i_n} \varepsilon_{j_1 \dots j_n}$, and in the integral $\int d^n \theta d^n \theta^* f(\theta, \theta^*)$ only the term $\propto \theta_1 \dots \theta_n \theta_1^* \dots \theta_n^*$ survives. Finally, we calculate the integral over a Gaussian:

$$\begin{aligned} \int d^n \theta d^n \theta^* e^{\theta^* B \theta} &= \int d^n \theta d^n \theta^* \frac{1}{n!} (\theta^* B \theta)^n \\ &= \frac{1}{n!} B_{i_1 j_1} \dots B_{i_n j_n} \int d^n \theta d^n \theta^* \theta_{i_1}^* \theta_{j_1} \dots \theta_{i_n}^* \theta_{j_n} \\ &= \frac{1}{n!} \varepsilon_{i_1 \dots i_n} \varepsilon_{j_1 \dots j_n} B_{i_1 j_1} \dots B_{i_n j_n} = \det B. \end{aligned} \quad (2.2.78)$$

If the exponent comes with a minus sign, this will produce a factor $(-1)^n$ which can be compensated by employing the integral measure $d^n \theta^* d^n \theta$:

$$\int d^n \theta^* d^n \theta e^{-\theta^* B \theta} = \det B. \quad (2.2.79)$$

The fact that we can express the determinant of a matrix as an integral over Grassmann variables will become extremely useful in a moment. Another useful relation is

$$\begin{aligned} \int d^n \theta^* d^n \theta e^{-(\theta^* B \theta + \eta^* \theta + \theta^* \eta)} &= e^{\eta^* B^{-1} \eta} \int d^n \theta^* d^n \theta e^{-(\theta^* + \eta^* B^{-1}) B (\theta + B^{-1} \eta)} \\ &= e^{\eta^* B^{-1} \eta} \det B, \end{aligned} \quad (2.2.80)$$

from where one can verify Eq. (2.2.68) for the tree-level quark propagator if one starts from the partition function of a free fermion Lagrangian.

The **gluon propagator** is given by

$$D^{\mu\nu}(x_1, x_2) = \langle 0 | T A^\mu(x_1) A^\nu(x_2) | 0 \rangle = \frac{i^2 \delta^2}{\delta J_\mu(x_1) \delta J_\nu(x_2)} \Big|_{J, \eta, \bar{\eta}=0} \frac{Z[J, \eta, \bar{\eta}]}{Z[0, 0, 0]}, \quad (2.2.81)$$

but from the discussion around Eq. (2.1.21) this expression does not yet make sense. The problem appears in the kinetic gluon term in the Lagrangian:

$$-\frac{1}{4} F_{\mu\nu}^a F_a^{\mu\nu} \cong \frac{1}{2} A_\mu^a (\square g^{\mu\nu} - \partial^\mu \partial^\nu) A_\nu^a + \dots \quad (2.2.82)$$

The tree-level inverse gluon propagator in momentum space is proportional to a transverse projector,

$$D_0^{-1}(p)^{\mu\nu} = ip^2 \left(g^{\mu\nu} - \frac{p^\mu p^\nu}{p^2} \right), \quad (2.2.83)$$

but projectors are not invertible: if P_+ and P_- are generic projectors that satisfy $P_{\pm}^2 = P_{\pm}$, $P_+P_- = 0$ and $P_+ + P_- = 1$, then a linear combination of them has the form

$$F = \alpha P_+ + \beta P_- \quad \Rightarrow \quad F^{-1} = \frac{P_+}{\alpha} + \frac{P_-}{\beta}. \quad (2.2.84)$$

For $\alpha = 0$ or $\beta = 0$, the inverse is not well-defined. Applied to Eq. (2.2.83), this means that the gluon propagator $D_0^{\mu\nu}(p)$ does not exist.

The problem appears in the gauge-boson sector and is structurally the same in QCD and QED. In fact, it already arises in the free theory by observing that A_0^μ has no canonically conjugate momentum. The conjugate momentum is defined by

$$\Pi_\mu = \frac{\partial \mathcal{L}}{\partial (\partial_0 A^\mu)} = F_{\mu 0}, \quad (2.2.85)$$

which entails $\Pi_0 = 0$, but this contradicts the equal-time commutation relations

$$[A_\mu(x), \Pi_\nu(y)]_{x^0=y^0} = ig_{\mu\nu} \delta^3(\mathbf{x} - \mathbf{y}). \quad (2.2.86)$$

Redundancy in the path integral. Another manifestation of the problem is that in the path integral over the gauge fields,

$$Z = \int \mathcal{D}A e^{iS[A]}, \quad (2.2.87)$$

we integrate over redundant degrees of freedom that are connected by gauge transformations. This is easy to see in QED: If we split the gauge fields into transverse and longitudinal parts $A^\mu = A_T^\mu + A_L^\mu$, with $\partial_\mu A_T^\mu = 0$ and $A_L^\mu = \partial^\mu \Theta$, we find that

$$\begin{aligned} (\square g^{\mu\nu} - \partial^\mu \partial^\nu) A_\nu^L &= \square \partial^\mu \Theta - \partial^\mu \square \Theta = 0 \\ \Rightarrow (\square g^{\mu\nu} - \partial^\mu \partial^\nu) A_\nu &= (\square g^{\mu\nu} - \partial^\mu \partial^\nu) A_\nu^T = \square A_T^\mu \end{aligned} \quad (2.2.88)$$

and thus A_L^μ drops out completely from the kinetic term in the QED Lagrangian:

$$-\frac{1}{4} F_{\mu\nu} F^{\mu\nu} = \frac{1}{2} A_\mu (\square g^{\mu\nu} - \partial^\mu \partial^\nu) A_\nu = \frac{1}{2} A_T^\mu \square A_T^\mu. \quad (2.2.89)$$

Now, if we perform a gauge transformation

$$A'^\mu = U A^\mu U^\dagger + \frac{i}{g} U (\partial^\mu U^\dagger) = A^\mu + \frac{1}{g} \partial^\mu \varepsilon = A_T^\mu + \partial^\mu \left(\Theta + \frac{\varepsilon}{g} \right) \quad (2.2.90)$$

and compare to $A'^\mu = A_T'^\mu + \partial^\mu \Theta'$, we see that the gauge transformation only affects A_L^μ :

$$A_T'^\mu = A_T^\mu, \quad \Theta' = \Theta + \frac{\varepsilon}{g}. \quad (2.2.91)$$

In other words, the path integral $\mathcal{D}A$ overcounts physically equivalent degrees of freedom, namely the longitudinal field components that emerge from each other by gauge transformations.

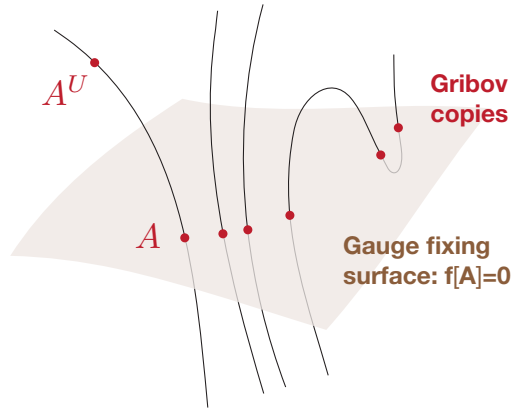


FIG. 2.11: Gauge orbits and gauge-fixing surface.

In QED the problem is cured by the **Gupta-Bleuler method**: The first step is to add a gauge-fixing term as a Lagrange multiplier with a gauge parameter ξ ,

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} - \frac{1}{2\xi}(\partial_\mu A^\mu)^2 \stackrel{\text{p.l.}}{\cong} \frac{1}{2}A_\mu \left(\square g^{\mu\nu} - \partial^\mu \partial^\nu + \frac{1}{\xi} \partial^\mu \partial^\nu \right) A_\nu \quad (2.2.92)$$

so that the inverse propagator picks up a longitudinal part and becomes invertible:

$$D_F^{-1}(p)^{\mu\nu} = ip^2 \left[\left(g^{\mu\nu} - \frac{p^\mu p^\nu}{p^2} \right) + \frac{1}{\xi} \frac{p^\mu p^\nu}{p^2} \right]. \quad (2.2.93)$$

The second step is to impose the transversality condition $\langle \lambda | \partial_\mu A^\mu | \lambda \rangle$ for physical states $|\lambda\rangle$, which has the effect that the longitudinal and timelike photon polarizations cancel each other in S-matrix elements and external photon legs are always transverse. In fact, the need to preserve this feature when switching on interactions is the reason why we impose *local* gauge invariance also for fermions.

Faddeev-Popov gauge fixing. Unfortunately, the situation in QCD is more complicated because gauge transformations mix the transverse and longitudinal field components, and both of them contribute to the gluonic Lagrangian due to the three-gluon and four-gluon interactions.

The analogue of the Gupta-Bleuler method in the non-Abelian case is the **Faddeev-Popov method**. Let us denote a gauge transformation of the gluon field by $A \rightarrow A^U$, where U is some gauge transformation with gauge parameter ε . The basic idea is then to restrict the path integral

$$Z = \int \mathcal{D}A e^{iS[A]} \quad (2.2.94)$$

to the gauge-fixing surface

$$f[A] = \partial_\mu A^\mu = 0, \quad (2.2.95)$$

which is illustrated in Fig. 2.11. The gauge-fixing function $f[A]$ singles out a hypersurface of fixed gauge, so that each gauge field counts only once. In principle the condition $f[A]$ is arbitrary; the choice (2.2.95) corresponds to linear covariant gauges.

How would we implement such a condition? Presumably by inserting a delta function $\delta(f[A])$; however, this would modify the path integral. A better strategy is to insert a quantity that equals 1 but *contains* $\delta(f[A])$. Consider a one-dimensional example: suppose we have an integral over a variable ε (the ‘gauge transformation parameter’) and we want to restrict it to $f(\varepsilon) = 0$. Then the identity

$$\int d\varepsilon \left| \frac{df(\varepsilon)}{d\varepsilon} \right| \delta(f(\varepsilon)) = \int d\varepsilon |f'(\varepsilon_0)| \frac{\delta(\varepsilon - \varepsilon_0)}{|f'(\varepsilon_0)|} = 1 \quad (2.2.96)$$

holds as long as $f(\varepsilon) = 0$ has only one solution ε_0 .

Let us generalize this relation to infinitely many variables, where ε becomes a continuous function $\varepsilon(x)$. In this case, the gauge transformation reads

$$(A^U)^\mu = U A^\mu U^\dagger + \frac{1}{g} U (\partial^\mu U^\dagger), \quad U(x) = e^{i \sum_a \varepsilon_a(x) t_a} \quad (2.2.97)$$

and the gauge-fixing condition is

$$f_a[A^U] = \partial_\mu (A_a^U)^\mu = \partial_\mu \left(A_a^\mu + \frac{1}{g} D_{ab}^\mu \varepsilon_b \right), \quad (2.2.98)$$

where $D_{ab}^\mu = \partial^\mu \delta_{ab} - g f_{abc} A_c^\mu$ is the covariant derivative in the adjoint representation from Eq. (2.1.34). Then the analogue of Eq. (2.2.96) is the ‘functional unity’

$$\int \mathcal{D}U \det M[A] \delta(f[A^U]) = 1. \quad (2.2.99)$$

Here, the path integral

$$\mathcal{D}U = \lim_{n \rightarrow \infty} \prod_{i=1}^n \prod_a \varepsilon_a(x_i) \quad (2.2.100)$$

is the group volume, the δ -function is an infinite product of δ -functions at each space-time point x , and the **Faddeev-Popov operator** $M[A]$ is the functional derivative of the gauge-fixing condition with respect to the gauge transformation parameter:

$$\begin{aligned} M_{ab}[A](x, y) &= \left. \frac{\delta f_a[A^U](x)}{\delta \varepsilon_b(y)} \right|_{f[A^U]=0} \\ &= \frac{1}{g} \partial_\mu D_{ab}^\mu \delta^4(x - y) \\ &= \frac{1}{g} \delta_{ab} \square \delta^4(x - y) - f_{abc} \partial_\mu (A_c^\mu \delta^4(x - y)). \end{aligned} \quad (2.2.101)$$

Note that $M[A]$ is independent of the gauge transformation parameter. In QED, it is also independent of A^μ because the second term with f_{abc} disappears, and thus $\det M[A]$ factorizes from the path integral and can be pulled out.

Now we can insert (2.2.99) in the path integral,

$$Z = \int \mathcal{D}A \int \mathcal{D}U \det M[A] \delta(f[A^U]) e^{iS[A]}, \quad (2.2.102)$$

and, because Z is gauge-invariant, perform a gauge transformation $A^U \rightarrow A$. The gauge field measure $\mathcal{D}A$, the group measure $\mathcal{D}U$, the Faddeev-Popov determinant and the classical action $S[A]$ are all invariant under this operation, so that it merely amounts to replacing $\delta(f[A^U]) \rightarrow \delta(f[A])$. The integrand then no longer depends on U and the group integration $\mathcal{D}U$ factorizes; it produces an infinite constant which drops out whenever we normalize Z , for example when calculating correlation functions. The remaining δ -function restricts the integration over all fields to the hypersurface $f[A] = 0$. Each gauge orbit contributes only one field configuration and we have an integration over physically distinct fields.

The caveat here is that we have assumed the gauge-fixing condition to be unique, like in the one-dimensional example where we assumed that the equation $f(\varepsilon) = 0$ admits only one solution ε_0 . This is usually not the case due to **Gribov copies**: the gauge-fixing condition can intersect the gauge orbits more than once and is therefore not complete (cf. Fig. 2.11); there is a residual gauge freedom. Also in QED $\partial_\mu A^\mu = \square \Theta = 0$ does not fix Θ completely because Eq. (2.2.91) still allows us to perform residual gauge transformations as long as $\square \varepsilon = 0$, but this freedom can be removed by imposing appropriate boundary conditions on the fields. In QCD it has been suggested to restrict the gauge fields to the interior of the Gribov horizon where $\det M[A] > 0$, or possibly even further, and there are indications that the properties close to the Gribov horizon might be related to confinement.

The remaining question is what to do with the Faddeev-Popov determinant and the δ -function. It turns out that one can shuffle both quantities into the exponential and thus into the action $S[A]$. We can take care of the δ -function by changing the gauge fixing condition to $f[A] + \frac{\xi}{2} B = 0$, where $B(x)$ lives in the Lie algebra but does not depend on A . This does not affect the Faddeev-Popov determinant, but the functional integral Z_B now implicitly depends on B . Since any B leads to the same gauge-invariant physics, we can work with Z_B , $Z_{B'}$ or $\int \mathcal{D}B F(B) Z_B$; these are all equivalent. If we integrate over the functions $B(x)$ with some Gaussian weight, we can remove the δ -function in favor of a new term in the action:

$$\begin{aligned} Z &= \int \mathcal{D}B e^{-\frac{i\xi}{8} \int d^4x B^2(x)} \mathcal{D}A \det M[A] \delta\left(f[A] + \frac{\xi}{2} B\right) e^{iS[A]} \\ &= \int \mathcal{D}A \det M[A] e^{i\left(S[A] - \int d^4x \frac{f[A]^2}{2\xi}\right)}. \end{aligned} \quad (2.2.103)$$

With a linear covariant gauge, this provides just the same modification as the Gupta-Bleuler method in QED, Eq. (2.2.92):

$$-\frac{1}{4} F_{\mu\nu}^a F_a^{\mu\nu} - \frac{(\partial_\mu A_\nu^a)^2}{2\xi} \cong \frac{1}{2} A_\mu^a \left(\square g^{\mu\nu} - \partial^\mu \partial^\nu + \frac{1}{\xi} \partial^\mu \partial^\nu \right) A_\nu^a + \dots \quad (2.2.104)$$

As a result, the inverse gluon propagator is no longer transverse and can be inverted. ξ is the gauge parameter: $\xi = 0$ defines the Landau gauge, $\xi = 1$ the Feynman gauge, and in principle there are many other possible choices which differ not only by the gauge parameter but also by the gauge fixing condition (Coulomb gauge, axial gauge, light-cone gauge, maximal Abelian gauge etc.).

Finally, we want to shift the Faddeev-Popov determinant into the action as well. The trick is that the determinant of an operator can be written as a path integral over anticommuting Grassmann fields, cf. Eq. (2.2.78):

$$\det M[A] = \int \mathcal{D}c \mathcal{D}\bar{c} e^{\int d^4x \int d^4y \bar{c}_a(x) M_{ab}[A](x,y) c_b(y)}, \quad (2.2.105)$$

where the **Faddeev-Popov ghosts** $c_a(x)$, $\bar{c}_a(x)$ are scalar but Grassmann-valued fields. They carry the wrong Bose-Fermi statistics, but this is of no concern since they are anyway unphysical — they are just a consequence of fixing the gauge. Thus, the procedure is a generalization of the Gupta-Bleuler method since in QED the Faddeev-Popov determinant is independent of A^μ and can be pulled out of the path integral, so that only the δ -function remains, whereas the non-Abelian case also requires the dynamical inclusion of ghost fields.

The new gauge-fixing terms in the action read

$$\begin{aligned} S_{\text{GF}}[A, c, \bar{c}] &= - \int d^4x \frac{f[A]^2}{2\xi} - i \int d^4x \int d^4y \bar{c}_a(x) M_{ab}[A](x, y) c_b(y) \\ &\stackrel{\text{p.I.}}{\cong} \int d^4x \left[\frac{1}{2\xi} A_\mu^a \partial^\mu \partial^\nu A_\nu^a - \frac{i}{g} \bar{c}_a \square c_a - i f_{abc} (\partial_\mu \bar{c}_a) A_c^\mu c_b \right] \\ &\cong \int d^4x \left[\frac{1}{2\xi} A_\mu^a \partial^\mu \partial^\nu A_\nu^a + \bar{c}_a \square c_a - g f_{abc} (\partial_\mu \bar{c}_a) A_b^\mu c_c \right], \end{aligned} \quad (2.2.106)$$

where in the third line we absorbed the factor $-i/g$ into the antighost field. In conclusion, we obtained a longitudinal term in the gluon propagator, a massless ghost propagator, and a ghost-gluon three-point vertex with coupling constant g .

Reinstating the quarks and including all source terms, the final partition function of QCD takes the form

$$Z[J, \bar{\eta}, \eta, \bar{\sigma}, \sigma] = \int \mathcal{D}A \mathcal{D}\psi \mathcal{D}\bar{\psi} \mathcal{D}c \mathcal{D}\bar{c} e^{i(S[A, \psi, \bar{\psi}] + S_{\text{GF}}[A, c, \bar{c}] + S_C)}. \quad (2.2.107)$$

The source term contains the gluon source J^μ , quark sources η , $\bar{\eta}$ as well as ghost sources σ and $\bar{\sigma}$:

$$S_C = - \int d^4x (J_\mu A^\mu + \bar{\eta} \psi + \bar{\psi} \eta + \bar{\sigma} c + \bar{c} \sigma). \quad (2.2.108)$$

BRST symmetry. There is another, more economical way to arrive at Eq. (2.2.106), which is to impose **BRST invariance** of the action (Becchi, Rouet, Stora, Tyutin). A BRST transformation is defined as an infinitesimal gauge transformation (2.1.37) where the gauge parameter is a ghost field $c(x) = \sum_a c_a(x) \mathbf{t}_a$, i.e., where the c_a are scalar anticommuting Grassmann fields:

$$\delta\psi = ic\psi, \quad \delta\bar{\psi} = -i\bar{\psi}c, \quad \delta A_\mu = \frac{1}{g} D_\mu c, \quad \delta F_{\mu\nu} = i[c, F_{\mu\nu}]. \quad (2.2.109)$$

We further demand this transformation to be nilpotent ($\delta^2 = 0$), so that δ is also Grassmann-valued and anticommutes with c . Then it is straightforward to prove that any of the relations above fixes the transformation behavior of the ghost itself, for example:

$$\delta^2\psi = \delta(ic\psi) = i(\delta c)\psi - ic(\delta\psi) = (i\delta c + c^2)\psi \stackrel{!}{=} 0, \quad (2.2.110)$$

so we have

$$\delta c = ic^2 = ic_a c_b \mathbf{t}_a \mathbf{t}_b = ic_a c_b \left(\frac{1}{2} [\mathbf{t}_a, \mathbf{t}_b] + \frac{1}{2} \{\mathbf{t}_a, \mathbf{t}_b\} \right) = -\frac{1}{2} f_{abc} c_a c_b \mathbf{t}_c \quad (2.2.111)$$

because $c_a c_b = -c_b c_a$ is antisymmetric. (Note that the Grassmann nature of the ghost fields c_a leads to weird relations such as $[c, c] = 2c^2$.) Thus, in components the BRST transformation of the ghost field becomes $\delta c_a = -\frac{1}{2} f_{abc} c_b c_c$.

Applying δ increases the ghost number (the charge corresponding to a $U(1)$ symmetry of the ghost fields) by one unit; hence, when applied to the antighost, it must produce a scalar field with ghost number zero, the so-called Nakanishi-Lautrup field: $\delta \bar{c}_a =: -B_a/g$. Nilpotency of the antighost transformation then fixes $\delta B = 0$. The different treatment of c and \bar{c} implies that they are not conjugates of each other but truly independent fields.

Since the classical action $S[A, \psi, \bar{\psi}]$ is gauge invariant and BRST is a gauge transformation, it is also BRST invariant. The most general BRST-invariant action is then the sum of the classical action plus a term $S_{\text{GF}} = \delta \mathcal{O}$ which is a BRST variation itself, since in that case we have $\delta S_{\text{GF}} = \delta^2 \mathcal{O} = 0$. Adding this to the action means fixing a gauge; which gauge we get depends on \mathcal{O} . To recover (2.2.106), we contract the antighost with our earlier gauge-fixing condition $f[A] + \frac{\xi}{2} B$:

$$\begin{aligned} S_{\text{GF}} &= -g \delta \int d^4x \bar{c}_a \left(f_a[A] + \frac{\xi}{2} B_a \right) \\ &= \int d^4x B_a \left(f_a[A] + \frac{\xi}{2} B_a \right) + g \int d^4x \int d^4y \bar{c}_a(x) M_{ab}[A](x, y) c_b(y). \end{aligned} \quad (2.2.112)$$

Inserting the equations of motion for B_a , namely $f_a + \xi B_a = 0$, yields again Eq. (2.2.106); the same result follows from integrating over B_a in the path integral. Hence, imposing BRST invariance simultaneously generates gauge-fixing and ghost terms in the action.

Ward-Takahashi and Slavnov-Taylor identities. Correlation functions are not gauge invariant, but the gauge invariance of the generating functional (2.2.107) can be used to derive identities for them. As in the derivation of Dyson-Schwinger equations, a gauge transformation under the path integral is just a relabeling of fields, so Z is invariant; also the classical action is gauge invariant. The only gauge-dependent terms in (2.2.107) are then S_{GF} and S_C , and as in the derivation in Eq. (2.2.36), invariance of Z leads to the relation

$$\langle \delta S_{\text{GF}} + \delta S_C \rangle_J = 0, \quad (2.2.113)$$

which represents the generic form of a **Ward-Takahashi identity**. Then, employing Eqs. (2.2.31), (2.2.49) or (2.2.55), one can write this as a master equation for full (with sources J and derivatives $Z'[J]$), connected (J and $W'[J]$) or 1PI correlation functions (with sources φ and derivatives of the effective action, $\Gamma'[\varphi]$), see Sec. 3.1.2 for details.

In the case of non-Abelian gauge theories it is more convenient to exploit BRST invariance. Here δS_{GF} vanishes as well and only the BRST variations of the fields in the source term remain to be evaluated: $\langle \delta S_C \rangle_J = 0$. In the compact notation from earlier:

$$-\langle \delta S_C \rangle_J = \int d^4x \sum_i J_i \langle \delta \phi_i \rangle_J = \int d^4x \sum_i \frac{\delta \Gamma}{\delta \varphi_i} \langle \delta \phi_i \rangle_J = 0, \quad (2.2.114)$$

where the $\delta \phi_i$ are now nonlinear functions of the fields themselves, cf. Eq. (2.2.109). This leads to the **Slavnov-Taylor identities**.

2.3 Renormalization

We are now *almost* in a position to write down the Feynman rules of QCD. In our discussion so far we have still bypassed the problem of renormalization. The need for renormalization is related to the behavior of a theory at infinitely large momenta and in practice arises in the calculation of loop diagrams, which are usually UV-divergent. Below we will see that the problem can be dealt with by introducing a small number of renormalization constants and setting corresponding renormalization conditions, which makes all correlation functions finite. We will also see that *renormalizability* is a deep property of a QFT that can already be read off from the Lagrangian of the theory.

2.3.1 Feynman rules of QCD

Renormalization constants. A possible starting point when dealing with renormalization is to interpret all fields, masses and couplings in the Lagrangian (2.1.29) as ‘bare’ and unphysical, and define their renormalized versions by:

$$\psi_B = Z_\psi^{1/2} \psi, \quad A_B = Z_A^{1/2} A, \quad c_B = Z_c^{1/2} c, \quad m_B = Z_m m, \quad g_B = Z_g g. \quad (2.3.1)$$

The quantities without a subscript are the renormalized ones and they are related to the bare quantities by **renormalization constants**. Then the full Lagrangian of QCD including the gauge-fixing terms becomes

$$\begin{aligned} \mathcal{L}_{\text{QCD}} = & Z_\psi \bar{\psi} (i\not{\partial} - Z_m m) \psi + \frac{1}{2} A_\mu^a \left[Z_A (\square g^{\mu\nu} - \partial^\mu \partial^\nu) + \frac{1}{\xi} \partial^\mu \partial^\nu \right] A_\nu^a + Z_c \bar{c}_a \square c_a \\ & - Z_{3g} \frac{g}{2} f_{abc} (\partial^\mu A_a^\nu - \partial^\nu A_a^\mu) A_\mu^b A_\nu^c - Z_{4g} \frac{g^2}{4} f_{abe} f_{cde} A_a^\mu A_b^\nu A_\mu^c A_\nu^d \\ & + Z_\Gamma g \bar{\psi} \not{A} \psi - \tilde{Z}_\Gamma g f_{abc} (\partial_\mu \bar{c}_a) A_b^\mu c_c. \end{aligned} \quad (2.3.2)$$

The first line contains the tree-level quark, gluon and ghost propagators, the second line the three- and four-gluon interactions, and the third line the quark-gluon and ghost-gluon interaction vertices. The renormalization constants for the vertices are related to those in (2.3.1) by

$$Z_\Gamma = Z_g Z_A^{1/2} Z_\psi, \quad \tilde{Z}_\Gamma = Z_g Z_A^{1/2} Z_c, \quad Z_{3g} = Z_g Z_A^{3/2}, \quad Z_{4g} = Z_g^2 Z_A^2. \quad (2.3.3)$$

In principle we could have different renormalization constants for each term in the Lagrangian, but the Slavnov-Taylor identities ensure that this is not the case. Thus, we have five independent renormalization constants $Z_\psi, Z_A, Z_c, Z_m, Z_g$, which means that at some point we will need to set five *renormalization conditions*.

Moreover, the renormalization constants also enter in the **Feynman rules** since they are derived from the Lagrangian (2.3.2). In the following we write down the Feynman rules for the renormalized propagators and 1PI vertices of QCD.

Before doing so, we note that one could equivalently introduce renormalization constants in the language of **counterterms**:

$$Z_\psi = 1 + \delta Z_\psi, \quad Z_\psi Z_m m = m + \delta m, \quad Z_g g = g + \delta g, \quad \dots \quad (2.3.4)$$

In this way we would split the Lagrangian into two pieces, where the first has the original form but with renormalized fields and without renormalization constants, and the second contains the counterterms which generate new propagators and vertices with new Feynman rules. We will not follow this strategy here and instead absorb the renormalization constants directly in the Feynman rules. Also, note that the renormalization constants in the literature usually go by different names:

$$Z_\psi = Z_2, \quad Z_A = Z_3, \quad Z_c = \tilde{Z}_3, \quad Z_\Gamma = Z_{1f}, \quad \tilde{Z}_\Gamma = \tilde{Z}_1, \quad Z_{3g} = Z_1, \quad Z_{4g} = Z_4. \quad (2.3.5)$$

Quark propagator. The quark propagator is a Dirac matrix with indices α and β , it depends on one momentum p^μ , and it is a diagonal matrix δ_{ij} with $i, j = 1, 2, 3$ in color space. (We ignore flavor since it merely amounts to replicating terms in the Lagrangian.) Since we count the spin indices from the top of the arrow, i.e. from left to right, we also let the momentum flow from right to left. Writing $Z_m m = m_B$, where m is the renormalized current-quark mass, the inverse tree-level quark propagator from the Lagrangian is then given by (we suppress the color indices on the l.h.s.)

$$S_0^{-1}(p) = -iZ_\psi (\not{p} - m_B) \delta_{ij},$$

$$S_0(p) = \frac{i}{Z_\psi} \frac{\not{p} + m_B}{p^2 - m_B^2 + i\epsilon} \delta_{ij}. \quad (2.3.6)$$

Can we also write down a ‘Feynman rule’ for the *full* propagator $S(p)$? In general, any n -point correlation function can be expanded in a tensor basis

$$G_{\alpha\beta\dots}^{\mu\nu\dots}(p_1, \dots, p_n) = \sum_{i=1}^N f_i(p_1^2, p_2^2, \dots) \tau_i(p_1, \dots, p_n)_{\alpha\beta\dots}^{\mu\nu\dots}, \quad (2.3.7)$$

where the τ_i are Lorentz-covariant tensors that inherit the Lorentz and Dirac structure of G . The f_i are Lorentz-invariant **dressings functions** (‘form factors’), which depend on all possible Lorentz-invariant momentum variables — they contain the physical information encoded in the correlation function. Like G itself, the basis elements transform under finite-dimensional representations of the Lorentz group,

$$\tau_i(p'_1, \dots, p'_n)_{\alpha\beta\dots}^{\mu\nu\dots} = \Lambda^\mu_{\mu'} \Lambda^\nu_{\nu'} \dots D_{\alpha\alpha'}(\Lambda) \tau_i(p_1, \dots, p_n)_{\alpha'\beta'\dots}^{\mu'\nu'\dots} D_{\beta'\beta}^{-1}(\Lambda), \quad (2.3.8)$$

where Λ is the Lorentz transformation and $D(\Lambda)$ its spinor representation matrix (see Appendix B). The same formula holds for parity transformations if $D(\Lambda)$ is replaced by γ^0 . In practice, this means that the tensors are constructed by combining

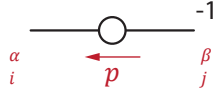
$$g^{\mu\nu}, \quad \varepsilon^{\mu\nu\alpha\beta}, \quad \mathbf{1}, \quad \gamma^\mu, \quad \gamma_5 \gamma^\mu, \quad \sigma^{\mu\nu}, \quad \gamma_5 \sigma^{\mu\nu} \quad (2.3.9)$$

with the four-momenta in the system.

The quark propagator depends on only one momentum,

$$S_{\alpha\beta}(p) = \sum_{i=1}^2 f_i(p^2) \tau_i(p)_{\alpha\beta}, \quad (2.3.10)$$

and from (2.3.9) we can only construct the two tensors $\mathbb{1}$ and \not{p} since those with γ_5 would have the wrong parity. Thus, the full quark propagator can be written as



$$S^{-1}(p) = -iA(p^2) (\not{p} - M(p^2)) \delta_{ij},$$

$$S(p) = \frac{i}{A(p^2)} \frac{\not{p} + M(p^2)}{p^2 - M(p^2)^2 + i\epsilon} \delta_{ij}. \quad (2.3.11)$$

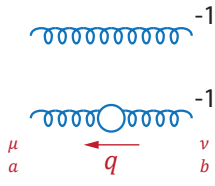
Here we defined the **quark mass function** $M(p^2)$, and the inverse of $A(p^2)$ is called the quark ‘wave-function renormalization’ $Z_f(p^2) = 1/A(p^2)$. If we knew these two functions for all $p^2 \in \mathbb{C}$ (recall the discussion around Fig. 2.5), we would know the full quark propagator in QCD. To project out the dressing functions, we take Dirac traces:

$$M(p^2)A(p^2) = -\frac{i}{4} \text{Tr} \{S^{-1}(p)\}, \quad A(p^2) = \frac{i}{4p^2} \text{Tr} \{\not{p} S^{-1}(p)\}. \quad (2.3.12)$$

Gluon propagator. The gluon propagator depends on two Lorentz indices μ, ν and one momentum q ; from this we can only form the two tensors $g^{\mu\nu}$ and $q^\mu q^\nu$. It is useful to define the transverse and longitudinal projectors as their linear combinations:

$$T_q^{\mu\nu} = g^{\mu\nu} - \frac{q^\mu q^\nu}{q^2}, \quad L_q^{\mu\nu} = \frac{q^\mu q^\nu}{q^2}. \quad (2.3.13)$$

Then the Feynman rules for the tree-level and full gluon propagator are



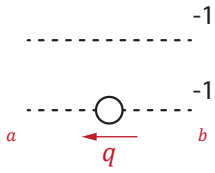
$$(D_0^{-1})^{\mu\nu}(q) = iq^2 \left[Z_A T_q^{\mu\nu} + \frac{1}{\xi} L_q^{\mu\nu} \right] \delta_{ab},$$

$$(D^{-1})^{\mu\nu}(q) = iq^2 \left[\frac{1}{Z(q^2)} T_q^{\mu\nu} + \frac{1}{\xi} L_q^{\mu\nu} \right] \delta_{ab}, \quad (2.3.14)$$

$$D^{\mu\nu}(q) = -\frac{i}{q^2 + i\epsilon} [Z(q^2) T_q^{\mu\nu} + \xi L_q^{\mu\nu}] \delta_{ab},$$

where $Z(q^2)$ is the **gluon dressing function**. (In principle the longitudinal part could also pick up a dressing, but the Slavnov-Taylor identity prevents this and ensures that the longitudinal part remains undressed.) The gluon is color-diagonal with $a, b = 1 \dots 8$.

Ghost propagator. The ghost propagator is scalar and thus the simplest case, since it has no tensor structure and there is only one ghost dressing function:



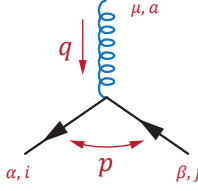
$$D_{G,0}^{-1}(q) = iq^2 Z_c \delta_{ab},$$

$$D_G^{-1}(q) = iq^2 G(q^2)^{-1} \delta_{ab}, \quad (2.3.15)$$

$$D_G(q) = -\frac{i}{q^2 + i\epsilon} G(q^2) \delta_{ab}.$$

Note that if we had not absorbed the minus sign into the antighost field in the third line of Eq. (2.2.106), the Feynman rules for the ghost propagator and ghost-gluon vertex would come with additional minus signs.

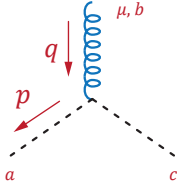
Quark-gluon vertex. The Feynman rules for the quark-gluon vertex are



$$\begin{aligned}\Gamma_0^\mu &= ig (\mathbf{t}_a)_{ij} Z_\Gamma \gamma^\mu, \\ \Gamma^\mu(p, q) &= ig (\mathbf{t}_a)_{ij} \sum_{i=1}^{12} f_i(p^2, q^2, p \cdot q) \tau_i^\mu(p, q).\end{aligned}\quad (2.3.16)$$

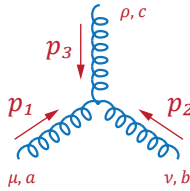
The full vertex becomes rather complicated since it depends on two independent momenta p and q . This leads to 12 possible tensors that are allowed by Lorentz covariance: $\gamma^\mu, p^\mu, q^\mu, [\gamma^\mu, \not{p}], \dots$, and the dressing functions depend on the three Lorentz invariants p^2, q^2 and $p \cdot q$. Since the vertex has a charge-conjugation symmetry, it is convenient to identify p with the average momentum between the incoming and outgoing quarks because this makes the symmetry manifest in the dressing functions (the dependence on $p \cdot q$ is then either even or odd).

Ghost-gluon vertex. The ghost-gluon vertex has no Dirac structure and therefore only two tensors p^μ and q^μ . In this case the tree-level vertex depends on the *outgoing* momentum p^μ because in the Lagrangian (2.3.2) the derivative acts on \bar{c}_a (i.e., the ghost and antighost fields are not related by charge conjugation):



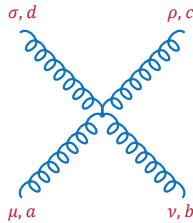
$$\begin{aligned}\Gamma_{\text{gh},0}^\mu(p) &= gf_{abc} \tilde{Z}_\Gamma p^\mu, \\ \Gamma_{\text{gh}}^\mu(p, q) &= gf_{abc} \sum_{i=1}^2 \tilde{f}_i(p^2, q^2, p \cdot q) \tau_i^\mu(p, q).\end{aligned}\quad (2.3.17)$$

Three-gluon vertex. Here things get a bit more cumbersome since the full vertex has 14 possible Lorentz tensors. The tree-level vertex (with $p_1 + p_2 + p_3 = 0$) reads:



$$\begin{aligned}\Gamma_{3g,0}^{\mu\nu\rho}(p_1, p_2, p_3) &= gf_{abc} Z_{3g} [(p_1 - p_2)^\rho g^{\mu\nu} \\ &\quad + (p_2 - p_3)^\mu g^{\nu\rho} + (p_3 - p_1)^\nu g^{\rho\mu}].\end{aligned}\quad (2.3.18)$$

Four-gluon vertex. In this case things get *really* cumbersome: The full vertex has 136 linearly independent Lorentz tensors and five color structures. The tree-level vertex is momentum-independent:



$$\begin{aligned}\Gamma_{4g,0}^{\mu\nu\rho\sigma} &= -ig^2 Z_{4g} \left[f_{abe} f_{cde} (g^{\mu\rho} g^{\nu\sigma} - g^{\nu\rho} g^{\mu\sigma}) \right. \\ &\quad + f_{ace} f_{bde} (g^{\mu\nu} g^{\rho\sigma} - g^{\nu\rho} g^{\mu\sigma}) \\ &\quad \left. + f_{ade} f_{cbe} (g^{\mu\rho} g^{\nu\sigma} - g^{\mu\nu} g^{\rho\sigma}) \right].\end{aligned}\quad (2.3.19)$$

One-loop perturbation theory. With the Feynman rules at hand, we are ready to set up perturbation theory. To avoid redundancies, we will do this for the 1PI correlation functions, i.e. we set up the perturbative expansion for the *inverse* propagators, in the same way as we wrote the Dyson-Schwinger equations in Eq. (2.2.62) and thereafter. Then we only need to work out the self-energy diagrams, whereas the expansion for the propagator is easily obtained from Eq. (2.2.63) if needed.

The DSE for the **quark propagator** has the generic form

$$S^{-1}(p) = S_0^{-1}(p) - i\Sigma(p), \quad (2.3.20)$$

where the quark self-energy $\Sigma(p)$ contains only one diagram at one-loop order:

If we write $\Sigma(p) = \Sigma_A(p^2) \not{p} - \Sigma_M(p^2)$ and insert Eqs. (2.3.6) and (2.3.11), we read off the relations for the two scalar dressing functions:

$$\begin{aligned} A(p^2) &= Z_\psi + \Sigma_A(p^2), \\ M(p^2)A(p^2) &= Z_\psi Z_m m + \Sigma_M(p^2). \end{aligned} \quad (2.3.21)$$

We will later see that the renormalization constants have the structure $Z = 1 + \mathcal{O}(g^2)$, and since the self-energy comes with a factor g^2 , the mass function up to $\mathcal{O}(g^2)$ is

$$M(p^2) = Z_m m + \Sigma_M(p^2) - m \Sigma_A(p^2). \quad (2.3.22)$$

We will work out the self-energy explicitly in Sec. 2.3.2.

The DSE for the **gluon propagator** is given by

$$(D^{-1})^{\mu\nu}(q) = (D_0^{-1})^{\mu\nu}(q) - i\Pi^{\mu\nu}(q), \quad (2.3.23)$$

where $\Pi^{\mu\nu}(q)$ is the gluon vacuum polarization. At one-loop order $\mathcal{O}(g^2)$ it consists of a quark loop, a gluon loop, a ghost loop, and a tadpole diagram:

We can split the vacuum polarization into two terms,

$$\Pi^{\mu\nu}(q) = \Pi(q^2) (q^2 g^{\mu\nu} - q^\mu q^\nu) + \tilde{\Pi}(q^2) g^{\mu\nu} = \Pi(q^2) q^2 T_q^{\mu\nu}, \quad (2.3.24)$$

where only the first survives because the Slavnov-Taylor identity entails $q_\mu \Pi^{\mu\nu}(q) = 0$ and therefore $\tilde{\Pi}(q^2) = 0$. (The one-loop result for $\tilde{\Pi}(q^2)$ indeed vanishes in dimensional regularization, but it is non-zero for a cutoff regulator which breaks gauge invariance.) Inserting Eq. (2.3.14) into the DSE, we see that there are no loop corrections for the longitudinal part, which is also why no renormalization is required for this term. The equation then simply becomes

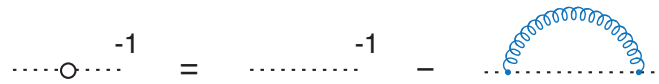
$$Z(q^2)^{-1} = Z_A - \Pi(q^2). \quad (2.3.25)$$

In the analogous case of QED, after renormalization $\Pi(q^2)$ becomes constant for $q^2 \rightarrow 0$, which means that the photon propagator has a massless $1/q^2$ pole and the photon remains massless also with interactions. In QCD, this is still what happens in perturbation theory but it may no longer be true non-perturbatively. Early ideas suggested a $1/q^4$ pole for the gluon ‘propagator’ $Z(q^2)/q^2$ since this would signal confinement: if one connects a quark and antiquark by a gluon, the three-dimensional Fourier transform of $1/|\mathbf{q}|^4$ leads to a potential $\propto |\mathbf{r}|$ in coordinate space simply by dimensional counting. Nowadays evidence from non-perturbative (lattice and functional) calculations in Landau gauge suggests that this is not what happens and that $Z(q^2)$ instead *vanishes* at $q^2 = 0$, either with a power q^2 (‘**massive**’ or ‘**decoupling**’ scenario) or higher (‘**scaling**’ scenario). As a result, $Z(q^2)/q^2$ becomes constant or even has a turnover in the infrared. Vice versa, $Z(q^2)^{-1}$ and therefore $\Pi(q^2)$ must be singular at $q^2 \rightarrow 0$, but the origin of this singularity is still under debate. Moreover, in the scaling scenario the infrared exponents for any quark-antiquark interaction diagram still match to produce a $1/q^4$ behavior (e.g., for the combination of a gluon propagator and two quark-gluon vertices)³, whereas in the massive scenario (which is supported by lattice calculations) this is not the case.

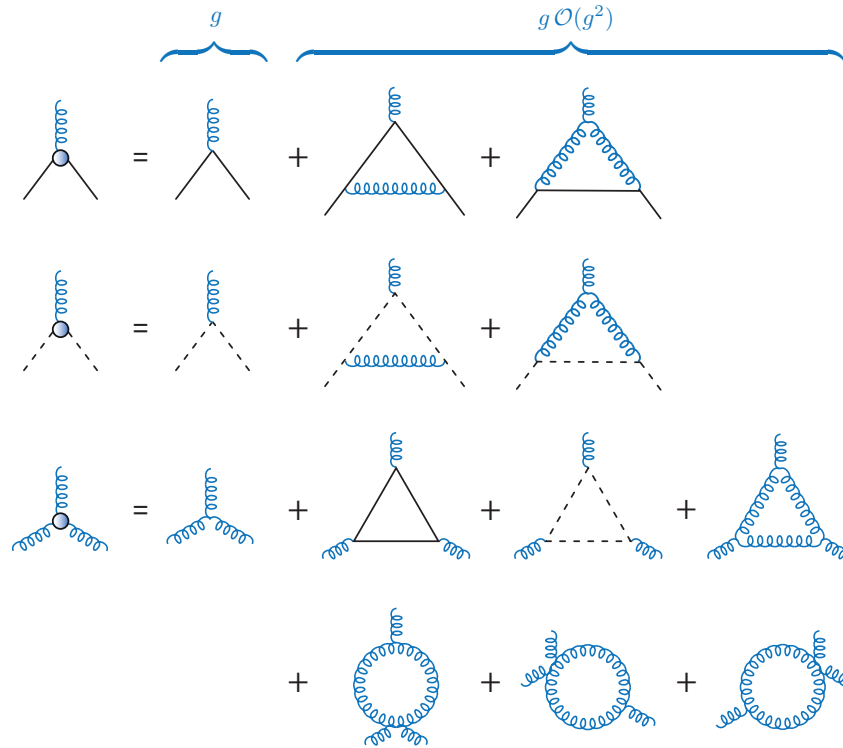
The DSE for the **ghost propagator** reads

$$D_G^{-1}(q) = D_{G,0}^{-1}(q) - iq^2 \Sigma_G(q^2) \quad \Rightarrow \quad G(q^2)^{-1} = Z_c - \Sigma_G(q^2), \quad (2.3.26)$$

where the perturbative expansion of the self-energy is analogous to the quark:

$$\text{---}\circ\text{---}^{-1} = \text{---}\text{---}^{-1} - \text{---}\text{---}\text{---}$$


Finally, the one-loop expressions of the quark-gluon, ghost-gluon and three-gluon **vertices** have the form (note that a factor g is implicit in the vertices):

$$\begin{array}{l} \underbrace{\phantom{\text{---}\circ\text{---}}}_{g} \quad \underbrace{\phantom{\text{---}\text{---}}}_{g \mathcal{O}(g^2)} \\ \text{---}\circ\text{---} = \text{---}\text{---} + \text{---}\text{---}\text{---} + \text{---}\text{---}\text{---} \\ \text{---}\circ\text{---} = \text{---}\text{---} + \text{---}\text{---}\text{---} + \text{---}\text{---}\text{---} \\ \text{---}\circ\text{---} = \text{---}\text{---}\text{---} + \text{---}\text{---}\text{---} + \text{---}\text{---}\text{---} \\ \phantom{\text{---}\circ\text{---}} + \text{---}\text{---}\text{---} + \text{---}\text{---}\text{---} + \text{---}\text{---}\text{---} \end{array}$$


³R. Alkofer, C. S. Fischer, F. J. Llanes-Estrada, Mod. Phys. Lett. A 23 (2008) 1105, [hep-ph/0607293](https://arxiv.org/abs/hep-ph/0607293).

2.3.2 Regularization and renormalization

In practice the diagrams we just drew are UV-divergent. The first step in dealing with this problem is regularization, which means to *isolate* the divergences. In the second step we *remove* the divergences; as we will see, there is a systematic procedure behind it, namely renormalization.

Feynman parameters. But first of all we must bring the integrals into a manageable form. To do so, we use the ‘Feynman trick’, where we write the quantity $1/(A_1 \dots A_n)$ as an integral over Feynman parameters x_1, \dots, x_n :

$$\frac{1}{A_1 \dots A_n} = \int \underbrace{dx_1 \dots dx_n}_{d\Omega_n} \delta\left(\sum_{i=1}^n x_i - 1\right) \frac{(n-1)!}{\left(\sum_{i=1}^n x_i A_i\right)^n}. \quad (2.3.27)$$

This is the generalization of the identity

$$\int_0^1 dx \int_0^1 dy \frac{\delta(x+y-1)}{(xA+yB)^2} = \int_0^1 dx \frac{1}{[xA+(1-x)B]^2} = -\frac{1}{A-B} \frac{1}{xA+(1-x)B} \Big|_0^1 = \frac{1}{AB}. \quad (2.3.28)$$

The integral measure $d\Omega_n$ for $n=2$ and $n=3$ is

$$\begin{aligned} \int d\Omega_2 \dots &= \int_0^1 dx \dots \Big|_{y=1-x}, \\ \int d\Omega_3 \dots &= \int_0^1 dx_1 \int_0^{1-x_1} dx_2 \dots \Big|_{x_3=1-x_1-x_2} = \frac{1}{2} \int_0^1 da \int_{-a}^a db \dots \end{aligned} \quad (2.3.29)$$

Here we set $a = x_1 + x_2 = 1 - x_3$ and $b = x_1 - x_2$, which is convenient since the integral over b is antisymmetric and thus only even terms in b survive. Similar expressions hold for $d\Omega_4, d\Omega_5$ etc.

Now consider a generic one-loop diagram L_n which has n propagators in the loop. If we write $A_i = (k + p_i)^2 - m_i^2 + i\epsilon$, where k is the loop momentum and the p_i are external momenta, its structure will always be the same irrespective of the theory:

$$L_n = \int \frac{d^4k}{(2\pi)^4} \frac{(\dots)}{\prod_{i=1}^n A_i} = (n-1)! \int d\Omega_n \int \frac{d^4k}{(2\pi)^4} \frac{(\dots)}{\left(\sum_{i=1}^n x_i A_i\right)^n}. \quad (2.3.30)$$

The numerator (\dots) can have Lorentz and Dirac indices and in general it also depends on k and p_i . If we define a new loop momentum l by

$$l = k + \sum_{i=1}^n x_i p_i, \quad (2.3.31)$$

then with $\sum_i x_i = 1$ it is easy to show that

$$\sum_{i=1}^n x_i A_i = l^2 - \Delta + i\epsilon, \quad \Delta = \left(\sum_i x_i p_i\right)^2 - \sum_i x_i (p_i^2 - m_i^2), \quad (2.3.32)$$

where Δ does not depend on l but only on the external momenta p_i and the Feynman parameters x_i . Thus we obtain

$$L_n = (n-1)! \int d\Omega_n \int \frac{d^4 l}{(2\pi)^4} \frac{(\dots)}{(l^2 - \Delta + i\epsilon)^n}. \quad (2.3.33)$$

Finally, we perform a Wick rotation (see Appendix C)

$$l_4 = il_0 \quad \Rightarrow \quad l^2 = -l_4^2 - \mathbf{l}^2 = -l_E^2, \quad \int d^4 l = -i \int d^3 l \int_{-\infty}^{\infty} dl_4 = i \int d^4 l_E \quad (2.3.34)$$

to arrive at the Euclidean integral

$$L_n = i(-1)^n (n-1)! \int d\Omega_n \int \frac{d^4 l_E}{(2\pi)^4} \frac{(\dots)}{(l_E^2 + \Delta)^n}. \quad (2.3.35)$$

Usually the hardest part is to work out the numerator, where we also have to express k in terms of l and the p_i through Eq. (2.3.31). In doing so, it will depend on powers of the loop momentum l^μ . What helps is that integrals over odd powers vanish by symmetry (replace $l^\mu \rightarrow -l^\mu$), e.g.

$$\int \frac{d^4 l_E}{(2\pi)^4} \frac{l^\mu}{(l_E^2 + \Delta)^2} = 0, \quad (2.3.36)$$

whereas even powers can always be reduced to integrals of the form

$$I_{nm} = \int \frac{d^4 l_E}{(2\pi)^4} \frac{(l_E^2)^m}{(l_E^2 + \Delta)^n}. \quad (2.3.37)$$

For example,

$$\int \frac{d^4 l_E}{(2\pi)^4} \frac{l^\mu l^\nu}{(l_E^2 + \Delta)^2} = -\frac{1}{4} g^{\mu\nu} \int \frac{d^4 l_E}{(2\pi)^4} \frac{l_E^2}{(l_E^2 + \Delta)^2} \quad (2.3.38)$$

because for $\mu \neq \nu$ the integral vanishes again by symmetry, whereas for $\mu = \nu$ it must be proportional to $g^{\mu\nu}$ by Lorentz invariance. The prefactors are then determined by contracting the indices on both sides, using $l^2 = -l_E^2$ and $\delta_\mu^\mu = 4$ (note that in d dimensions one has $\delta_\mu^\mu = d$, so the prefactor on the r.h.s. becomes $-1/d$).

As a consequence, the numerator under the integral in Eq. (2.3.35) can be written as $(\dots) = \sum_m (\dots)_m (l_E^2)^m$, and the integral becomes

$$L_n = i(-1)^n (n-1)! \int d\Omega_n \sum_m (\dots)_m I_{nm}. \quad (2.3.39)$$

Dimensional regularization. The remaining task is to work out the integrals I_{nm} , which are divergent for $n - m \leq 2$. The idea of **regularization** is to isolate the divergent pieces and write the expressions as a sum of finite and divergent terms. In the following we use dimensional regularization, where we generalize the $d^4 l$ integral to d dimensions:

$$I_{nm}^{(d)} = \frac{1}{M^{d-4}} \int \frac{d^d l_E}{(2\pi)^d} \frac{(l_E^2)^m}{(l_E^2 + \Delta)^n}. \quad (2.3.40)$$

To preserve the mass dimension, we put an (arbitrary) mass scale M in front of the integral. This seemingly innocuous operation has profound consequences, namely: *Regularization always introduces a scale*. When splitting the integrals into finite and divergent pieces, the finite terms still depend on this scale, which cannot be removed.

There are many different ways to regularize the theory: instead of dimensional regularization, which is convenient for perturbative calculations, we could also

- introduce a hard momentum cutoff $\int_0^\infty dl_E^2 \rightarrow \int_0^{\Lambda^2} dl_E^2$, which unfortunately breaks gauge invariance;
- use Pauli-Villars regularization, where we subtract each propagator by another propagator with a large mass M ,
- or use a lattice regularization, where we discretize spacetime and introduce a lattice spacing a .

In all these cases we end up with an arbitrary mass scale in the theory: the mass M in dimensional or Pauli-Villars regularization, the cutoff Λ , or the inverse lattice spacing $1/a$. Later we will see that we can trade the dependence on this scale for a dependence on an arbitrary renormalization point. Even for the massless QCD Lagrangian, which has no intrinsic scale and is therefore scale invariant, regularization introduces a scale. (And fortunately so, because if we were to compute the hadron spectrum of massless QCD, we would otherwise expect all hadrons to be massless since nothing sets the scale.) This is also called **anomalous breaking of scale invariance**, since an anomaly is a symmetry of the classical action that is broken at the quantum level.

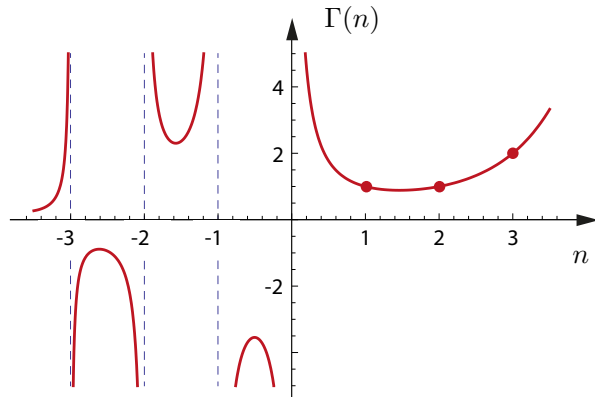
Moving on with dimensional regularization, we do not repeat the calculation for the integral (2.3.40) (which can be found in QFT textbooks) but only quote its result:

$$I_{nm}^{(d)} = \frac{1}{M^{d-4}} \frac{1}{(4\pi)^{d/2}} \frac{1}{\Gamma(n)} \frac{1}{\Delta^{n-m-d/2}} \frac{\Gamma(\frac{d}{2} + m)}{\Gamma(\frac{d}{2})} \Gamma(n - m - \frac{d}{2}) \quad (2.3.41)$$

$$\stackrel{d=4-\varepsilon}{=} \frac{1}{(4\pi)^2} \frac{\Gamma(m + 2 - \frac{\varepsilon}{2})}{\Gamma(n) \Gamma(2 - \frac{\varepsilon}{2})} \frac{1}{\Delta^{n-m-2}} \left(\frac{4\pi M^2}{\Delta} \right)^{\varepsilon/2} \Gamma(n - m - 2 + \frac{\varepsilon}{2}).$$

$\Gamma(n)$ is the Gamma function, which provides an analytic continuation of the result for arbitrary values of d . It has the properties

- $\Gamma(n) = \int_0^\infty dx x^{n-1} e^{-x}$,
- $\Gamma(n) = (n-1)!$ for $n \in \mathbb{N}_+$,
- $\Gamma(n)$ has poles at $n = 0, -1, -2, \dots$
- $\Gamma(n+1) = n \Gamma(n)$,
- $\Gamma'(1) = -\gamma = -0.5772\dots$ is the Euler-Mascheroni constant.



For $\varepsilon \rightarrow 0$ and thus $d \rightarrow 4$, one can see that (2.3.41) is divergent for $n - m - 2 \leq 0$. In this case we can use

$$\Gamma\left(\frac{\varepsilon}{2}\right) = \frac{2}{\varepsilon} - \gamma + \mathcal{O}(\varepsilon), \quad x^{\varepsilon/2} = 1 + \frac{\varepsilon}{2} \ln x + \mathcal{O}(\varepsilon^2) \quad (2.3.42)$$

to obtain the convergent integrals

$$\{I_{30}, I_{40}, I_{41}, \dots\} = \frac{1}{(4\pi)^2} \left\{ \frac{1}{2\Delta}, \frac{1}{6\Delta^2}, \frac{1}{3\Delta}, \dots \right\}. \quad (2.3.43)$$

The divergent integrals are given by

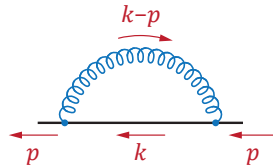
$$\{I_{20}, I_{31}, I_{42}, \dots\} = \frac{1}{(4\pi)^2} \left\{ D, D - \frac{1}{2}, D - \frac{5}{6}, \dots \right\} \quad (2.3.44)$$

with

$$D = \frac{2}{\varepsilon} - \gamma + \ln \frac{4\pi M^2}{\Delta} + \mathcal{O}(\varepsilon). \quad (2.3.45)$$

Be careful with the limit $\varepsilon \rightarrow 0$ for the divergent terms: also $\mathcal{O}(\varepsilon)$ terms must be kept in the calculation since they combine with the $1/\varepsilon$ term to give a finite contribution. In conclusion, we have managed to split the integrals into divergent pieces, where the UV divergences appear in the form of $1/\varepsilon$ terms, and finite pieces which depend on the arbitrary mass scale M .

Quark self-energy. Let us work out a concrete example, namely the quark self-energy from Eq. (2.3.20):



Using the Feynman rules, it reads explicitly:

$$\begin{aligned} i\Sigma(p) &= \int \frac{d^4 k}{(2\pi)^4} (ig\gamma_\mu) S_0(k) (ig\gamma_\nu) D_0^{\mu\nu}(k-p) \left(\sum_a \mathbf{t}_a \mathbf{t}_a \right)_{ij} \\ &= -g^2 C_F \delta_{ij} \int \frac{d^4 k}{(2\pi)^4} \frac{\gamma^\mu (\not{k} + m) \gamma_\mu}{[(k-p)^2 + i\epsilon][k^2 - m^2 + i\epsilon]}. \end{aligned} \quad (2.3.46)$$

Here we employed the gluon propagator in Feynman gauge ($\xi = 1$), the color factor is $C_F = (N_c^2 - 1)/(2N_c)$, and we ignored the renormalization constants multiplying the self-energy since they have the structure $Z = 1 + \mathcal{O}(g^2)$ and will thus only contribute to higher orders in perturbation theory.

The integral is of the form (2.3.30) with $p_1 = -p$, $m_1 = 0$, $p_2 = 0$ and $m_2 = m$. Therefore, we have

$$\begin{aligned} l &= k + \sum_i x_i p_i = k - xp, \\ \Delta &= x^2 p^2 - xp^2 - (1-x)(-m^2) = (1-x)(m^2 - xp^2). \end{aligned} \quad (2.3.47)$$

Denoting the denominator by $\mathcal{N} = \gamma^\mu (\not{k} + m) \gamma_\mu$ and removing the unit matrix δ_{ij} in color space, we can immediately use the result (2.3.35):

$$i\Sigma(p) = -g^2 C_F \frac{i}{M^{d-4}} \int_0^1 dx \int \frac{d^d l_E}{(2\pi)^d} \frac{\mathcal{N}}{(l_E^2 + \Delta)^2}, \quad (2.3.48)$$

which we already generalized to d dimensions.

To work out the numerator, use $\{\gamma^\mu, \gamma^\nu\} = 2g^{\mu\nu}$ and $\gamma^\mu \gamma_\mu = \delta_\mu^\mu = d$ in d dimensions. This gives

$$\begin{aligned} \mathcal{N} &= \gamma^\mu (\not{k} + m) \gamma_\mu = -\gamma^\mu \gamma_\mu \not{k} + 2\not{k} + m\gamma^\mu \gamma_\mu \\ &= (2-d)\not{k} + md \\ &= (2-d)\not{l} + (2-d)x\not{p} + md. \end{aligned} \quad (2.3.49)$$

The first term in the last line is odd in l^μ , so it vanishes after integration according to Eq. (2.3.36), whereas the remainder is independent of the loop momentum and can be pulled out of the integral. We obtain

$$i\Sigma(p) = -ig^2 C_F \int_0^1 dx I_{20}^{(d)} [(2-d)x\not{p} + md], \quad (2.3.50)$$

and if we split the self-energy into $\Sigma(p) = \Sigma_A(p^2)\not{p} - \Sigma_M(p^2)$ we can read off the scalar expressions:

$$\begin{aligned} \Sigma_A(p^2) &= g^2 C_F (d-2) \int dx x I_{20}^{(d)}, \\ \Sigma_M(p^2) &= g^2 C_F md \int dx I_{20}^{(d)}. \end{aligned} \quad (2.3.51)$$

Setting $d = 4 - \varepsilon$ and taking $\varepsilon \rightarrow 0$, with $I_{20} = D/(4\pi)^2$ and $\alpha = g^2/(4\pi)$, we finally arrive at

$$\begin{aligned} \Sigma_A(p^2) &= \frac{\alpha}{2\pi} C_F \int dx x \left(\frac{2}{\varepsilon} - \gamma + \ln \frac{4\pi M^2}{\Delta} - 1 \right), \\ \Sigma_M(p^2) &= \frac{\alpha m}{\pi} C_F \int dx \left(\frac{2}{\varepsilon} - \gamma + \ln \frac{4\pi M^2}{\Delta} - \frac{1}{2} \right). \end{aligned} \quad (2.3.52)$$

In conclusion, we have split the quark self-energy into divergent and finite pieces. But what are we supposed to do with the divergences — throw them away? How would that make any sense? Surprisingly enough, this is indeed what eventually has to happen, but there is a deeper underlying reason which can be understood in the course of renormalization.

Renormalization. The basic idea is the following and can be motivated from QED. There, the full fermion propagator should have a pole at $p^2 = m^2$, where it returns to a free propagator but with the *physical* mass m . Thus we could impose

$$S(p) \xrightarrow{p^2=m^2} \frac{i(\not{p} + m)}{p^2 - m^2 + i\epsilon} \quad \Rightarrow \quad \begin{aligned} A(p^2 = m^2) &\stackrel{!}{=} 1 \\ M(p^2 = m^2) &\stackrel{!}{=} m. \end{aligned} \quad (2.3.53)$$

These are two conditions, where one fixes the pole position and the other the residue of the propagator. They correspond to an **onshell renormalization**; likewise, we would demand that the photon dressing function becomes $Z(q^2 = 0) = 1$ at the onshell point.

In QCD it would not make much sense to impose such conditions, since there are no free quarks and gluons due to confinement. Fortunately, it turns out that these renormalization conditions are arbitrary and thus we can generalize them to an arbitrary **renormalization point** μ :

$$S(p) \xrightarrow{p^2=\mu^2} \frac{i(\not{p} + m)}{p^2 - m^2 + i\epsilon} \Big|_{p^2=\mu^2} \quad \Rightarrow \quad \begin{aligned} A(p^2 = \mu^2) &\stackrel{!}{=} 1, \\ M(p^2 = \mu^2) &\stackrel{!}{=} m, \end{aligned} \quad (2.3.54)$$

$$Z(q^2 = \mu^2) \stackrel{!}{=} 1, \quad G(q^2 = \mu^2) \stackrel{!}{=} 1, \quad \Gamma_{\text{gh}}^\mu(p^2 = \mu^2) \stackrel{!}{=} g f_{abc} p^\mu.$$

Here we imposed five conditions, four for the quark, gluon and ghost dressing functions and one for the ghost-gluon vertex. (We could have chosen any other vertex, and in fact we could have even chosen different renormalization points for each correlation function, but let's keep matters simple.)

The effect is that these five conditions determine the five renormalization constants in Eq. (2.3.1). For example, for the quark propagator we obtain according to (2.3.21):

$$\begin{aligned} A(\mu^2) = Z_\psi + \Sigma_A(\mu^2) &\stackrel{!}{=} 1 & \Rightarrow & \quad Z_\psi = 1 - \Sigma_A(\mu^2), \\ M(\mu^2)A(\mu^2) = Z_\psi Z_m m + \Sigma_M(\mu^2) &\stackrel{!}{=} m & \Rightarrow & \quad Z_\psi Z_m = 1 - \frac{\Sigma_M(\mu^2)}{m}. \end{aligned} \quad (2.3.55)$$

This is, by the way, also the reason why we could set the renormalization constants attached to the one-loop quark self-energy in Eq. (2.3.46) to 1, since the remaining contributions would only enter at higher loop orders. As a result, the renormalized dressing functions are finite because the $1/\epsilon$ divergences drop out:

$$\begin{aligned} A(p^2) &= 1 + \Sigma_A(p^2) - \Sigma_A(\mu^2) = 1 + \frac{\alpha}{2\pi} C_F \int dx x \ln \frac{m^2 - x\mu^2}{m^2 - xp^2}, \\ M(p^2)A(p^2) &= m + \Sigma_M(p^2) - \Sigma_M(\mu^2) = m + \frac{\alpha m}{\pi} C_F \int dx \ln \frac{m^2 - x\mu^2}{m^2 - xp^2}. \end{aligned} \quad (2.3.56)$$

The resulting mass function up to $\mathcal{O}(\alpha)$ is

$$M(p^2) = m \left[1 + \frac{\alpha}{\pi} C_F \int dx \left(1 - \frac{x}{2} \right) \ln \frac{m^2 - x\mu^2}{m^2 - xp^2} \right] + \mathcal{O}(\alpha^2), \quad (2.3.57)$$

which for $p^2, \mu^2 \gg m^2$ becomes (we will return to this at the end of Sec. 2.3.3)

$$M(p^2) \approx m \left[1 - \frac{3\alpha}{4\pi} C_F \ln \frac{p^2}{\mu^2} \right]. \quad (2.3.58)$$

We could repeat the procedure to determine the one-loop results for the remaining propagators and vertices and in all cases the $1/\varepsilon$ divergences would drop out as well. As a result, imposing five renormalization conditions determines the five renormalization constants Z_i and removes *all* divergences from the theory (we will better see how this works below). The resulting correlation functions are finite but depend on the arbitrary renormalization point μ , which replaces the dependence on the arbitrary mass scale M . The renormalization constants Z_i are still divergent since they absorb the $1/\varepsilon$ terms, but they drop out in all observables that can be calculated from the theory.

In this way, the mass $m(\mu)$ is a parameter of the theory which has to be taken from experiment. In QED we could set the physical mass of the electron by onshell renormalization ($p^2 = m^2$), because this is where the electron propagator has a pole. In QCD, the current-quark masses must be specified at some suitable renormalization scale where theory predictions can be compared to experiment. This scale should also be spacelike ($\mu^2 < 0$) to avoid branch-point singularities that appear in the loop diagrams. High-energy scattering experiments with hadrons probe quarks and gluons at large spacelike momenta, which is also where the QCD coupling is small and perturbation theory applicable.

The coupling $g(\mu)$, on the other hand, is not truly a parameter but sets the scale: so far we have been working in arbitrary units, but to connect to GeV units we must set the coupling $\alpha(\mu^2)$ at a given momentum scale. Different values of $\alpha(\mu^2)$ then merely rescale the system, which means that the *running* of the coupling $\alpha(\mu^2)$ is an inherent property of the theory itself. Therefore, the parameters of QCD are a scale, where the coupling takes a specific value, and the current-quark masses at that scale — these must be taken from experiment.

The choice of a **renormalization scheme** reflects the arbitrariness in the specification of $m(\mu)$ and $g(\mu)$:

- Imposing *overall* renormalization conditions of the form (2.3.54) defines a momentum subtraction (**MOM**) scheme. This is convenient for nonperturbative calculations since at no point in the previous discussion we *needed* to resort to a perturbative expansion: Eq. (2.3.21) can equally be viewed as the Dyson-Schwinger equation for the full self-energy, which is nonperturbative and exact.
- Alternatively, one can explicitly subtract the divergent $1/\varepsilon$ terms order by order in perturbation theory, which defines the **MS scheme** (minimal subtraction). In that case our definition of the renormalization scale μ is no longer available and $M = \mu$ takes its place instead, since it is not cancelled by the subtraction anymore.
- Another possibility is to subtract not only the divergences but *all* terms that are independent of M ; this defines the $\overline{\text{MS}}$ **scheme** (modified minimal subtraction).

As a consequence, the masses and couplings depend not only on the renormalization point but also on the renormalization scheme. For example, the Particle Data Group (PDG) quotes the current-quark masses in the $\overline{\text{MS}}$ scheme at a renormalization scale $\mu = 2$ GeV. The quantities obtained in different schemes are related to each other by finite terms, and the invariance in the choice of μ , $m(\mu)$ and $g(\mu)$ leads to the concept of the **renormalization group**. At the end of the day, all physical observables must be independent of the renormalization point and scheme.

Renormalizability. So far we have only considered one explicit diagram. Do the singularities *always* cancel? Let's consider the action for a generic ϕ^p theory:

$$S = - \int d^4x \left[\frac{1}{2} \phi (\square + m^2) \phi + \lambda_p \phi^p \right], \quad (2.3.59)$$

where we suppress the renormalization constants for simplicity. Now count the mass dimensions of the quantities that appear in the action:

$$[S] = 0, \quad [d^4x] = -4 \quad \Rightarrow \quad [\mathcal{L}] = 4, \quad [\phi] = 1, \quad [\phi^p] = p, \quad [\lambda_p] = 4 - p. \quad (2.3.60)$$

From here we can infer the dimensions of the 1PI n -point functions in momentum space:

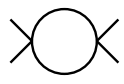
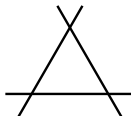
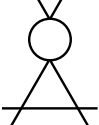
$$\begin{aligned} \Gamma_2 &= \text{---} \bigcirc \text{---}^{-1} = -i(p^2 - m^2) + \dots \quad \Rightarrow \quad [\Gamma_2] = 2, \\ \Gamma_4 &= \text{---} \bigcirc \text{---} \quad \Rightarrow \quad [\Gamma_4] = 0, \\ \Gamma_6 &= \text{---} \bigcirc \text{---} \quad \Rightarrow \quad [\Gamma_6] = -2, \end{aligned} \quad (2.3.61)$$

because Γ_{n+2} follows from Γ_n after taking two functional derivatives $\delta^2/\delta\phi^2$. Thus, the mass dimension of Γ_n is $[\Gamma_n] = 4 - n$.

On the other hand, we can also count the dimension of Γ_n in some given order in perturbation theory. To do so, we count the number of loops L (each comes with dimension four), the number of internal propagators I (each with dimension -2), and the number of vertices (each with dimension $[\lambda_p]$):

$$[\Gamma_n] = 4L - 2I + [\lambda_p] V. \quad (2.3.62)$$

For example in ϕ^4 theory, where $[\lambda_4] = 0$:

	$L = 1$ $I = 2$ $[\Gamma_4] = 0$		$L = 1$ $I = 3$ $[\Gamma_6] = -2$		$L = 2$ $I = 5$ $[\Gamma_6] = -2$
---	--	---	---	---	---

Obviously this is consistent.

Now, the quantity $D = 4L - 2I$ also tells us how badly divergent a given diagram will be: if the number of loops L beats the number of propagators I it will diverge; if there are many propagators in a loop it will converge. D is called the **superficial degree of divergence**: if $D < 0$ the diagram converges, if $D \geq 0$ it diverges. The first diagram above has $D = 0$ and diverges logarithmically. The second has $D = -2$ and is convergent; the third has $D = -2$ but unfortunately it is still divergent because it contains a divergent subdiagram (the one on the left). Hence the name 'superficial' degree of divergence:

- a diagram with $D \geq 0$ can still be finite due to cancellations,
- a diagram with $D < 0$ can be divergent if it contains divergent subdiagrams,
- tree-level diagrams have $D = 0$ but they are finite.

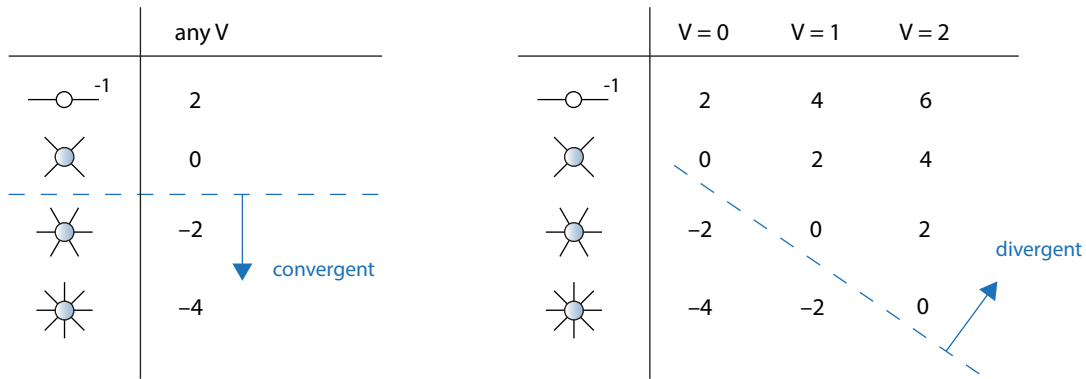


FIG. 2.12: Degree of divergence D in ϕ^4 theory (left) and ϕ^6 theory (right).

Let us ignore these subtleties for a moment and assume that D counts the actual degree of divergence. Then from Eq. (2.3.62) the degree of divergence of a given Γ_n in ϕ^p theory (with $\lambda_p = \lambda$) is

$$D = [\Gamma_n] - [\lambda] V. \quad (2.3.63)$$

The mass dimension $[\Gamma_n]$ is fixed, so depending on the mass dimension $[\lambda]$ of the coupling, D can rise or fall with higher orders in perturbation theory (expressed by V). Take ϕ^4 theory in the left panel of Fig. 2.12, where $[\lambda] = 0$ and $D = [\Gamma_n]$ is independent of V . In this case there are only two divergent n -point functions, namely the inverse propagator and the four-point function. These are also the ones with a tree-level term in the Lagrangian; they are called the **primitively divergent n -point functions** .

One can indeed show that the analysis goes through in general, also for divergent subdiagrams, which is known as the **BPHZ theorem** (Bogoliubov, Parasiuk, Hepp, Zimmermann). The reason is that the Z_i factors in front of the diagrams (which we can neglect at one-loop) cancel the divergences at higher orders. Take for example the two rightmost diagrams below Eq. (2.3.62): both contribute to the six-point function, one with $V = 3$ and the other with $V = 4$. The $V = 3$ diagram carries factors $Z = 1 + \delta Z$, where δZ contributes at higher order to the $V = 4$ graph. The sum of all contributions at a given order cancels the divergences. Here it is especially useful to employ the counterterm language, because the subdivergences will cancel with the counterterms at each order in perturbation theory.

On the other hand, the same analysis for ϕ^6 theory, where $[\lambda] = -2$ and thus $D = [\Gamma_n] + 2V$, gives the result in the right panel of Fig. 2.12: if we go high enough in perturbation theory, eventually *every* n -point function will diverge!

This leads to the notion of **renormalizability**: a theory is renormalizable if only a *finite* number of Green functions have $D \geq 0$, so that only a finite number of renormalization conditions are necessary to remove the divergences from the theory. From Eq. (2.3.63) this is equivalent to the following statement:

$$\text{A theory is renormalizable if } [\lambda] \geq 0.$$

Thus, the coupling must either be dimensionless or have a positive mass dimension (in the latter case the theory is called *super-renormalizable*).

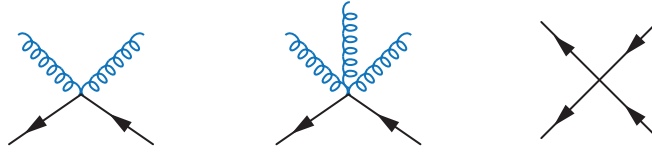


FIG. 2.13: Examples for non-renormalizable interactions constructed from fermions and gauge bosons. From left to right, the diagrams carry mass dimensions 5, 6 and 6.

A **non-renormalizable theory** has a coupling with negative mass dimension: in that case every n -point function eventually becomes divergent. Here we would need new renormalization conditions at each order in perturbation theory, and eventually infinitely many, so we must specify infinitely many constants from outside — the theory thereby loses its predictive power.

On the other hand, we will see in Sec. 4.4 that non-renormalizable theories are still perfectly acceptable **low-energy effective theories** since higher loop diagrams also come with higher momentum powers. For example, chiral perturbation theory is a non-renormalizable low-energy expansion of QCD; the non-renormalizable Fermi theory of weak interactions is the low-energy limit of the electroweak theory. In this sense, non-renormalizable theories are merely ‘less fundamental’ since they are not applicable at all energy scales.

Another caveat is that all considerations above are based on perturbation theory. For example, the Einstein-Hilbert action in quantum gravity defines a non-renormalizable gauge theory, which is also the reason why it is not considered as a part of the Standard Model and which has spurred developments e.g. in string theory. There is still the possibility that a non-renormalizable theory becomes *non-perturbatively* renormalizable, i.e., it ‘renormalizes itself’ by developing nontrivial UV fixed points. This leads to the concept of **asymptotic safety**, and there are indications that this is what could be at play in quantum gravity.

In any case, a renormalizable QFT contains only a small number of superficially divergent amplitudes, namely those with a tree-level counterpart in the Lagrangian, and therefore it only needs a finite number of renormalization constants. The good news is that we can read off a theory’s renormalizability directly from its Lagrangian: we just need to look at the mass dimension of the coupling constant. For a scalar ϕ^p theory only ϕ^3 and ϕ^4 interactions are renormalizable whereas those with $p > 4$ are not. Likewise, the QCD Lagrangian is renormalizable, whereas diagrams such as in Fig. 2.13 are not: with $[\psi] = 3/2$ and $[A] = 1$, their mass dimensions are greater than 4, and to compensate this we would need to attach couplings with negative mass dimensions. Renormalizability restricts the possible forms of interactions dramatically!

2.3.3 β function and running coupling

Callan-Symanzik equation. Consider again a generic field theory with a field ϕ or several fields ϕ_i . Then the bare and renormalized fields are related by $\phi_B = Z_\phi^{1/2} \phi$ in analogy to Eq. (2.3.1). Since this implies

$$\frac{\delta^n \Gamma}{\delta \varphi_B^n} = Z_\phi^{-n/2} \frac{\delta^n \Gamma}{\delta \varphi^n}, \quad (2.3.64)$$

we can read off how a renormalized 1PI correlation function ($\Gamma^n = \delta^n \Gamma / \delta \varphi^n$), which depends on a set of momenta $\{p_i\}$, the renormalized coupling g , the renormalized mass m and the renormalization point μ , is related to its bare counterpart ($\Gamma_B^n = \delta^n \Gamma / \delta \varphi_B^n$):

$$\Gamma^n(\{p_i\}, g, m, \mu) = Z_\phi^{n/2} \Gamma_B^n(\{p_i\}, g_B, m_B). \quad (2.3.65)$$

The bare quantities cannot depend on the renormalization scale μ . If we apply the derivative $\mu d/d\mu$ and use $d\Gamma_B^n/d\mu = 0$, we obtain the **Callan-Symanzik equation**:

$$\left(\underbrace{\mu \frac{\partial}{\partial \mu}}_{\beta(g)} + \underbrace{\mu \frac{dg}{d\mu} \frac{\partial}{\partial g}}_{m\gamma_m(g)} + \underbrace{\mu \frac{dm}{d\mu} \frac{\partial}{\partial m}}_{\gamma(g)} \right) \Gamma^n = \mu \frac{n}{2} Z_\phi^{n/2-1} \frac{dZ_\phi}{d\mu} \Gamma_B^n = n \underbrace{\frac{\mu}{2} \frac{d \ln Z_\phi}{d\mu}}_{\gamma(g)} \Gamma^n. \quad (2.3.66)$$

Here we defined the β function $\beta(g)$, the anomalous mass dimension $\gamma_m(g)$, and the anomalous dimension of the field $\gamma(g)$; they determine the respective change of the coupling, the mass and the field renormalization under a change of the renormalization scale. For n-point functions that depend on more than one field we would have to include a separate $\gamma(g)$ for each of them.

The Callan-Symanzik equation entails that a shift of the renormalization scale can be compensated by an appropriate shift of the coupling, the mass and the fields. Suppose for the moment that $\gamma(g) = 0$, so that Z_ϕ is independent of μ . We also set $m = 0$ to simplify the discussion. The l.h.s of the equation then implies $d\Gamma^n/d\mu = 0$, i.e. also the renormalized n-point function is μ -independent. A change of the renormalization point can then always be compensated by a shift of the coupling:

$$\Gamma^n(\{p_i\}, g(\mu), \mu) = \Gamma^n(\{p_i\}, g(\mu_0), \mu_0). \quad (2.3.67)$$

Moreover, the Callan-Symanzik equation also allows us to compensate the *momentum* dependence of an n-point function by a change in its coupling. Consider an n-point function with mass dimension D ; it can be written as

$$\Gamma^n(\{p_i\}, g(\mu), \mu) = \mu^D f\left(\left\{\frac{p_i}{\mu}\right\}, g(\mu)\right) = \mu_0^D f\left(\left\{\frac{p_i}{\mu_0}\right\}, g(\mu_0)\right), \quad (2.3.68)$$

where the function f is dimensionless. The first equality is simply a dimensional argument, and the second follows from Eq. (2.3.67) since the expression is independent of μ . Now replace all momenta $p_i \rightarrow \lambda p_i$, where $\lambda = \mu/\mu_0$:

$$f\left(\lambda \left\{\frac{p_i}{\mu_0}\right\}, g(\mu_0)\right) = \lambda^D f\left(\left\{\frac{p_i}{\mu_0}\right\}, g(\lambda\mu_0)\right). \quad (2.3.69)$$

Hence, at a fixed renormalization point μ_0 , a uniform rescaling of momenta can be compensated by an according shift of the coupling on which the Green function depends. If we dropped our simplifications $\gamma(g) = 0$ and $m = 0$, the equation would pick up a scaling factor that depends on $\gamma(g)$, and the renormalized mass would obtain a scaling factor $\sim \gamma_m(g)$, hence the name ‘anomalous dimensions’.

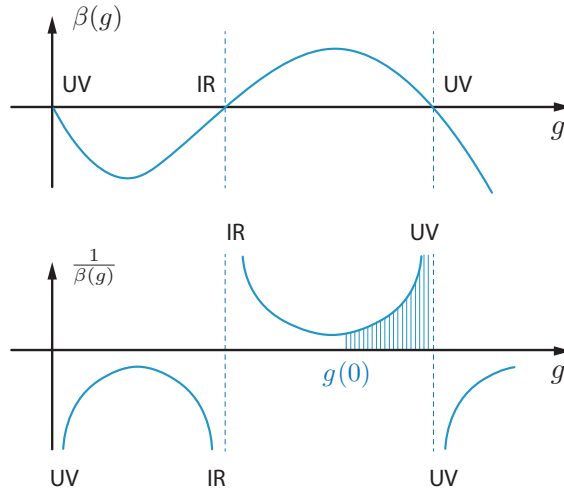


FIG. 2.14: Possible shape of the β function and its inverse that appears in Eq. (2.3.70).

β function. The β function of a theory, $\beta(g) = \mu dg/d\mu$, encodes the change of the running coupling with the momentum scale. If we change the scale from μ_0 to μ and define the dimensionless variable $t = \ln(\mu/\mu_0) \in [-\infty, \infty]$, which entails $\mu d/d\mu = d/dt$, then the change from the coupling $g(0)$ at μ_0 to $g(t)$ at μ is given by

$$\beta(g) = \frac{dg(t)}{dt} \quad \Rightarrow \quad \int_{g(0)}^{g(t)} \frac{dg}{\beta(g)} = \int_0^t dt' = t, \quad (2.3.70)$$

which can be solved for $g(t)$ if $\beta(g)$ is known.

To understand this equation better, let us study possible shapes of the β function (Fig. 2.14). The values g^* where $\beta(g^*) = 0$ are **fixed points** under a renormalization-group evolution because the coupling does not change in the vicinity of g^* ($dg/dt = 0$). Eq. (2.3.70) entails that the l.h.s. must diverge for $t \rightarrow \pm\infty$: this happens when $g(t)$ runs into the fixed point nearest to $g(0)$, or when it goes to infinity because there is no zero of $\beta(g)$ to approach. Whether the fixed point corresponds to $t \rightarrow \infty$ or $t \rightarrow -\infty$ depends on the sign of the β function and the integration direction:

- An **ultraviolet (UV) fixed point** ($t \rightarrow +\infty$) implies $g(t) > g(0)$ and $\beta > 0$ or $g(t) < g(0)$ and $\beta < 0$;
- An **infrared (IR) fixed point** ($t \rightarrow -\infty$) implies $g(t) > g(0)$ and $\beta < 0$ or $g(t) < g(0)$ and $\beta > 0$.

The origin $g = 0$ is always a fixed point since $\beta(0) = 0$. A theory is called

- **asymptotically free** if $g = 0$ is a UV fixed point, because then the coupling becomes small for $t \rightarrow \infty$ (as we will see below, this is the case for QCD);
- **infrared stable** if $g = 0$ is an IR fixed point (e.g. QED, ϕ^4 theory).

The domains separated by fixed points correspond to different theories, unless there are several couplings in the theory (in which case one ends up with a multidimensional phase diagram).

Calculation of the β function. In the following we sketch the calculation of the β function in QCD (for which Gross, Politzer and Wilczek received the Nobel Prize in 2004). We start with the relation $g_B = Z_g g$ from Eq. (2.3.1), where the bare coupling g_B does not depend on μ . In four dimensions g is dimensionless, but since we want to employ dimensional regularization we must work out the dimension $[g]$ of the coupling in $d = 4 - \varepsilon$ dimensions. Because the action remains dimensionless and the spacetime integral is $d^d x$, we have

$$[\mathcal{L}] = d, \quad [\psi] = \frac{d-1}{2}, \quad [A] = \frac{d-2}{2} \quad \Rightarrow \quad [g] = [\mathcal{L}] - [\bar{\psi} A \psi] = \frac{\varepsilon}{2}. \quad (2.3.71)$$

Thus we write $g_B = Z_g g \mu^{\varepsilon/2}$, where g is the dimensionless coupling in arbitrary dimensions (this is equivalent to putting factors μ^ε in front of loop integrals such as Eq. (2.3.48)). The β function then becomes

$$\begin{aligned} \beta(g) &= \frac{dg}{dt} = \mu \frac{d}{d\mu} \left(\frac{g_B}{Z_g \mu^{\varepsilon/2}} \right) = \mu \left(-\frac{\varepsilon}{2} \frac{g_B}{Z_g \mu^{\varepsilon/2+1}} - \frac{1}{Z_g^2} \frac{dZ_g}{d\mu} \frac{g_B}{\mu^{\varepsilon/2}} \right) \\ &= - \left(\frac{\varepsilon}{2} + \frac{d}{dt} \ln Z_g \right) g. \end{aligned} \quad (2.3.72)$$

To proceed, we must calculate the g dependence of Z_g . From Eq. (2.3.3) we see that Z_g appears in all vertices in the Lagrangian in combination with other renormalization constants, so we could obtain it from any of the combinations $\{Z_A, Z_\psi, Z_\Gamma\}$, $\{Z_A, Z_c, \tilde{Z}_\Gamma\}$, $\{Z_A, Z_{3g}\}$ or $\{Z_A, Z_{4g}\}$. In the first case, we must calculate the one-loop diagrams for the gluon propagator, the quark propagator and the quark-gluon vertex. Because the renormalization constants absorb the infinities, the simplest option is to use the $\overline{\text{MS}}$ scheme where they only absorb the $1/\varepsilon$ terms and nothing else. For example, for the quark propagator we have from Eqs. (2.3.21) and (2.3.52):

$$A(p^2) = Z_\psi + \Sigma_A(p^2) = Z_\psi + \frac{\alpha}{2\pi} C_F \int dx x \left(\frac{2}{\varepsilon} - \gamma + \ln \frac{4\pi M^2}{\Delta} - 1 \right). \quad (2.3.73)$$

In our earlier MOM scheme we demanded $A(\mu^2) = 1$, which led to $Z_\psi = 1 - \Sigma_A(\mu^2)$, whereas in the $\overline{\text{MS}}$ scheme we only subtract the infinities:

$$Z_\psi = 1 - \frac{\alpha}{2\pi} C_F \int dx x \frac{2}{\varepsilon} = 1 - \frac{g^2}{(4\pi)^2} \frac{2C_F}{\varepsilon}. \quad (2.3.74)$$

In the same way one computes Z_A and Z_Γ , where only the highest momentum powers in the loop integrals contribute since only those produce the divergences and thus the $1/\varepsilon$ terms. Putting everything together, the one-loop result for Z_g becomes

$$Z_g = 1 - \frac{b}{\varepsilon} g^2 + \dots, \quad b = \frac{\beta_0}{(4\pi)^2}, \quad \beta_0 = 11 - \frac{2}{3} N_f, \quad (2.3.75)$$

where N_f is the number of flavors. Inserting this back into Eq. (2.3.72) gives

$$\begin{aligned} \frac{d}{dt} \ln Z_g &= -\frac{2b}{\varepsilon} g \beta(g) + \dots \Rightarrow \beta(g) = - \left(\frac{\varepsilon}{2} - \frac{2b}{\varepsilon} g \beta(g) \right) g \\ &\Rightarrow \beta(g) \left(1 - \frac{2b}{\varepsilon} g^2 \right) = -\frac{\varepsilon g}{2} \\ &\Rightarrow \beta(g) = -\frac{\varepsilon g}{2} - b g^3 + \dots \end{aligned} \quad (2.3.76)$$

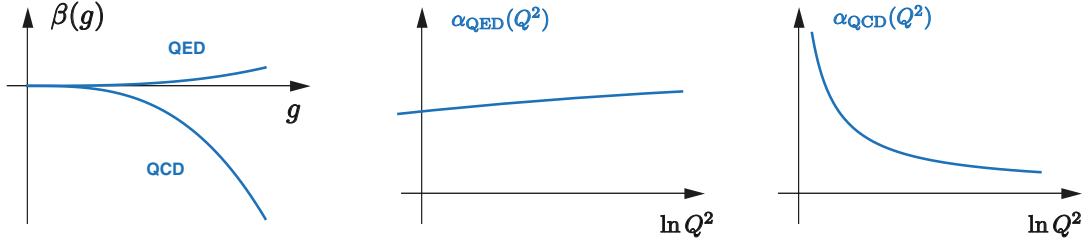


FIG. 2.15: β function in QCD and QED (left) and resulting shapes of the running coupling.

For $\varepsilon \rightarrow 0$, we obtain the result

$$\beta(g) = -bg^3 + \mathcal{O}(g^5). \quad (2.3.77)$$

The negative sign of the β function at $g \rightarrow 0$ shows that QCD is indeed an asymptotically free theory, i.e., $g(t)$ becomes small at large momenta $t \rightarrow \infty$. Note that this is only true for $\beta_0 > 0$, which entails $N_f \leq 16$; for more than 16 flavors we would lose asymptotic freedom. The lowest-order coefficients at $\mathcal{O}(g^3)$ and $\mathcal{O}(g^5)$ are independent of the renormalization scheme, whereas higher-order terms are not.

Running coupling. If we put the result for $\beta(g)$ back into Eq. (2.3.70), we obtain the running coupling of QCD:

$$\int_{g(0)}^{g(t)} \frac{dg}{-bg^3} = \frac{1}{2b} \left(\frac{1}{g(t)^2} - \frac{1}{g(0)^2} \right) = t \quad \Rightarrow \quad g(t)^2 = \frac{g(0)^2}{1 + 2bt g(0)^2}, \quad (2.3.78)$$

or equivalently $\alpha(t) = g(t)^2/(4\pi) = \alpha(0)/[1 + \frac{\beta_0}{4\pi} \alpha(0) 2t]$. Writing $2t = \ln(\mu^2/\mu_0^2)$, this expression has a pole at $\mu^2 = \Lambda_{\text{QCD}}^2$ defined by

$$\alpha(0) = \frac{1}{\frac{\beta_0}{4\pi} \ln \frac{\mu_0^2}{\Lambda_{\text{QCD}}^2}} \quad \Rightarrow \quad \alpha(t) = \frac{1}{\frac{\beta_0}{4\pi} \ln \frac{\mu_0^2}{\Lambda_{\text{QCD}}^2} + \ln \frac{\mu^2}{\mu_0^2}} = \frac{1}{\frac{\beta_0}{4\pi} \ln \frac{\mu^2}{\Lambda_{\text{QCD}}^2}}. \quad (2.3.79)$$

From the Callan-Symanzik equation we can interpret the dependence on μ^2 as a dependence on q^2 . Actually we should have started from large *spacelike* (‘Euclidean’) momenta $q^2 = -Q^2 < 0$, because this is the momentum region where we can compare to experiment and where $\alpha(Q^2)$ is guaranteed to be free of singularities. As long as μ^2 and μ_0^2 are also spacelike, this does not change the formulas and we arrive at

$$\alpha(Q^2) = \frac{1}{\frac{\beta_0}{4\pi} \ln(Q^2/\Lambda_{\text{QCD}}^2)}. \quad (2.3.80)$$

At large momenta where $\alpha(Q^2)$ becomes small, quarks and gluons behave as asymptotically free particles and we can apply perturbation theory. On the other hand, this also means that the coupling increases at small momenta and perturbation theory will eventually fail. In that region, nonperturbative effects related to the formation of bound states become important.

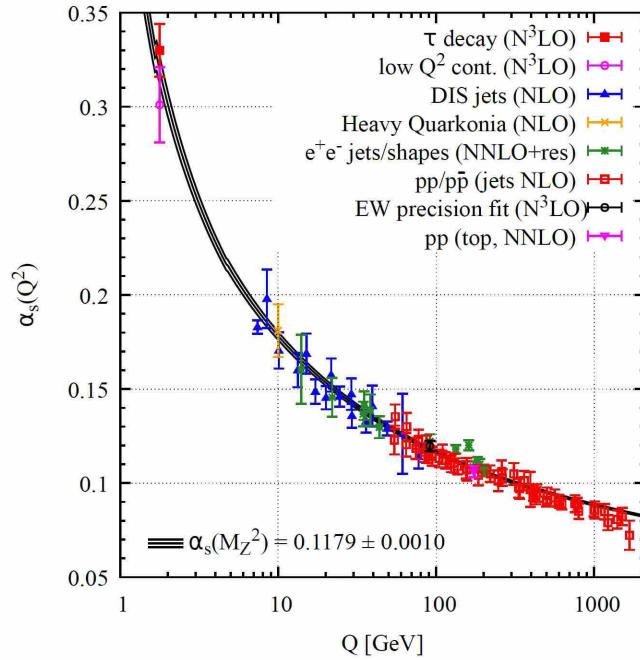


FIG. 2.16: Overview of $\alpha(Q^2)$ measurements from the PDG, taken from P. A. Zyla et al., *Prog. Theor. Exp. Phys.* **2020**, 083C01 (2020).

The analogous calculation in QED gives $\beta_0 = -4/3$ so that $\beta(g \rightarrow 0)$ is positive: QED is infrared stable and the coupling grows with increasing momenta. It actually grows very slowly (Fig. 2.15), so that perturbation theory works very well over many orders of magnitude.

Λ_{QCD} is the scale where perturbation theory definitely breaks down since it produces an unphysical Landau pole in the perturbative expansion. Eq. (2.3.80) and its refinements at higher loop orders allow one to convert the running coupling at a given scale — see Fig. 2.16 for the current world average of $\alpha(M_Z^2)$ — to a value for Λ_{QCD} , which therefore depends on the order in perturbation theory, the renormalization scheme, and the number of active flavors at the scale where the coupling is probed (due to the N_f dependence in β_0). Comparison of $\alpha(Q^2)$ at four-loop order with experimental results yields the value $\Lambda_{\overline{\text{MS}}}^{N_f=5} = 210(14)$ MeV [PDG 2018].

Alternative calculation of the running coupling. Another way to compute the running coupling is to start from the *finite* quantities (i.e., the renormalized propagators and vertices) instead of the divergent ones (the renormalization constants). To do so, note that the renormalization constants do not only relate the renormalized with the bare fields, but also the renormalized dressing functions of the propagators and vertices with their bare counterparts, cf. Eq. (2.3.65). For the gluon and ghost propagator and the ghost-gluon vertex this reads:

$$Z_B(q^2) = Z_A Z(q^2), \quad G_B(q^2) = Z_c G(q^2), \quad \Gamma_{\text{gh}}(q^2) = \tilde{Z}_\Gamma \Gamma_{\text{gh}}^B(q^2). \quad (2.3.81)$$

Here, $\Gamma_{\text{gh}}(q^2) = \tilde{f}_1(q^2, q^2, q^2)$ is the dressing function attached to the tree-level tensor of the ghost-gluon vertex in Eq. (2.3.17). We also have $g_B = Z_g g$ and thus $\alpha_B = Z_g^2 \alpha$.

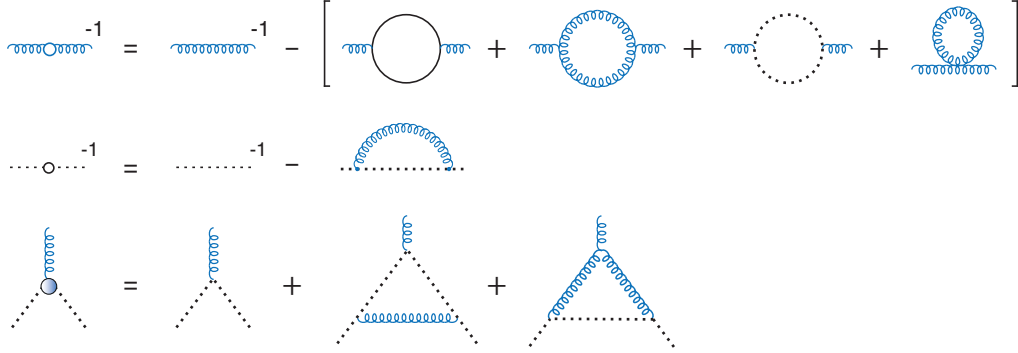


FIG. 2.17: One-loop diagrams for the gluon and ghost propagator and the ghost-gluon vertex.

We can use Eq. (2.3.3) to find combinations that stay unrenormalized, i.e. for which $F_B = F$, since only those can contain observable information. One such combination is the ‘running coupling from the ghost-gluon vertex’:

$$\bar{\alpha}(q^2) = \alpha Z(q^2) G^2(q^2) \Gamma_{\text{gh}}^2(q^2) = \frac{\tilde{Z}_\Gamma^2}{\underbrace{Z_g^2 Z_A Z_c^2}_{=1}} \bar{\alpha}_B(q^2). \quad (2.3.82)$$

The bare quantities are individually divergent but the divergences must cancel in the combination.

To determine $\bar{\alpha}(q^2)$, one must calculate the one-loop diagrams in Fig. 2.17 for the gluon propagator, the ghost propagator and the ghost-gluon vertex. The results for a general gauge parameter ξ are

$$\begin{aligned} Z(q^2) &= 1 - \frac{\alpha}{4\pi} \left[\frac{N_c}{2} \left(\frac{13}{3} - \xi \right) - \frac{4}{3} T_F N_f \right] \ln \frac{q^2}{\mu^2}, \\ G(q^2) &= 1 - \frac{\alpha}{4\pi} \left[N_c \frac{3 - \xi}{4} \right] \ln \frac{q^2}{\mu^2}, \\ \Gamma_{\text{gh}}(q^2) &= 1 - \frac{\alpha}{4\pi} \left[N_c \frac{\xi}{2} \right] \ln \frac{q^2}{\mu^2}. \end{aligned} \quad (2.3.83)$$

The first term in the bracket for $Z(q^2)$ is the sum of the gluon and ghost loop (the tadpole drops out). The second term comes from the quark loop, where the color trace is $T_F = 1/2$ in the fundamental representation and we set $q^2, \mu^2 \gg m^2$. Taking the squares of $G(q^2)$ and $\Gamma_{\text{gh}}(q^2)$, the terms in the brackets add up to

$$\beta_0 = \frac{N_c}{2} \left(\frac{13}{3} - \xi \right) - \frac{4}{3} T_F N_f + N_c \frac{3 - \xi}{2} + N_c \xi = \frac{11}{3} N_c - \frac{4}{3} T_F N_f, \quad (2.3.84)$$

where the dependence on the gauge parameter ξ has dropped out. This is identical to the result (2.3.75) and the resulting running coupling at one-loop order is

$$\bar{\alpha}(q^2) = \alpha \left(1 - \frac{\alpha}{4\pi} \beta_0 \ln \frac{q^2}{\mu^2} + \dots \right) \approx \frac{\alpha}{1 + \alpha \frac{\beta_0}{4\pi} \ln \frac{q^2}{\mu^2}}. \quad (2.3.85)$$

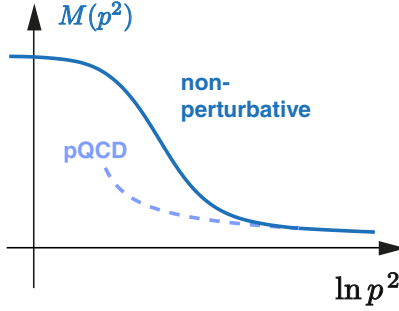


FIG. 2.18: Qualitative shape of the quark mass function from perturbation theory and non-perturbative calculations.

In QED, in the absence of gauge-boson self-interactions and ghosts, the relations in Eq. (2.3.3) reduce to $Z_g Z_A^{1/2} = 1$ and $Z_\Gamma = Z_\psi$, so the analogous definition of the running coupling is $\bar{\alpha}(q^2) = \alpha Z(q^2)$. In that case only the diagram with the fermion loop in the photon propagator survives, which yields $\beta_0 = -4/3$ for one species of fermions. In Eq. (2.3.84) one can see how the screening effect from the quark loop, which gives a negative contribution to β_0 for $N_f \leq 16$, is overwhelmed by the ‘antiscreening’ from the remaining diagrams involving gluons and ghosts.

Running quark mass. The **quark mass function** is another combination that stays unrenormalized since Eq. (2.3.11) entails

$$iS^{-1}(p) = A(p^2) (\not{p} - M(p^2)) = Z_\psi iS_B^{-1}(p) = Z_\psi A_B(p^2) (\not{p} - M_B(p^2)) \quad (2.3.86)$$

and thus $M(p^2) = M_B(p^2)$. We already worked out the one-loop result for the mass function in Eq. (2.3.58). If we define the anomalous mass dimension γ_m as

$$\gamma_m = \frac{3C_F}{\beta_0} = \frac{4}{11 - \frac{2}{3}N_f}, \quad (2.3.87)$$

then we can write to one-loop order for $\alpha \ll 1$:

$$\begin{aligned} M(p^2) &= m \left[1 - \frac{3\alpha}{4\pi} C_F \ln \frac{p^2}{\mu^2} \right] = m \left[1 - \gamma_m \frac{\beta_0}{4\pi} \alpha \ln \frac{p^2}{\mu^2} \right] \\ &= m \left[1 + \frac{\beta_0}{4\pi} \alpha \ln \frac{p^2}{\mu^2} \right]^{-\gamma_m} = m \left[\frac{\alpha(p^2)}{\alpha(\mu^2)} \right]^{\gamma_m} \\ &= m \left[\frac{\frac{1}{2} \ln(\mu^2/\Lambda_{\text{QCD}}^2)}{\frac{1}{2} \ln(p^2/\Lambda_{\text{QCD}}^2)} \right]^{\gamma_m} = \frac{\hat{m}}{\left[\frac{1}{2} \ln(p^2/\Lambda_{\text{QCD}}^2) \right]^{\gamma_m}}. \end{aligned} \quad (2.3.88)$$

This gives the one-loop running of the quark mass function at large p^2 . It is also independent of the gauge parameter ξ , whereas the result for $A(p^2)$ is

$$A(p^2) = 1 - \xi C_F \frac{\alpha}{4\pi} \ln \frac{p^2}{\mu^2}. \quad (2.3.89)$$

Unfortunately, QCD perturbation theory turns out to be of limited use in this case because for light quarks the biggest contribution to the mass function $M(p^2)$ is generated **non-perturbatively** by spontaneous chiral symmetry breaking (Fig. 2.18). We will return to this point in Sec. 4.2.

From Eq. (2.3.83) we can also read off the anomalous dimensions for the gluon and ghost propagators and the ghost-gluon vertex. Writing

$$Z(q^2) \propto \frac{1}{[\ln(q^2/\Lambda^2)]^{\gamma_{\text{gl}}}}, \quad G(q^2) \propto \frac{1}{[\ln(q^2/\Lambda^2)]^{\gamma_{\text{gh}}}}, \quad \Gamma_{\text{gh}}(q^2) \propto \frac{1}{[\ln(q^2/\Lambda^2)]^{\gamma_{\text{gh-gl}}}} \quad (2.3.90)$$

we find

$$\gamma_{\text{gl}} = \frac{1}{\beta_0} \left[\frac{N_c}{2} \left(\frac{13}{3} - \xi \right) - \frac{4}{3} T_F N_f \right], \quad \gamma_{\text{gh}} = \frac{1}{\beta_0} \left[N_c \frac{3 - \xi}{4} \right], \quad \gamma_{\text{gh-gl}} = \frac{1}{\beta_0} \left[N_c \frac{\xi}{2} \right], \quad (2.3.91)$$

where $\gamma_{\text{gl}} + 2\gamma_{\text{gh}} + 2\gamma_{\text{gh-gl}} = 1$. In Landau-gauge Yang-Mills theory ($\xi = 0$, $N_f = 0$) this reduces to

$$\gamma_{\text{gl}} = \frac{13}{22}, \quad \gamma_{\text{gh}} = \frac{9}{44}, \quad \gamma_{\text{gh-gl}} = 0. \quad (2.3.92)$$

Further reading

General QFT textbooks:

- M. Maggiore, *A Modern Introduction to Quantum Field Theory*. Oxford University Press, 2005
- M. E. Peskin and D. V. Schroeder, *An Introduction to Quantum Field Theory*. Perseus, 1995
- S. Pokorski, *Gauge Field Theories*. Cambridge University Press, 1987
- S. Weinberg, *The Quantum Theory of Fields. Vol. I + II*. Cambridge University Press, 1995/96

Classical equations of motion:

- V. P. Nair, *Quantum Field Theory: A Modern Perspective*. Springer, 2005
- R. A. Bertlmann, *Anomalies in Quantum Field Theory*. Oxford University Press, 1996

Spectral representation and analyticity:

- V. Gribov, *Strong Interactions of Hadrons at High Energies*. Cambridge University Press, 2009

Lattice QCD:

- C. Gattringer and C. B. Lang, *Quantum Chromodynamics on the Lattice*. Springer, 2010

Quantum equations of motion:

- C. D. Roberts and A. G. Williams, *Dyson-Schwinger equations and their applications to hadron physics*. Prog. Part. Nucl. Phys. **33** (1994) 477, [arXiv:hep-ph/9403224](#)
- R. Alkofer and L. von Smekal, *The infrared behavior of QCD Green's functions: Confinement, dynamical symmetry breaking, and hadrons as relativistic bound states*. Phys. Rept. **353** (2001) 281. [arXiv:hep-ph/0007355](#)
- E. S. Swanson, *A Primer on Functional Methods and the Schwinger-Dyson Equations*. [arXiv:1008.4337 \[hep-ph\]](#)

Faddeev-Popov gauge fixing and BRST symmetry:

- R. A. Bertlmann, *Anomalies in Quantum Field Theory*. Oxford University Press, 1996
- N. Vandersickel and D. Zwanziger, *The Gribov problem and QCD dynamics*. Phys.Rept. **520** (2012) 175-251. [arXiv:1202.1491 \[hep-th\]](#)

Confinement:

- J. Greensite, *An Introduction to the Confinement Problem*. Springer, 2011

Callan-Symanzik equation and β function:

- S. Pokorski, *Gauge Field Theories*. Cambridge University Press, 1987
- T. P. Cheng and L. F. Li, *Gauge Theory of Elementary Particle Physics*. Oxford University Press, 1984

Asymptotic safety:

- M. Reuter and F. Saueressig, *Quantum Einstein Gravity*. New J. Phys. **14** (2012) 055022. [arXiv:1202.2274 \[hep-th\]](#)

Chapter 3

Hadrons

In the previous chapter we mostly ignored the flavor structure of the QCD Lagrangian because it was less relevant for the properties of quarks and gluons compared to the color structure. Vice versa, QCD's local gauge invariance does not tell us much about the systematics of the hadron spectrum except that hadrons are color singlets and can be constructed from $q\bar{q}$ and qqq (and also more complicated combinations). Here we will turn the wheel around and focus exclusively on the global flavor symmetries of QCD, in particular chiral symmetry, which becomes important and leads to effects that are observable (or conspicuously missing) in the mass spectrum.

3.1 Flavor symmetries and currents

Noether theorem. The Noether theorem states that any continuous symmetry transformation that leaves the classical action invariant implies the existence of a conserved current, where the corresponding charge is a constant of motion. This is true for space-time symmetries and, in our context, global symmetries (but effectively also for local ones in the sense that each local symmetry has an underlying global symmetry). Let us exemplify the statement for a generic field theory with action $S = \int d^4x \mathcal{L}(\phi_i, \partial_\mu \phi_i)$. Consider a global transformation

$$\phi'_i = D_{ij}(\varepsilon) \phi_j = \left(e^{i \sum_a \varepsilon_a \mathbf{t}_a} \right)_{ij} \phi_j = \phi_i + \delta \phi_i \quad (3.1.1)$$

of the fields under some Lie group G , where ε_a are the group parameters, the \mathbf{t}_a with $[\mathbf{t}_a, \mathbf{t}_b] = if_{abc} \mathbf{t}_c$ are the generators of the Lie algebra in the representation to which the ϕ_i belong, and $D(\varepsilon)$ are the representation matrices. Compute the variation of the action with respect to the group parameter ε_a inside a spacetime volume V and for solutions of the classical equations of motion:

$$\begin{aligned} \delta S &= \int_V d^4x \delta \mathcal{L} = \int_V d^4x \sum_i \left[\frac{\partial \mathcal{L}}{\partial \phi_i} \delta \phi_i + \frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi_i)} \delta (\partial_\mu \phi_i) \right] \\ &= \int_V d^4x \left[\underbrace{\partial_\mu \left(\sum_i \frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi_i)} \delta \phi_i \right)}_{=: -\sum_a \varepsilon_a j_a^\mu} + \sum_i \underbrace{\left(\frac{\partial \mathcal{L}}{\partial \phi_i} - \partial_\mu \frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi_i)} \right)}_{= 0 \text{ for classical solutions}} \delta \phi_i \right]. \end{aligned} \quad (3.1.2)$$

Here we considered a fixed volume V where the fields do not vanish on the surface, so the surface term does not vanish automatically. The second bracket vanishes for solutions of the Euler-Lagrange equations. Hence, if the classical action is invariant under the symmetry and thus $\delta S = 0$, there is one **conserved current** for each generator of the symmetry group when evaluated along the classical trajectories:

$$\partial_\mu j_a^\mu = 0 \quad \text{with} \quad j_a^\mu = -i \sum_{ij} \frac{\partial \mathcal{L}}{\partial (\partial_\mu \phi_i)} (t_a)_{ij} \phi_j. \quad (3.1.3)$$

We can further use Gauss' law to convert the volume to a surface integral. For two spacelike hypersurfaces σ_1 and σ_2 , if the surface term at spatial infinity is zero, this yields

$$\int_V d^4x \partial_\mu j_a^\mu = \int_{\partial V} d\sigma_\mu j_a^\mu = \left[\int_{\sigma_2} - \int_{\sigma_1} \right] d\sigma_\mu j_a^\mu = 0. \quad (3.1.4)$$

In particular for two spacelike surfaces at fixed time t , the integral over a surface in four dimensions becomes a three-dimensional volume integral, where the four-vector $d\sigma_\mu = (d^3x, \mathbf{0})$ points in the time direction, and therefore we find a **conserved charge** for each generator t_a of the symmetry group:

$$Q_a(t) = \int d^3x j_a^0(x) = \text{const.} \quad (3.1.5)$$

Note that the currents and charges are still well-defined if the fields do not obey the classical equations of motion (then the second parenthesis in Eq. (3.1.2) is nonzero) or if the symmetry is classically broken (then the action is not invariant, $\delta S \neq 0$). In these cases the currents and charges are not conserved:

$$\partial_\mu j_a^\mu \neq 0, \quad \frac{d}{dt} Q_a(t) \neq 0. \quad (3.1.6)$$

Quantization. When the classical field theory is quantized, the fields $\phi_i(x)$, currents $j_a^\mu(x)$ and charges $Q_a(t)$ become operators on the state space of the theory. As we will see later, the (anti-) commutation relations of the fields imply that the charges satisfy the same commutator relations as the generators of the symmetry group,

$$[Q_a, Q_b] = if_{abc} Q_c, \quad (3.1.7)$$

so they form a representation of the Lie algebra on the state space (the **charge algebra**). This relation remains intact even if the charges are time-dependent, i.e., if the symmetry is broken. The charges can be used to construct a representation of the group on the state space under which the field operators transform,

$$U = e^{i \sum_a \varepsilon_a Q_a}, \quad |\lambda'\rangle = U |\lambda\rangle, \quad U \phi_i U^{-1} = D_{ij}^{-1} \phi_j, \quad (3.1.8)$$

which implements the classical relation (3.1.1) at the level of expectation values:

$$\langle \lambda'_1 | \phi_i | \lambda'_2 \rangle = \langle \lambda_1 | U^{-1} \phi_i U | \lambda_2 \rangle = D_{ij} \langle \lambda_1 | \phi_j | \lambda_2 \rangle. \quad (3.1.9)$$

(Note that later we will not always be consistent in the notation and denote the transformation matrices of the classical fields by U instead of D while leaving the operators $\exp(i \sum_a \varepsilon_a Q_a)$ unnamed.) By expanding $U \approx 1 + i \sum_a \varepsilon_a Q_a$ and $D \approx 1 + i \sum_a \varepsilon_a \mathbf{t}_a$, Eq. (3.1.8) entails

$$[Q_a, \phi_i] = -(\mathbf{t}_a)_{ij} \phi_j. \quad (3.1.10)$$

If the symmetry leaves the vacuum invariant, $U|0\rangle = |0\rangle$, then all generators Q_a must annihilate the vacuum: $Q_a|0\rangle = 0$, and we find

$$\langle 0|\phi_i|0\rangle = D_{ij} \langle 0|\phi_j|0\rangle = \langle 0|\phi_i|0\rangle + i \sum_a \varepsilon_a (\mathbf{t}_a)_{ij} \langle 0|\phi_j|0\rangle. \quad (3.1.11)$$

Thus, the vacuum expectation values must vanish for those directions ε_a that do not leave the ϕ_i invariant, which is the usual ‘**Wigner-Weyl realization**’ of a symmetry: $\langle 0|\phi_i|0\rangle = 0$. Later we will study examples where the ϕ_i can be composite fields (such as $\bar{\psi}\psi$) or also collections of different fields (e.g. σ and π_a in the σ model).

The **Heisenberg equations of motion**, on the other hand, are a consequence of translation invariance $\partial_\mu F(\phi) = i[P_\mu, F(\phi)]$, which holds for arbitrary polynomials of the fields including the charges $Q_a(t)$:

$$\frac{dQ_a}{dt} = i [H_{\text{QCD}}, Q_a]. \quad (3.1.12)$$

Therefore, if the charges are conserved, they commute with the Hamiltonian and have a common eigenvalue spectrum. In other words, the mass spectrum of the theory can be labeled by the irreducible representations of the symmetry group, which will lead to the flavor multiplets of hadrons.

In addition to the explicit breaking of a symmetry, there are also other possibilities how classical symmetries can be broken at the quantum level:

- **Spontaneous symmetry breaking:** Here the classical action is still invariant under the global symmetry and the currents are conserved, $\partial_\mu j_a^\mu = 0$, but the vacuum and the correlation functions of the theory lose this symmetry and develop condensates $\langle 0|\phi_i|0\rangle \neq 0$. As a consequence, $U|0\rangle \neq |0\rangle$ and there are charges which do not annihilate the vacuum: $Q_a|0\rangle \neq 0$ (we will refine these statements in Sec. 4.2). For each such charge there is a **massless Goldstone boson**. Spontaneous symmetry breaking is a dynamical effect due to the dynamics inherent in the theory, so one may as well turn the argument around and argue that it happens *because* the dynamics contains massless long-range interactions. The QCD example is chiral symmetry or, more precisely, the group $SU(N_f)_A$ for vanishing quark masses.

- **Anomalous symmetry breaking:** Also here the classical action is invariant, but the symmetry is broken at the quantum level due to regularization, i.e. if there is no regulator that preserves the classical symmetry. We already mentioned the anomalous breaking of scale invariance; other typical candidates are again axial symmetries: in dimensional regularization, γ_5 has no natural extension to $d \neq 4$ dimensions; a Pauli-Villars regulator breaks chiral symmetry explicitly due to a mass term, etc. In Sec. 4.3 we will study the $U(1)_A$ anomaly in QCD. In contrast to spontaneous symmetry breaking, also the currents are no longer conserved and pick up additional terms so that $\partial_\mu j_a^\mu \neq 0$.

3.1.1 Symmetries of the QCD Lagrangian

In order to discuss QCD's flavor symmetries, we only need to consider the quark parts of the QCD Lagrangian since only the quark fields carry flavor labels:

$$\mathcal{L} = \bar{\psi} (i\not{\partial} - M) \psi + g \bar{\psi} \not{A} \psi \quad \text{with} \quad \psi_{\alpha,i}(x), \quad \bar{\psi}_{\alpha,i}(x). \quad (3.1.13)$$

In the following the index $i = 1 \dots N_f$ denotes the flavor and we suppress the color indices. For simplicity we also work with unrenormalized quantities and discuss renormalization when necessary. The spinor fields $\bar{\psi}_{\alpha,i}(x)$, $\psi_{\alpha,i}(x)$ transform under the fundamental representation of $SU(N_f)$:

$$\psi'(x) = U \psi(x), \quad \bar{\psi}'(x) = \bar{\psi}(x) U^\dagger \quad \text{with} \quad U = e^{i \sum_a \varepsilon_a \mathbf{t}_a}. \quad (3.1.14)$$

The \mathbf{t}_a are the $SU(N_f)$ generators, e.g., the Pauli matrices $\mathbf{t}_a = \tau_a/2$ for two flavors and Gell-Mann matrices $\mathbf{t}_a = \lambda_a/2$ for three flavors (see Appendix A). In the two-flavor case, the diagonal quark mass matrix in the Lagrangian has the form

$$M = \begin{pmatrix} m_u & 0 \\ 0 & m_d \end{pmatrix} = \frac{m_u + m_d}{2} \mathbf{1} + (m_u - m_d) \mathbf{t}_3, \quad (3.1.15)$$

whereas in the three-flavor case it is given by $M = \text{diag}(m_u, m_d, m_s)$ or

$$M = \frac{m_u + m_d + m_s}{3} \mathbf{1} + (m_u - m_d) \mathbf{t}_3 + \frac{m_u + m_d - 2m_s}{\sqrt{3}} \mathbf{t}_8. \quad (3.1.16)$$

Flavor symmetries. Consider the following global transformations of the quark fields:

$$\begin{aligned} e^{i \sum_a \varepsilon_a \mathbf{t}_a} &\in SU(N_f)_V, & e^{i\varepsilon} &\in U(1)_V, \\ e^{i\gamma_5 \sum_a \varepsilon_a \mathbf{t}_a} &\in SU(N_f)_A, & e^{i\gamma_5 \varepsilon} &\in U(1)_A, \end{aligned} \quad (3.1.17)$$

where $SU(N_f)_V$ denotes the usual transformation from Eq. (3.1.14). The subscripts V and A indicate that these transformations will induce vector and axialvector currents. The axial transformations involve γ_5 matrices and in the $U(1)$ cases ε is just a number. The infinitesimal transformations of the quark and antiquark fields read

$$SU(N_f)_V : \quad \delta\psi = \sum_a \varepsilon_a \mathbf{t}_a i\psi, \quad \delta\bar{\psi} = -i\bar{\psi} \sum_a \varepsilon_a \mathbf{t}_a, \quad (3.1.18)$$

$$SU(N_f)_A : \quad \delta\psi = \gamma_5 \sum_a \varepsilon_a \mathbf{t}_a i\psi, \quad \delta\bar{\psi} = i\bar{\psi} \gamma_5 \sum_a \varepsilon_a \mathbf{t}_a. \quad (3.1.19)$$

$$U(1)_V : \quad \delta\psi = \varepsilon i\psi, \quad \delta\bar{\psi} = -i\bar{\psi} \varepsilon, \quad (3.1.20)$$

$$U(1)_A : \quad \delta\psi = \varepsilon \gamma_5 i\psi, \quad \delta\bar{\psi} = i\bar{\psi} \gamma_5 \varepsilon. \quad (3.1.21)$$

Note the positive signs for the $\delta\bar{\psi}$ terms in the axial cases, which follow from the anticommutation of γ_5 and γ_0 in obtaining $\bar{\psi} = \psi^\dagger \gamma_0$:

$$(\gamma_5 \mathbf{t}_a i\psi)^\dagger \gamma_0 = -i\psi^\dagger \mathbf{t}_a \gamma_5 \gamma_0 = +i\bar{\psi} \mathbf{t}_a \gamma_5. \quad (3.1.22)$$

We will make frequent use of the following **quark bilinears**:

$$j_a^\Gamma(x) := \bar{\psi}(x) \Gamma \mathbf{t}_a \psi(x), \quad j^\Gamma(x) := \bar{\psi}(x) \Gamma \psi(x), \quad (3.1.23)$$

where $\Gamma \in \{\gamma^\mu, \gamma^\mu \gamma_5, \mathbf{1}, i\gamma_5\}$ are vector, axialvector, scalar and pseudoscalar Dirac matrices. We denote the corresponding vector, axialvector, scalar and pseudoscalar currents or densities $j_{(a)}^\Gamma(x)$ by ¹

$$\gamma^\mu \rightarrow V_{(a)}^\mu(x), \quad \gamma^\mu \gamma_5 \rightarrow A_{(a)}^\mu(x), \quad \mathbf{1} \rightarrow S_{(a)}(x), \quad i\gamma_5 \rightarrow P_{(a)}(x). \quad (3.1.24)$$

These quantities are all hermitian, e.g.

$$P^\dagger = (\bar{\psi} i\gamma_5 \psi)^\dagger = -i\psi^\dagger \gamma_5 \gamma_0 \psi = +i\bar{\psi} \gamma_5 \psi = P. \quad (3.1.25)$$

In the following we investigate the symmetry transformations $U(1)_V \times SU(N_f)_V \times SU(N_f)_A \times U(1)_A$ in detail.

■ **U(1)_V**: The action is invariant under a global phase transformation $\psi' = e^{i\varepsilon} \psi$. The corresponding flavor-singlet vector current according to Eq. (3.1.2) is

$$V^\mu = - \left[\frac{\partial \mathcal{L}}{\partial(\partial_\mu \psi_{\alpha,i})} (i\psi_{\alpha,i}) + \frac{\partial \mathcal{L}}{\partial(\partial_\mu \bar{\psi}_{\alpha,i})} (-i\bar{\psi}_{\alpha,i}) \right] = \bar{\psi} \gamma^\mu \psi, \quad (3.1.26)$$

where we used (cf. Eq. (2.1.40))

$$\frac{\partial \mathcal{L}}{\partial(\partial_\mu \psi)} = i\bar{\psi} \gamma^\mu, \quad \frac{\partial \mathcal{L}}{\partial(\partial_\mu \bar{\psi})} = 0. \quad (3.1.27)$$

Current conservation $\partial_\mu V^\mu = 0$ can be verified by inserting the solutions of the classical Dirac equations of motion from Eq. (2.1.41), where A^μ is the gluon field,

$$\not{\partial} \psi = (g \not{A} - \mathbf{M}) i\psi, \quad \bar{\psi} \overleftarrow{\not{\partial}} = -i\bar{\psi} (g \not{A} - \mathbf{M}), \quad (3.1.28)$$

and thus $\partial_\mu V^\mu = \bar{\psi} \not{\partial} \psi + \bar{\psi} \overleftarrow{\not{\partial}} \psi = 0$. The conserved charge is

$$Q^V(t) = \int d^3x \bar{\psi} \gamma^0 \psi = \int d^3x \psi^\dagger \psi = \text{const.} \quad (3.1.29)$$

and reflects fermion number conservation, because in the quantum field theory it counts the number of quarks minus antiquarks in the state. If we define $n_q = (\#q) - (\#\bar{q})$ for each flavor, then the eigenvalue of Q^V (which we also call Q^V) is the baryon number. For three flavors:

$$B = \frac{Q^V}{3} = \frac{n_u + n_d + n_s}{3}, \quad (3.1.30)$$

and the $U(1)_V$ symmetry entails **baryon number conservation**.

¹Here is a clash of notation: A^μ denotes both the axialvector current and the gluon field. Fortunately we won't be dealing with gluons for a while, and if so we will use the gluon field-strength tensor $F^{\mu\nu}$ instead. Unless stated otherwise, from now on A^μ will refer to an axialvector current.

■ $\mathbf{SU}(N_f)_V$: is explicitly broken by the mass matrix $\mathbf{M} \neq m \mathbf{1}$ since $U^\dagger \mathbf{M} U \neq \mathbf{M}$. We can still write down the currents, one for each generator of the group, and compute their divergences:

$$V_a^\mu = \bar{\psi} \gamma^\mu \mathbf{t}_a \psi, \quad \partial_\mu V_a^\mu = i \bar{\psi} [\mathbf{M}, \mathbf{t}_a] \psi. \quad (3.1.31)$$

The action is only invariant if all quark masses are identical. In that case the $(N_f^2 - 1)$ vector currents are conserved, $\partial_\mu V_a^\mu = 0$, and so are the corresponding charges:

$$Q_a^V(t) = \int d^3x \psi^\dagger \mathbf{t}_a \psi = \text{const}. \quad (3.1.32)$$

Because the diagonal generators (\mathbf{t}_3 in the two-flavor and $\mathbf{t}_3, \mathbf{t}_8$ in the three-flavor case) commute with each other and hence also with the mass matrix \mathbf{M} , the corresponding **isospin** and **hypercharge** currents

$$\begin{aligned} V_3^\mu &= \bar{\psi} \gamma^\mu \mathbf{t}_3 \psi = \frac{1}{2} (\bar{u} \gamma^\mu u - \bar{d} \gamma^\mu d), \\ V_8^\mu &= \bar{\psi} \gamma^\mu \mathbf{t}_8 \psi = \frac{1}{2\sqrt{3}} (\bar{u} \gamma^\mu u + \bar{d} \gamma^\mu d - 2\bar{s} \gamma^\mu s) \end{aligned} \quad (3.1.33)$$

are *always* conserved, even if $\mathbf{M} \neq m \mathbf{1}$. In combination with the vector-singlet current $V^\mu = \bar{\psi} \gamma^\mu \psi$ from Eq. (3.1.26), this implies that the flavor-diagonal vector currents $\bar{u} \gamma^\mu u$, $\bar{d} \gamma^\mu d$ and $\bar{s} \gamma^\mu s$ are individually conserved, which reflects flavor conservation in QCD. The corresponding charges are the third component of the isospin I_3 and the hypercharge Y :

$$I_3 = Q_3^V = \frac{n_u - n_d}{2}, \quad Y = \frac{2}{\sqrt{3}} Q_8^V = \frac{n_u + n_d - 2n_s}{3}. \quad (3.1.34)$$

This is what allows us to arrange hadrons in $\{I_3, Y\}$ multiplets even if the underlying flavor symmetry is broken due to the unequal quark masses. From the eigenvalues of B , I_3 and Y we obtain

$$Y = B + S, \quad Q = I_3 + \frac{Y}{2} = \frac{2}{3} n_u - \frac{1}{3} n_d - \frac{1}{3} n_s, \quad (3.1.35)$$

where $S = -n_s$ is the strangeness and Q the electric charge of the state. The relation $Q = I_3 + Y/2$ is the **Gell-Mann-Nishijima formula**.

The remaining flavor-changing vector currents have divergences proportional to quark-mass differences; if we go back to the two-flavor case with $m_u \neq m_d$ and use instead of $\mathbf{t}_{1,2} = \tau_{1,2}/2$ the generators

$$\mathbf{t}_+ = \mathbf{t}_1 + i\mathbf{t}_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad \mathbf{t}_- = \mathbf{t}_1 - i\mathbf{t}_2 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad (3.1.36)$$

we obtain

$$\partial_\mu V_\pm^\mu = i \bar{\psi} [\mathbf{M}, \mathbf{t}_\pm] \psi = i(m_u - m_d) \begin{cases} \bar{u}d \\ -\bar{d}u. \end{cases} \quad (3.1.37)$$

■ $SU(N_f)_A$: is explicitly broken by the mass matrix $M \neq 0$:

$$A_a^\mu = \bar{\psi} \gamma^\mu \gamma_5 \mathbf{t}_a \psi, \quad \partial_\mu A_a^\mu = i \bar{\psi} \{M, \mathbf{t}_a\} \gamma_5 \psi. \quad (3.1.38)$$

Even if all quark masses are equal, there remains a non-zero contribution proportional to the quark mass:

$$\partial_\mu A_a^\mu = 2m \bar{\psi} i \gamma_5 \mathbf{t}_a \psi = 2m P_a. \quad (3.1.39)$$

This is the **PCAC relation** (‘partially conserved axialvector current’): the divergence of the axialvector current is proportional to a pseudoscalar density. This equation will become extremely useful later. Using (3.1.36) in the two-flavor case, we obtain

$$\begin{aligned} \partial_\mu A_+^\mu &= i(m_u + m_d) \bar{u} \gamma_5 d, \\ \partial_\mu A_-^\mu &= i(m_u + m_d) \bar{d} \gamma_5 u, \\ \partial_\mu A_3^\mu &= im_u \bar{u} \gamma_5 u - im_d \bar{d} \gamma_5 d, \end{aligned} \quad (3.1.40)$$

which are the creation operators for the three pions π^+ , π^- and π^0 .

On the other hand, in the **chiral limit** where $M = 0$, Eq. (3.1.38) entails that the axial currents and the corresponding axial charges are conserved:

$$\partial_\mu A_a^\mu = 0 \quad \Rightarrow \quad Q_a^A(t) = \int d^3x \psi^\dagger \gamma_5 \mathbf{t}_a \psi = \text{const.} \quad (3.1.41)$$

Since the vector currents are conserved as well in that case, we have an enlarged flavor symmetry, namely **chiral symmetry**: $SU(N_f)_V \times SU(N_f)_A \simeq SU(N_f)_L \times SU(N_f)_R$. Later we will see that the $SU(N_f)_A$ part is spontaneously broken at the quantum level; nevertheless all relations for the currents remain valid.

Because of the spontaneous breaking of the axial part, in QCD the V/A terminology is more useful than the L/R notation — in contrast to the electroweak theory, where left- and right-handed fermions enter asymmetrically in the Lagrangian. Nevertheless, let us collect some relations that will become useful later. We define the chiral projectors

$$P_\pm := \frac{1}{2} (\mathbb{1} \pm \gamma_5) \quad \Rightarrow \quad P_\omega = P_\omega^\dagger, \quad \sum_\omega P_\omega = \mathbb{1}, \quad P_\omega P_{\omega'} = \delta_{\omega\omega'} P_\omega, \quad (3.1.42)$$

where chirality is denoted by the index $\omega = +$ (R , right-handed) or $\omega = -$ (L , left-handed). The projectors can be used to define right- and left-handed spinors:

$$\psi_\omega = P_\omega \psi, \quad \bar{\psi}_\omega = (P_\omega \psi)^\dagger \gamma_0 = \psi^\dagger P_\omega \gamma_0 = \bar{\psi} P_{-\omega}, \quad \psi = \sum_\omega \psi_\omega. \quad (3.1.43)$$

Now consider the product of infinitesimal vector and axialvector transformations:

$$U_V U_A = e^{i\varepsilon_V} e^{i\varepsilon_A \gamma_5} = 1 + i\varepsilon_V + i\varepsilon_A \gamma_5 + \dots = 1 + i \sum_\omega \varepsilon_\omega P_\omega = \sum_\omega U_\omega P_\omega. \quad (3.1.44)$$

Here we abbreviated $\varepsilon_{V,A} = \sum_a \varepsilon_a^{V,A} \mathbf{t}_a$ and defined $\varepsilon_\pm = \varepsilon_V \pm \varepsilon_A$ and $U_\omega = e^{i\varepsilon_\omega}$, which are all just flavor matrices. As a consequence, the left- and right-handed spinors transform as

$$\psi'_\omega = P_\omega \psi' = P_\omega U_V U_A \psi = P_\omega \sum_{\omega'} U_{\omega'} P_{\omega'} \psi = U_\omega \psi_\omega. \quad (3.1.45)$$

Therefore, they transform *separately* under $SU(N_f)_L \times SU(N_f)_R$, with independent group parameters ε_a^R and ε_a^L :

$$\psi'_\omega = U_\omega \psi_\omega, \quad \bar{\psi}'_\omega = \bar{\psi}_\omega U_\omega^\dagger, \quad U_\omega = e^{i \sum_a \varepsilon_a^\omega \mathbf{t}_a}, \quad U_\omega^\dagger = U_\omega^{-1}. \quad (3.1.46)$$

Now let us cast the currents $\bar{\psi} \Gamma \psi$ with $\Gamma \in \{\gamma^\mu, \gamma^\mu \gamma_5, \mathbb{1}, \gamma_5, \sigma^{\mu\nu}\}$ in the L/R notation:

$$\bar{\psi} \Gamma \psi = \sum_{\omega} \bar{\psi} \Gamma \mathbb{P}_{\omega}^2 \psi = \begin{cases} \sum_{\omega} \bar{\psi} \mathbb{P}_{-\omega} \Gamma \mathbb{P}_{\omega} \psi = \sum_{\omega} \bar{\psi}_{\omega} \Gamma \psi_{\omega} & \dots \Gamma \in \{\gamma^\mu, \gamma^\mu \gamma_5\}, \\ \sum_{\omega} \bar{\psi} \mathbb{P}_{\omega} \Gamma \mathbb{P}_{\omega} \psi = \sum_{\omega} \bar{\psi}_{-\omega} \Gamma \psi_{\omega} & \dots \Gamma \in \{\mathbb{1}, \gamma_5, \sigma^{\mu\nu}\}. \end{cases} \quad (3.1.47)$$

This means that for the currents constructed from the Dirac matrices γ^μ and $\gamma^\mu \gamma_5$ only the diagonal terms survive ($LL + RR$), whereas for the remaining ones only the mixed terms survive ($LR + RL$). How do these transform under chiral symmetry? The diagonal ones are invariant,

$$\sum_{\omega} \bar{\psi}'_{\omega} \Gamma \psi'_{\omega} = \sum_{\omega} \bar{\psi}_{\omega} U_{\omega}^{\dagger} \Gamma U_{\omega} \psi_{\omega} = \sum_{\omega} \bar{\psi}_{\omega} \Gamma \psi_{\omega}, \quad (3.1.48)$$

because U_{ω} is just a flavor matrix and $U_{\omega}^{\dagger} U_{\omega} = 1$. The off-diagonal currents, on the other hand, are not invariant because $U_{-\omega}^{\dagger} U_{\omega} \neq 1$:

$$\sum_{\omega} \bar{\psi}'_{-\omega} \Gamma \psi'_{\omega} = \sum_{\omega} \bar{\psi}_{\omega} \Gamma U_{-\omega}^{\dagger} U_{\omega} \psi_{\omega} \neq \sum_{\omega} \bar{\psi}_{-\omega} \Gamma \psi_{\omega}. \quad (3.1.49)$$

As a consequence, the massless Lagrangian $\bar{\psi} i \not{D} \psi$ is chirally invariant, whereas a mass term $\bar{\psi} \psi$ breaks chiral symmetry since it mixes left- and right-handed components. The Lagrangian (3.1.13) takes the form

$$\mathcal{L} = \sum_{\omega} (\bar{\psi}_{\omega} i \not{D} \psi_{\omega} - \bar{\psi}_{-\omega} \mathbf{M} \psi_{\omega}). \quad (3.1.50)$$

From the global $SU(N_f) \times SU(N_f)$ transformations we can define $2 \times (N_f^2 - 1)$ currents and charges, which are only conserved in the chiral limit:

$$j_{a,\omega}^{\mu} = \bar{\psi}_{\omega} \gamma^{\mu} \mathbf{t}_a \psi_{\omega}, \quad Q_{a,\omega} = \int d^3 x \psi_{\omega}^{\dagger} \mathbf{t}_a \psi_{\omega}. \quad (3.1.51)$$

Inserting the Dirac equations for ψ_{ω} and $\bar{\psi}_{\omega}$, their divergences for $\mathbf{M} \neq 0$ become

$$\partial_{\mu} j_{a,\omega}^{\mu} = i (\bar{\psi}_{-\omega} \mathbf{M} \mathbf{t}_a \psi_{\omega} - \bar{\psi}_{\omega} \mathbf{t}_a \mathbf{M} \psi_{-\omega}). \quad (3.1.52)$$

The vector and axialvector currents from Eqs. (3.1.31) and (3.1.38) and corresponding charges are linear combinations of these, with $V = R + L$ and $A = R - L$:

$$\begin{aligned} V_a^{\mu} &= \bar{\psi} \gamma^{\mu} \mathbf{t}_a \psi = \sum_{\omega} \bar{\psi}_{\omega} \gamma^{\mu} \mathbf{t}_a \psi_{\omega} = \sum_{\omega} j_{a,\omega}^{\mu}, \\ A_a^{\mu} &= \bar{\psi} \gamma^{\mu} \gamma_5 \mathbf{t}_a \psi = \sum_{\omega} \bar{\psi}_{\omega} \gamma^{\mu} \underbrace{\gamma_5}_{\mathbb{P}_+ - \mathbb{P}_-} \mathbf{t}_a \psi_{\omega} = j_{a,+}^{\mu} - j_{a,-}^{\mu}. \end{aligned} \quad (3.1.53)$$

■ **$U(1)_A$** : is classically conserved for $\mathbf{M} = 0$, but not preserved after quantization which leads to the $U(1)_A$ anomaly. The divergence of the axialvector singlet current picks up an anomalous contribution whose origin and consequences we will discuss in Sec. 4.3:

$$A^{\mu} = \bar{\psi} \gamma^{\mu} \gamma_5 \psi, \quad \partial_{\mu} A^{\mu} = 2i \bar{\psi} \mathbf{M} \gamma_5 \psi + \frac{g^2 N_f}{(4\pi)^2} \tilde{F}_a^{\mu\nu} F_{\mu\nu}^a. \quad (3.1.54)$$

3.1.2 Symmetry relations at the quantum level

Current and charge algebra. The symmetry relations we have discussed so far hold for the classical currents and charges. When we quantize the theory, the quark fields become operators on the state space which satisfy the anticommutation relations (2.2.67):

$$\begin{aligned} \{\psi_{\alpha i}(x), \psi_{\beta j}^\dagger(y)\}_{x^0=y^0} &= \delta^3(\mathbf{x} - \mathbf{y}) \delta_{\alpha\beta} \delta_{ij}, \\ \{\psi_{\alpha i}(x), \psi_{\beta j}(y)\}_{x^0=y^0} &= \{\psi_{\alpha i}^\dagger(x), \psi_{\beta j}^\dagger(y)\}_{x^0=y^0} = 0. \end{aligned} \quad (3.1.55)$$

As a consequence, also the currents in Eq. (3.1.23) become operators,

$$j_a^\Gamma(x) = \bar{\psi}(x) \Gamma \mathbf{t}_a \psi(x), \quad j^\Gamma(x) = \bar{\psi}(x) \Gamma \psi(x), \quad (3.1.56)$$

which satisfy the equal-time commutation relations

$$\begin{aligned} [j_a^\Gamma(x), j_b^{\Gamma'}(y)]_{x^0=y^0} &= \left[i f_{abc} j_c^{\Gamma^+}(x) + d_{abc} j_c^{\Gamma^-}(x) + \frac{\delta_{ab}}{N} j^{\Gamma^-}(x) \right] \delta^3(\mathbf{x} - \mathbf{y}), \\ [j_a^\Gamma(x), j^{\Gamma'}(y)]_{x^0=y^0} &= 2 j_a^{\Gamma^-}(x) \delta^3(\mathbf{x} - \mathbf{y}) \end{aligned} \quad (3.1.57)$$

with $\Gamma_\pm = \frac{1}{2}(\Gamma\gamma^0\Gamma' \pm \Gamma'\gamma^0\Gamma)$. These relations are valid independently of whether the currents are conserved or not. Moreover, for spacelike distances the commutators vanish, which ensures causality from Eq. (2.2.1):

$$[j_a^\Gamma(x), j_b^{\Gamma'}(y)] = [j_a^\Gamma(x), j^{\Gamma'}(y)] = 0 \quad \text{for} \quad (x - y)^2 < 0. \quad (3.1.58)$$

The proof is straightforward. We write

$$[j_a^\Gamma(x), j_b^{\Gamma'}(y)]_{x^0=y^0} = (\gamma_0 \Gamma)_{\alpha\beta} (\mathbf{t}_a)_{ij} (\gamma_0 \Gamma')_{\gamma\delta} (\mathbf{t}_b)_{kl} [\psi_{\alpha i}^\dagger(x) \psi_{\beta j}(x), \psi_{\gamma k}^\dagger(y) \psi_{\delta l}(y)]_{x^0=y^0} \quad (3.1.59)$$

and use the identity

$$[AB, CD] = A\{B, C\}D - C\{A, D\}B - \frac{\{A, C\}[B, D] + [A, C]\{B, D\}}{2} \quad (3.1.60)$$

together with the anticommutation relations (3.1.55) for the quark fields. The terms with $\{A, C\}$ and $\{B, D\}$ vanish and the commutator on the r.h.s. of (3.1.59) becomes

$$[\dots]_{x^0=y^0} = \delta^3(\mathbf{x} - \mathbf{y}) \left[\delta_{\beta\gamma} \delta_{jk} \psi_{\alpha i}^\dagger(x) \psi_{\delta l}(y) - \delta_{\alpha\delta} \delta_{il} \psi_{\gamma k}^\dagger(y) \psi_{\beta j}(x) \right]_{x^0=y^0}. \quad (3.1.61)$$

Since $\mathbf{x} = \mathbf{y}$ and $x^0 = y^0$ entails $x = y$, then in combination with the Dirac and flavor matrices the full commutator is

$$[j_a^\Gamma(x), j_b^{\Gamma'}(y)]_{x^0=y^0} = \delta^3(\mathbf{x} - \mathbf{y}) \bar{\psi}(x) (\Gamma\gamma_0\Gamma' \mathbf{t}_a \mathbf{t}_b - \Gamma'\gamma_0\Gamma \mathbf{t}_b \mathbf{t}_a) \psi(x). \quad (3.1.62)$$

With

$$AX - BY = \frac{A+B}{2}(X-Y) + \frac{A-B}{2}(X+Y), \quad \Gamma_\pm = \frac{\Gamma\gamma^0\Gamma' \pm \Gamma'\gamma^0\Gamma}{2} \quad (3.1.63)$$

and the (anti-) commutation relations (A.1.2) and (A.1.7) for the $SU(N)$ generators

$$[\mathbf{t}_a, \mathbf{t}_b] = i f_{abc} \mathbf{t}_c, \quad \{\mathbf{t}_a, \mathbf{t}_b\} = \frac{1}{N} \delta_{ab} + d_{abc} \mathbf{t}_c \quad (3.1.64)$$

we arrive at the result in Eq. (3.1.57). We note that for commutators involving spatial components of the currents these relations must be taken with some caution because additional **Schwinger terms** may arise on the r.h.s., which are derivatives of δ -functions of the form $\partial_i \delta^3(\mathbf{x} - \mathbf{y})$.

Some examples of (3.1.57) involving temporal current components are:

- For $\Gamma, \Gamma' \in \{\gamma^0, \gamma^0\gamma_5\}$ we find

$$\Gamma_+ = \begin{cases} \gamma^0 & \dots \Gamma = \Gamma', \\ \gamma^0\gamma_5 & \dots \Gamma \neq \Gamma', \end{cases} \quad \Gamma_- = 0 \quad (3.1.65)$$

which leads to the so-called ‘**local current algebra**’:

$$\begin{aligned} [V_a^0(x), V_b^0(y)]_{x^0=y^0} &= if_{abc} V_c^0(x) \delta^3(\mathbf{x} - \mathbf{y}), \\ [V_a^0(x), A_b^0(y)]_{x^0=y^0} &= if_{abc} A_c^0(x) \delta^3(\mathbf{x} - \mathbf{y}), \\ [A_a^0(x), A_b^0(y)]_{x^0=y^0} &= if_{abc} V_c^0(x) \delta^3(\mathbf{x} - \mathbf{y}). \end{aligned} \quad (3.1.66)$$

The time components V_a^0 and A_a^0 form a closed algebra since they obey equal-time commutation relations with the structure constants of the Lie algebra, and the Dirac δ -functions additionally ensure that all commutators vanish for $x \neq y$. If we further integrate over $\int d^3x$ and $\int d^3y$, we obtain the corresponding **charge algebra**:

$$[Q_a^V, Q_b^V] = [Q_a^A, Q_b^A] = if_{abc} Q_c^V, \quad [Q_a^V, Q_b^A] = if_{abc} Q_c^A. \quad (3.1.67)$$

Therefore, the charges are the generators of the symmetry group when acting on the state space. Actually, because the $SU(N_f)_A$ symmetry is spontaneously broken, the axial charges are not well-defined in the chiral limit and it is more practical to work directly with the time components of the currents.

- For $\Gamma = \gamma^0\gamma_5$ and $\Gamma' = i\gamma_5$ we find $\Gamma_+ = 0$, $\Gamma_- = -i$ and therefore

$$[Q_a^A, P_b(x)] = -i \left[\frac{\delta_{ab}}{N_f} S(x) + d_{abc} S_c(x) \right], \quad (3.1.68)$$

where $S(x) = \bar{\psi}(x)\psi(x)$ is the scalar density. Its vacuum expectation value is the scalar quark condensate which turns out to be nonzero due to spontaneous chiral symmetry breaking, and later we will use this relation for proving Goldstone’s theorem and deriving the Gell-Mann-Oakes-Renner relation.

■ Using the relation $[AB, C] = A\{B, C\} - \{A, C\}B$, we can similarly obtain the commutation relations of the currents with the quark fields,

$$\begin{aligned} [j_a^\Gamma(x), \psi(y)]_{x^0=y^0} &= -(\mathbf{t}_a \gamma^0 \Gamma \psi(x)) \delta^3(\mathbf{x} - \mathbf{y}), \\ [j_a^\Gamma(x), \bar{\psi}(y)]_{x^0=y^0} &= (\bar{\psi}(x) \Gamma \gamma^0 \mathbf{t}_a) \delta^3(\mathbf{x} - \mathbf{y}), \end{aligned} \quad (3.1.69)$$

for example for the vector currents ($\Gamma = \gamma^0$):

$$\begin{aligned} [V_a^0(x), \psi(y)]_{x^0=y^0} &= -\mathbf{t}_a \psi(x) \delta^3(\mathbf{x} - \mathbf{y}), \\ [V_a^0(x), \bar{\psi}(y)]_{x^0=y^0} &= \bar{\psi}(x) \mathbf{t}_a \delta^3(\mathbf{x} - \mathbf{y}). \end{aligned} \quad (3.1.70)$$

Integrating over $\int d^3x$, we get the commutation relations of the charges with the fields:

$$[Q_a^V(x_0), \psi(x)] = -\mathbf{t}_a \psi(x), \quad [Q_a^V(x_0), \bar{\psi}(x)] = \bar{\psi}(x) \mathbf{t}_a. \quad (3.1.71)$$

$$\partial_\mu \left[\text{diagram} \right] = (\partial_\mu j^\mu) \text{diagram} + \text{diagram with } \otimes \text{ on top} + \text{diagram with } \otimes \text{ on left} + \text{diagram with } \otimes \text{ on bottom} + \text{diagram with } \otimes \text{ on right}$$

FIG. 3.1: Generic form of a Ward-Takahashi identity from Eq. (3.1.74).

Ward-Takahashi identities. Ultimately we would like to turn the classical symmetry relations into identities for the correlation functions of the QFT. These are the **Ward-Takahashi identities (WTIs)**, which relate the n -point and $(n+1)$ -point functions of the theory with each other. They are usually derived in the path-integral approach (and we will come back to this below), but it is somewhat more transparent to work them out using canonical quantization.

Consider two generic field operators $j^\mu(x)$, $\varphi(y)$ at different spacetime points. The divergence of their time-ordered product with respect to x (with $\partial_\mu^x = \partial/\partial x^\mu$) is

$$\begin{aligned} \partial_\mu^x (\mathbb{T} j^\mu(x) \varphi(y)) &= \partial_\mu^x [\Theta(x^0 - y^0) j^\mu(x) \varphi(y) + \Theta(y^0 - x^0) \varphi(y) j^\mu(x)] \\ &= \mathbb{T} (\partial_\mu j^\mu(x)) \varphi(y) + \delta(x^0 - y^0) [j^0(x), \varphi(y)]. \end{aligned} \quad (3.1.72)$$

The first term comes from the derivative of $j^\mu(x)$ (simply resum the time orderings) and the second one results from differentiating the step functions:

$$\partial_\mu^x \Theta(x^0 - y^0) = -\partial_\mu^x \Theta(y^0 - x^0) = \delta(x^0 - y^0) \delta_{0\mu}. \quad (3.1.73)$$

Eq. (3.1.72) is quite general and retains its structure for products of n different fields (which can also be fermionic). In the general case one has to write down all possible time orderings; the time-ordering of $n+1$ distinct space-time points leads to $(n+1)!$ terms, each of which includes products of n step functions. If fermion fields are involved, the individual time-ordered terms pick up minus signs arising from the anticommutativity. In either case, the final result is the same:

$$\begin{aligned} \partial_\mu^x (\mathbb{T} j^\mu(x) \varphi_1(x_1) \dots \varphi_n(x_n)) &= \mathbb{T} (\partial_\mu j^\mu(x)) \varphi_1(x_1) \dots \varphi_n(x_n) \\ &+ \sum_{k=1}^n \delta(x^0 - x_k^0) \mathbb{T} \varphi_1(x_1) \dots [j^0(x), \varphi_k(x_k)] \dots \varphi_n(x_n). \end{aligned} \quad (3.1.74)$$

If we take its vacuum expectation value $\langle 0 | \dots | 0 \rangle$, it relates the $(n+1)$ -point function, where one leg corresponds to the external current, to the n -point functions since the commutators in the second row are proportional to the fields, cf. Eq. (3.1.69). This is the generic form of a Ward-Takahashi identity and illustrated in Fig. 3.1. Current conservation (or its absence) only enters in the first term on the r.h.s., which vanishes if the current is conserved.

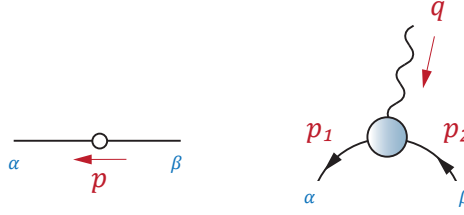


FIG. 3.2: Quark propagator and three-point function in Eqs. (3.1.75) and (3.1.76).

Let us apply this result to QCD. The quark propagator is the two-point function

$$S_{\alpha\beta}(x_1, x_2) = \langle 0 | \mathbb{T} \psi_\alpha(x_1) \bar{\psi}_\beta(x_2) | 0 \rangle. \quad (3.1.75)$$

How the quark couples to a vector, axialvector, scalar or pseudoscalar current (e.g. photons, Z -bosons, pions, ...) is encoded in the three-point function (see Fig. 3.2)

$$G_{a,\alpha\beta}^\Gamma(x, x_1, x_2) := \langle 0 | \mathbb{T} j_a^\Gamma(x) \psi_\alpha(x_1) \bar{\psi}_\beta(x_2) | 0 \rangle, \quad (3.1.76)$$

with $j^\Gamma \in \{V^\mu, A^\mu, S, P\}$. This is the *full* three-point function, which is the same as the 1PI vertex with external quark propagators attached (i.e., to obtain the vertex, multiply with inverse quark propagators from the left and right).

- The **quark-vector vertex** describes the coupling of the quark to a vector current. An example is the quark-photon vertex, which is the fundamental quantity that appears whenever a quark inside a hadron interacts with a photon.

- The **quark-axialvector vertex** encodes its coupling to an axialvector current (e.g., the W - and Z -boson interactions with quarks are linear combinations of vector and axialvector vertices).

In the vector and axialvector cases, the two- and three-point functions above are related by a WTI which follows from Eq. (3.1.74):

$$\begin{aligned} \partial_\mu^x G^\mu(x, x_1, x_2) &= \langle 0 | \mathbb{T} (\partial_\mu j^\mu(x)) \psi(x_1) \bar{\psi}(x_2) | 0 \rangle \\ &+ \delta(x^0 - x_1^0) \langle 0 | \mathbb{T} [j^0(x), \psi(x_1)] \bar{\psi}(x_2) | 0 \rangle \\ &+ \delta(x^0 - x_2^0) \langle 0 | \mathbb{T} \psi(x_1) [j^0(x), \bar{\psi}(x_2)] | 0 \rangle. \end{aligned} \quad (3.1.77)$$

If we insert the respective commutator (3.1.69) for each type of current, we reproduce the quark propagator together with a flavor generator and a δ -function. Employing vector current conservation $\partial_\mu V_a^\mu = 0$ and the PCAC relation $\partial_\mu A_a^\mu = 2mP_a$ (for equal quark masses), we obtain the **vector** and **axialvector WTIs**:

$$\partial_\mu^x G_V^\mu(x, x_1, x_2) = -\delta^4(x - x_1) \mathbf{t}_a S(x_1, x_2) + \delta^4(x - x_2) S(x_1, x_2) \mathbf{t}_a, \quad (3.1.78)$$

$$\begin{aligned} \partial_\mu^x G_A^\mu(x, x_1, x_2) &= 2m G_P(x, x_1, x_2) \\ &- \delta^4(x - x_1) \gamma_5 \mathbf{t}_a S(x_1, x_2) - \delta^4(x - x_2) S(x_1, x_2) \mathbf{t}_a \gamma_5. \end{aligned} \quad (3.1.79)$$

The quark propagator is a diagonal matrix in flavor space but with different entries for different flavors, so it does not necessarily commute with all flavor generators \mathbf{t}_a .

These relations become more transparent in momentum space, where the derivative with respect to x becomes a contraction with the momentum $q = p_1 - p_2$:

$$iq_\mu G_V^\mu(p_1, p_2) = S(p_1) \mathbf{t}_a - \mathbf{t}_a S(p_2), \quad (3.1.80)$$

$$iq_\mu G_A^\mu(p_1, p_2) = 2m G_P(p_1, p_2) - S(p_1) \mathbf{t}_a \gamma_5 - \gamma_5 \mathbf{t}_a S(p_2). \quad (3.1.81)$$

In other words, the effect of classical symmetries in the QFT is that they constrain the longitudinal (better: *non-transverse*) parts of $(n+1)$ -point functions with respect to q^μ from the corresponding n -point functions. The vector WTI can be solved to obtain the most general form of the vertex that is compatible with vector current conservation, apart from further transverse terms with respect to the momentum q^μ (more below). The axialvector WTI relates the longitudinal part of the axialvector vertex with the pseudoscalar vertex and the quark propagator. Here we considered only the flavor-octet axial current A_a^μ ; in the flavor-singlet channel we would have an additional term from the axial anomaly. Similar relations can be derived for higher n -point functions.

To work out the Fourier transforms, note that n -point correlation functions only depend on $n-1$ spacetime coordinates due to translation invariance: $G(x_1, \dots, x_n) = G(x_1 - X, \dots, x_n - X)$. For example, for a two-point function $S(x_1, x_2)$ we can define total and relative coordinates by

$$x_1 = X + \frac{z}{2}, \quad x_2 = X - \frac{z}{2} \quad \Leftrightarrow \quad X = \frac{x_1 + x_2}{2}, \quad z = x_1 - x_2. \quad (3.1.82)$$

From the behavior of the field operators under translations, Eqs. (2.2.10–2.2.11), we find

$$S(x_1, x_2) = \langle 0 | \mathbb{T} \phi(X + \frac{z}{2}) \phi(X - \frac{z}{2}) | 0 \rangle = \langle 0 | \mathbb{T} \phi(\frac{z}{2}) \phi(-\frac{z}{2}) | 0 \rangle = S(z). \quad (3.1.83)$$

Then the Fourier transform becomes

$$\int d^4 x_1 \int d^4 x_2 e^{i(p_1 x_1 - p_2 x_2)} S(x_1, x_2) = \int d^4 X \int d^4 z e^{i(PX + pz)} S(z) = (2\pi)^4 \delta^4(P) S(p), \quad (3.1.84)$$

where $P = p_1 - p_2$, $p = (p_1 + p_2)/2$ and the δ -function ensures $P = 0$, $p_1 = p_2 = p$ so that

$$S(z) = \int \frac{d^4 p}{(2\pi)^4} e^{-ipz} S(p), \quad S(p) = \int d^4 z e^{ipz} S(z). \quad (3.1.85)$$

For a three-point function $G(x, x_1, x_2)$ we add the coordinate $x = X - y$ to (3.1.82). Translation invariance implies

$$G(x, x_1, x_2) = \langle 0 | \mathbb{T} \phi(-y) \phi(\frac{z}{2}) \phi(-\frac{z}{2}) | 0 \rangle = G(y, z) \quad (3.1.86)$$

and the Fourier transform becomes

$$\begin{aligned} \int d^4 x \int d^4 x_1 \int d^4 x_2 e^{i(p_1 x_1 - p_2 x_2 - qx)} G(x, x_1, x_2) &= \int d^4 X \int d^4 z \int d^4 y e^{i(PX + pz + qy)} G(y, z) \\ &= (2\pi)^4 \delta^4(P) G(p, q), \end{aligned}$$

where $P = p_1 - p_2 - q$, the average momentum is $p = (p_1 + p_2)/2$, and the δ -function ensures $q = p_1 - p_2$. In the same way we can work out

$$\int d^4 x_1 \int d^4 x_2 \int d^4 x e^{i(p_1 x_1 - p_2 x_2 - qx)} \begin{bmatrix} G^\mu(x, x_1, x_2) \\ \partial_\mu^x G^\mu(x, x_1, x_2) \\ \delta^4(x - x_1) S(x_1, x_2) \\ \delta^4(x - x_2) S(x_1, x_2) \end{bmatrix} = (2\pi)^4 \delta^4(P) \begin{bmatrix} G^\mu(p_1, p_2) \\ iq_\mu G^\mu(p_1, p_2) \\ S(p_2) \\ S(p_1) \end{bmatrix}$$

to arrive at Eq. (3.1.80). Note that we use the notation $G(p_1, p_2)$ and $G(p, q)$ interchangeably to keep things transparent, and we employ the same symbol for $G(y, z)$ in coordinate space, but this does not mean that G is the same function of the arguments.

Let us rewrite the vector WTI (3.1.80) for the **vector vertex** $\Gamma_V^\mu(p_1, p_2)$ defined by

$$G_V^\mu(p_1, p_2) = S(p_1) \Gamma_V^\mu(p_1, p_2) S(p_2). \quad (3.1.87)$$

If we multiply with $S(p_1)^{-1}$ from the left and $S(p_2)^{-1}$ from the right and denote $p_1 = p + q/2$, $p_2 = p - q/2$, the WTI becomes

$$q_\mu \Gamma^\mu(p, q) = iS(p + \frac{q}{2})^{-1} - iS(p - \frac{q}{2})^{-1}, \quad (3.1.88)$$

where we dropped the flavor matrices for simplicity. Inserting the decomposition (2.3.11) for the quark propagator, we obtain

$$\begin{aligned} q_\mu \Gamma^\mu(p, q) &= A(p_1^2) (\not{p}_1 - M(p_1^2)) - A(p_2^2) (\not{p}_2 - M(p_2^2)) \\ &= (A(p_1^2) - A(p_2^2)) \not{p} + \frac{A(p_1^2) + A(p_2^2)}{2} \not{q} - (B(p_1^2) - B(p_2^2)) \\ &= \Sigma_A \not{q} + 2p \cdot q (\Delta_A \not{p} - \Delta_B). \end{aligned} \quad (3.1.89)$$

Here we defined $B(p^2) = A(p^2) M(p^2)$ and the average and difference quotient

$$\Sigma_A = \frac{A(p_1^2) + A(p_2^2)}{2}, \quad \Delta_A = \frac{A(p_1^2) - A(p_2^2)}{p_1^2 - p_2^2} = \frac{A(p_1^2) - A(p_2^2)}{2p \cdot q}, \quad (3.1.90)$$

which are regular for $q^\mu \rightarrow 0$, and likewise for Δ_B . As a result, we can split off the momentum q^μ and read off the **Ball-Chiu vertex**

$$\Gamma_{\text{BC}}^\mu(p, q) = \Sigma_A \gamma^\mu + 2p^\mu (\Delta_A \not{p} - \Delta_B). \quad (3.1.91)$$

For a tree-level vertex with the replacements $A(p^2) \rightarrow Z_\psi$ and $M(p^2) \rightarrow m_B$, this expression becomes $\Gamma_{\text{BC}}^\mu \rightarrow Z_\psi \gamma^\mu$ as expected.

Instead of the singlet and octet currents V^μ and V_a^μ , we could also consider linear combinations of them such as the **electromagnetic current**, which couples to the quarks through the quark charge matrix \mathbf{Q} , e.g. for three flavors:

$$\mathbf{Q} = \begin{pmatrix} q_u & 0 & 0 \\ 0 & q_d & 0 \\ 0 & 0 & q_s \end{pmatrix} = \mathbf{t}_3 + \frac{\mathbf{t}_8}{\sqrt{3}} \quad \Rightarrow \quad V_{\text{em}}^\mu(x) = \bar{\psi} \gamma^\mu \mathbf{Q} \psi = V_3^\mu + \frac{1}{\sqrt{3}} V_8^\mu. \quad (3.1.92)$$

The corresponding vertex is the **quark-photon vertex**, which satisfies the same WTI (3.1.80) if we replace $\mathbf{t}_a \rightarrow \mathbf{Q}$ and thus has the same form as above. In other words, once we know the quark propagator, we already know a great deal about the quark-photon vertex from symmetries alone. The full vertex can be written as $\Gamma^\mu = \Gamma_{\text{BC}}^\mu + \Gamma_T^\mu$ with $q_\mu \Gamma_T^\mu = 0$, where the transverse part is not constrained and carries the dynamics such as the vector-meson poles (more in Sec. 3.1.3).

Here one can also see that it is actually not the *longitudinal* part that is constrained by the WTI, because if we had split the vertex into $\Gamma^\mu = q^\mu \Gamma_L + \Gamma_T^\mu$ with $q_\mu \Gamma_T^\mu = 0$, we would have obtained

$$\Gamma_L = \frac{1}{q^2} \left[iS(p + \frac{q}{2})^{-1} - iS(p - \frac{q}{2})^{-1} \right] = \frac{1}{q^2} [\Sigma_A \not{q} + 2p \cdot q (\Delta_A \not{p} - \Delta_B)], \quad (3.1.93)$$

which is singular for $q^\mu \rightarrow 0$ (and violates the **Ward identity** $\Gamma^\mu(p, 0) = i \partial S(p)^{-1} / \partial p_\mu$ which follows from Eq. (3.1.88)). One can systematically work out the WTI constraints for n -point functions by constructing ‘minimal’ tensor bases that are free of kinematic singularities and constraints.

WTIs from the path integral. As already mentioned around Eq. (2.2.113), WTIs can also be derived in the path integral formalism. This is relatively straightforward to do for the **Abelian local** $U(1)$ gauge invariance in QED. The partition function in QED has the same form as in QCD,

$$Z[J, \bar{\eta}, \eta] = \int \mathcal{D}[A, \psi, \bar{\psi}] e^{i(S[A, \psi, \bar{\psi}] + S_{\text{GF}}[A] + S_{\text{C}}[A, \psi, \bar{\psi}])}, \quad (3.1.94)$$

except there are no ghosts because the Faddeev-Popov determinant is independent of the photon field A^μ and can be pulled out of the path integral. The resulting gauge-fixing and source terms read

$$S_{\text{GF}} + S_{\text{C}} = \int d^4x \left[\frac{1}{2\xi} A^\mu \partial_\mu \partial_\nu A^\nu - J_\mu A^\mu - \bar{\psi} \eta - \bar{\eta} \psi \right]. \quad (3.1.95)$$

If we keep the sources fixed, then a gauge transformation is just a relabeling of fields under the integral and leaves the generating functional invariant. Since the QED action is gauge invariant, and assuming that the integral measure remains invariant as well, this only affects the gauge-fixing and source terms:

$$\begin{aligned} Z[J, \bar{\eta}, \eta] &= \int \mathcal{D}A' \mathcal{D}\psi' \mathcal{D}\bar{\psi}' e^{i(S[A', \psi', \bar{\psi}'] + S_{\text{GF}}[A'] + S_{\text{C}}[A', \psi', \bar{\psi}'])} \\ &= \int \mathcal{D}A \mathcal{D}\psi \mathcal{D}\bar{\psi} e^{i(S[A, \psi, \bar{\psi}] + S_{\text{GF}}[A] + S_{\text{C}}[A, \psi, \bar{\psi}])} e^{i(\delta S_{\text{GF}} + \delta S_{\text{C}})} \\ &= Z[J, \bar{\eta}, \eta] \left\langle e^{i(\delta S_{\text{GF}} + \delta S_{\text{C}})} \right\rangle_J \Rightarrow \langle \delta S_{\text{GF}} + \delta S_{\text{C}} \rangle_J = 0. \end{aligned} \quad (3.1.96)$$

Inserting the infinitesimal gauge transformations (2.1.37) in the Abelian case,

$$\delta\psi = i\varepsilon\psi, \quad \delta\bar{\psi} = -i\bar{\psi}\varepsilon, \quad \delta A_\mu = \frac{1}{g}\partial_\mu\varepsilon, \quad (3.1.97)$$

and taking partial integrations to factor out $\varepsilon(x)$, we obtain

$$\langle \delta S_{\text{GF}} + \delta S_{\text{C}} \rangle_J = \int d^4x \varepsilon(x) \left\langle \frac{1}{g} \partial_\mu \left(J^\mu - \frac{1}{\xi} \square A^\mu \right) + i(\bar{\psi} \eta - \bar{\eta} \psi) \right\rangle_J = 0. \quad (3.1.98)$$

Since $\varepsilon(x)$ is arbitrary, the integrand must vanish too. At this point there is also no longer a need to distinguish the classical fields from the averaged fields in the notation (ϕ versus φ in Eqs. (2.2.44) and (2.2.55)), so we simply write $\langle A^\mu \rangle_J = A^\mu$, $\langle \bar{\psi} \rangle_J = \bar{\psi}$ and $\langle \psi \rangle_J = \psi$:

$$\partial_\mu \left(J^\mu - \frac{1}{\xi} \square A^\mu \right) + ig(\bar{\psi} \eta - \bar{\eta} \psi) = 0. \quad (3.1.99)$$

With the effective action $\Gamma[A, \psi, \bar{\psi}] = W[J, \bar{\eta}, \eta] + S_{\text{C}}$ we can use (2.2.44) to transform this relation into a generating WTI for connected n -point functions by writing

$$A^\mu = -\frac{\delta W}{\delta J_\mu}, \quad \bar{\psi} = \frac{\delta W}{\delta \eta}, \quad \psi = -\frac{\delta W}{\delta \bar{\eta}}. \quad (3.1.100)$$

Here we took into account the Grassmann nature of the sources η and $\bar{\eta}$, i.e., η has to be permuted to the derivative operator which gives a minus sign. If we take further derivatives with respect to $\bar{\eta}$ and η , we arrive at the WTI for the connected three-point function. Alternatively, we can use (2.2.45) and convert the equation into a generating WTI for 1PI n -point functions by writing

$$J^\mu = \frac{\delta\Gamma}{\delta A_\mu}, \quad \bar{\eta} = -\frac{\delta\Gamma}{\delta\bar{\psi}}, \quad \eta = \frac{\delta\Gamma}{\delta\psi}. \quad (3.1.101)$$

In this case Eq. (3.1.99) becomes

$$\partial_\mu \left(\frac{\delta\Gamma}{\delta A_\mu} - \frac{1}{\xi} \square A^\mu \right) + ig \left(\bar{\psi} \frac{\delta\Gamma}{\delta\bar{\psi}} + \frac{\delta\Gamma}{\delta\psi} \psi \right) = 0. \quad (3.1.102)$$

Taking two further derivatives with respect to $\bar{\psi}$ and ψ and setting all fields to zero yields the WTI for the 1PI fermion-photon vertex:

$$\partial_\mu \left(\frac{1}{g} \frac{\delta^3\Gamma}{\delta\psi(x_1) \delta\bar{\psi}(x_2) \delta A_\mu(x)} \right) = \frac{i \delta^2\Gamma}{\delta\psi(x_1) \delta\bar{\psi}(x_2)} [\delta^4(x-x_1) - \delta^4(x-x_2)]. \quad (3.1.103)$$

This is identical to Eq. (3.1.88) and says that the divergence of the vertex equals the difference of the inverse quark propagators. We can also take a derivative of Eq. (3.1.102) with respect to A^μ , which yields the WTI for the inverse photon propagator:

$$\partial_\mu \frac{\delta^2\Gamma}{\delta A_\mu(x) \delta A_\nu(y)} = \frac{1}{\xi} \square \partial^\nu \delta^4(x-y). \quad (3.1.104)$$

In momentum space, this entails that the longitudinal part of the propagator remains undressed, which proves our statement below Eq. (2.3.14) in the Abelian theory:

$$q_\mu (D^{-1})^{\mu\nu}(q) = iq^2 \frac{q^\nu}{\xi}. \quad (3.1.105)$$

In principle one can derive WTIs also for **non-Abelian local** gauge symmetries, but they become very cumbersome and it is more convenient to use BRST invariance to obtain relations of the form (2.2.114). The quickest way to generate the **Slavnov-Taylor identities** is to apply the BRST transformation directly to the correlation functions, which must also be BRST-invariant since already the QCD action including the gauge-fixing terms is BRST-invariant. As an example, we derive the QCD version of Eq. (3.1.105) by starting from the BRST transformations in Eq. (2.2.109),

$$\delta\psi = ic\psi, \quad \delta\bar{\psi} = -i\bar{\psi}c, \quad \delta A_a^\mu = \frac{1}{g} D_{ab}^\mu c_b, \quad \delta c_a = -\frac{1}{2} f_{abc} c_b c_c, \quad \delta\bar{c}_a = -\frac{B_a}{g} = \frac{\partial_\mu A_a^\mu}{g\xi},$$

where we inserted the equations of motion $f_a[A] + \xi B_a = 0$ for B_a . Now consider the quantity

$$\frac{\partial}{\partial x^\mu} \delta \langle A_a^\mu(x) \bar{c}_b(y) \rangle_J = \frac{1}{g} \langle \partial_\mu D_{ac}^\mu c_c(x) \bar{c}_b(y) \rangle_J + \frac{1}{g\xi} \partial_\mu^x \partial_\nu^y \langle A_a^\mu(x) A_b^\nu(y) \rangle_J. \quad (3.1.106)$$

For vanishing sources the l.h.s. is zero. In momentum space, the second term on the right is the contraction of the gluon propagator with $q_\mu q_\nu$. For the first term we insert the DSE for the ghost propagator obtained from Eq. (2.2.40):

$$\left\langle \frac{\delta S}{\delta\bar{c}_a(x)} \bar{c}_b(y) \right\rangle = \langle \partial_\mu D_{ac}^\mu c_c(x) \bar{c}_b(y) \rangle = i\delta^4(x-y) \delta_{ab}. \quad (3.1.107)$$

Thus, in momentum space we arrive at $q_\mu q_\nu D_{ab}^{\mu\nu}(q) = -i\xi \delta_{ab}$, which states that the longitudinal part of the gluon propagator remains undressed also with interactions.

Now what if we are instead interested in **global flavor symmetries**? Let us first check QED with a global $U(1)$ symmetry instead of a local one. In that case $\delta A^\mu = 0$ since ε is a constant, and we can no longer eliminate the integral in Eq. (3.1.98) but get instead:

$$\langle \delta S_{\text{GF}} + \delta S_C \rangle_J = i\varepsilon \int d^4x \langle \bar{\psi} \eta - \bar{\eta} \psi \rangle_J = 0. \quad (3.1.108)$$

This equation is correct but not very useful: In the context of Eq. (3.1.88) it only tells us that the integrated equation vanishes – or in momentum space, that the difference of propagators on the right-hand side vanishes if their momenta are equal ($q^\mu = 0$).

We can cure the problem by tricking the path integral into believing that it deals with a *local* symmetry instead of a global one. Suppose we start from the free quark Lagrangian in Eq. (3.1.13) without the quark-gluon vertex, which it is not relevant for the discussion:

$$\mathcal{L} = \bar{\psi} (i\cancel{\partial} - m) \psi, \quad Z[\eta, \bar{\eta}] = \int \mathcal{D}[\psi, \bar{\psi}] e^{i(S[\psi, \bar{\psi}] + S_C)} \quad (3.1.109)$$

with source terms $S_C = -\int d^4x (\bar{\psi} \eta + \bar{\eta} \psi)$. The action $S[\psi, \bar{\psi}]$ is invariant under the global $SU(N_f)_V \times U(1)_V$ symmetry. We consider $U(1)_V$ for simplicity, whose flavor-singlet current $V^\mu = \bar{\psi} \gamma^\mu \psi$ is conserved. The idea is to add more source terms to the action and define appropriate gauge transformations for the source fields, so that the *total* action including all sources becomes *locally* gauge invariant with respect to $U(1)_V$. This means we need a covariant derivative; from Eq. (3.1.109) we only need to add a term $\bar{\psi} B \psi = V \cdot B$ to establish local $U(1)$ gauge invariance:

$$Z[B, \eta, \bar{\eta}] = \int \mathcal{D}[\psi, \bar{\psi}] e^{i(S[\psi, \bar{\psi}] + V \cdot B + S_C)}. \quad (3.1.110)$$

Here, B plays the role of the gauge field but it is a **‘background field’** since it does not appear in the path integral measure and thus does not change the content of the QFT. From Eq. (2.1.37) we have $\delta B_\mu = \partial_\mu \varepsilon$ because we are dealing with an Abelian gauge symmetry (we set the irrelevant new coupling to 1). As a result, the sum $S[\psi, \bar{\psi}] + V \cdot B$ is locally gauge invariant. Finally, we also make S_C gauge invariant in itself by imposing appropriate gauge transformations $\delta \eta = i\varepsilon \eta$ and $\delta \bar{\eta} = -i\varepsilon \bar{\eta}$ for the source fields.

Now start from $Z[B, \eta, \bar{\eta}]$ and perform a gauge transformation to primed quantities $\{\psi', \bar{\psi}', B', \eta', \bar{\eta}'\}$. The total action is gauge-invariant and the path integral measure as well, so that also the partition function is invariant: $Z[B, \eta, \bar{\eta}] = Z[B', \eta', \bar{\eta}']$. Next, relabel the fields ψ and $\bar{\psi}$ back to unprimed quantities and work out the transformation of B , η and $\bar{\eta}$ only:

$$\begin{aligned} & \int d^4x \langle V \cdot \delta B - \bar{\psi} \delta \eta - \delta \bar{\eta} \psi \rangle_J \\ &= \int d^4x [\langle V^\mu \rangle_J \partial_\mu \varepsilon - i\varepsilon (\langle \bar{\psi} \rangle_J \eta - \bar{\eta} \langle \psi \rangle_J)] \\ &= - \int d^4x \varepsilon(x) [\partial_\mu \langle V^\mu \rangle_J + i (\langle \bar{\psi} \rangle_J \eta - \bar{\eta} \langle \psi \rangle_J)] = 0. \end{aligned} \quad (3.1.111)$$

Once again, because $\varepsilon(x)$ is arbitrary, we arrive at

$$\partial_\mu \langle V^\mu \rangle_J + i (\langle \bar{\psi} \rangle_J \eta - \bar{\eta} \langle \psi \rangle_J) = 0. \quad (3.1.112)$$

In order to arrive at Eq. (3.1.78) including connected Green functions, replace

$$\langle V^\mu \rangle_J = -\frac{\delta W}{\delta B_\mu}, \quad \langle \bar{\psi} \rangle_J = \frac{\delta W}{\delta \eta}, \quad \langle \psi \rangle_J = -\frac{\delta W}{\delta \bar{\eta}} \quad (3.1.113)$$

and perform a partial integration. Since $\varepsilon(x)$ is again arbitrary one can remove the integral, and the resulting master WTI becomes

$$\partial_\mu \frac{\delta W}{\delta B_\mu} = i \left(\frac{\delta W}{\delta \eta} \eta + \bar{\eta} \frac{\delta W}{\delta \bar{\eta}} \right). \quad (3.1.114)$$

It has the same form as in our first attempt (3.1.108) except that now we have a new correlation function $\delta W/\delta B$ that incorporates the current. The vector WTI (3.1.78) follows from applying two further derivatives with respect to η and $\bar{\eta}$ and setting the sources to zero.

Renormalization of currents. So far we have only dealt with bare currents that we derived from the bare Lagrangian (3.1.13). However, if we included renormalization constants for the vector and axialvector currents, the current-algebra relations (3.1.66) would fix both of them to $Z^2 = Z = 1$. Hence, these currents stay unrenormalized, which entails

$$\begin{aligned} V_B^\mu &= (\bar{\psi} \gamma^\mu \psi)_B = Z_2 (\bar{\psi} \gamma^\mu \psi)_R = V_R^\mu, \\ A_B^\mu &= (\bar{\psi} \gamma^\mu \gamma_5 \psi)_B = Z_2 (\bar{\psi} \gamma^\mu \gamma_5 \psi)_R = A_R^\mu. \end{aligned} \quad (3.1.115)$$

On the other hand, those relations do not give us closed equations for the scalar and pseudoscalar densities. In that case we can exploit the fact that their divergences are proportional to the quark masses, e.g., from the PCAC relation:

$$\partial_\mu A_B^\mu = (2mP)_B \stackrel{!}{=} (2mP)_R = \partial_\mu A_R^\mu \quad \Rightarrow \quad P_B = \frac{1}{Z_m} P_R, \quad (3.1.116)$$

and consequently

$$P_B = (\bar{\psi} \gamma_5 \psi)_B = Z_2 (\bar{\psi} \gamma_5 \psi)_R = \frac{1}{Z_m} P_R. \quad (3.1.117)$$

The same result follows for the scalar density. In summary, the renormalized currents are (we drop the label 'R'):

$$\begin{aligned} V^\mu &= Z_2 \bar{\psi} \gamma^\mu \psi, & P &= Z_2 Z_m \bar{\psi} \gamma_5 \psi, \\ A^\mu &= Z_2 \bar{\psi} \gamma^\mu \gamma_5 \psi, & S &= Z_2 Z_m \bar{\psi} \psi. \end{aligned} \quad (3.1.118)$$

3.1.3 Extracting hadrons from QCD

We have not yet talked about how we can actually extract **hadron properties** from QCD. How would you calculate the mass of a hadron in a QFT? In quantum mechanics the answer is clear: define a potential V and solve the Schrödinger equation $H\psi = E\psi$ to obtain the energy spectrum of the system. Once you know the wave function ψ , you can calculate matrix elements for observables. But what becomes of the Schrödinger equation in QFT? Earlier we argued that the well-defined objects in a QFT are the correlation functions and that they encode the full content of the theory. Therefore, they should also carry any possible information on hadrons. But how can we extract that information?

We already mentioned in Sec. 2.2.1 that hadrons are contained in the state space of QCD: $|\pi\rangle$, $|N\rangle$, ... are one-particle states with a well-defined mass, momentum \mathbf{p} and other quantum numbers that reflect the symmetries of QCD (angular momentum, parity, flavor, etc.). As a consequence, they enter in the completeness relation

$$\mathbb{1} = \sum_{\lambda} \frac{1}{(2\pi)^3} \int \frac{d^3p}{2E_p} |\lambda\rangle\langle\lambda|, \quad (3.1.119)$$

where the Lorentz-invariant integral weight implements the condition that each hadron is on its mass shell ($p^2 = m_{\lambda}^2$ or $E_p^2 = \mathbf{p}^2 + m_{\lambda}^2$). One should keep in mind that the state space of QCD is *enormous*: it can contain (unphysical) colored states, colorless ‘one-particle’ bound states like mesons and baryons but also glueballs, multiquark and multi-hadron states – also the C^{14} nucleus should be somewhere buried in it.

Hadrons generate poles. In principle, the extraction of hadron properties from QCD is based on the **spectral representation** (2.2.7), which is also closely related to the experimental situation. When we derive it for a two-point function, it tells us that the onshell states in the completeness relation produce poles in the propagator at $p^2 = m_{\lambda}^2$, where m_{λ} is the physical mass of the state, and the multiparticle states produce cuts which start at $p^2 = 4m_{\lambda}^2$ and extend to infinity. Unfortunately this does not quite work out in QCD: when we insert the completeness relation into a quark or gluon two-point function, a colored quark cannot create a colorless hadron. In other words, quark and gluon propagators cannot produce physical hadron poles.

Fortunately, the spectral representation is not limited to two-point functions. In general one can show that for a given correlation function

$$G(x_1, \dots, x_r) = \langle 0 | \mathbb{T} \phi(x_1) \dots \phi(x_n) \phi(y_1) \dots \phi(y_r) | 0 \rangle \quad (3.1.120)$$

each one-particle state $|\lambda(p)\rangle$ with onshell momentum $p^2 = m_{\lambda}^2$ produces a pole, where the correlation function factorizes at the pole:

$$G(x_1, \dots, x_r) = \int \frac{d^4p}{(2\pi)^4} e^{-ipz} \left[\frac{i \Psi(\{x_i\}, p) \Psi^{\dagger}(\{y_j\}, p)}{p^2 - m_{\lambda}^2 + i\epsilon} + \text{finite} \right]. \quad (3.1.121)$$

This is true as long as the residues $\Psi(\{x_i\}, p)$ at the poles are nonzero. These residues are the transition elements between the vacuum and the onshell hadron and they are called **Bethe-Salpeter wave functions (BSWFs)**.

The proof goes as follows. We start from a general $(n+r)$ -point function

$$G(x_1, \dots, y_r) = \langle 0 | \mathbb{T} \phi(x_1) \dots \phi(x_n) \phi(y_1) \dots \phi(y_r) | 0 \rangle. \quad (3.1.122)$$

Because we want to insert the completeness relation, we are only interested in the time orderings where all $x_i^0 > y_j^0$. We can separate this contribution by writing

$$G(x_1, \dots, y_r) = \Theta_{xy} \langle 0 | \mathbb{T} \{ \phi(x_1) \dots \phi(x_n) \} \mathbb{T} \{ \phi(y_1) \dots \phi(y_r) \} | 0 \rangle + (\dots), \quad (3.1.123)$$

where $\Theta_{xy} := \theta(\min(x_1^0, \dots, x_n^0) - \max(y_1^0, \dots, y_r^0))$ and (\dots) contains the remaining time orderings. Inserting the completeness relation, this becomes

$$G(x_1, \dots, y_r) = \sum_{\lambda} \frac{1}{(2\pi)^3} \int \frac{d^3 k}{2E_{\lambda}} \Theta_{xy} \langle 0 | \mathbb{T} \phi(x_1) \dots \phi(x_n) | \lambda \rangle \langle \lambda | \mathbb{T} \phi(y_1) \dots \phi(y_r) | 0 \rangle + (\dots), \quad (3.1.124)$$

where $|\lambda\rangle$ is an onshell state with momentum \mathbf{k} and energy $E_{\lambda} = \sqrt{\mathbf{k}^2 + m_{\lambda}^2}$.

For the correlation functions of the theory, which are vacuum-to-vacuum transition matrix elements, translation invariance entails that they do not depend on the total coordinate (see Eq. (3.1.83)). For a vacuum-to-hadron amplitude, the behavior of the field operators and one-particle states under translations $U(1, a)$,

$$U(1, a) \psi_{\alpha}(x) U(1, a)^{-1} = \psi_{\alpha}(x + a), \quad U(1, a) |\lambda(p)\rangle = e^{ip \cdot a} |\lambda(p)\rangle, \quad U(1, a) |0\rangle = |0\rangle \quad (3.1.125)$$

entails that the dependence on the total coordinate can only enter through a phase. That is, if we write $x_i = X + x'_i$ and $y_i = Y + y'_i$, we can factor out the dependence on X and Y :

$$\begin{aligned} \langle 0 | \mathbb{T} \phi(x_1) \dots \phi(x_n) | \lambda \rangle &= \langle 0 | \mathbb{T} U(1, X) \phi(x'_1) U(1, X)^{-1} \dots U(1, X) \phi(x'_n) U(1, X)^{-1} | \lambda \rangle \\ &= \langle 0 | \mathbb{T} \phi(x'_1) \dots \phi(x'_n) | \lambda \rangle e^{-ikX} = \Psi(\{x_i\}, \mathbf{k}) e^{-ikX}, \end{aligned} \quad (3.1.126)$$

where the Bethe-Salpeter wave function $\Psi(\{x_i\}, \mathbf{k})$ only depends on $n-1$ coordinates and no longer on X . (For example, if we set $X = x_1$ it only depends on $x'_2 \dots x'_n$.) Likewise,

$$\langle \lambda | \mathbb{T} \phi(y_1) \dots \phi(y_r) | 0 \rangle = \langle \lambda | \mathbb{T} \phi(y'_1) \dots \phi(y'_r) | 0 \rangle e^{ikY} = \Psi^{\dagger}(\{y_j\}, \mathbf{k}) e^{ikY}. \quad (3.1.127)$$

Denoting $z = X - Y$, we also have

$$\min(x_1^0, \dots, x_n^0) - \max(y_1^0, \dots, y_r^0) = X^0 - Y^0 + \min(x'_1{}^0, \dots, x'_n{}^0) - \max(y'_1{}^0, \dots, y'_r{}^0) =: z^0 + \Delta \quad (3.1.128)$$

and the full correlation function becomes

$$G(x_1, \dots, y_r) = \sum_{\lambda} \frac{1}{(2\pi)^3} \int \frac{d^3 k}{2E_{\lambda}} \theta(z^0 + \Delta) e^{-ikz} \Psi(\{x_i\}, \mathbf{k}) \Psi^{\dagger}(\{y_j\}, \mathbf{k}) + (\dots). \quad (3.1.129)$$

Now we use the following representation of the step function:

$$\theta(x) = \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} \frac{i}{\omega + i\epsilon} e^{-i\omega x} \quad (3.1.130)$$

and take the Fourier transform with respect to z :

$$\begin{aligned} \int d^4 z e^{ipz} G(x_1, \dots, y_r) &= \sum_{\lambda} \int_{-\infty}^{\infty} d\omega \frac{i}{\omega + i\epsilon} e^{-i\omega\Delta} \int \frac{d^3 k}{2E_{\lambda}} \underbrace{\frac{1}{(2\pi)^4} \int d^4 z e^{i(p-k)z} e^{-i\omega z^0}}_{\delta^3(\mathbf{p}-\mathbf{k}) \delta(p^0 - E_{\lambda} - \omega)} \\ &\quad \times \Psi(\{x_i\}, \mathbf{k}) \Psi^{\dagger}(\{y_j\}, \mathbf{k}) + (\dots). \end{aligned} \quad (3.1.131)$$

When integrating over $d^3 k$, the δ -function sets $\mathbf{k} = \mathbf{p}$ and thus $E_{\lambda} = \sqrt{\mathbf{p}^2 + m_{\lambda}^2}$, so we arrive at

$$\int d^4 z e^{ipz} G(x_1, \dots, y_r) = \sum_{\lambda} \frac{i}{p^0 - E_{\lambda} + i\epsilon} \frac{e^{-i(p^0 - E_{\lambda})\Delta}}{2E_{\lambda}} \Psi(\{x_i\}, \mathbf{p}) \Psi^{\dagger}(\{y_j\}, \mathbf{p}) + (\dots). \quad (3.1.132)$$

Furthermore, we can write

$$\frac{i}{p_0 - E_\lambda + i\epsilon} \frac{e^{-i(p^0 - E_\lambda)\Delta}}{2E_\lambda} = \frac{i(p_0 + E_\lambda)}{p_0^2 - E_\lambda^2 + i\epsilon} \frac{e^{-i(p^0 - E_\lambda)\Delta}}{2E_\lambda} \approx \frac{i}{p^2 - m_\lambda^2 + i\epsilon}, \quad (3.1.133)$$

where we approximated $p_0 \approx E_\lambda$ in the vicinity of the pole. Thus we arrive at the final result

$$\int d^4 z e^{ipz} G(x_1, \dots, y_r) \xrightarrow{p^2 \rightarrow m_\lambda^2} \frac{i \Psi(\{x_i\}, p) \Psi^\dagger(\{y_j\}, p)}{p^2 - m_\lambda^2 + i\epsilon}. \quad (3.1.134)$$

Note that n and r can be different, which means we can squeeze in the completeness relation at *any* position in a matrix element $\langle 0 | \mathbb{T} \phi(x_1) \phi(x_2) \dots | 0 \rangle$. As long as the BSWFs on both sides are non-zero so that there is a non-vanishing overlap with the onshell state $|\lambda\rangle$, this will produce a pole at $p^2 = m_\lambda^2$ in the form of a Feynman propagator.

Applied to QCD, this means that even though hadrons are color singlets and cannot produce poles in elementary two-point functions, they still generate poles in *higher* n -point functions, e.g. the quark-antiquark **four-point function** in Fig. 3.3:

$$G_{\alpha\beta\gamma\delta}(x_1, x_2, x_3, x_4) = \langle 0 | \mathbb{T} \psi_\alpha(x_1) \bar{\psi}_\beta(x_2) \psi_\gamma(x_3) \bar{\psi}_\delta(x_4) | 0 \rangle. \quad (3.1.135)$$

Inserting a complete set of states will produce meson poles because a composite operator $\psi\bar{\psi}$ can produce color singlet quantum numbers ($3 \otimes \bar{3} = 1 \oplus 8$). In fact, the four-point function encodes the *complete* meson spectrum that is compatible with the flavor quantum numbers of the quarks. The corresponding BSWF of a meson reads

$$\begin{aligned} \langle 0 | \mathbb{T} \psi_\alpha(x_1) \bar{\psi}_\beta(x_2) | \lambda, a \rangle &= \langle 0 | \mathbb{T} \psi_\alpha\left(\frac{z}{2}\right) \bar{\psi}_\beta\left(-\frac{z}{2}\right) | \lambda, a \rangle e^{-ip \cdot x} \\ &= \Psi_{\alpha\beta}^a(z, p) e^{-ip \cdot x}, \end{aligned} \quad (3.1.136)$$

where we set the total coordinate as $x = (x_1 + x_2)/2$ and the relative one by $z = x_1 - x_2$, so that $x_1 = x + \frac{z}{2}$ and $x_2 = x - \frac{z}{2}$. Since the flavor quantum numbers of mesons are related to the $SU(N_f)$ generators \mathbf{t}_a , we attached a flavor index a . Likewise, we would find baryon poles in the analogous quark six-point function and we could write down the analogous BSWF with three quark fields.

The BSWFs are not truly ‘wave functions’ in the quantum-mechanical sense since they transform under finite-dimensional, non-unitary representations of the Lorentz group and thus one cannot directly extract probability information from them. Depending on the total angular momentum J of the onshell hadron $|\lambda\rangle$, after splitting off polarization vectors (for $J = 1$ states), Dirac spinors (for $J = \frac{1}{2}$ states) etc., they can be expanded in tensor bases just like the correlation functions in Eq. (2.3.7):

$$\Psi_{\alpha\beta\dots}^{\mu\nu\dots}(\{q_i\}, p) = \sum_{i=1}^N f_i(q_1^2, q_2^2, q_1 \cdot p, \dots, p^2 = m_\lambda^2) \tau_i(\{q_i\}, p)_{\alpha\beta\dots}^{\mu\nu\dots}. \quad (3.1.137)$$

For example, Lorentz covariance and parity invariance settle the general structure of the BSWF for a pseudoscalar meson in momentum space:

$$\Psi_{\alpha\beta}^a(q, p) = [\gamma_5 (f_1 + f_2 \not{p} + f_3 \not{q} + f_4 [\not{q}, \not{p}])]_{\alpha\beta} \mathbf{t}_a. \quad (3.1.138)$$

Here p is the total momentum of the meson and q the relative momentum between the quark and antiquark. The $f_i(q^2, q \cdot p, p^2 = m_\lambda^2)$ are the Lorentz-invariant dressing functions which depend on all invariant momentum variables, and they contain the information about the meson in question.

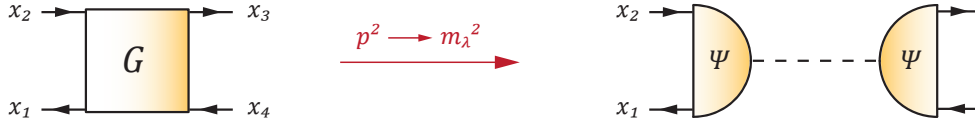


FIG. 3.3: Quark four-point function (3.1.135) and its separability at a particular meson pole according to Eq. (3.1.121). The dashed line is the Feynman propagator.

Current correlators. So how do we compute a hadron mass in practice? It appears that in order to extract the mass of a meson, we need to calculate the four-point function in Eq. (3.1.135) (or any other n -point function that creates meson poles) in some nonperturbative way and look for the poles in this quantity. While this is true in principle, it would also be a rather cumbersome endeavor: four-point functions are complicated objects, and moreover the ones above are not gauge-invariant since they contain quark fields with uncontracted color indices.

The advantage of Eq. (3.1.121) is that it is completely general and applies to *any* correlation function that has a non-vanishing overlap with the state $|\lambda\rangle$, in particular also those with *composite* operators. This is where the currents we defined in Eq. (3.1.23) become useful: Instead of working with the four-point function directly, we can simplify the problem by setting $x_1 = x_2 = x$ and $x_3 = x_4 = y$ and contracting the quark pairs with Dirac and flavor matrices $\mathbf{t}_a \Gamma_{\beta\alpha}$ and $\Gamma'_{\delta\gamma} \mathbf{t}_b$, where $\Gamma \in \{\gamma^\mu, \gamma^\mu \gamma_5, \mathbf{1}, i\gamma_5, \dots\}$. In this way we obtain **current correlators**

$$\begin{aligned} & (\mathbf{t}_a)_{ji} \Gamma_{\beta\alpha} \Gamma'_{\delta\gamma} (\mathbf{t}_b)_{lk} \langle 0 | \mathbb{T} \psi_{\alpha i}(x) \bar{\psi}_{\beta j}(x) \psi_{\gamma k}(y) \bar{\psi}_{\delta l}(y) | 0 \rangle \\ &= \langle 0 | \mathbb{T} \{ \bar{\psi}(x) \Gamma \mathbf{t}_a \psi(x) \} \{ \bar{\psi}(y) \Gamma' \mathbf{t}_b \psi(y) \} | 0 \rangle \\ &= \langle 0 | \mathbb{T} j_a^\Gamma(x) j_b^{\Gamma'}(y) | 0 \rangle, \end{aligned} \quad (3.1.139)$$

which are visualized in the upper panel in Fig. 3.4 and have the form

$$\langle 0 | \mathbb{T} P_a(x) P_b(y) | 0 \rangle, \quad \langle 0 | \mathbb{T} V_a^\mu(x) V_b^\nu(y) | 0 \rangle, \quad \langle 0 | \mathbb{T} A_a^\mu(x) A_b^\nu(y) | 0 \rangle, \quad \text{etc.} \quad (3.1.140)$$

These are again two-point functions and can be viewed as effective meson propagators since they contain the composite fields P_a , V_a^μ , A_a^μ , etc. This is also a convenient way to filter the overwhelming information from the state space, because poles will only emerge from those states that coincide with the quantum numbers of the currents: a PP correlator produces pseudoscalar-meson poles, a VV correlator vector-meson poles and so on. Another advantage is that, in contrast to the four-point function with elementary quark field operators, the current correlators are also gauge-invariant since they contain gauge-invariant, local products of quark fields.

The pole residues of the current correlators are the BSWFs for $x_1 = x_2 = x$, i.e., $z = 0$ (which in momentum space means integration over the relative momentum), and contracted with the Dirac-flavor structures (i.e., taking Dirac and flavor traces):

$$\begin{aligned} -(\mathbf{t}_a)_{ji} \Gamma_{\beta\alpha} \langle 0 | \mathbb{T} \psi_{\alpha i}(x) \bar{\psi}_{\beta j}(x) | \lambda, b \rangle &= -\text{Tr} \left\{ \mathbf{t}_a \Gamma \Psi^b(0, p) \right\} e^{-ip \cdot x} \\ &= \langle 0 | j_a^\Gamma(x) | \lambda, b \rangle = \langle 0 | j_a^\Gamma(0) | \lambda, b \rangle e^{-ip \cdot x}. \end{aligned} \quad (3.1.141)$$

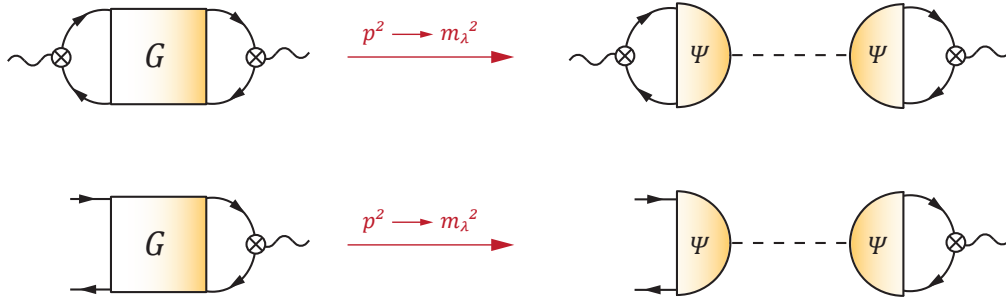


FIG. 3.4: Current correlators from Eq. (3.1.140) and three-point functions from (3.1.146). The symbol \otimes represents the Dirac-flavor matrix $\Gamma \mathbf{t}_a$.

This is the vacuum-to-hadron transition element of the corresponding current. Take for example $\Gamma = \gamma^\mu \gamma_5$ and $i\gamma_5$, which produce axialvector and pseudoscalar currents, respectively. This restricts $|\lambda, a\rangle$ to pseudoscalar and axialvector mesons (for the moment we consider pseudoscalars only):

$$\langle 0 | A_a^\mu(x) | \lambda, b \rangle = \delta_{ab} i p^\mu f_\lambda e^{-ip \cdot x}, \quad \langle 0 | P_a(x) | \lambda, b \rangle = \delta_{ab} r_\lambda e^{-ip \cdot x}. \quad (3.1.142)$$

The first quantity encodes the transition from a pseudoscalar meson to an axialvector current. By translation invariance the dependence on x goes into the phase, and the remainder is a Lorentz vector which can only depend on the onshell momentum p^μ with $p^2 = m_\lambda^2$, so the only possible tensor is p^μ . Since we also take the flavor trace of two generators, the only structure in flavor space is $\sim \delta_{ab}$, cf. (A.1.6). The remaining constant f_λ is the **electroweak decay constant** of the pseudoscalar meson: For example, the pion ($\lambda = \pi$) decays weakly into leptons ($\pi^+ \rightarrow W^+ \rightarrow \mu^+ + \nu_\mu$), so this defines the pion's electroweak decay constant f_π . The analogue r_λ for the pseudoscalar density is not directly measurable but will be useful in the following.

From here we can immediately derive a very useful relation. If we apply the PCAC relation (3.1.39) for equal quark masses, $\partial_\mu A_a^\mu(x) = 2m P_a(x)$, we obtain

$$f_\lambda m_\lambda^2 = 2m r_\lambda, \quad (3.1.143)$$

which is valid for all flavor non-singlet pseudoscalar mesons (in the singlet case there would be an additional term from the anomaly.) For example, it relates the pion decay constant and pion mass with the pseudoscalar transition matrix element r_π . If we go to the chiral limit and set $m = 0$, then the equation tells us that either the pion mass m_π or its decay constant f_π must vanish. This already resembles the Gell-Mann-Oakes-Renner (GMOR) relation, but so far we know nothing about spontaneous chiral symmetry breaking! The essence of the Goldstone theorem, which we will prove in Sec. 4.2, is that the pion decay constant f_π does *not* vanish in the chiral limit as a consequence of spontaneous chiral symmetry breaking, and thus the pion must be massless. Vice versa, the decay constants of excited pions ($m_\lambda \neq 0$) must vanish.

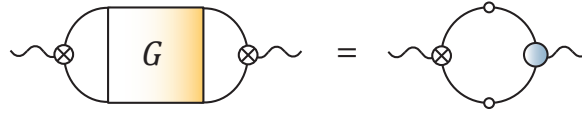


FIG. 3.5: Current correlator in terms of the quark propagator and corresponding vertex.

Since the current correlators are two-point functions, we can derive the spectral representation like in Eq. (2.2.15), e.g. for the pseudoscalar correlator ($z = x - y$),

$$\begin{aligned}
 \langle 0 | \mathbb{T} P_a(x) P_b(y) | 0 \rangle &= \Theta(z^0) \langle 0 | P_a(x) P_b(y) | 0 \rangle + \Theta(-z^0) \langle 0 | P_b(y) P_a(x) | 0 \rangle \\
 &= \sum_{\lambda} \left[\int \frac{d^3 p}{2E_{\mathbf{p}}} \frac{\Theta(z^0) e^{-ipz} + \Theta(-z^0) e^{ipz}}{(2\pi)^3} \right] r_{\lambda}^2 \delta_{ab} \\
 &= \sum_{\lambda} D_F(z, m_{\lambda}) r_{\lambda}^2 \delta_{ab} = \int \frac{d^4 p}{(2\pi)^4} e^{-ipz} \sum_{\lambda} \frac{i r_{\lambda}^2 \delta_{ab}}{p^2 - m_{\lambda}^2 + i\varepsilon},
 \end{aligned} \tag{3.1.144}$$

or also mixed correlators:

$$\begin{aligned}
 \langle 0 | \mathbb{T} A_a^{\mu}(x) P_b(y) | 0 \rangle &= \sum_{\lambda} \left[\int \frac{d^3 p}{2E_{\mathbf{p}}} \frac{\Theta(z^0) e^{-ipz} - \Theta(-z^0) e^{ipz}}{(2\pi)^3} \right] i p^{\mu} f_{\lambda} r_{\lambda} \delta_{ab} \\
 &= -\frac{\partial}{\partial z_{\mu}} \sum_{\lambda} D_F(z, m_{\lambda}) f_{\lambda} r_{\lambda} \delta_{ab} = -\int \frac{d^4 p}{(2\pi)^4} e^{-ipz} \sum_{\lambda} \frac{p^{\mu} f_{\lambda} r_{\lambda} \delta_{ab}}{p^2 - m_{\lambda}^2 + i\varepsilon}.
 \end{aligned} \tag{3.1.145}$$

The sum over λ only goes over states which have an overlap with the pseudoscalar density, i.e., the pseudoscalar mesons, and in principle we should generalize the formulas to spectral densities which include the multiparticle contributions.

Since the result (3.1.121) is general, it also applies to three-point functions such as the ones in Eq. (3.1.76):

$$G_{a,\alpha\beta}^{\Gamma}(x, x_1, x_2) = \langle 0 | \mathbb{T} j_a^{\Gamma}(x) \psi_{\alpha}(x_1) \bar{\psi}_{\beta}(x_2) | 0 \rangle. \tag{3.1.146}$$

This is just the four-point function contracted on one side only, as shown in Fig. 3.4. On the other hand, it is the vertex with quark propagators attached, e.g. in the vector case: $G_V^{\mu}(p_1, p_2) = S(p_1) \Gamma_V^{\mu}(p_1, p_2) S(p_2)$. This means that the vertex must also contain meson poles, but since its non-transverse part is fixed by the WTI these poles can only appear in the transverse part. Hence the quark-photon vertex must have transverse vector-meson poles, which is the origin of ‘**vector-meson dominance**’: when a photon couples to a quark, it fluctuates into ρ, ω, \dots mesons.

With the same reasoning we can write the current correlator as in Fig. 3.5, since it is identical to the quark loop diagram with dressed quark propagators and the corresponding vertex. The VV correlator is also called **hadronic vacuum polarization** because it encodes the QCD contributions to the photon propagator. This quantity is experimentally accessible in the process $e^+e^- \rightarrow$ hadrons, and it is the biggest QCD contribution to the anomalous magnetic moment of the muon where the current Standard Model prediction deviates from experiment by $3 \dots 4\sigma$.

Finally, one can make repeated use of Eq. (3.1.121) also for higher n -point functions. An example is the quantity

$$\langle 0 | \mathbb{T} \psi_\alpha(x_1) \bar{\psi}_\beta(x_2) j_c^\Gamma(x) \psi_\gamma(x_3) \bar{\psi}_\delta(x_4) | 0 \rangle, \quad (3.1.147)$$

where one can insert the completeness relation both to the left and the right of the current operator to produce BSWFs on either side. As a consequence, at the double pole location this becomes

$$\frac{i \Psi_{\alpha\beta}^a(\{q_i\}, p)}{p^2 - m_\lambda^2 + i\varepsilon} \langle \lambda, a | j_c^\Gamma(0) | \lambda', b \rangle \frac{i \bar{\Psi}_{\gamma\delta}^{\dagger b}(\{q'_i\}, p')}{p'^2 - m_{\lambda'}^2 + i\varepsilon}. \quad (3.1.148)$$

The residue $\langle \lambda, a | j_c^\Gamma(0) | \lambda', b \rangle$ defines a hadron's **current matrix element**, such as for example the electromagnetic current matrix element $\langle \pi | V_{\text{em}}^\mu(0) | \pi \rangle$ which describes the coupling of the photon to a pion. The analogous n -point function for baryons contains $\langle N | V_{\text{em}}^\mu(0) | N \rangle$ which describes the electromagnetic coupling to the nucleon. The tensor decompositions of these matrix elements in analogy to Eq. (3.1.137) encode the various measurable **form factors** of hadrons: electromagnetic, axial, pseudoscalar, scalar form factors, etc., and we will return to them in Sec. 4.5.

Lattice QCD. Current correlators are frequently used in **lattice QCD** to compute the hadron spectrum. From the general formula (2.2.24) that relates a correlation function with the path integral, a current correlator can be calculated from

$$G(x - y) = \langle 0 | \mathbb{T} j_1(x) j_2(y) | 0 \rangle = \frac{\int \mathcal{D}\phi e^{iS[\phi]} j_1(x) j_2(y)}{\int \mathcal{D}\phi e^{iS[\phi]}}. \quad (3.1.149)$$

We can write a generic Euclidean correlator in momentum space as

$$G_E(p_E) = \sum_\lambda \frac{R_\lambda}{p_E^2 + m_\lambda^2}, \quad (3.1.150)$$

where $p_E^2 = \mathbf{p}^2 + p_4^2$, $E_\lambda^2 = \mathbf{p}^2 + m_\lambda^2$, and as usual the sum over λ is formal but suppose it contains one or a few isolated bound state poles at $p_E^2 = -m_\lambda^2$. When we take a Fourier transform to $z_4 > 0$, then a pole in momentum space shows up as an exponential decay in Euclidean time:

$$G_E(z_4, \mathbf{p}) = \sum_\lambda R_\lambda \int \frac{dp_4}{2\pi} \frac{e^{ip_4 z_4}}{p_4^2 + E_\lambda^2} = \sum_\lambda R_\lambda \frac{e^{-E_\lambda z_4}}{2E_\lambda}. \quad (3.1.151)$$

At large Euclidean times the mass m_0 of the ground state will dominate the sum, so one can extract m_0 from

$$- \lim_{z_4 \rightarrow \infty} \frac{1}{z_4} \ln G_E(z_4, \mathbf{p} = 0) = m_0. \quad (3.1.152)$$

The discretization of spacetime in lattice QCD and the restriction to a finite volume comes with a number of technical subtleties. For example, in a finite volume the multiparticle continuum turns into a series of discrete poles in p_E^2 (scattering states), which means that the energy levels computed on the lattice are not directly related to the masses of unstable hadrons above open thresholds. The formalism that allows one to relate the energy levels in a finite box to the pole positions of resonances in the complex plane is called the **Luescher method**.



FIG. 3.6: Bethe-Salpeter equation for the four-point function and corresponding homogeneous equation for the Bethe-Salpeter wave function.

Bethe-Salpeter equations. Another way to extract hadron observables from QCD is to start from elementary correlation functions such as the four-point function (3.1.135) and write down a **Bethe-Salpeter equation (BSE)** for it. It has the schematic form shown in Fig. 3.6:

$$G = G_0 + G_0 K G, \quad (3.1.153)$$

which is also called Dyson equation. Each multiplication stands for a four-dimensional integration in momentum space, so this is an integral equation for G , where G_0 is the disconnected part and K the kernel of the equation. The structure of the equation is similar to Eq. (2.2.65), where each step in

$$G = G_0 + G_0 K G = G_0 + G_0 K G_0 + G_0 K G_0 K G = \dots \quad (3.1.154)$$

is exact and gives $G^{-1} = G_0^{-1} - K$, whereas a perturbative series would only converge to that result if K is ‘small’ enough. The kernel can be modelled (e.g. by a ladder approximation which amounts to a gluon exchange between quark and antiquark) but in principle also systematically expanded in terms of the underlying correlation functions such as the quark and gluon propagators, three-point vertices, etc.

If G admits hadronic poles, then at the pole location it factorizes according to Eq. (3.1.121) and one arrives at the homogeneous BSE for the Bethe-Salpeter wave function:

$$\Psi = G_0 K \Psi. \quad (3.1.155)$$

Thus, also here one does not actually need to calculate the four-point function directly in order to extract the pole locations. In practice the homogeneous BSE is an eigenvalue equation because it has the form $(G_0 K) \Psi_\lambda = \eta_\lambda \Psi_\lambda$, where the η_λ are the eigenvalues of $G_0 K$. The masses of ground and excited states can then be read off from the conditions $\eta_\lambda(p^2 = m_\lambda^2) = 1$. If the poles lie above thresholds, then they move into the complex plane onto higher Riemann sheets, but this does not invalidate the equation which still holds at the resonance pole location through analytic continuation, i.e., the above condition has solutions in the complex plane of p^2 . Analogous BSEs can be derived for baryons (the three-body versions are also called **Faddeev equations**) and higher multiquark systems.

3.2 Hadron spectrum

We have studied the flavor structure of the QCD Lagrangian and its group-theoretical implications for hadron properties as well as for currents and for n -point functions. Now it is time for a reality check, because in principle the various symmetries of the Lagrangian should be reflected in the hadron spectrum:

■ **SU(3) color gauge invariance:** Hadrons must be colorless, so they can only appear in the singlet representation of $SU(3)_c$. Color singlets can be obtained by combining quarks and antiquarks to **mesons** or three quarks to **baryons**:

$$\mathbf{3} \otimes \bar{\mathbf{3}} = \mathbf{1} \oplus \mathbf{8}, \quad \mathbf{3} \otimes \mathbf{3} \otimes \mathbf{3} = \mathbf{1} \oplus \mathbf{8} \oplus \mathbf{8} \oplus \mathbf{10}. \quad (3.2.1)$$

Color singlets also arise from combining two (or more) gluons, which leads to the notion of **glueballs**:

$$\mathbf{8} \otimes \mathbf{8} = \mathbf{1} \oplus \mathbf{8} \oplus \mathbf{8} \oplus \mathbf{10} \oplus \bar{\mathbf{10}} \oplus \mathbf{27}. \quad (3.2.2)$$

The product representations of $SU(N)$ are easiest to work out using Young diagrams (see Appendix A.3). Moreover, color singlets also appear in higher patterns of these combinations such as **tetraquarks**,

$$\mathbf{3} \otimes \bar{\mathbf{3}} \otimes \mathbf{3} \otimes \bar{\mathbf{3}} = (\mathbf{1} \oplus \mathbf{8}) \otimes (\mathbf{1} \oplus \mathbf{8}) = \mathbf{1} \oplus \mathbf{1} \oplus \mathbf{8} \oplus \mathbf{8} \oplus \mathbf{8} \oplus \mathbf{8} \oplus \mathbf{10} \oplus \bar{\mathbf{10}} \oplus \mathbf{27}, \quad (3.2.3)$$

pentaquarks ($\mathbf{3} \otimes \mathbf{3} \otimes \mathbf{3} \otimes \mathbf{3} \otimes \bar{\mathbf{3}}$), **hybrid mesons** ($\mathbf{3} \otimes \bar{\mathbf{3}} \otimes \mathbf{8}$), and so on. If we change the number of colors N_c , the nature of a ‘hadron’ will change as well (see Table A.2 in the appendix):

$$N_c \otimes \bar{N}_c = \mathbf{1} \oplus \dots, \quad \underbrace{N_c \otimes \dots \otimes N_c}_{N_c \text{ times}} = \mathbf{1} \oplus \dots, \quad (3.2.4)$$

which means that mesons still survive as $q\bar{q}$ states while baryons become bound states of N_c quarks instead of three.

■ **Flavor symmetries:** The usual $SU(N_f)_V$ flavor symmetry allows us to classify hadrons in flavor multiplets, where in contrast to color all combinations are allowed. In the three-flavor case, mesons form flavor singlets and octets whereas baryons come in singlets, octets and decuplets (see Fig. 3.7):

$$\mathbf{3} \otimes \bar{\mathbf{3}} = \mathbf{1} \oplus \mathbf{8}, \quad \mathbf{3} \otimes \mathbf{3} \otimes \mathbf{3} = \mathbf{1} \oplus \mathbf{8} \oplus \mathbf{8} \oplus \mathbf{10}. \quad (3.2.5)$$

The states within these multiplets are labeled by the third isospin component I_3 and the hypercharge Y (or equivalently, the strangeness), which are conserved quantum numbers even if the $SU(3)$ flavor symmetry is broken. In fact, the observation that hadrons appear in $SU(3)$ octet, decuplet and singlet representations but *not* in the fundamental one was the starting point for the development of the quark model.

The $U(1)_V$ symmetry, on the other hand, corresponds to the **baryon number**. Since there is no quantum number that distinguishes mesons from glueballs, tetraquarks or hybrids, these are strictly speaking all mesons ($B = 0$), whereas pentaquarks are technically baryons ($B = 1$).

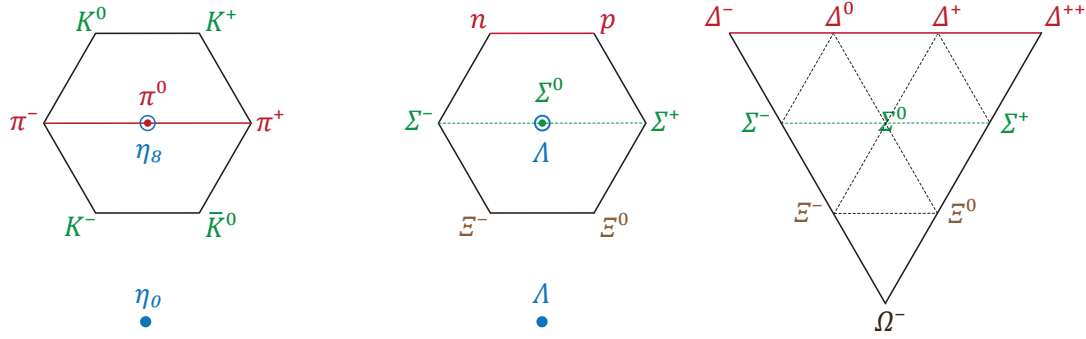


FIG. 3.7: $SU(3)_f$ meson singlet and octet (for 0^{-+} states); baryon singlet, octet and decuplet.

■ **Poincaré invariance:** The invariance of the QCD action under the Poincaré group gives us two quantum numbers to label the states, namely the eigenvalues of its Casimir operators: the **total angular momentum** (‘spin’) J and the **mass** M (see Appendix B). Together with parity invariance of the strong interaction, this allows us to arrange hadrons according to their J^P quantum numbers. We find scalar (0^+), pseudoscalar (0^-), vector (1^-), axialvector (1^+), tensor (2^+) mesons and more, whereas the possible J^P values for baryons are $\frac{1}{2}^\pm$, $\frac{3}{2}^\pm$, $\frac{5}{2}^\pm$, etc.

■ **Charge-conjugation invariance:** Charge conjugation exchanges a particle with its antiparticle and therefore reverses n_q for all flavors, the number of quarks minus antiquarks: $U_c |n_u, n_d, n_s, \dots\rangle = |-n_u, -n_d, -n_s, \dots\rangle$. Since B , I_3 , Y and Q are then reversed as well, only states for which all these additive quantum numbers vanish can be C -parity eigenstates. These are the neutral flavorless mesons, which are their own antiparticles and can be classified according to J^{PC} . Applying U_c twice reverts the state back to its original one ($U_c^2 = 1$), so its possible eigenvalues are $C = \pm 1$. From the transformation properties of the quark fields,

$$U_c \psi_\alpha U_c^{-1} = \eta_c \bar{\psi}_\beta C_{\beta\alpha}, \quad U_c \bar{\psi}_\alpha U_c^{-1} = \eta_c^* C_{\alpha\beta} \psi_\beta \quad C = i\gamma^2 \gamma^0, \quad (3.2.6)$$

together with their anticommutativity, one can show that the Lagrangian is charge-conjugation invariant (η_c is a phase factor). One can also work out the transformation behavior of the currents:

$$S \rightarrow S, \quad P \rightarrow P, \quad V^\mu \rightarrow -V^\mu, \quad A^\mu \rightarrow A^\mu. \quad (3.2.7)$$

Therefore, the mesons that are created by these currents carry the quantum numbers $J^{PC} = 0^{++}$, 0^{-+} , 1^{--} and 1^{+-} , respectively.

Experimentally, hadrons do indeed come in $J^{P(C)}$ multiplets. For three flavors, in each $J^{P(C)}$ channel one finds $SU(3)_f$ octets and singlets for mesons as well as octets, decuplets and singlets for baryons (which can mix, see below). The corresponding states are distinguished by their quantum numbers I_3 and Y . In addition, the multiplets form ground states and **radial excitations**, which are distinguished by the remaining ‘quantum number’ M , i.e., their mass. In the following we discuss the current experimental status on the hadron spectrum.

3.2.1 Mesons

$SU(3)$ multiplets. Let us start with the meson spectrum obtained from three light quark flavors u , d and s . We first discuss the $SU(3)$ multiplets and resulting flavor wave functions. A vector ψ that transforms under the **fundamental representation** of $SU(3)$ satisfies $\psi' = U\psi$. In a given basis $|j\rangle$ with $\langle i|j\rangle = \delta_{ij}$, this implies

$$\psi = \sum_k \psi_k |k\rangle \quad \Rightarrow \quad \psi'_i = \sum_j U_{ij} \psi_j, \quad U|j\rangle = \sum_k U_{kj} |k\rangle, \quad (3.2.8)$$

where $\langle i|U|j\rangle = U_{ij}$ are the matrix elements in that basis, such that

$$\psi' = U\psi = \sum_j \psi_j U|j\rangle = \sum_{j,k} U_{kj} \psi_j |k\rangle = \sum_k \psi'_k |k\rangle. \quad (3.2.9)$$

The same relations hold for the generators, where for later convenience we attach a hat to distinguish the basis-independent operators from their matrix elements:

$$\hat{\mathbf{t}}_a |j\rangle = \sum_k (\mathbf{t}_a)_{kj} |k\rangle, \quad \langle i|\hat{\mathbf{t}}_a|j\rangle = (\mathbf{t}_a)_{ij}. \quad (3.2.10)$$

In the fundamental representation the generators are proportional to the Gell-Mann matrices. The two Cartan generators

$$\mathbf{t}_3 = \frac{1}{2} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \mathbf{Y} = \frac{2}{\sqrt{3}} \mathbf{t}_8 = \frac{1}{3} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix} \quad (3.2.11)$$

correspond to the third isospin component I_3 and the hypercharge Y and label the states inside the multiplet. From Eq. (3.2.10) we can work out their eigenvalues:

$$\begin{aligned} \hat{\mathbf{t}}_3 |u\rangle &= \frac{1}{2}|u\rangle, & \hat{\mathbf{Y}} |u\rangle &= \frac{1}{3}|u\rangle, \\ \hat{\mathbf{t}}_3 |d\rangle &= -\frac{1}{2}|d\rangle, & \hat{\mathbf{Y}} |d\rangle &= \frac{1}{3}|d\rangle, \\ \hat{\mathbf{t}}_3 |s\rangle &= 0, & \hat{\mathbf{Y}} |s\rangle &= -\frac{2}{3}|s\rangle, \end{aligned} \quad (3.2.12)$$

or we can read them off directly from the matrices \mathbf{t}_3 and \mathbf{Y} using a Cartesian basis for $|u\rangle$, $|d\rangle$ and $|s\rangle$. The eigenvalues (I_3, Y) define the **weight vectors**,

$$\left(\frac{1}{2}, \frac{1}{3}\right) \dots u, \quad \left(-\frac{1}{2}, \frac{1}{3}\right) \dots d, \quad \left(0, -\frac{2}{3}\right) \dots s, \quad (3.2.13)$$

from where we can draw the triplet in the (I_3, Y) plane (left panel in Fig. 3.8). The remaining generators

$$\hat{\mathbf{t}}_{\pm} = \hat{\mathbf{t}}_1 \pm i\hat{\mathbf{t}}_2, \quad \hat{\mathbf{u}}_{\pm} = \hat{\mathbf{t}}_6 \pm i\hat{\mathbf{t}}_7, \quad \hat{\mathbf{v}}_{\pm} = \hat{\mathbf{t}}_4 \pm i\hat{\mathbf{t}}_5 \quad (3.2.14)$$

are ladder operators which lead away from the origin in the (I_3, Y) plane and connect these states with each other, cf. Fig. 3.8:

$$\begin{aligned} \hat{\mathbf{t}}_+ |d\rangle &= |u\rangle, & \hat{\mathbf{u}}_+ |s\rangle &= |d\rangle, & \hat{\mathbf{v}}_+ |s\rangle &= |u\rangle, \\ \hat{\mathbf{t}}_- |u\rangle &= |d\rangle, & \hat{\mathbf{u}}_- |d\rangle &= |s\rangle, & \hat{\mathbf{v}}_- |u\rangle &= |s\rangle. \end{aligned} \quad (3.2.15)$$

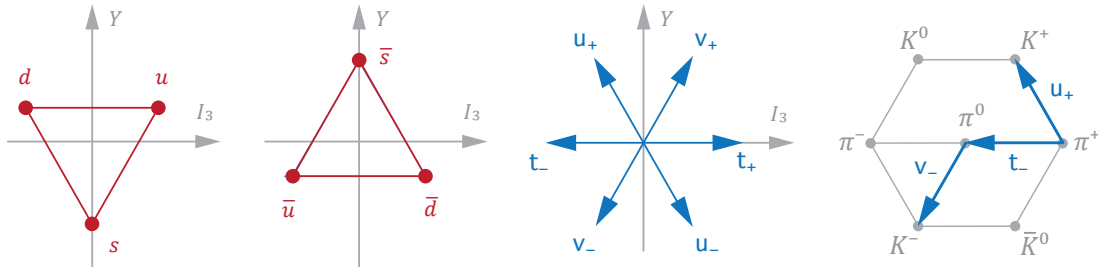


FIG. 3.8: Weight diagrams for the $SU(3)$ fundamental triplet and antitriplet in the (I_3, Y) plane and ladder operators. Right: Construction of the octet using ladder operators.

A vector in the **antitriplet representation** $\bar{\mathbf{3}}$ transforms as

$$\psi'^{\dagger} = \psi^{\dagger} U^{\dagger} \Leftrightarrow \psi'_i{}^* = U_{ij}^* \psi_j^*. \quad (3.2.16)$$

In this case the generators are given by $-\mathbf{t}_a^*$, which satisfy the same commutation relations as the \mathbf{t}_a ,

$$[-\mathbf{t}_a^*, -\mathbf{t}_b^*] = if_{abc} (-\mathbf{t}_c^*), \quad (3.2.17)$$

and thus define another three-dimensional representation (the conjugate representation of the group). For the group $SU(2)$, the generators \mathbf{t}_a and $-\mathbf{t}_a^*$ are related by a unitary transformation and hence equivalent ($SU(2)$ representations are pseudoreal), but this is no longer true for $SU(N)$ with $N > 2$.

Writing the basis as $|\bar{j}\rangle$, we have

$$\hat{\mathbf{t}}_a |\bar{j}\rangle = \sum_k (-\mathbf{t}_a^*)_{kj} |\bar{k}\rangle, \quad \langle \bar{i} | \hat{\mathbf{t}}_a |\bar{j}\rangle = (-\mathbf{t}_a^*)_{ij} \quad (3.2.18)$$

which entails

$$\begin{aligned} \hat{\mathbf{t}}_3 |\bar{u}\rangle &= -\frac{1}{2} |\bar{u}\rangle, & \hat{\mathbf{Y}} |\bar{u}\rangle &= -\frac{1}{3} |\bar{u}\rangle, \\ \hat{\mathbf{t}}_3 |\bar{d}\rangle &= \frac{1}{2} |\bar{d}\rangle, & \hat{\mathbf{Y}} |\bar{d}\rangle &= -\frac{1}{3} |\bar{d}\rangle, \\ \hat{\mathbf{t}}_3 |\bar{s}\rangle &= 0, & \hat{\mathbf{Y}} |\bar{s}\rangle &= \frac{2}{3} |\bar{s}\rangle. \end{aligned} \quad (3.2.19)$$

The weight vectors (I_3, Y) are

$$\left(-\frac{1}{2}, -\frac{1}{3}\right) \dots \bar{u}, \quad \left(\frac{1}{2}, -\frac{1}{3}\right) \dots \bar{d}, \quad \left(0, \frac{2}{3}\right) \dots \bar{s} \quad (3.2.20)$$

and produce the inverted triangle in Fig. 3.8. The ladder operators work as before except the representation matrices of $\hat{\mathbf{t}}_{\pm}$ are $(-\mathbf{t}_1^*) \pm i(-\mathbf{t}_2^*) = -\mathbf{t}_{\mp}^*$ and not $-\mathbf{t}_{\pm}^*$, and similarly for the remaining ones, because due to the complex conjugation these are antilinear operators. As a result,

$$\begin{aligned} \hat{\mathbf{t}}_+ |\bar{u}\rangle &= -|\bar{d}\rangle, & \hat{\mathbf{u}}_+ |\bar{d}\rangle &= -|\bar{s}\rangle, & \hat{\mathbf{v}}_+ |\bar{u}\rangle &= -|\bar{s}\rangle, \\ \hat{\mathbf{t}}_- |\bar{d}\rangle &= -|\bar{u}\rangle, & \hat{\mathbf{u}}_- |\bar{s}\rangle &= -|\bar{d}\rangle, & \hat{\mathbf{v}}_- |\bar{s}\rangle &= -|\bar{u}\rangle. \end{aligned} \quad (3.2.21)$$

Appendix A collects more information on the irreducible representations of $SU(N)$.

0^-	1^-	I	I_3	S		
π^+	ρ^+	1	1	0	$u\bar{d}$	\mathbf{t}_+
π^0	ρ^0	1	0	0	$\frac{1}{\sqrt{2}}(u\bar{u} - d\bar{d})$	$\sqrt{2}\mathbf{t}_3$
π^-	ρ^-	1	-1	0	$d\bar{u}$	\mathbf{t}_-
K^+	K^{*+}	1/2	1/2	1	$u\bar{s}$	\mathbf{v}_+
K^0	K^{*0}	1/2	-1/2	1	$d\bar{s}$	\mathbf{u}_+
\bar{K}^0	\bar{K}^{*0}	1/2	1/2	-1	$s\bar{d}$	\mathbf{u}_-
K^-	K^{*-}	1/2	-1/2	-1	$s\bar{u}$	\mathbf{v}_-
η_8	ω_8	0	0	0	$\frac{1}{\sqrt{6}}(u\bar{u} + d\bar{d} - 2s\bar{s})$	$\sqrt{2}\mathbf{t}_8$
η_0	ω_0	0	0	0	$\frac{1}{\sqrt{3}}(u\bar{u} + d\bar{d} + s\bar{s})$	$\frac{1}{\sqrt{3}}\mathbf{1}$

TABLE 3.1: Normalized $SU(3)_f$ flavor wave functions for mesons.

Flavor wave functions for mesons. Next, we construct the irreducible $\mathbf{1}$ (singlet) and $\mathbf{8}$ (octet) representations along the lines of the discussion in App. A.3: We build the product wave functions as tensors of mixed rank $(1,1)$ that transform under the reducible representation $\mathbf{3} \otimes \bar{\mathbf{3}} = \mathbf{1} \oplus \mathbf{8}$, and by orthogonalizing them we single out the irreducible components in the end.

The simplest construction principle is the one via ladder operators illustrated in Fig. 3.8. If we start from $|\pi^+\rangle = |u\bar{d}\rangle$, then from Eqs. (3.2.12) and (3.2.19) the eigenvalues of $\hat{\mathbf{t}}_3$ and \hat{Y} are

$$\begin{aligned}\hat{\mathbf{t}}_3 |u\bar{d}\rangle &= (\mathbf{t}_3 \otimes \mathbf{1} + \mathbf{1} \otimes (-\mathbf{t}_3^*)) |u\bar{d}\rangle = |u\bar{d}\rangle, \\ \hat{Y} |u\bar{d}\rangle &= (\mathbf{Y} \otimes \mathbf{1} + \mathbf{1} \otimes (-\mathbf{Y}^*)) |u\bar{d}\rangle = 0,\end{aligned}\tag{3.2.22}$$

so the weight vector for $|u\bar{d}\rangle$ is $(I_3, Y) = (1, 0)$. If we apply the ladder operators $\hat{\mathbf{t}}_-$ and $\hat{\mathbf{u}}_+$, we obtain from Eqs. (3.2.15) and (3.2.21):

$$\hat{\mathbf{t}}_- |u\bar{d}\rangle = |d\bar{d}\rangle - |u\bar{u}\rangle \sim |\pi^0\rangle, \quad \hat{\mathbf{u}}_+ |u\bar{d}\rangle = -|u\bar{s}\rangle \sim |K^+\rangle, \quad \text{etc.}\tag{3.2.23}$$

These states are then normalized so that e.g. $\langle \pi^+ | \pi^+ \rangle = 1$. The remaining two states with $I_3 = 0$ and $Y = 0$ are constructed such that $|\eta_0\rangle$ is a singlet and $|\eta_8\rangle$ is orthogonal to $|\pi^0\rangle$ and $|\eta_0\rangle$. The resulting flavor wave functions are collected in Table 3.1.

Note that in a Cartesian basis the flavor wave functions are 3×3 matrices which are proportional to the $SU(N_f)$ generators: the π^+ wave function is $u \otimes \bar{d} = \mathbf{t}_+$, etc. Thus, the flavor wave functions for mesons can already be read off from the generators of the group, which goes back to Eq. (3.1.40): Since the currents and charges define representations of their algebra on the state space, the flavor content of the generators is inherited by the mesons that they create out of the vacuum. This is also the reason why we attached the group generators to the Bethe-Salpeter wave function (3.1.138), which should be read as the combinations that appear in Table 3.1.

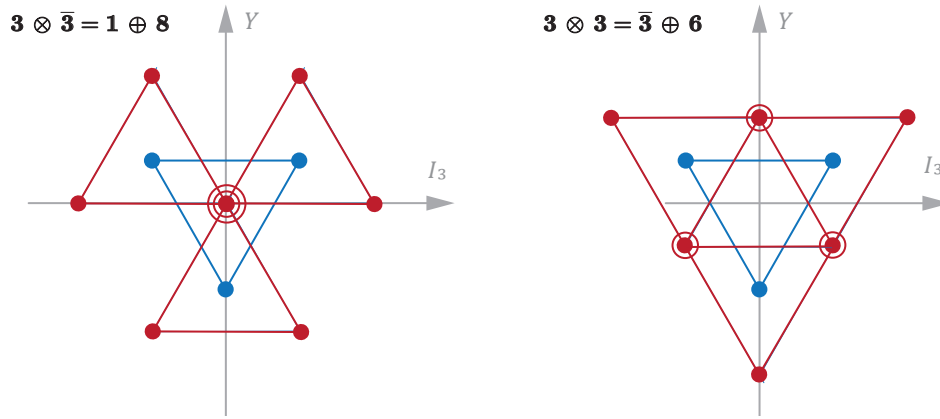


FIG. 3.9: Construction of weight diagrams by superimposing multiplets.

Yet another construction principle is shown in Fig. 3.9. Because the quantum numbers (I_3, Y) are additive, one can simply superimpose multiplets to arrive at the product states. For $\mathbf{3} \otimes \bar{\mathbf{3}} = \mathbf{1} \oplus \mathbf{8}$, draw the triangle defined by $\mathbf{3}$ and add another $\bar{\mathbf{3}}$ at each corner of that triangle. This gives nine states for the corresponding values of (I_3, Y). Likewise, for $\mathbf{3} \otimes \mathbf{3} = \bar{\mathbf{3}} \oplus \mathbf{6}$, add another triplet $\mathbf{3}$ at each corner of $\mathbf{3}$ to arrive at the product states.

Mixing. Unfortunately, the identification of Table 3.1 with physical states only works out in the limit of exact $SU(3)$ flavor symmetry. If the quark masses $m_u = m_d = m_s$ are identical, the Lagrangian is invariant under $SU(3)_V$. As a consequence,

- All states in the multiplet have the same mass;
- All vector currents and charges are conserved;
- Not only the third isospin component I_3 and the hypercharge Y are conserved, but also the Casimirs of $SU(2)$ and $SU(3)$, which are the isospin I and the quantum numbers (p, q) that distinguish the multiplets (see Appendix A.2);
- The states π^0 (with $I = 1$) and η_8, η_0 (with $I = 0$), which have the same I_3 and Y , differ in at least one quantum number I or (p, q) .

If $SU(3)_V$ is broken due to unequal quark masses, the states in the multiplets are no longer mass-degenerate and the $SU(3)$ Casimirs are no longer good quantum numbers. However, I_3 and Y are still conserved and commute with the Hamiltonian, so they can still be used to label the states. As a consequence, mesons carrying the same I_3 and S can mix with each other. This concerns for example the pseudoscalar mesons $\{\pi^0, \eta_8, \eta_0\}$ and the vector mesons $\{\rho^0, \omega_8, \omega_0\}$ which carry $I_3 = S = 0$: their flavor wave functions can mix with each other, and the mixed states are those that appear in the physical spectrum.

In principle the mixing effect can already be seen from the flavor matrix elements of the quark mass matrix \mathbf{M} . Suppose we could write down an effective Hamiltonian of the form

$$H = H_0 + \mathbf{M}, \quad (3.2.24)$$

e.g. in the quark model or derived from some effective Lagrangian, where H_0 is flavor-independent and the quark mass operator is

$$\begin{aligned} \mathbf{M} = & (m_u |u\rangle\langle u| + m_d |d\rangle\langle d| + m_s |s\rangle\langle s|) \otimes \mathbb{1} \\ & + \mathbb{1} \otimes (m_u |\bar{u}\rangle\langle \bar{u}| + m_d |\bar{d}\rangle\langle \bar{d}| + m_s |\bar{s}\rangle\langle \bar{s}|). \end{aligned} \quad (3.2.25)$$

Applied to the flavor wave functions in Table 3.1, we find

$$\begin{aligned} \langle \pi^\pm | \mathbf{M} | \pi^\pm \rangle &= \langle \pi^0 | \mathbf{M} | \pi^0 \rangle = m_u + m_d, \\ \langle \eta_0 | \mathbf{M} | \eta_0 \rangle &= \frac{2}{3} (m_u + m_d + m_s), \\ \langle \eta_8 | \mathbf{M} | \eta_8 \rangle &= \frac{1}{3} (m_u + m_d + 4m_s), \end{aligned} \quad (3.2.26)$$

where the off-diagonal matrix elements are zero except for

$$\begin{aligned} \langle \eta_0 | \mathbf{M} | \eta_8 \rangle &= \frac{\sqrt{2}}{3} (m_u + m_d - 2m_s), \\ \langle \pi_0 | \mathbf{M} | \eta_8 \rangle &= \frac{1}{\sqrt{2}} \langle \pi_0 | \mathbf{M} | \eta_0 \rangle = \frac{1}{\sqrt{3}} (m_u - m_d). \end{aligned} \quad (3.2.27)$$

Because $m_u \approx m_d$, isospin symmetry is still approximately realized and the flavor breaking mostly comes from the strange-quark mass. Hence, the isospin I related to the Casimir of $SU(2)$ is approximately still a good quantum number, which leaves only a mixing for η_0 and η_8 .

If we denote the flavor states generically by ψ_8 and ψ_0 and the mixed ones by ψ and ψ' , we can define a mixing angle:

$$\begin{pmatrix} \psi \\ \psi' \end{pmatrix} = \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} \psi_8 \\ \psi_0 \end{pmatrix} \xrightarrow{\text{ideal}} \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & \sqrt{2} \\ -\sqrt{2} & 1 \end{pmatrix} \begin{pmatrix} \psi_8 \\ \psi_0 \end{pmatrix}. \quad (3.2.28)$$

In the case of **'ideal mixing'** we have $\cos \theta = 1/\sqrt{3}$, which leads to a separation into $SU(2)$ flavor wave functions, i.e., one state made of light quarks and another one made of strange quarks only:

$$\psi = \frac{1}{\sqrt{2}} (u\bar{u} + d\bar{d}), \quad \psi' = s\bar{s}. \quad (3.2.29)$$

These diagonalize the mass matrix,

$$\langle \psi | \mathbf{M} | \psi \rangle = m_u + m_d, \quad \langle \psi' | \mathbf{M} | \psi' \rangle = 2m_s, \quad \langle \psi | \mathbf{M} | \psi' \rangle = 0, \quad (3.2.30)$$

and we find

$$\langle \psi | \mathbf{M} | \psi \rangle + \langle \psi' | \mathbf{M} | \psi' \rangle = \langle \psi_8 | \mathbf{M} | \psi_8 \rangle + \langle \psi_0 | \mathbf{M} | \psi_0 \rangle = m_u + m_d + 2m_s. \quad (3.2.31)$$

The actual mixing angles in the various meson channels are dynamical effects and have to be inferred from experiment (or computed by theory).

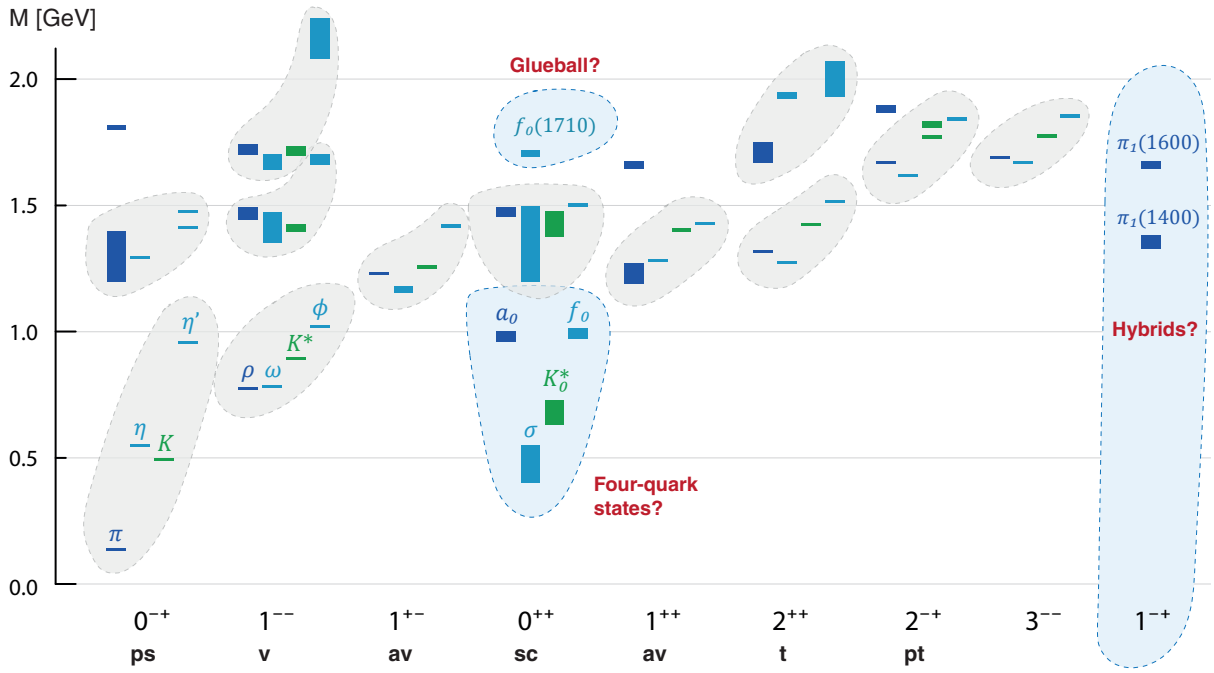


FIG. 3.10: Light and strange meson spectrum from the PDG (<https://pdglive.lbl.gov>).

Experimental spectrum. Now let us compare our expectations with the experimental spectrum. Fig. 3.10 shows the light meson spectrum from the PDG, where the bars are the quoted mass ranges. In each J^{PC} channel there are ground states and radial excitations. Each blob encloses a presumptive ‘nonet’ (i.e., octet plus singlet), where the states in light (dark) blue are those with $I = 0$ ($I = 1$) and the ones in green are the kaons. The naming scheme is as follows:

$$\begin{aligned}
 PC = -+ : & \quad \{\pi, \eta, \eta'\}_J, & PC = ++ : & \quad \{a, f, f'\}_J, \\
 -- : & \quad \{\rho, \omega, \phi\}_J, & +- : & \quad \{b, h, h'\}_J,
 \end{aligned}
 \tag{3.2.32}$$

where the subscript J is dropped for 0^{-+} and $1^{- -}$ states. In addition, kaon-like states with $J^P = 0^+, 1^-, 2^+, 3^-, \dots$ are denoted by K^* .

A good channel to start with are the **vector mesons**, since this sets the prototype regarding expectations. Here we observe

$$m_\rho \approx m_\omega \quad \text{and} \quad m_\phi - m_{K^*} \approx m_{K^*} - m_\omega.
 \tag{3.2.33}$$

Suppose we have isospin symmetry ($m_u = m_d$) and ideal mixing, so that the ω is only made of u/d quarks and the ϕ is a pure $s\bar{s}$ state like in Eq. (3.2.29). If the dynamics were of the form (3.2.24), where the mass differences are entirely due to the different strange and u/d masses, then we would find:

$$\begin{aligned}
 m_\rho = m_\omega &= M_0 + 2m_u, \\
 m_{K^*} &= M_0 + m_u + m_s, \\
 m_\phi &= M_0 + 2m_s.
 \end{aligned}
 \tag{3.2.34}$$

M	I	S	0^{-+}	1^{--}	1^{+-}	0^{++}	1^{++}
8	1	0	π (138) π (1300) π (1800)	ρ (770) ρ (1450) ρ (1700)	b_1 (1235)	a_0 (980) a_0 (1450)	a_1 (1260) a_1 (1640)
8, 1	0	0	η (548) η' (958) η (1295) η (1405) η (1475)	ω (782) ϕ (1020) ω (1420) ω (1650) ϕ (1680) ϕ (2170)	h_1 (1170) h_1 (1415)	f_0 (500) f_0 (980) f_0 (1370) f_0 (1500) f_0 (1710)	f_1 (1285) f_1 (1420)
8	$\frac{1}{2}$	± 1	K (495)	K^* (892) K^* (1410) K^* (1680)	K_1 (1270)	K_0^* (700) K_0^* (1430)	K_1 (1400)
M	I	S	2^{++}	2^{-+}	3^{--}	4^{++}	1^{-+}
8	1	0	a_2 (1320) a_2 (1700)	π_2 (1670) π_2 (1880)	ρ_3 (1690)	a_4 (1970)	π_1 (1400) π_1 (1600)
8, 1	0	0	f_2 (1270) f_2' (1525) f_2 (1950) f_2 (2010) f_2 (2300) f_2 (2340)	η_2 (1645) η_2 (1870)	ω_3 (1670) ϕ_3 (1850)	f_4 (2050)	
8	$\frac{1}{2}$	± 1	K_2^* (1430)	K_2 (1770) K_2 (1820)	K_3^* (1780)	K_4^* (2045)	

TABLE 3.2: Well-established light and strange mesons in terms of J^{PC} , isospin I and strangeness S (PDG 2020, <https://pdglive.lbl.gov>). Mesons with $I = S = 0$ belonging to different multiplets ($M = \mathbf{1}$ or $\mathbf{8}$) can mix with each other, and in principle also the neutral members of the $I = 1$ states, so in these cases an identification with flavor-octet or singlet states is not possible. Note also that C parity is only a good quantum number for neutral mesons.

M_0 is some flavor-independent mass that depends on J^{PC} and the radial quantum number (otherwise they would be the same for each multiplet). Then Eq. (3.2.33) with the values from Table 3.2 yields $m_s - m_u \approx 120$ MeV, and we have the relation

$$m_\omega + m_\phi = 2m_{K^*} \quad (3.2.35)$$

which is realized to good extent in nature. Such empirical mass formulas are called **Gell-Mann-Okubo relations**.

In any case, for vector mesons ideal mixing seems to be well realized since the masses of $\{\rho, \omega\} \rightarrow K^* \rightarrow \phi$ differ roughly by one unit of the strange-quark mass. As one can see in Fig. 3.10, the pattern is (to a lesser extent) still visible in the 1^{+-} , 1^{++} , 2^{++} and some other channels, but there are two channels where the mass ordering does not work at all: the pseudoscalars 0^{-+} and the scalars 0^{++} . Apparently there are further mechanisms at play to which we turn now.

No parity doublets. In the chiral limit $m_u = m_d = m_s = 0$ the Lagrangian is invariant under a $SU(3)_V \times SU(3)_A$ chiral symmetry. In that case all mesons within a given J^{PC} multiplet would become mass-degenerate, and we expect **parity doublets** for mesons with same J but different P . Suppose $|\lambda\rangle$ is an eigenstate of the Hamiltonian with positive parity, so that

$$H|\lambda\rangle = E|\lambda\rangle, \quad U_P|\lambda\rangle = +|\lambda\rangle. \quad (3.2.36)$$

Because the axial charge switches sign under parity,

$$\begin{aligned} U_P Q_A U_P^{-1} &= \int d^3x U_P \psi_\alpha^\dagger(x) U_P^{-1} (\gamma_5)_{\alpha\beta} U_P \psi_\beta(x) U_P^{-1} \\ &= \int d^3x \psi^\dagger(t, -\mathbf{x}) \gamma^0 \gamma_5 \gamma^0 \psi(t, -\mathbf{x}) \\ &= - \int d^3x \psi^\dagger(x) \gamma_5 \psi(x) = -Q_A, \end{aligned} \quad (3.2.37)$$

the state $|\lambda'\rangle = Q_A |\lambda\rangle$ carries negative parity:

$$U_P |\lambda'\rangle = U_P Q_A |\lambda\rangle = -Q_A U_P |\lambda\rangle = -|\lambda'\rangle. \quad (3.2.38)$$

If chiral symmetry holds, the axial charge commutes with the Hamiltonian, $[Q_A, H] = 0$, and therefore $|\lambda'\rangle$ is an eigenstate with the same mass:

$$H |\lambda'\rangle = H Q_A |\lambda\rangle = Q_A H |\lambda\rangle = E |\lambda'\rangle. \quad (3.2.39)$$

Thus, chiral symmetry entails that the masses of the pseudoscalars (0^{-+}) are degenerate with the scalars (0^{++}), vector mesons (1^{--}) with axialvectors (1^{+-}), and so on.

Now, the three-flavor chiral symmetry is explicitly broken because $m_s \gg m_u \approx m_d$, but the two-flavor $SU(2)_V \times SU(2)_A$ symmetry should still approximately work since u and d quarks are almost massless. Hence we should still see remnants of this pattern in the spectrum. We do not, though: the pion is almost massless in contrast to its scalar partner and the vector mesons are much lighter than the axialvector ones.

On the other hand, the fact that $SU(2)_V$ still works out well (mesons with same isospin I have about the same mass) tells us that something is wrong with the $SU(N_f)_A$ part. Combined with the unnaturally light pseudoscalar mesons, these are the typical symptoms of a **spontaneous symmetry breaking** of $SU(N_f)_A$, which would produce $N_f^2 - 1$ massless Goldstone bosons in the chiral limit. The three pions are indeed almost massless ($m_\pi \approx 140$ MeV); the kaons are heavier but they also contain one strange quark, so they should feel the impact of explicit chiral symmetry breaking more strongly than the pions.

In Sec. 4.2 we will derive the Gell-Mann-Oakes-Renner relation, which states that the *squared* pseudoscalar meson masses are proportional to the quark masses. Based on this, we would interpret matrix elements such as $\langle \pi | M | \pi \rangle = m_u + m_d$ to be proportional to m_π^2 instead of m_π (this becomes explicit in chiral perturbation theory). With Eq. (3.2.31), the analogue of the Gell-Mann-Okubo relation (3.2.35) then becomes

$$m_\eta^2 + m_{\eta'}^2 = m_{\eta_0}^2 + m_{\eta_8}^2 = 2m_K^2. \quad (3.2.40)$$

$\eta - \eta'$ mixing. If chiral symmetry is indeed spontaneously broken, it should break all axial symmetries, $SU(3)_A$ and $U(1)_A$. Hence we would expect nine Goldstone bosons, including the pions, the kaons and both the η_8 and η_0 . Suppose we had ideal mixing: then η and η' would be the analogues of ω and ϕ in the vector channel,

$$\eta = \frac{1}{\sqrt{2}}(u\bar{u} + d\bar{d}), \quad \eta' = s\bar{s}, \quad (3.2.41)$$

where the η is mass-degenerate with the pion and the η' (as a pure $s\bar{s}$ state) would acquire mass due to the explicit chiral symmetry breaking, similarly to the kaons. Assuming that Eq. (3.2.40) holds, then with $m_\eta = m_\pi$ we expect to find $m_{\eta'} \sim 690$ MeV. Going away from ideal mixing, the masses of η and η' should move in opposite directions: if $m_\eta > m_\pi$, we should find $m_{\eta'} \lesssim 690$ MeV. However, this is not realized at all: the η is heavier than the kaon ($m_\eta \approx 550$ MeV) and the η' mass is almost twice as large ($m_{\eta'} \approx 960$ MeV).

This argument and other ones lead us to believe that the $U(1)_A$ symmetry may not have been realized to begin with. It is still satisfied by the QCD Lagrangian in the chiral limit, but the classical symmetry is *anomalously* broken at the quantum level because it does not survive the quantization — this is the $U(1)_A$ **anomaly**. We already anticipated in Eq. (3.1.54) that the divergence of the axial current picks up an extra term. If we work out Eq. (3.1.143) in the isosinglet case, we obtain

$$f_{\eta_0} m_{\eta_0}^2 = 2 \frac{m_u + m_d + m_s}{3} r_{\eta_0} + \frac{g^2 N_f}{(4\pi)^2} \langle 0 | \tilde{F}_a^{\mu\nu}(0) F_{\mu\nu}^a(0) | \eta_0 \rangle. \quad (3.2.42)$$

Even if we set all quark masses to zero, the right-hand side of this equation remains nonzero and therefore also the η_0 remains massive in the chiral limit. Through a mixing with η_8 , the extra term contributes to both η and η' masses. Another manifestation of this is the **Witten-Veneziano relation**, where χ_T is the so-called topological susceptibility:

$$m_\eta^2 + m_{\eta'}^2 = 2m_K^2 + \frac{4N_f \chi_T}{f_\pi^2}. \quad (3.2.43)$$

Missing exotics. Another observation in Fig. (3.10) is that not all J^{PC} quantum numbers appear: The ‘**exotic**’ quantum numbers 0^{--} , 0^{+-} , 1^{-+} and 2^{+-} are absent from the light-meson spectrum, with the exception of the higher-lying $\pi_1(1400)$ and $\pi_1(1600)$ in the 1^{-+} channel.

The absence of exotic mesons can be understood from the nonrelativistic quark model. So far we have labeled $q\bar{q}$ states according to their J^{PC} eigenvalues. Now assume that the total spin S of the $q\bar{q}$ pair ($S = 0$ or $S = 1$) and its intrinsic orbital angular momentum $L = 0, 1, 2, \dots$ are also good quantum numbers. Then from the angular-momentum addition rules we have $|L - S| \leq J \leq L + S$, and we can motivate the following two relations:

$$P = (-1)^{L+1} \quad \text{and} \quad C = (-1)^{L+S}. \quad (3.2.44)$$

The first arises from the observation that a $q\bar{q}$ pair has intrinsic parity -1 and its spatial wave function has parity $(-1)^L$; parity is multiplicative, hence the factor $(-1)^{L+1}$. Charge conjugation exchanges quark and antiquark, so the value of C can be deduced by exchanging $q \leftrightarrow \bar{q}$ and then interchanging their positions and spins. The symmetry of the spin states is $(-1)^{S+1}$ because $S = 0$ is antisymmetric ($|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle$) and $S = 1$ symmetric ($|\uparrow\uparrow\rangle, |\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle, |\downarrow\downarrow\rangle$). The factor $(-1)^L$ is as before, and a minus sign comes from interchanging the fermions. The combined operation gives $C = -(-1)^L (-1)^{S+1} = (-1)^{L+S}$.

The first relation above says that states with alternating L have alternating parity, and the second one entails that once L and S are specified, C (and thus J^{PC}) is fixed as well. These rules are quite efficient for cataloguing the possible J^{PC} combinations:

- $L = 0$ are orbital ground states (s waves) and should therefore correspond to the lightest mesons. According to (3.2.44) they must have negative parity. $S = 0$ gives us the pseudoscalars 0^{-+} ; from $S = 1$ we obtain the vector mesons 1^{--} .
- $L = 1$ are orbital excitations (p waves) with positive parity. From $S = 0$ we obtain the axialvectors 1^{+-} and from $S = 1$ we get scalar (0^{++}), axialvector (1^{++}) and tensor mesons (2^{++}).
- From $L = 2$ (d waves) we obtain further vectors (1^{--}) plus states with $J = 2$ and $J = 3$, and so on for higher L .

The resulting mass ordering is shown in Fig. 3.11. Pseudoscalars and vectors are the lightest mesons because they are in an orbital s wave. Since they carry different quark spin S , their mass splitting is generated by spin-spin interactions between quark and antiquark. This is called ‘**hyperfine splitting**’ because of its analogy to the hydrogen atom, where the hyperfine structure is caused by the coupling between electron and proton spin. All other mesons are orbitally excited because they carry higher L . For the lowest-lying states we should thus expect a mass pattern

$$\{0^{-+}\} < \{1^{--}\} < \{1^{+-}\} \lesssim \{0^{++}, 1^{++}, 2^{++}\} \lesssim \dots \quad (3.2.45)$$

which is also how we arranged the columns in Fig. 3.10. Also frequently used is the **spectroscopic notation** $^{2S+1}L_J$ with $L = S, P, D, F, \dots$, where this becomes

$$\{^1S_0\} < \{^3S_1\} < \{^1P_1\} \lesssim \{^3P_0, ^3P_1, ^3P_2\} \lesssim \dots \quad (3.2.46)$$

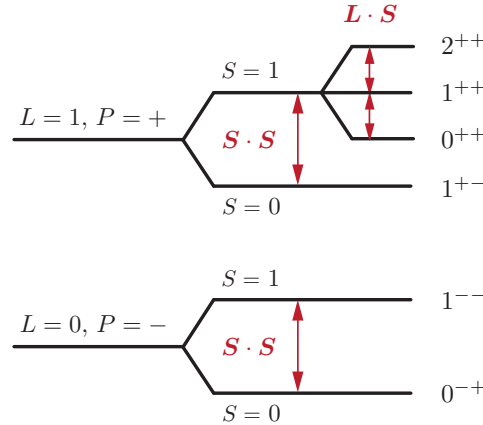


FIG. 3.11: Expected J^{PC} level ordering for mesons.

Interestingly, the analysis forbids exotic quantum numbers 0^{--} , 0^{+-} , 1^{-+} and 2^{+-} . If such states were observed, we would then conclude that they are not made of $q\bar{q}$ but something else. The only known examples in the light-meson spectrum are the $\pi_1(1400)$ and $\pi_1(1600)$ with 1^{-+} . They are candidates for **hybrid mesons**, i.e., states with valence quarks and gluons, because the combination $q\bar{q}g$ produces the quantum numbers 1^{-+} naturally (among others). From the general formula (3.1.121) we could look for hybrid mesons in higher n -point functions with gluonic content or, equivalently, current correlators of the form

$$\langle 0 | T j(x) j(y) | 0 \rangle \quad \text{e.g. with} \quad j(x) = \bar{\psi}(x) \not{D} \psi(x), \quad (3.2.47)$$

where $\bar{\psi} \not{D} \psi$ is the simplest gauge-invariant combination that involves gluon fields. Such calculations are being done in lattice QCD and they find indeed additional states which do not show up when using $\bar{\psi}\psi$ operators only.²

On the other hand, Eq. (3.1.121) states that a meson pole will appear in *any* correlation function that has non-zero overlap with the state. It turns out that Bethe-Salpeter wave functions of the form (3.1.136) do not vanish for exotic quantum numbers even though they only contain quark and antiquark operators. As a consequence, poles with exotic quantum numbers can also show up in the quark-antiquark four-point function in Fig. 3.3. This goes back to the observation that the relations (3.2.44) are nonrelativistic, because P and C are conserved quantum numbers whereas L and S are not. Only J corresponds to a Casimir operator of the Poincaré group; S and L are not Poincaré-invariant and can mix in different reference frames. This is also why only J^{PC} should be used to label multiplets. For example, from the nonrelativistic analysis above the pion should carry $L=0$, but the pion's relativistic BSWF in Eq. (3.1.138) also contains $L=1$ components, namely the tensors \not{q} and $[\not{q}, \not{p}]$ which depend on the relative momentum q and thus correspond to p waves in the pion's rest frame. (This is analogous to the 'lower components' in Dirac spinors which come about through relativity.) Similarly, at the level of BSWFs, exotic mesons are not generally forbidden as $q\bar{q}$ states but merely do not survive the non-relativistic limit.

²J. J. Dudek, Phys.Rev.D 84 (2011) 074023, [arXiv:1106.5515](https://arxiv.org/abs/1106.5515) [hep-ph].

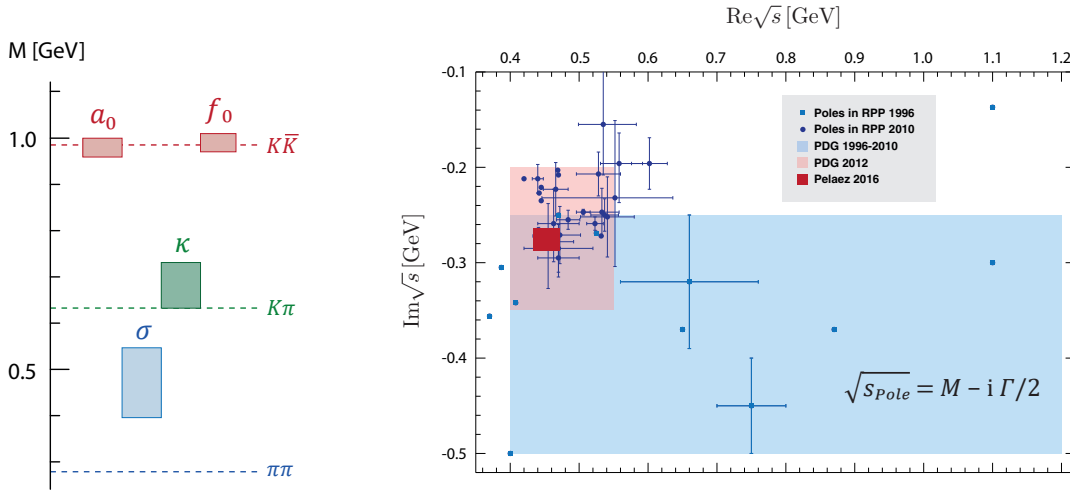


FIG. 3.12: Left: Mass ordering of the light scalar mesons. Right: σ pole location in the complex \sqrt{s} plane (adapted from J. Pelaez, Phys. Rept. 658, 1 (2016)).

Scalar mesons. Another curious case in Fig. 3.10 is the lowest-lying multiplet of scalar 0^{++} mesons. They do not fit into the mass ordering (3.2.45), which would be (roughly) realized if we simply *removed* them from the spectrum. Also the mass ordering inside the multiplet is far from ‘ideal’: the isosinglet $f_0(500)$ or σ meson is the lightest state, followed by the $K_0^*(700)$ or κ and the almost degenerate $a_0(980)$ and $f_0(980)$. This does not make much sense given the flavor content: why would a_0 and f_0 be mass-degenerate if one is made of light quarks and the other is the $s\bar{s}$ state?

Such arguments initiated the idea that the 0^{++} ground states may not be actual $q\bar{q}$ states but rather **tetraquarks** in the form of diquark-antidiquark combinations.³ Two quarks can form a diquark through $\mathbf{3} \otimes \mathbf{3} = \bar{\mathbf{3}} \oplus \mathbf{6}$, cf. Fig. 3.9, where it turns out that the color-antitriplet channel $\bar{\mathbf{3}}$ is attractive but the sextet channel $\mathbf{6}$ is repulsive. Hence one can write the combination (3.2.3) also differently:

$$(\mathbf{3} \otimes \mathbf{3}) \otimes (\bar{\mathbf{3}} \otimes \bar{\mathbf{3}}) = (\bar{\mathbf{3}} \oplus \mathbf{6}) \otimes (\mathbf{3} \oplus \bar{\mathbf{6}}) = (\bar{\mathbf{3}} \otimes \mathbf{3}) \oplus \dots = \mathbf{1} \oplus \mathbf{8} \oplus \dots \quad (3.2.48)$$

In flavor space, the antitriplet $\bar{\mathbf{3}}$ corresponds to antisymmetric flavor wave functions $[ud] = ud - du$, $[us]$ and $[ds]$ (up to normalization). Combining a diquark with an antidiquark then produces a singlet and an octet, except with different flavor content: The isoscalar σ is made of light quarks only, whereas both a_0 and f_0 contain $s\bar{s}$ which would make them mass-degenerate. Since the σ and κ lie above the $\pi\pi$ and $K\pi$ thresholds, respectively, they can then simply fall apart without the need for exchanging gluons which would turn them into broad resonances. In fact, the large decay width of the σ has prohibited a precise determination of its pole location until recently (see Fig. 3.12). There has been a long history of support for the non- $q\bar{q}$ nature of the light scalar mesons, although their internal decomposition (diquark-antidiquark, meson-meson, or possible $q\bar{q}$ admixture) is still under debate.⁴

³R.L. Jaffe, Phys. Rev. D 15, 267 (1977).

⁴For a review, see J. Pelaez, Phys. Rept. 658, 1 (2016), [arXiv:1510.00653](https://arxiv.org/abs/1510.00653) [hep-ph].

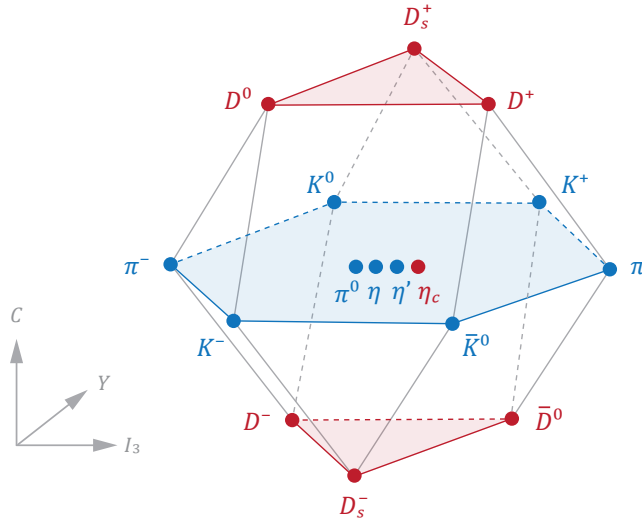


FIG. 3.13: $SU(4)_f$ multiplet arrangement for the pseudoscalar mesons.

Mesons with charm. Let us turn to the spectrum of heavy mesons. In order to include charm quarks, we should start from the group $SU(4)_f$ which has 15 generators. The $SU(4)_f$ symmetry of the Lagrangian is badly broken by the large charm-quark mass, but like in the three-flavor case discussed in Sec. 3.1.1 the diagonal currents corresponding to the three Cartan generators are still conserved. They are related to the quantum numbers I_3 , Y and C (charm) which label the states (Fig. 3.13). On top of the light and strange sector, this leads to additional D and D_s mesons:

$$I = \frac{1}{2}, S = 0, C = \pm 1: \quad \{D^+, D^0, \bar{D}^0, D^-\} = \{c\bar{d}, c\bar{u}, u\bar{c}, d\bar{c}\}, \quad (3.2.49)$$

$$I = 0, S = C = \pm 1: \quad \{D_s^+, D_s^-\} = \{c\bar{s}, s\bar{c}\}. \quad (3.2.50)$$

The separation into different multiplets ($4 \otimes \bar{4} = \mathbf{1} \oplus \mathbf{15}$) is not useful because due to the broken symmetry the states in the center will mix; ideal mixing amounts to the usual separation into $\frac{1}{\sqrt{2}}(u\bar{u} \pm d\bar{d})$, $s\bar{s}$ and $c\bar{c}$.

Let us focus on the **charmonium** spectrum consisting of $c\bar{c}$ (Fig. 3.14). The first charmonium state to be discovered was the J/ψ , which owes its double name to the simultaneous discovery by two independent collaborations in November 1974 ('November revolution'). The J/ψ and its excitations are vector particles with $J^{PC} = 1^{--}$, so they can be directly produced from a photon in e^+e^- collisions. The naming scheme for the remaining J^{PC} channels is as follows:

$$PC = -+ : \eta_c, \quad -- : \psi, \quad ++ : \chi_c, \quad +- : h_c. \quad (3.2.51)$$

Since charm quarks are heavy ($m_c \gg \Lambda_{\text{QCD}}$), relativity and chiral symmetry no longer play a major role, which are the two main effects that complicate the light meson spectrum. Thus, effective theories such as **heavy-quark effective theory (HQET)** and **nonrelativistic QCD (NRQCD)** can be used to study heavy-quark physics. In fact, already non-relativistic quark potential models provide an efficient description of the charmonium spectrum.

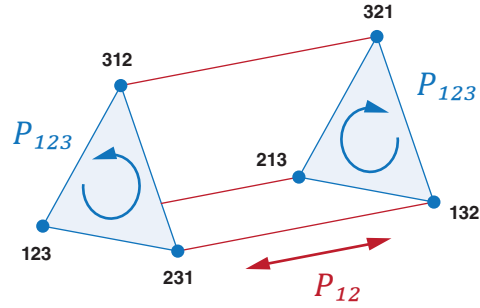


FIG. 3.15: Cayley graph for the permutation group S_3 . Any permutation can be reconstructed from a transposition P_{12} and a cyclic permutation P_{123} .

3.2.2 Baryons

Let us come to the baryon sector. Baryons are fermions, so their angular momentum takes half-integer values: $J^P = 1/2^\pm, 3/2^\pm, 5/2^\pm$, and so on. If we start again with three flavors u , d and s , then because of

$$\mathbf{3} \otimes \mathbf{3} \otimes \mathbf{3} = \mathbf{1} \oplus \mathbf{8} \oplus \mathbf{8} \oplus \mathbf{10} \quad (3.2.52)$$

in principle each J^P channel can contain $SU(3)$ flavor octets, decuplets and singlets. These are shown in Fig. 3.7, and they can come in the form of ground states and radial excitations. As before, $SU(3)$ flavor breaking entails that baryons with the same I_3 and strangeness S can mix.

The construction of the flavor wave functions is a bit different from the case of mesons since we must combine three quarks instead of a quark and an antiquark. What helps is that baryons satisfy the Pauli principle, i.e., in the flavor-symmetric limit their total (Bethe-Salpeter) wave function

$$\Psi = \text{Dirac} \times \text{Flavor} \times \text{Color} \quad (3.2.53)$$

must be totally antisymmetric under exchange of any two quarks. Here, ‘Dirac’ is a shorthand for the full spatial and spin (or momentum and spin) contribution that transforms under the Lorentz group, or the rotation group in the non-relativistic case. We can then arrange each part in irreducible representations of the permutation group S_3 , with definite symmetry, and figure out in the end which symmetry states are allowed in the combination.

For example concerning the color part, the singlet in $\mathbf{3} \otimes \mathbf{3} \otimes \mathbf{3} = \mathbf{1} \oplus \dots$ is totally antisymmetric (see further below). If we use the quark color basis states $C_1 = R$, $C_2 = G$ and $C_3 = B$, then the color wave function is given by ε_{ijk} :

$$RGB + GBR + BRG - GRB - BGR - RBG = \varepsilon_{ijk} C_i C_j C_k. \quad (3.2.54)$$

For the flavor part we must also cast the remaining combinations in Eq. (3.2.52) in permutation-group multiplets, i.e., we must classify them into *simultaneous* irreducible representations of $SU(3)$ and the permutation group S_3 .

Permutation group S_3 . The permutation group S_3 consists of $3! = 6$ elements. The group manifold can be visualized by the **Cayley graph** in Fig. 3.15: any permutation of an object ψ_{123} can be reconstructed from a transposition P_{12} and a cyclic permutation P_{123} . The former interchanges the indices $1 \leftrightarrow 2$ and the latter is a cyclic permutation $1 \rightarrow 2, 2 \rightarrow 3, 3 \rightarrow 1$. The group elements acting on ψ_{123} are given by

$$\begin{aligned} 1, & & P_{12}, \\ P_{13} P_{12} = P_{123}, & & P_{23} = P_{12} P_{123}, \\ P_{23} P_{12} = P_{123}^2, & & P_{13} = P_{12} P_{123}^2, \end{aligned} \quad (3.2.55)$$

for example

$$P_{13} \psi_{123} = P_{12} P_{123}^2 \psi_{123} = P_{12} P_{123} \psi_{231} = P_{12} \psi_{312} = \psi_{321}, \quad (3.2.56)$$

and they are represented by paths along the Cayley graph.

To find the **irreducible representations** of S_3 , we define the combinations

$$\psi_1^\pm = \frac{\psi_{123} \pm \psi_{213}}{2}, \quad \psi_2^\pm = \frac{\psi_{231} \pm \psi_{132}}{2}, \quad \psi_3^\pm = \frac{\psi_{312} \pm \psi_{321}}{2}. \quad (3.2.57)$$

You can convince yourself that applying P_{12} and P_{123} to them amounts to

$$P_{12} \psi_i^\pm = \pm \psi_i^\pm, \quad P_{123} \psi_i^\pm = \frac{\psi_j^+ + \psi_j^- \pm (\psi_k^+ - \psi_k^-)}{2}, \quad (3.2.58)$$

where $\{i, j, k\}$ is a cyclic permutation of $\{1, 2, 3\}$. If we further define

$$\mathcal{S} = \psi_1^+ + \psi_2^+ + \psi_3^+, \quad \mathcal{A} = \psi_1^- + \psi_2^- + \psi_3^-, \quad (3.2.59)$$

then we see that

$$P_{12} \mathcal{S} = \mathcal{S}, \quad P_{123} \mathcal{S} = \mathcal{S}, \quad P_{12} \mathcal{A} = -\mathcal{A}, \quad P_{123} \mathcal{A} = \mathcal{A}. \quad (3.2.60)$$

Since any permutation can be reconstructed from P_{12} and P_{123} , the combinations \mathcal{S} and \mathcal{A} only transform into themselves, so they form irreducible one-dimensional subspaces under the permutation group. \mathcal{S} is invariant under permutations, so it is a symmetric **singlet**. The antisymmetric singlet (**'antisinglet'**) \mathcal{A} is totally antisymmetric under exchange of any two indices and thus picks up a minus sign under a transposition. The remaining four combinations can be grouped into **doublets**,

$$\mathcal{D}_1 = \begin{bmatrix} \psi_2^- - \psi_3^- \\ \frac{1}{\sqrt{3}} (\psi_2^+ + \psi_3^+ - 2\psi_1^+) \end{bmatrix}, \quad \mathcal{D}_2 = \begin{bmatrix} -\frac{1}{\sqrt{3}} (\psi_2^- + \psi_3^- - 2\psi_1^-) \\ \psi_2^+ - \psi_3^+ \end{bmatrix}, \quad (3.2.61)$$

which also transform into themselves and therefore define a two-dimensional subspace:

$$P_{12} \mathcal{D}_j = \mathbf{M}_{12}^\top \mathcal{D}_j, \quad P_{123} \mathcal{D}_j = \mathbf{M}_{123}^\top \mathcal{D}_j. \quad (3.2.62)$$

The representation matrices (\mathbf{M}^\top denotes the matrix transpose) are given by

$$\mathbf{M}_{12} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \mathbf{M}_{123} = \frac{1}{2} \begin{pmatrix} -1 & -\sqrt{3} \\ \sqrt{3} & -1 \end{pmatrix}, \quad (3.2.63)$$

from where all other ones can be reconstructed through Eq. (3.2.55), e.g.:

$$\begin{aligned} P_{23} \mathcal{D}_j &= P_{12} P_{123} \mathcal{D}_j = P_{12} \left(M_{123}^\top \mathcal{D}_j \right) = M_{123}^\top P_{12} \mathcal{D}_j \\ &= M_{123}^\top M_{12}^\top \mathcal{D}_j = (M_{12} M_{123})^\top \mathcal{D}_j. \end{aligned} \quad (3.2.64)$$

The upper (lower) components of the doublets are antisymmetric (symmetric) under transpositions P_{12} and we denote them by

$$\mathcal{D}_j = \begin{bmatrix} a_j \\ s_j \end{bmatrix}. \quad (3.2.65)$$

In the language of Young diagrams (see Appendix A.3), the irreducible representations of S_3 correspond to

$$\begin{array}{ccc} \mathcal{S} & \mathcal{D}_j & \mathcal{A} \\ \begin{array}{|c|c|c|} \hline \square & \square & \square \\ \hline \end{array} & \begin{array}{|c|c|} \hline \square & \square \\ \hline \square & \\ \hline \end{array} & \begin{array}{|c|} \hline \square \\ \hline \square \\ \hline \square \\ \hline \end{array} \end{array}$$

In practice we will also need the tensor products of S_3 multiplets. Given two sets of singlets $\mathcal{S}, \mathcal{S}'$, antisinglets $\mathcal{A}, \mathcal{A}'$ and doublets $\mathcal{D} = [a, s]$, $\mathcal{D}' = [a', s']$, there are 16 possible combinations

$$\{\mathcal{S}, \mathcal{A}, a, s\} \times \{\mathcal{S}', \mathcal{A}', a', s'\} \quad (3.2.66)$$

which we can again arrange into multiplets. Clearly, the products of two singlets ($\mathcal{S}\mathcal{S}'$) or antisinglets ($\mathcal{A}\mathcal{A}'$) must be singlets. The inner product $\mathcal{D} \cdot \mathcal{D}'$ of two doublets is also a singlet and invariant under any permutation because the representation matrices $M \in \{M_{12}, M_{123}\}$ are orthogonal:

$$(M^\top \mathcal{D}) \cdot (M^\top \mathcal{D}') = \mathcal{D}_k (M M^\top)_{kl} \mathcal{D}'_l = \mathcal{D} \cdot \mathcal{D}'. \quad (3.2.67)$$

Therefore, there are three possibilities for constructing singlets in the product space:

$$\mathcal{S}\mathcal{S}', \quad \mathcal{A}\mathcal{A}', \quad \mathcal{D} \cdot \mathcal{D}' = aa' + ss'. \quad (3.2.68)$$

Antisinglets are obtained from

$$\mathcal{S}\mathcal{A}', \quad \mathcal{A}\mathcal{S}', \quad \mathcal{D} \wedge \mathcal{D}' := as' - sa', \quad (3.2.69)$$

where we defined an antisymmetric wedge product, and doublets are formed by

$$\begin{array}{ll} \mathcal{S}\mathcal{D}', & \mathcal{A}(\varepsilon \mathcal{D}'), \\ \mathcal{S}'\mathcal{D}, & \mathcal{A}'(\varepsilon \mathcal{D}), \end{array} \quad \mathcal{D} * \mathcal{D}' := \begin{bmatrix} as' + sa' \\ aa' - ss' \end{bmatrix}, \quad (3.2.70)$$

where

$$\varepsilon = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \Rightarrow \varepsilon \mathcal{D} = \begin{bmatrix} s \\ -a \end{bmatrix}. \quad (3.2.71)$$

This covers all 16 possibilities. You can easily check this using Eqs. (3.2.60) and (3.2.62) written down for a, s, a' and s' : the singlets stay invariant under permutations, the antisinglets pick up a minus sign for odd permutations, and the doublets transform under M_{12} and M_{123} .

	uuu	uud	duu	ddd	uus	uds	dds	ssu	ssd	sss
\mathcal{S}	Δ^{++}	Δ^+	Δ^0	Δ^-	Σ^+	Σ^0	Σ^-	Ξ^0	Ξ^-	Ω^-
\mathcal{D}_1		p	n		Σ^+	Σ^0	Σ^-	Ξ^0	Ξ^-	
\mathcal{D}_2						Λ^0				
\mathcal{A}						Λ^0				

TABLE 3.3: $SU(3)_f$ flavor wave functions for baryons.

Flavor wave functions for baryons. Suppose u , d and s denote flavor vectors that transform under the fundamental representation of $SU(3)$, for example in a Cartesian basis. Combining three of them gives $3 \times 3 \times 3 = 27$ possible combinations, which would transform under the 27-dimensional reducible representation of $SU(3)$. The irreducible representations contained in $\mathbf{3} \otimes \mathbf{3} \otimes \mathbf{3} = \mathbf{1} \oplus \mathbf{8} \oplus \mathbf{8} \oplus \mathbf{10}$ differ by their symmetry, so we must find the combined irreducible representations of $SU(3)$ and S_3 .

To construct flavor wave functions from the tensor products of three flavor vectors, e.g. for a baryon with flavor content uud such as the proton or the Δ^+ , we write

$$\begin{aligned}
\psi_{123} &= u_i u_j d_k = (uud)_{ijk}, & \psi_{213} &= u_j u_i d_k = (uud)_{ijk}, \\
\psi_{231} &= u_j u_k d_i = (duu)_{ijk}, & \psi_{132} &= u_i u_k d_j = (udu)_{ijk}, \\
\psi_{312} &= u_k u_i d_j = (udu)_{ijk}, & \psi_{321} &= u_k u_j d_i = (duu)_{ijk}.
\end{aligned} \tag{3.2.72}$$

The combinations in Eq. (3.2.57) become

$$\psi_1^+ = uud, \quad \psi_1^- = 0, \quad \psi_2^\pm = \pm \psi_3^\pm = \pm \frac{ud \pm du}{2} u, \tag{3.2.73}$$

so we arrive at the multiplets

$$\begin{aligned}
\mathcal{S}(uud) &= uud + udu + duu, & \mathcal{D}_1(uud) &= \left[\begin{array}{c} duu - udu \\ \frac{1}{\sqrt{3}}(udu + duu - 2uud) \end{array} \right], \\
\mathcal{A}(uud) &= 0, & \mathcal{D}_2(uud) &= 0.
\end{aligned} \tag{3.2.74}$$

Apart from overall normalization, $\mathcal{S}(uud)$ is the flavor wave function of the Δ^+ and $\mathcal{D}_1(uud)$ is that of the proton. Had we started from ddu instead of uud , we would have obtained the wave functions for the Δ^0 and the neutron (replace $u \leftrightarrow d$ in the equation above). The combination uuu returns only a singlet (Δ^{++}), and from uds we get everything: \mathcal{S} , \mathcal{A} and two doublets.

If we take all 10 combinations with different flavor content into account (uuu , ddd , sss , uud , uus , ddu , dds , ssu , ssd , uds), the permutation group gives us

- 10 singlets, which form the flavor decuplet with Δ , Σ , Ξ and Ω ,
- 8 doublets which form the flavor octet, including proton, neutron, Σ , Ξ and Λ ,
- and one antisinglet from uds , the flavor singlet for Λ .

These are just the irreducible representations of $SU(3)_f$: decuplet, octet and singlet. The resulting states are collected in Table 3.3 and written explicitly in Table 3.4.

		I	I_3	S	
uud	p	$1/2$	$1/2$	0	$\frac{1}{\sqrt{2}} \left[\begin{array}{c} udu - duu \\ -\frac{1}{\sqrt{3}} (udu + duu - 2uud) \end{array} \right]$
udd	n	$1/2$	$-1/2$	0	$\frac{1}{\sqrt{2}} \left[\begin{array}{c} udd - dud \\ \frac{1}{\sqrt{3}} (dud + udd - 2ddu) \end{array} \right]$
uus	Σ^+	1	1	-1	$\frac{1}{\sqrt{2}} \left[\begin{array}{c} usu - suu \\ -\frac{1}{\sqrt{3}} (usu + suu - 2uus) \end{array} \right]$
uds	Σ^0	1	0	-1	$\frac{1}{2} \left[\begin{array}{c} sud - usd + sdu - dsu \\ \frac{1}{\sqrt{3}} (sud + usd + sdu + dsu - 2uds - 2dus) \end{array} \right]$
dds	Σ^-	1	-1	-1	$\frac{1}{\sqrt{2}} \left[\begin{array}{c} dsd - sdd \\ -\frac{1}{\sqrt{3}} (dsd + sdd - 2dds) \end{array} \right]$
uds	Λ^0	0	0	-1	$\frac{1}{2} \left[\begin{array}{c} \frac{1}{\sqrt{3}} (2uds - 2dus + usd - dsu + sdu - sud) \\ usd - dsu + sud - sdu \end{array} \right]$
uss	Ξ^0	$1/2$	$1/2$	-2	$\frac{1}{\sqrt{2}} \left[\begin{array}{c} uss - sus \\ \frac{1}{\sqrt{3}} (sus + uss - 2ssu) \end{array} \right]$
dss	Ξ^-	$1/2$	$-1/2$	-2	$\frac{1}{\sqrt{2}} \left[\begin{array}{c} dss - sds \\ \frac{1}{\sqrt{3}} (sds + dss - 2ssd) \end{array} \right]$
uuu	Δ^{++}	$3/2$	$3/2$	0	uuu
uud	Δ^+	$3/2$	$1/2$	0	$\frac{1}{\sqrt{3}} (uud + udu + duu)$
udd	Δ^0	$3/2$	$-1/2$	0	$\frac{1}{\sqrt{3}} (udd + dud + ddu)$
ddd	Δ^-	$3/2$	$3/2$	0	ddd
uus	Σ^+	1	1	-1	$\frac{1}{\sqrt{3}} (uus + usu + suu)$
uds	Σ^0	1	0	-1	$\frac{1}{\sqrt{6}} (uds + sud + dsu + dus + usd + sdu)$
dds	Σ^-	1	-1	-1	$\frac{1}{\sqrt{3}} (dds + dsd + sdd)$
uss	Ξ^0	$1/2$	$1/2$	-2	$\frac{1}{\sqrt{3}} (uss + sus + ssu)$
dss	Ξ^-	$1/2$	$-1/2$	-2	$\frac{1}{\sqrt{3}} (dss + sds + ssd)$
sss	Ω^-	0	0	-3	sss
uds	Λ^0	0	0	-1	$\frac{1}{\sqrt{6}} (uds + sud + dsu - dus - usd - sdu)$

TABLE 3.4: $SU(3)_f$ flavor wave functions for octet, decuplet and singlet baryons.

Including charm as a fourth flavor, we can immediately extend the construction to $SU(4)_f$ which would give us 20 singlets, 20 doublets and 4 antisinglets:

$$\mathbf{4} \otimes \mathbf{4} \otimes \mathbf{4} = \mathbf{20}_S \oplus \mathbf{20}_{M_A} \oplus \mathbf{20}_{M_S} \oplus \mathbf{4}_A. \quad (3.2.75)$$

In the $SU(2)_f$ case, on the other hand, we get four singlets (the four Δ baryons) and two doublets (proton and neutron):

$$\mathbf{2} \otimes \mathbf{2} \otimes \mathbf{2} = \mathbf{4}_S \oplus \mathbf{2}_{M_A} \oplus \mathbf{2}_{M_S}. \quad (3.2.76)$$

If we identified $SU(2)$ with spin instead of flavor, this would give us the three-spinor wave functions, e.g. the four symmetric ones:

$$|\uparrow\uparrow\uparrow\rangle, \quad \frac{1}{\sqrt{3}}(|\uparrow\uparrow\downarrow\rangle + |\uparrow\downarrow\uparrow\rangle + |\downarrow\uparrow\uparrow\rangle), \quad \frac{1}{\sqrt{3}}(|\downarrow\downarrow\uparrow\rangle + |\downarrow\uparrow\downarrow\rangle + |\uparrow\downarrow\downarrow\rangle), \quad |\downarrow\downarrow\downarrow\rangle. \quad (3.2.77)$$

Full baryon wave function. The remaining question is what the Dirac part in Eq. (3.2.53) looks like. From the above analysis we conclude that even without knowing its explicit form, we can also arrange it into permutation group multiplets \mathcal{S} , \mathcal{A} , \mathcal{D}_1 and \mathcal{D}_2 to write

$$\Psi = \{\mathcal{S}, \mathcal{A}, \mathcal{D}_1, \mathcal{D}_2\}_D \times \{\mathcal{S}, \mathcal{A}, \mathcal{D}_1, \mathcal{D}_2\}_F \times \mathcal{A}_C \stackrel{!}{=} \mathcal{A}_{\text{total}}. \quad (3.2.78)$$

Because color is antisymmetric, the Dirac-flavor part must be symmetric. This leaves the three possible combinations in Eq. (3.2.68):

$$\mathcal{A}_{\text{total}} = \begin{cases} (\mathcal{D}_D \cdot \mathcal{D}_F) \mathcal{A}_C & \text{(octet),} \\ (\mathcal{S}_D \mathcal{S}_F) \mathcal{A}_C & \text{(decuplet),} \\ (\mathcal{A}_D \mathcal{A}_F) \mathcal{A}_C & \text{(singlet).} \end{cases} \quad (3.2.79)$$

That is, flavor octet baryons come with a mixed-symmetric Dirac part, decuplet baryons with a symmetric and flavor-singlet baryons with an antisymmetric Dirac part.

In principle the Dirac part can be constructed from the Bethe-Salpeter wave function

$$\langle 0 | \mathbb{T} \psi_\alpha(x_1) \psi_\beta(x_2) \psi_\gamma(x_3) | \lambda \rangle \quad (3.2.80)$$

in analogy to Eqs. (3.1.136–3.1.137): In momentum space it has the structure

$$\Psi_{\alpha\beta\gamma}(p_1, p_2, p_3) = \sum_{i=1}^N f_i(q_1^2, q_2^2, q_1 \cdot p, q_2 \cdot p, q_1 \cdot q_2, p^2 = m_\lambda^2) \tau_i(p_1, p_2, p_3)_{\alpha\beta\gamma}, \quad (3.2.81)$$

where p is the onshell momentum of the baryon, q_1 and q_2 are the two relative momenta in the system, and the τ_i form a linearly independent and complete tensor basis. These can be grouped into S_3 multiplets, so that the symmetry properties are inherited by the dressing functions f_i .

It is instructive to go back to the **nonrelativistic quark model**, like we did in the discussion of mesons below Eq. (3.2.44). In that case J^P , L and S are good quantum numbers. The Dirac parts are taken to be the direct product of $O(3)$ spatial and $SU(2)$ spin wave functions. The combination of three spins $\frac{1}{2} \otimes \frac{1}{2} \otimes \frac{1}{2}$ in a three-quark qqq state only permits total spin $S = \frac{1}{2}$ or $S = \frac{3}{2}$. The corresponding wave functions are those in Eq. (3.2.77), i.e., there are four permutation-group singlets \mathcal{S}_S (subscript S for spin) with $S = 3/2$ and two doublets \mathcal{D}_S with spin $S = 1/2$.

The $SU(2)$ spin states ($\mathcal{D}_S, \mathcal{S}_S$) are then combined with the $SU(3)$ flavor states ($\mathcal{D}_F, \mathcal{S}_F, \mathcal{A}_F$) into spin-flavor multiplets according to Eqs. (3.2.68–3.2.70):

$$\begin{aligned}
 \blacksquare \text{ 56-plet } \mathcal{S}_{SF}: & \quad \mathcal{D}_S \cdot \mathcal{D}_F && \rightarrow 2 \times 8 = 16 \text{ spin-flavor states,} \\
 & \quad \mathcal{S}_S \mathcal{S}_F && \rightarrow 4 \times 10 = 40, \\
 \blacksquare \text{ 70-plet } \mathcal{D}_{SF}: & \quad \mathcal{D}_S * \mathcal{D}_F && \rightarrow 2 \times 8 = 16, \\
 & \quad \mathcal{D}_S \mathcal{S}_F && \rightarrow 2 \times 10 = 20, \\
 & \quad (\varepsilon \mathcal{D}_S) \mathcal{A}_F && \rightarrow 2 \times 1 = 2, \\
 & \quad \mathcal{S}_S \mathcal{D}_F && \rightarrow 4 \times 8 = 32, \\
 \blacksquare \text{ 20-plet } \mathcal{A}_{SF}: & \quad \mathcal{D}_S \wedge \mathcal{D}_F && \rightarrow 2 \times 8 = 16, \\
 & \quad \mathcal{S}_S \mathcal{A}_F && \rightarrow 4 \times 1 = 4.
 \end{aligned}$$

This is the ‘ $SU(6)$ -symmetric’ quark model classification, since $SU(2)_{\text{spin}} \times SU(3)_{\text{flavor}} \sim SU(6)_{\text{spin-flavor}}$ and in $SU(6)$ one has

$$6 \otimes 6 \otimes 6 = 56_S \oplus 70_{M_A} \oplus 70_{M_S} \oplus 20_A. \quad (3.2.82)$$

Keep in mind, however, that different spin polarizations do not correspond to different particles: only the number of states in a flavor multiplet counts the number of baryons we expect to find in the spectrum.

The **spatial wave functions** can depend on the total coordinate \mathbf{R} and the relative (Jacobi) coordinates $\boldsymbol{\rho}$ and $\boldsymbol{\lambda}$:

$$\mathbf{R} = \frac{\mathbf{x}_1 + \mathbf{x}_2 + \mathbf{x}_3}{\sqrt{3}}, \quad \boldsymbol{\rho} = \frac{\mathbf{x}_1 - \mathbf{x}_2}{\sqrt{2}}, \quad \boldsymbol{\lambda} = \frac{\mathbf{x}_1 + \mathbf{x}_2 - 2\mathbf{x}_3}{\sqrt{6}}. \quad (3.2.83)$$

After removing the center-of-mass motion induced by \mathbf{R} , the spatial wave functions $\phi(\boldsymbol{\rho}, \boldsymbol{\lambda})$ only depend on the relative coordinates. These can be arranged in a permutation-group doublet, since from Eq. (3.2.62) one can verify

$$P_{12} \begin{bmatrix} \boldsymbol{\rho} \\ \boldsymbol{\lambda} \end{bmatrix} = M_{12}^T \begin{bmatrix} \boldsymbol{\rho} \\ \boldsymbol{\lambda} \end{bmatrix}, \quad P_{123} \begin{bmatrix} \boldsymbol{\rho} \\ \boldsymbol{\lambda} \end{bmatrix} = M_{123}^T \begin{bmatrix} \boldsymbol{\rho} \\ \boldsymbol{\lambda} \end{bmatrix}. \quad (3.2.84)$$

From a doublet \mathcal{D} one can construct further multiplets such as the $O(3)$ invariants

$$\mathcal{S} = \mathcal{D} \cdot \mathcal{D} = \boldsymbol{\rho}^2 + \boldsymbol{\lambda}^2, \quad \mathcal{D}' = \mathcal{D} * \mathcal{D} = \begin{bmatrix} 2\boldsymbol{\rho} \cdot \boldsymbol{\lambda} \\ \boldsymbol{\rho}^2 - \boldsymbol{\lambda}^2 \end{bmatrix}, \quad (3.2.85)$$

and in this way also the spatial wave function $\phi(\boldsymbol{\rho}, \boldsymbol{\lambda})$ can produce permutation-group singlets, doublets and antisingslets.

The spatial wave functions are usually set up in a spherical harmonic oscillator basis

$$\phi_L(\boldsymbol{\rho}, \boldsymbol{\lambda}) = \sum_{n_\rho, l_\rho, n_\lambda, l_\lambda} c_{n_\rho l_\rho n_\lambda l_\lambda}^L [\phi_{n_\rho l_\rho}(\boldsymbol{\rho}) \otimes \phi_{n_\lambda l_\lambda}(\boldsymbol{\lambda})]_L, \quad (3.2.86)$$

where both internal motions support radial ($n_\alpha > 0$) and orbital ($l_\alpha > 0$) excitations. With $n = n_\rho + n_\lambda$ and $l = l_\rho + l_\lambda$, the total orbital angular momentum is constructed from $L = |l_\rho - l_\lambda| \dots l_\rho + l_\lambda$ and the parity of the state is $P = (-1)^l$. This yields excitations bands for the ‘band quantum number’ $N = 2n + l$ corresponding to the same energy. The resulting spatial wave functions can be arranged into permutation-group multiplets \mathcal{S}_O (subscript O for orbital), \mathcal{D}_O and \mathcal{A}_O , which are finally combined with the spin-flavor wave functions to yield the totally symmetric combinations

$$\begin{aligned} \mathcal{S}_O \mathcal{S}_{SF} &\rightarrow \mathcal{S}_O \mathcal{D}_S \cdot \mathcal{D}_F, \mathcal{S}_O \mathcal{S}_S \mathcal{S}_F, \\ \mathcal{D}_O \cdot \mathcal{D}_{SF} &\rightarrow \mathcal{D}_O \cdot (\mathcal{D}_S * \mathcal{D}_F), \mathcal{D}_O \cdot (\mathcal{D}_S \mathcal{S}_F), \mathcal{D}_O \cdot (\varepsilon \mathcal{D}_S) \mathcal{A}_F, \mathcal{D}_O \cdot (\mathcal{S}_S \mathcal{D}_F), \\ \mathcal{A}_O \mathcal{A}_{SF} &\rightarrow \mathcal{A}_O (\mathcal{D}_S \wedge \mathcal{D}_F), \mathcal{A}_O \mathcal{S}_S \mathcal{A}_F. \end{aligned}$$

Since the spatial wave functions carry definite L and P and the spin wave functions definite S , their combination $J = |L - S| \dots L + S$ determines J^P . The resulting states and their flavor assignments are listed in Table 3.5.

One can see that the **ground states** ($N = 0$) correspond to flavor octet baryons with $J = \frac{1}{2}^+$ and decuplet baryons with $\frac{3}{2}^+$. We could have inferred this directly from Eq. (3.2.79): For ground states the orbital wave functions are spatially symmetric, i.e., permutation-group singlets, so the different Dirac multiplets can only come from the spin. Ground states have $L = 0$ and thus $J = S$, so the only possible combinations are

$$\mathcal{A}_{\text{total}} \sim \begin{cases} (\mathcal{D}_S \cdot \mathcal{D}_F) \mathcal{A}_C & (J = \frac{1}{2}^+, \text{ octet}), \\ (\mathcal{S}_S \mathcal{S}_F) \mathcal{A}_C & (J = \frac{3}{2}^+, \text{ decuplet}). \end{cases} \quad (3.2.87)$$

Because there is no antisymmetric spin wave function \mathcal{A}_S , the flavor-singlet baryons Λ^0 cannot appear as ground states.

What Table 3.5 also shows is that the quark model predicts *a lot* of states. While the ‘bands’ for $N = 0$ and $N = 1$ can be identified with experimentally known baryons, already the $N = 2$ and especially the $N = 3$ states have not all been observed. This is the so-called **missing resonances problem**, which could have several explanations:

- We simply have not found them yet. Excited baryons (generically called N^*) have traditionally been extracted from $N\pi$ scattering ($N\pi \rightarrow N\pi$), but if they did not strongly couple to $N\pi$ it would be hard to see their peaks in experimental cross sections (remember Eq. (3.1.121)). Recent photoproduction experiments (e.g. $\gamma N \rightarrow N\pi$) have indeed found new states, but the spectrum as of today is still quite sparse compared to what the quark model predicts.
- If two quarks inside a baryon clustered to a **diquark**, this would freeze internal excitation degrees of freedom and we should see fewer states in the spectrum.
- The assumptions we made (nonrelativistic quark model, harmonic oscillator) are simply too drastic to provide a realistic description of light baryons.

N	L^P	O	SF	S	F	J^P			
0	0^+	\mathcal{S}_0	56	$\frac{1}{2}$	8	$\frac{1}{2}^+$			
				$\frac{3}{2}$	10	$\frac{3}{2}^+$			
	1	1^-	\mathcal{D}_0	70	$\frac{1}{2}$	8	$\frac{1}{2}^-, \frac{3}{2}^-$		
					$\frac{1}{2}$	10	$\frac{1}{2}^-, \frac{3}{2}^-$		
					$\frac{1}{2}$	1	$\frac{1}{2}^-, \frac{3}{2}^-$		
					$\frac{3}{2}$	8	$\frac{1}{2}^-, \frac{3}{2}^-, \frac{5}{2}^-$		
					$\frac{1}{2}$	8	$\frac{1}{2}^-, \frac{3}{2}^-, \frac{5}{2}^-$		
					$\frac{3}{2}$	10	$\frac{1}{2}^-, \frac{3}{2}^-, \frac{5}{2}^-$		
	2	0^+	\mathcal{S}_0	56	$\frac{1}{2}$	8	$\frac{1}{2}^+$		
					$\frac{3}{2}$	10	$\frac{3}{2}^+$		
					$\frac{5}{2}$	10	$\frac{5}{2}^+$		
			\mathcal{D}_0	70	$\frac{1}{2}$	8	$\frac{1}{2}^+$		
$\frac{1}{2}$					10	$\frac{1}{2}^+$			
$\frac{1}{2}$					1	$\frac{1}{2}^+$			
1^+		\mathcal{A}_0	20	$\frac{1}{2}$	8	$\frac{1}{2}^+, \frac{3}{2}^+$			
				$\frac{3}{2}$	1	$\frac{1}{2}^+, \frac{3}{2}^+, \frac{5}{2}^+$			
				$\frac{3}{2}$	8	$\frac{3}{2}^+, \frac{5}{2}^+$			
				2^+	\mathcal{S}_0	56	$\frac{1}{2}$	8	$\frac{3}{2}^+, \frac{5}{2}^+$
							$\frac{3}{2}$	10	$\frac{1}{2}^+, \frac{3}{2}^+, \frac{5}{2}^+, \frac{7}{2}^+$
							$\frac{5}{2}$	10	$\frac{3}{2}^+, \frac{5}{2}^+$
\mathcal{D}_0	70	$\frac{1}{2}$	8	$\frac{3}{2}^+, \frac{5}{2}^+$					
		$\frac{1}{2}$	10	$\frac{3}{2}^+, \frac{5}{2}^+$					
		$\frac{1}{2}$	1	$\frac{3}{2}^+, \frac{5}{2}^+$					
\mathcal{A}_0	20	$\frac{3}{2}$	8	$\frac{1}{2}^+, \frac{3}{2}^+, \frac{5}{2}^+, \frac{7}{2}^+$					
		$\frac{3}{2}$	1	$\frac{1}{2}^+, \frac{3}{2}^+, \frac{5}{2}^+, \frac{7}{2}^+$					
		$\frac{5}{2}$	8	$\frac{1}{2}^+, \frac{3}{2}^+, \frac{5}{2}^+, \frac{7}{2}^+$					
3	1^-	\mathcal{S}_0	56	$\frac{1}{2}$	8	$\frac{1}{2}^-, \frac{3}{2}^-$			
				$\frac{3}{2}$	10	$\frac{1}{2}^-, \frac{3}{2}^-, \frac{5}{2}^-$			
				$\frac{5}{2}$	10	$\frac{1}{2}^-, \frac{3}{2}^-, \frac{5}{2}^-$			
		\mathcal{D}_0	70	$\frac{1}{2}$	8	$\frac{1}{2}^-, \frac{3}{2}^-$			
				$\frac{1}{2}$	10	$\frac{1}{2}^-, \frac{3}{2}^-$			
				$\frac{1}{2}$	1	$\frac{1}{2}^-, \frac{3}{2}^-$			
				$\frac{3}{2}$	8	$\frac{1}{2}^-, \frac{3}{2}^-, \frac{5}{2}^-$			
				$\frac{3}{2}$	8	$\frac{1}{2}^-, \frac{3}{2}^-, \frac{5}{2}^-$			
				$\frac{5}{2}$	10	$\frac{1}{2}^-, \frac{3}{2}^-, \frac{5}{2}^-$			
	\mathcal{A}_0	20	$\frac{1}{2}$	8	$\frac{1}{2}^-, \frac{3}{2}^-$				
			$\frac{3}{2}$	1	$\frac{1}{2}^-, \frac{3}{2}^-, \frac{5}{2}^-$				
			$\frac{5}{2}$	8	$\frac{1}{2}^-, \frac{3}{2}^-, \frac{5}{2}^-$				
	2^-	\mathcal{D}_0	70	$\frac{1}{2}$	8	$\frac{3}{2}^-, \frac{5}{2}^-$			
				$\frac{1}{2}$	10	$\frac{3}{2}^-, \frac{5}{2}^-$			
				$\frac{1}{2}$	1	$\frac{3}{2}^-, \frac{5}{2}^-$			
				$\frac{3}{2}$	8	$\frac{1}{2}^-, \frac{3}{2}^-, \frac{5}{2}^-, \frac{7}{2}^-$			
				$\frac{3}{2}$	8	$\frac{1}{2}^-, \frac{3}{2}^-, \frac{5}{2}^-, \frac{7}{2}^-$			
				$\frac{5}{2}$	10	$\frac{1}{2}^-, \frac{3}{2}^-, \frac{5}{2}^-, \frac{7}{2}^-$			
\mathcal{A}_0		20	$\frac{1}{2}$	8	$\frac{3}{2}^-, \frac{5}{2}^-$				
			$\frac{3}{2}$	1	$\frac{1}{2}^-, \frac{3}{2}^-, \frac{5}{2}^-, \frac{7}{2}^-$				
			$\frac{5}{2}$	8	$\frac{1}{2}^-, \frac{3}{2}^-, \frac{5}{2}^-, \frac{7}{2}^-$				
3^-	\mathcal{S}_0	56	$\frac{1}{2}$	8	$\frac{5}{2}^-, \frac{7}{2}^-$				
			$\frac{3}{2}$	10	$\frac{3}{2}^-, \frac{5}{2}^-, \frac{7}{2}^-, \frac{9}{2}^-$				
			$\frac{5}{2}$	10	$\frac{3}{2}^-, \frac{5}{2}^-, \frac{7}{2}^-, \frac{9}{2}^-$				
			\mathcal{D}_0	70	$\frac{1}{2}$	8	$\frac{5}{2}^-, \frac{7}{2}^-$		
					$\frac{1}{2}$	10	$\frac{5}{2}^-, \frac{7}{2}^-$		
					$\frac{1}{2}$	1	$\frac{5}{2}^-, \frac{7}{2}^-$		
	\mathcal{A}_0	20	$\frac{3}{2}$	8	$\frac{3}{2}^-, \frac{5}{2}^-, \frac{7}{2}^-, \frac{9}{2}^-$				
			$\frac{3}{2}$	1	$\frac{3}{2}^-, \frac{5}{2}^-, \frac{7}{2}^-, \frac{9}{2}^-$				
			$\frac{5}{2}$	8	$\frac{3}{2}^-, \frac{5}{2}^-, \frac{7}{2}^-, \frac{9}{2}^-$				

TABLE 3.5: Quark-model classification of light and strange baryons up to $N \leq 3$.

It is also amusing to think about a world without color. The Δ^{++} carries three up quarks (uuu) and has all spins aligned ($\uparrow\uparrow\uparrow$), which does not yield a totally antisymmetric wave function — which was historically one of the motivations for introducing the color degree of freedom. If we wanted to respect the Pauli principle *without* color, then Eq. (3.2.69) provides us with the following options:

$$\mathcal{A}_{\text{total}} = \begin{cases} \mathcal{D}_D \wedge \mathcal{D}_F & (\text{octet}), \\ \mathcal{A}_D \mathcal{S}_F & (\text{decuplet}), \\ \mathcal{S}_D \mathcal{A}_F & (\text{singlet}). \end{cases} \quad (3.2.88)$$

With the $SU(2)$ spin wave functions \mathcal{D}_S and \mathcal{S}_S (but no \mathcal{A}_S) we could still construct nucleons but not Δ baryons, at least not as ground states.

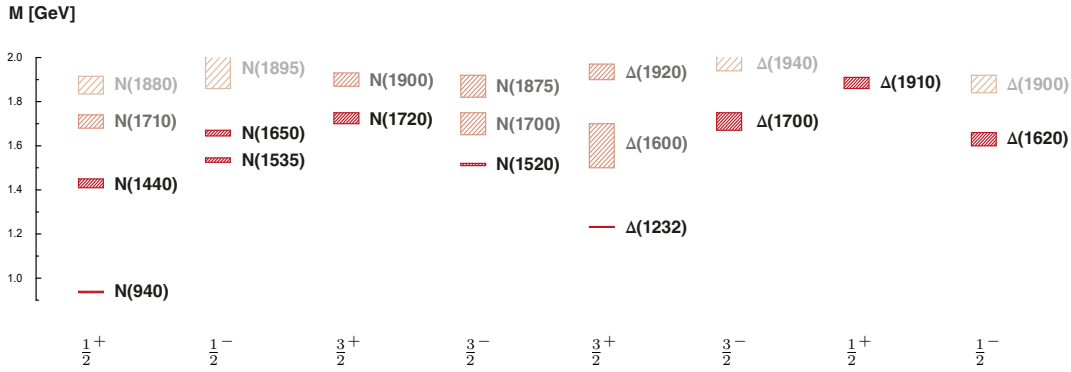


FIG. 3.16: Light baryon spectrum for $J^P = \frac{1}{2}^\pm$ and $\frac{3}{2}^\pm$ from the PDG.

Light baryons. Let us have a look at the experimental spectrum of light and strange baryons (Table 3.6). In contrast to mesons, the naming scheme is the same for different J^P channels, i.e., all states with $I = \frac{1}{2}$ and $S = 0$ are called nucleons, all states with $I = \frac{3}{2}$ and $S = 0$ are called Δ baryons, etc. From the point of view of the Poincaré group, each J^P channel contains one ‘ground state’ plus radial excitations. Due to $SU(3)_V$ breaking, multiplets with the same I_3 and S can mix. This affects the baryons containing strange quarks (the **hyperons**): the Λ states (uds) can be mixtures of **8** and **1** and the Σ and Ξ states can be mixtures of **8** and **10**.

The well established states are the ground states that are also predicted by the quark model: the octet baryons with $J^P = \frac{1}{2}^+$ and the decuplet baryons with $J^P = \frac{3}{2}^+$. Their lightest members are the nucleon (proton and neutron) and the **$\Delta(1232)$** resonance. Since they carry different three-quark spin S (see Table 3.5), their mass difference of about 300 MeV can be understood as a hyperfine splitting due to spin-dependent interactions. The Δ resonance decays almost exclusively into $N\pi$ and thus appears as a prominent peak in $N\pi$ scattering.

The $N = 1$ band in Table 3.5 can still be identified with experimental states, e.g. in the nucleon channel the $(\frac{1}{2}, \mathbf{8})$ states would correspond to the $N(1535)$ and $N(1520)$, where the former is the parity partner of the nucleon (see also Fig. 3.16). For the higher-lying states the quark-model identification becomes more problematic: the $N = 2$ band already overpredicts the positive-parity spectrum for $\frac{3}{2}^+$ states, and the $N = 3$ band contains over 20 negative-parity states which have not been seen in experiments.

An open question concerns the **Roper** resonance $N(1440)$, which is the first radial excitation of the nucleon but has properties that are incompatible with the quark model; for example, its mass is lower than that of the $N(1535)$. The Roper has been suggested to be a **dynamically generated resonance**, in the sense that the interactions between nucleons and pions could generate additional states on top of qqq configurations. This ties in with the ‘**meson cloud**’ picture, where baryons are thought to be surrounded by clouds of light pseudoscalar mesons which change their properties. A similar case is the $\Lambda(1405)$ with $J^P = \frac{1}{2}^-$ which is also not well described by quark models. From a microscopic point of view, such effects would signal a multiquark admixture for light baryons similarly to the meson spectrum. For these reasons, a thorough understanding of light baryons from QCD remains an open problem.

M	I	S	$\frac{1}{2}^+$	$\frac{3}{2}^+$	$\frac{5}{2}^+$	$\frac{7}{2}^+$	$\frac{9}{2}^+$	$\frac{11}{2}^+$	$\frac{13}{2}^+$
8	$\frac{1}{2}$	0	N (939) N (1440) N (1710) N (1880) N (2100) N (2300)	N (1720) N (1900)	N (1680) N (1860) N (2000)	N (1990)	N (2220)		N (2700)
10	$\frac{3}{2}$	0	Δ (1910)	Δ (1232) Δ (1600) Δ (1920)	Δ (1905) Δ (2000)	Δ (1950)	Δ (2300)	Δ (2420)	
8, 1	0	-1	Λ (1116) Λ (1600) Λ (1810)	Λ (1890)	Λ (1820) Λ (2110)	Λ (2085)	Λ (2350)		
8, 10	1	-1	Σ (1193) Σ (1660) Σ (1880)	Σ (1385)	Σ (1915)	Σ (2030)			
8, 10	$\frac{1}{2}$	-2	Ξ (1318)	Ξ (1530)					
10	0	-3		Ω (1672)					

M	I	S	$\frac{1}{2}^-$	$\frac{3}{2}^-$	$\frac{5}{2}^-$	$\frac{7}{2}^-$	$\frac{9}{2}^-$	$\frac{11}{2}^-$	$\frac{13}{2}^-$
8	$\frac{1}{2}$	0	N (1535) N (1650) N (1895)	N (1520) N (1700) N (1875) N (2120)	N (1675) N (2060) N (2570)	N (2190)	N (2250)	N (2600)	
10	$\frac{3}{2}$	0	Δ (1620) Δ (1900)	Δ (1700) Δ (1940)	Δ (1930)	Δ (2200)	Δ (2400)		Δ (2750)
8, 1	0	-1	Λ (1380) Λ (1405) Λ (1670) Λ (1800)	Λ (1520) Λ (1690)	Λ (1830)	Λ (2100)			
8, 10	1	-1	Σ (1750) Σ (1900)	Σ (1670) Σ (1910)	Σ (1775)				
8, 10	$\frac{1}{2}$	-2	Ξ (1690)	Ξ (1820)					
10	0	-3							

TABLE 3.6: Light and strange baryon spectrum in terms of J^P , isospin I and strangeness S from the PDG 2020 (<https://pdglive.lbl.gov>). Only established states (two-, three- and four-star resonances) are included. The ground states are shown in color.

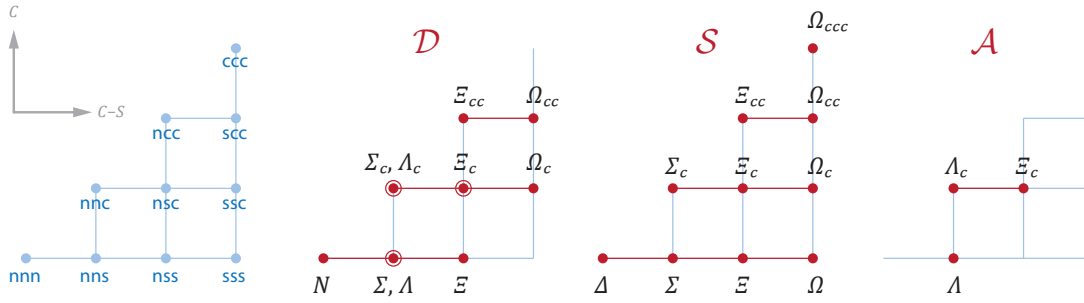


FIG. 3.17: Charmed baryon multiplets in the $\{C - S, C\}$ plane. The left figure shows the quark content for each state, where n stands for light quarks.

	uuc	udc	ddc	usc	dsc	ssc	ucc	dcc	scc	ccc
\mathcal{S}	Σ_c^{++}	Σ_c^+	Σ_c^0	Ξ_c^+	Ξ_c^0	Ω_c^0	Ξ_{cc}^{++}	Ξ_{cc}^+	Ω_{cc}^+	Ω_{ccc}^+
\mathcal{D}_1	Σ_c^{++}	Σ_c^+	Σ_c^0	Ξ_c^+	Ξ_c^0	Ω_c^0	Ξ_{cc}^{++}	Ξ_{cc}^+	Ω_{cc}^+	
\mathcal{D}_2		Λ_c^+		Ξ_c^+	Ξ_c^0					
\mathcal{A}		Λ_c^+		Ξ_c^+	Ξ_c^0					

TABLE 3.7: $SU(3)_f$ flavor wave functions for baryons.

Charmed baryons. The extension of Eqs. (3.2.72–3.2.74) to construct the flavor wave functions of charmed baryons is straightforward: start with a given quark content like uuc and work out the multiplets. This produces new singlets, doublets and antisingslets, which are collected in Table 3.7 and add to the former ones to yield

$$4 \otimes 4 \otimes 4 = 20_S \oplus 20_{M_A} \oplus 20_{M_S} \oplus 4_A. \quad (3.2.89)$$

The resulting flavor multiplets are shown in Fig. 3.17 and contain the $SU(3)$ octet, decuplet and singlet as their bottom levels. As before, because the $SU(4)_f$ symmetry is badly broken, states with the same quark content (the same I_3 , strangeness S and number of charm quarks C) will mix.

For singly-charmed baryons, the multiplet partners of the $J^P = \frac{1}{2}^+$ octet and $\frac{3}{2}^+$ decuplet baryons are experimentally established, along with a few other states with different J^P and some whose quantum numbers have not yet been determined. So far there is evidence for only one doubly charmed Ξ_{cc}^{++} baryon; presumably these would have a very different structure from light baryons and resemble a heavy ‘double-star’ system with an attached light ‘planet’.

Pentaquarks? Another type of baryon made of charm quarks was recently observed by the LHCb collaboration, who found several peaks in the $J/\psi p$ spectrum in the 4300...4500 MeV region. Since this implies a minimal quark content $uudc\bar{c}$, it would be the first experimental evidence for **pentaquarks**. The proximity of those peaks to the $\Sigma_c \bar{D}$ and $\Sigma_c \bar{D}^*$ thresholds suggests a molecular explanation in terms of meson-baryon molecules, in analogy to exotic meson candidates in the charmonium sector.

References and further reading

Symmetries and Ward-Takahashi identities:

- M. Maggiore, *A Modern Introduction to Quantum Field Theory*. Oxford University Press, 2005
- R. A. Bertlmann, *Anomalies in Quantum Field Theory*. Oxford University Press, 1996.
- S. Pokorski, *Gauge Field Theories*. Cambridge University Press, 1987
- P. H. Frampton, *Gauge Field Theories*. Wiley-VCH, 2008.
- V. P. Nair, *Quantum Field Theory: A Modern Perspective*. Springer, 2005.
- S. Scherer, *Introduction to Chiral Perturbation Theory*, Adv. Nucl. Phys. 27 (2003) 277, [arXiv:hep-ph/0210398](https://arxiv.org/abs/hep-ph/0210398)

Hadrons produce poles:

- S. Weinberg, *The Quantum Theory of Fields. Vol. I*. Cambridge University Press, 1995/96 (Sec. 10.2)
- M. D. Schwartz, *Quantum Field Theory and the Standard Model*. Cambridge University Press, 2014 (Sec. 24.3)

Lattice QCD calculations of hadron masses:

- Z. Fodor, C. Hoelbling, *Light Hadron Masses from Lattice QCD*, Rev. Mod. Phys. 84 (2012) 449, [arXiv:1203.4789](https://arxiv.org/abs/1203.4789) [[hep-lat](#)]
- R. A. Briceno, J. J. Dudek, R. D. Young, *Scattering processes and resonances from lattice QCD*, Rev. Mod. Phys. 90 (2018) 2, 025001 [arXiv:1706.06223](https://arxiv.org/abs/1706.06223) [[hep-lat](#)]

Bethe-Salpeter equations:

- P. Maris, C. D. Roberts, *π and K meson Bethe-Salpeter amplitudes*, Phys. Rev. C 56 (1997) 3369, [arXiv:nuc1-th/9708029](https://arxiv.org/abs/nuc1-th/9708029)
- G. Eichmann, H. Sanchis-Alepuz, R. Williams, R. Alkofer, C. S. Fischer, *Baryons as relativistic three-quark bound states*, Prog. Part. Nucl. Phys. 91 (2016) 1, [arXiv:1606.09602](https://arxiv.org/abs/1606.09602) [[hep-ph](#)]

Flavor wave functions:

- H. Georgi, *Lie Algebras in Particle Physics*. Westview Press, 1999
- A. Hosaka and H. Toki, *Quarks, Baryons and Chiral Symmetry*. World Scientific, 2001
- F. E. Close, *An Introduction to Quarks and Partons*. Academic Press, 1979
- T. P. Cheng and L. F. Li, *Gauge Theory of Elementary Particle Physics*. Oxford University Press, 1984
- F. Halzen and A. D. Martin, *Quarks and Leptons: An Introductory Course in Modern Particle Physics*. Wiley, 1984.

Reviews on light exotic mesons:

- J. Pelaez, *From controversy to precision on the sigma meson: a review on the status of the non-ordinary $f_0(500)$ resonance*, Phys. Rept. 658, 1 (2016), [arXiv:1510.00653](https://arxiv.org/abs/1510.00653) [[hep-ph](#)].
- C. A. Meyer, E. S. Swanson, *Hybrid Mesons*, Prog. Part. Nucl. Phys. 82 (2015), 21 [arXiv:1502.07276](https://arxiv.org/abs/1502.07276) [[hep-ph](#)]
- W. Ochs, *The Status of Glueballs*, J. Phys. G 40 (2013) 043001, [arXiv:1301.5183](https://arxiv.org/abs/1301.5183) [[hep-ph](#)]

Quark model classification of baryons:

- A. Hosaka and H. Toki, *Quarks, Baryons and Chiral Symmetry*. World Scientific, 2001
- E. Klempt and B. C. Metsch, *Multiplet classification of light-quark baryons*. Eur. Phys. J. A (2012) **48**,
- [PDG review on the quark model](#), P.A. Zyla et al., Prog. Theor. Exp. Phys. 2020, 083C01 (2020)

Reviews on light baryons:

- E. Klempt, J.-M. Richard, *Baryon spectroscopy*, Rev. Mod. Phys. 82 2010 (1095), [arXiv:0901.2055 \[hep-ph\]](#)
- V. Crede, W. Roberts, *Progress Toward Understanding Baryon Resonances*, Reports on Progress in Physics 76 (2013) 7, [arXiv:1302.7299 \[nucl-ex\]](#)

Reviews on heavy exotics:

- H.-X. Chen, W. Chen, X. Liu, S.-L. Zhu, *The hidden-charm pentaquark and tetraquark states*, Phys. Rept. 639, 1-121 (2016), [arXiv:1601.02092 \[hep-ph\]](#)
- R. F. Lebed, R. E. Mitchell, E. S. Swanson, *Heavy-Quark QCD Exotica*, Prog. Part. Nucl. Phys. 93, 143 (2017), [arXiv:1610.04528 \[hep-ph\]](#)
- A. Esposito, A. Pilloni, A. D. Polosa, *Multiquark Resonances*, Phys. Rept. 668 (2017) 1, [arXiv:1611.07920 \[hep-ph\]](#)
- F.-K. Guo, C. Hanhart, U.-G. Meißner, Q. Wang, Q. Zhao, B.-S. Zou, *Hadronic molecules*, Rev. Mod. Phys. 90, 015004 (2018), [arXiv:1705.00141 \[hep-ph\]](#)
- A. Ali, J. S. Lange, S. Stone, *Exotics: Heavy Pentaquarks and Tetraquarks*, Prog. Part. Nucl. Phys. 97 (2017) 123, [arXiv:1706.00610 \[hep-ph\]](#)
- S. L. Olsen, T. Skwarnicki, D. Zieminska, *Non-Standard Heavy Mesons and Baryons, an Experimental Review*, [arXiv:1708.04012 \[hep-ph\]](#)
- Y. R. Liu, H.-X. Chen, W. Chen, X. Liu, S.-L. Zhu, *Pentaquark and Tetraquark states*, Prog. Part. Nucl. Phys. 107 (2019) 237-320, [arXiv:1903.11976 \[hep-ph\]](#)
- N. Brambilla et al., *The XYZ states: experimental and theoretical status and perspectives*, Physics Reports 873 (2020) 1-154, [arXiv:1907.07583 \[hep-ex\]](#)

Chapter 4

Low-energy QCD phenomenology

4.1 Quark potential models

Among the earliest approaches to QCD have been **quark potential models** describing quarks that move nonrelativistically within a hadron. The assumption is that the QCD interactions dress each quark with a cloud of virtual gluons and $q\bar{q}$ pairs and in this way transform it into a **constituent quark**, whose dynamical ‘constituent mass’ is so large that it becomes nonrelativistic. The energy levels and wave functions of hadrons are then obtained by solving a nonrelativistic Schrödinger equation using some assumed potential. While this strategy seems acceptable for truly massive quarks like b quarks, it becomes questionable for light u , d and s quarks where relativity and chiral symmetry complicate the dynamics (in fact, in Sec. 4.2 we will see how the dynamical breaking of chiral symmetry motivates the emergence of such constituent masses). Nevertheless, nonrelativistic quark models provide a framework for describing both ground and excited hadronic states and they have proven very useful for a basic understanding of their properties, in particular also for distinguishing ‘ordinary’ versus ‘exotic’ hadrons. In addition, with experimental indications for multi-quark states, quark models have seen a revival in recent years.

The basic idea is to write down a Hamiltonian for a system of n quarks and/or antiquarks, where the interquark potential is typically the sum of two-body interactions:

$$H = \sum_{i=1}^n \frac{\mathbf{p}_i^2}{2m_i} + \sum_{i<j} V(\mathbf{r}_{ij}), \quad \mathbf{r}_{ij} = \mathbf{r}_i - \mathbf{r}_j. \quad (4.1.1)$$

The Schrödinger equation

$$H \Psi_\lambda = E_\lambda \Psi_\lambda \quad (4.1.2)$$

then determines the binding energy E_λ and wave function Ψ_λ of a hadronic state $|\lambda\rangle$, whose mass is given by $M_\lambda = \sum_i m_i + E_\lambda$. Although there is no unique specification for the interquark potential V , it typically contains a spin- and flavor-dependent short-range potential, a spin- and flavor-independent long-range confining potential, basis mixing in the meson and baryon sectors, and relativistic corrections.

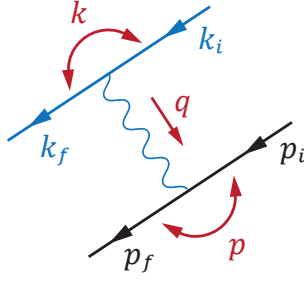


FIG. 4.1: One-photon exchange diagram between two fermions.

Breit-Fermi interaction. A useful starting point for constructing the potential $V(\mathbf{r}_{ij})$ is the nonrelativistic expansion of the **one-gluon exchange** interaction in a $q\bar{q}$ or qq system. The template for this is the analogous one-photon exchange amplitude in QED, e.g. between the electron and the proton in a hydrogen atom (Fig. 4.1), which is identical apart from the coupling, the masses and color factors:

$$\mathcal{M}_{\sigma\sigma'\lambda\lambda'}(q, p, k) = \frac{e^2}{q^2} \bar{u}_{\sigma'}(p_f) \gamma^\mu u_\sigma(p_i) \bar{u}_{\lambda'}(k_f) \gamma_\mu u_\lambda(k_i). \quad (4.1.3)$$

We define three independent momenta q , p and k by

$$\begin{aligned} p_i &= p - \frac{q}{2}, & k_i &= k + \frac{q}{2}, & \Leftrightarrow & & q &= p_f - p_i = k_i - k_f. \\ p_f &= p + \frac{q}{2}, & k_f &= k - \frac{q}{2}, & & & & \end{aligned} \quad (4.1.4)$$

In the standard representation, the onshell spinors have the form

$$u_\sigma(p) = \sqrt{\frac{E_p + m}{2E_p}} \begin{bmatrix} \xi_\sigma \\ \boldsymbol{\alpha} \cdot \boldsymbol{\tau} \xi_\sigma \end{bmatrix}, \quad \boldsymbol{\alpha} = \frac{\mathbf{p}}{E_p + m}, \quad E_p = \sqrt{\mathbf{p}^2 + m^2}, \quad (4.1.5)$$

where m is the mass of the respective particle. The Pauli spinors satisfy $\xi_{\sigma'}^\dagger \xi_\sigma = \delta_{\sigma'\sigma}$, but note that we included a factor $1/\sqrt{2E_p}$ which corresponds to the ‘nonrelativistic’ Dirac spinor normalization $u_{\sigma'}^\dagger(p) u_\sigma(p) = \delta_{\sigma'\sigma}$.

To work out the non-relativistic expansion of the amplitude (4.1.3), we expand it in the three-momenta \mathbf{q} , \mathbf{p} and \mathbf{k} . The spinors become

$$u_\sigma(p) \approx \begin{bmatrix} \left(1 - \frac{\mathbf{p}^2}{8m^2}\right) \xi_\sigma \\ \frac{\mathbf{p} \cdot \boldsymbol{\tau}}{2m} \xi_\sigma \end{bmatrix}, \quad \bar{u}_{\sigma'}(p) = u_{\sigma'}^\dagger(p) \gamma^0 \approx \left[\xi_{\sigma'}^\dagger \left(1 - \frac{\mathbf{p}^2}{8m^2}\right), -\xi_{\sigma'}^\dagger \frac{\mathbf{p} \cdot \boldsymbol{\tau}}{2m} \right] \quad (4.1.6)$$

and the expansion of the factor $1/q^2$ in front of the amplitude yields

$$q^2 \approx -\mathbf{q}^2 + \frac{(\mathbf{p} \cdot \mathbf{q})(\mathbf{q} \cdot \mathbf{k})}{m_1 m_2} \Rightarrow \frac{1}{q^2} \approx -\frac{1}{\mathbf{q}^2} \left(1 + \frac{(\mathbf{p} \cdot \hat{\mathbf{q}})(\hat{\mathbf{q}} \cdot \mathbf{k})}{m_1 m_2}\right), \quad (4.1.7)$$

where $\hat{\mathbf{q}} = \mathbf{q}/|\mathbf{q}|$. Employing the γ -matrices in the standard representation, the resulting amplitude is

$$\begin{aligned} \mathcal{M}(\mathbf{q}, \mathbf{p}, \mathbf{k}) \approx & -\frac{4\pi\alpha}{\mathbf{q}^2} \left[1 - \frac{\mathbf{q}^2}{8} \left(\frac{1}{m_1^2} + \frac{1}{m_2^2} \right) - \frac{\mathbf{p} \cdot \mathbf{k} - (\mathbf{p} \cdot \hat{\mathbf{q}})(\mathbf{k} \cdot \hat{\mathbf{q}})}{m_1 m_2} - \frac{(\mathbf{q} \times \mathbf{s}_1) \cdot (\mathbf{q} \times \mathbf{s}_2)}{m_1 m_2} \right. \\ & \left. + \frac{i\mathbf{s}_1 \cdot (\mathbf{q} \times \mathbf{p})}{2m_1^2} - \frac{i\mathbf{s}_2 \cdot (\mathbf{q} \times \mathbf{k})}{2m_2^2} - \frac{i\mathbf{s}_1 \cdot (\mathbf{q} \times \mathbf{k}) - i\mathbf{s}_2 \cdot (\mathbf{q} \times \mathbf{p})}{m_1 m_2} \right] \end{aligned} \quad (4.1.8)$$

with $\alpha = e^2/(4\pi)$. Here we suppressed the polarization indices by dropping the unit matrices $\delta_{\sigma'\sigma} \delta_{\lambda'\lambda}$ in the notation and introducing the spin vectors

$$(\mathbf{s}_1)_{\sigma'\sigma} = \chi_{\sigma'}^\dagger \frac{\boldsymbol{\tau}}{2} \chi_\sigma, \quad (\mathbf{s}_2)_{\lambda'\lambda} = \chi_{\lambda'}^\dagger \frac{\boldsymbol{\tau}}{2} \chi_\lambda. \quad (4.1.9)$$

Note that all instances of \mathbf{p} and \mathbf{k} in the above expression are transverse in \mathbf{q} , so we could as well replace $\mathbf{p} \rightarrow \mathbf{p}_{i,f}$ and $\mathbf{k} \rightarrow \mathbf{k}_{i,f}$ which only differ by $\pm\mathbf{q}/2$. Finally, we take the Fourier transform with respect to \mathbf{q} using

$$\int \frac{d^3q}{(2\pi)^3} \frac{e^{i\mathbf{q}\cdot\mathbf{r}}}{\mathbf{q}^2} \begin{bmatrix} 1 \\ \mathbf{q} \\ \mathbf{q}^2 \\ q_i q_j \\ \frac{q_i q_j}{\mathbf{q}^2} \end{bmatrix} = \frac{1}{4\pi r} \begin{bmatrix} 1 \\ \frac{i\mathbf{r}}{r^2} \\ 4\pi r \delta^3(\mathbf{r}) \\ \frac{1}{r^2} (\delta_{ij} - 3\hat{r}_i \hat{r}_j) + \frac{4\pi r}{3} \delta^3(\mathbf{r}) \delta_{ij} \\ \frac{1}{2} (\delta_{ij} - \hat{r}_i \hat{r}_j) \end{bmatrix}, \quad (4.1.10)$$

where $\hat{\mathbf{r}} = \mathbf{r}/r$ and $r = |\mathbf{r}|$, to arrive at the final expression below.

The resulting **Breit-Fermi interaction** is the three-dimensional Fourier transform of the one-photon exchange amplitude (4.1.3) in the non-relativistic limit:

$$\begin{aligned} V(\mathbf{r}, \mathbf{p}, \mathbf{k}) &= \int \frac{d^3q}{(2\pi)^3} e^{i\mathbf{q}\cdot\mathbf{r}} \mathcal{M}(\mathbf{q}, \mathbf{p}, \mathbf{k}) \\ &= \alpha \left[-\frac{1}{r} + T_d + \frac{T_o}{r} + \frac{8\pi}{3} \delta^3(\mathbf{r}) T_{ss} + \frac{T_{\text{ten}} + T_{\text{so}}^{(1)} + T_{\text{so}}^{(2)}}{r^3} \right]. \end{aligned} \quad (4.1.11)$$

It is identical for a fermion-fermion and fermion-antifermion system, and apart from the color factor (which we discuss below) it can be directly carried over to QCD to establish a non-relativistic qq and $q\bar{q}$ interaction potential between (onshell) quarks.

■ The first three terms in Eq. (4.1.11) are the Coulomb term, the Darwin term and the orbit-orbit interaction, which are all spin-independent:

$$T_d = \frac{\pi}{2} \left(\frac{1}{m_1^2} + \frac{1}{m_2^2} \right) \delta^3(\mathbf{r}), \quad T_o = \frac{\mathbf{p} \cdot \mathbf{k} + (\mathbf{p} \cdot \hat{\mathbf{r}})(\mathbf{k} \cdot \hat{\mathbf{r}})}{2m_1 m_2}. \quad (4.1.12)$$

■ The next two terms constitute the **hyperfine interaction**, which consists of a spin-spin contact term and a tensor force:

$$T_{ss} = \frac{\mathbf{s}_1 \cdot \mathbf{s}_2}{m_1 m_2}, \quad T_{\text{ten}} = \frac{3(\mathbf{s}_1 \cdot \hat{\mathbf{r}})(\mathbf{s}_2 \cdot \hat{\mathbf{r}}) - \mathbf{s}_1 \cdot \mathbf{s}_2}{m_1 m_2}. \quad (4.1.13)$$

It arises from a magnetic dipole-dipole interaction, since the Fourier transform of the term $(\mathbf{q} \times \mathbf{s}_1) \cdot (\mathbf{q} \times \mathbf{s}_2)$ is proportional to

$$(\mathbf{s}_1 \times \nabla) \cdot (\mathbf{s}_2 \times \nabla) \frac{1}{r} = \mathbf{s}_1 \cdot \left[\nabla \times \left((\mathbf{s}_2 \times \nabla) \frac{1}{r} \right) \right] \propto \mathbf{s}_1 \cdot \mathbf{B}_2, \quad (4.1.14)$$

where \mathbf{B}_2 is the magnetic field produced by the magnetic dipole moment \mathbf{s}_2 . Spin-spin interactions of the form $\mathbf{s}_1 \cdot \mathbf{s}_2$ in the Hamiltonian induce level splittings between states with different spin, which leads to mass formulas of the form

$$H = \dots + c \mathbf{s}_1 \cdot \mathbf{s}_2 + \dots \quad \Rightarrow \quad M = \dots + c \langle \mathbf{s}_1 \cdot \mathbf{s}_2 \rangle + \dots \quad (4.1.15)$$

If we write the total spin of a two-fermion system as $\mathbf{S} = \mathbf{s}_1 + \mathbf{s}_2$, we have

$$\mathbf{S}^2 = \mathbf{s}_1^2 + \mathbf{s}_2^2 + 2\mathbf{s}_1 \cdot \mathbf{s}_2 \quad \Rightarrow \quad \langle \mathbf{s}_1 \cdot \mathbf{s}_2 \rangle = \frac{S(S+1) - \frac{3}{2}}{2}, \quad (4.1.16)$$

which generates the mass splittings between pseudoscalar ($S = 0$) and vector mesons ($S = 1$). For baryons, the hyperfine interaction is responsible for the dominant mass splittings between the ground-state octet ($S = \frac{1}{2}$) and decuplet baryons ($S = \frac{3}{2}$).

■ The remaining two terms constitute the **spin-orbit interaction**:

$$T_{\text{so}}^{(1)} = \frac{\mathbf{s}_1 \cdot (\mathbf{r} \times \mathbf{p})}{2m_1^2} - \frac{\mathbf{s}_2 \cdot (\mathbf{r} \times \mathbf{k})}{2m_2^2}, \quad T_{\text{so}}^{(2)} = \frac{\mathbf{s}_2 \cdot (\mathbf{r} \times \mathbf{p}) - \mathbf{s}_1 \cdot (\mathbf{r} \times \mathbf{k})}{m_1 m_2}. \quad (4.1.17)$$

For example, for $m_1 = m_2 = m$ and in the frame where $\mathbf{k} = -\mathbf{p}$, this yields

$$T_{\text{so}}^{(1)} + T_{\text{so}}^{(2)} = \frac{3(\mathbf{s}_1 + \mathbf{s}_2) \cdot \mathbf{L}}{2m^2} = \frac{3\mathbf{L} \cdot \mathbf{S}}{2m^2}, \quad \mathbf{L} = \mathbf{r} \times \mathbf{p}. \quad (4.1.18)$$

The resulting mass splittings can be estimated from $\mathbf{J}^2 = (\mathbf{L} + \mathbf{S})^2$ as

$$\langle \mathbf{L} \cdot \mathbf{S} \rangle = \frac{1}{2} (J(J+1) - L(L+1) - S(S+1)). \quad (4.1.19)$$

For orbital ground states ($L = 0$ and thus $J = S$) the spin-orbit interaction does not contribute. For $L = S \neq 0$, it leads to splittings between states with different J , e.g. for the charmonium states $\{\chi_{c0}, \chi_{c1}, \chi_{c2}\}$ with $J^{PC} = \{0, 1, 2\}^{++}$, which carry $S = L = 1$ and only differ in their total angular momentum J (see Fig. 3.14):

$$\langle \mathbf{L} \cdot \mathbf{S} \rangle = \frac{1}{2} (J(J+1) - 4) = \{-2, -1, 2\}. \quad (4.1.20)$$

For baryons, spin-orbit interactions are generally small and usually neglected. (In the hydrogen atom with $m_e \ll m_p$, the spin-orbit force beats the spin-spin interaction and gives rise to the atomic **fine structure** effects.)

It turns out that Eq. (4.1.11) is quite general. If we had started from a different amplitude with scalar, pseudoscalar, axialvector or tensor particle exchanges, we could still split the resulting potential into spin-independent and spin-dependent terms:

$$V = V_0 + (\dots) + V_{\text{ss}} T_{\text{ss}} + V_{\text{ten}} T_{\text{ten}} + V_{\text{so}}^{(1)} T_{\text{so}}^{(1)} + V_{\text{so}}^{(2)} T_{\text{so}}^{(2)}. \quad (4.1.21)$$

The spin-dependent terms can be expressed through derivatives of $V_0(r)$, e.g., for a general vector-exchange potential one finds

$$V_{\text{ss}} = \frac{2}{3} \Delta V_0, \quad V_{\text{ten}} = \frac{1}{3} \left(\frac{V_0'}{r} - V_0'' \right), \quad V_{\text{so}}^{(1)} = V_{\text{so}}^{(2)} = \frac{V_0'}{r}, \quad (4.1.22)$$

where Δ is the Laplace operator and a Coulomb potential corresponds to $V_0 = -1/r$. A scalar exchange only leads to a spin-orbit interaction,

$$V_{\text{so}}^{(1)} = -\frac{V_0'}{r}, \quad V_{\text{ss}} = V_{\text{ten}} = V_{\text{so}}^{(2)} = 0, \quad (4.1.23)$$

whereas a pseudoscalar exchange does not produce spin-dependent terms at all. Axialvector and tensor exchanges only contribute to V_{ss} .

Color factors. In order to apply the Breit-Fermi potential to QCD, where the photon is replaced by a gluon, we must also work out the color algebra since gluons couple to quarks through the $SU(3)_c$ generators \mathbf{t}_a . If we denote the generators in an arbitrary $SU(3)$ representation by $\hat{\mathbf{t}}_a$ and write

$$\mathbf{t}^2 = \sum_a \hat{\mathbf{t}}_a^2 = C(R), \quad (4.1.24)$$

then the Casimir in a general $SU(3)$ representation is given by

$$C(R) = \frac{3p + 3q + p^2 + pq + q^2}{3}, \quad (4.1.25)$$

where p and q are the quantum numbers that label the multiplets, cf. Eq. (B.2.7) in the appendix. This yields

$$C(\mathbf{1}) = 0, \quad C(\mathbf{3}) = C(\bar{\mathbf{3}}) = \frac{4}{3}, \quad C(\mathbf{6}) = C(\bar{\mathbf{6}}) = \frac{10}{3}, \quad C(\mathbf{8}) = 3, \quad \text{etc.} \quad (4.1.26)$$

If we write

$$\mathbf{t}_1 \cdot \mathbf{t}_2 = \sum_a \hat{\mathbf{t}}_a \otimes \hat{\mathbf{t}}_a, \quad (4.1.27)$$

where 1 and 2 refer to the particles on which they act, then the generator in the product space is $\mathbf{t}_{12} = \mathbf{t}_1 + \mathbf{t}_2$ with $\mathbf{t}_{12}^2 = C_{12}$, $\mathbf{t}_1^2 = C_1$ and $\mathbf{t}_2^2 = C_2$. As a consequence,

$$\mathbf{t}_{12}^2 = \mathbf{t}_1^2 + \mathbf{t}_2^2 + 2\mathbf{t}_1 \cdot \mathbf{t}_2 \quad \Rightarrow \quad \mathbf{t}_1 \cdot \mathbf{t}_2 = \frac{C_{12} - C_1 - C_2}{2}. \quad (4.1.28)$$

For an attractive color potential we must have $\mathbf{t}_1 \cdot \mathbf{t}_2 < 0$. Because quarks and antiquarks live in the $\mathbf{3}$ and $\bar{\mathbf{3}}$ representations, one has $C_1 = C_2 = \frac{4}{3}$ both for $q\bar{q}$ and qq systems which entails $C_{12} < \frac{8}{3}$ for an attractive potential. From Eq. (4.1.26) this only leaves color-singlet mesons and color-antitriplet diquarks (or color-triplet antiquarks):

$$C_{12}(\mathbf{1}) = 0 \quad \Rightarrow \quad \mathbf{t}_1 \cdot \mathbf{t}_2 = -\frac{4}{3}, \quad C_{12}(\bar{\mathbf{3}}) = \frac{4}{3} \quad \Rightarrow \quad \mathbf{t}_1 \cdot \mathbf{t}_2 = -\frac{2}{3}. \quad (4.1.29)$$

Thus, the $q\bar{q}$ color interaction in the color-singlet meson channel is maximally attractive, whereas in the color-octet channel with $C_{12}(\mathbf{8}) = 3$ it is repulsive. Likewise, the interaction between two quarks in the $\bar{\mathbf{3}}$ channel is attractive, with a color factor half as strong as for mesons, whereas in the $\mathbf{6}$ channel it is repulsive. This is relevant for the binding of baryons, since from

$$\mathbf{3} \otimes \mathbf{3} \otimes \mathbf{3} = (\bar{\mathbf{3}} \oplus \mathbf{6}) \otimes \mathbf{3} = \mathbf{1} \oplus \mathbf{8} \oplus \mathbf{8} \oplus \mathbf{10} \quad (4.1.30)$$

a $\bar{\mathbf{3}}$ diquark can bind together with the remaining quark to form a color-singlet baryon. One can furthermore show that

$$f_{abc}(\mathbf{t}_a)_{il}(\mathbf{t}_b)_{jm}(\mathbf{t}_c)_{kn}\varepsilon_{lmn} = 0. \quad (4.1.31)$$

This is the color factor stemming from a three-gluon vertex (f_{abc}) that is connected to three quarks, which are then contracted with the antisymmetric color wave function of a baryon (ε_{lmn}). Hence, the leading three-body force in a baryon mediated by a three-gluon vertex vanishes! This suggests that the internal structure of baryons is dominated by two-body forces in the attractive $\bar{\mathbf{3}}$ diquark channels.

Potential in QCD. What would the interquark potential look like in QCD?

- At short distances, by means of asymptotic freedom, we expect a **Coulomb**-like potential $V(r) \propto -1/r$ mediated by the massless gluon. Here we could take over the Breit-Fermi interaction (4.1.11), replace $\alpha \rightarrow \alpha_s$ and attach a color factor $4/3$ for the $q\bar{q}$ interaction and $2/3$ for the qq interaction (the minus signs are already implicit, e.g. in the Coulomb term), or just start with the Coulomb term alone.
- At large distances, we expect a linear **confinement** potential $V_{\text{conf}} = \sigma r$, where σ is called the string tension. This is motivated from several angles, including the observed mass orderings (Regge phenomenology) and lattice calculations of the Wilson loop in pure Yang-Mills theory.

Two examples of how to interpolate between a single gluon exchange at short distances and confinement at large distances are the **Cornell** and **Richardson** potentials:

$$\begin{aligned} V_C(r) &= -\frac{4\alpha_s}{3r} + \sigma r, \\ V_R(r) &= -\frac{4}{3} \frac{(4\pi)^2}{\beta_0} \text{FT} \left[\frac{1}{\mathbf{q}^2 \ln(1 + \mathbf{q}^2/\Lambda^2)} \right], \end{aligned} \quad (4.1.32)$$

where the latter also incorporates asymptotic freedom in terms of a logarithmic running of the coupling (FT denotes the Fourier transform).

From the discussion around Eq. (4.1.21), the generic structure of the potential is not limited to a gluon exchange. For example, we could distribute the confinement potential between a scalar and vector exchange potential by a parameter $0 \leq \xi \leq 1$,

$$V_0 = V_S + V_V = (1 - \xi) \sigma r + \left(\xi \sigma r - \frac{a}{r} \right), \quad (4.1.33)$$

where $\xi = 0$ corresponds to scalar confinement and $\xi = 1$ to vector confinement. According to Eqs. (4.1.22–4.1.23), this yields

$$V = V_0 + (\dots) + \frac{2}{3} \Delta V_V T_{\text{ss}} + \frac{1}{3} \left(\frac{V'_V}{r} - V''_V \right) T_{\text{ten}} + \frac{V'_V - V'_S}{r} T_{\text{so}}^{(1)} + \frac{V'_V}{r} T_{\text{so}}^{(2)}. \quad (4.1.34)$$

Fits to the charmonium spectrum based on this expression have suggested that confinement is predominantly **scalar**.

One should keep in mind that the concept of a *potential* does not account for the full dynamics as it assumes interactions to be instantaneous. More generally, one expects a large-distance behavior $\propto r$ for the $q\bar{q}$ four-point function in Eq. (3.1.135), since this is the quantity related to the Wilson loop (for infinitely heavy static quarks). The dynamical origin of confinement is still under debate and has been attributed to **center vortices**, or to the formation of color-electric flux tubes and the condensation of color-magnetic monopoles in the **dual superconductor** picture. Diagrammatically, it is conceivable that confinement may only arise from complicated combinations of gluon exchanges. On the other hand, the full quark-gluon vertex in QCD has a more general structure than the γ^μ part, cf. Eq. (2.3.16), and a gluon exchange with a full propagator and full vertices also contains scalar parts. In Landau gauge, the scaling solution mentioned below Eq. (2.3.25) indeed generates a $q\bar{q}$ interaction $\propto 1/q^4$ which leads to a linear rise in coordinate space.

Quark models for baryons. In order to construct a Hamiltonian (4.1.36) for baryons with $n = 3$, let us assume equal constituent masses $m_i = m$. It is convenient to introduce a center-of-mass coordinate \mathbf{R} and two relative coordinates $\boldsymbol{\rho}$ and $\boldsymbol{\lambda}$,

$$\mathbf{R} = \frac{\mathbf{r}_1 + \mathbf{r}_2 + \mathbf{r}_3}{\sqrt{3}}, \quad \boldsymbol{\rho} = \frac{\mathbf{r}_1 - \mathbf{r}_2}{\sqrt{2}}, \quad \boldsymbol{\lambda} = \frac{\mathbf{r}_1 + \mathbf{r}_2 - 2\mathbf{r}_3}{\sqrt{6}}, \quad (4.1.35)$$

because in this way one can remove the center-of-mass motion (which would have led to spurious excitations) to arrive at

$$H = \frac{\mathbf{p}_\rho^2}{2m} + \frac{\mathbf{p}_\lambda^2}{2m} + \sum_{i < j} V(\mathbf{r}_{ij}). \quad (4.1.36)$$

We could start with a harmonic oscillator potential $V(\mathbf{r}_{ij}) = k\mathbf{r}_{ij}^2/2$, which because of $\sum_{i < j} \mathbf{r}_{ij}^2 = 3(\boldsymbol{\lambda}^2 + \boldsymbol{\rho}^2)$ leads to two independent spherical harmonic oscillators

$$H = \left(\frac{\mathbf{p}_\rho^2}{2m} + \frac{3k}{2} \boldsymbol{\rho}^2 \right) + \left(\frac{\mathbf{p}_\lambda^2}{2m} + \frac{3k}{2} \boldsymbol{\lambda}^2 \right) \quad (4.1.37)$$

with frequency $\omega_0 = \sqrt{3k/m}$. The resulting baryon spectrum is then $E_N = E_0 + N\omega_0$, where E_0 is the ground-state energy, $N = 2n + l$, $n = n_\rho + n_\lambda$ and $l = l_\rho + l_\lambda$, and n_α and l_α are the radial and orbital excitations of the oscillators. The total angular momentum $\mathbf{J} = \mathbf{L} + \mathbf{S}$ is the sum of the total quark spin $\mathbf{S} = \sum_i \mathbf{s}_i$ and the orbital angular momentum $\mathbf{L} = \mathbf{l}_\rho + \mathbf{l}_\lambda$, which takes the values $L = |l_\rho - l_\lambda| \dots l_\rho + l_\lambda$. The parity of a given state is $P = (-1)^l$. For example, the ground state is given by

$$\phi_{\text{grd}} = \left(\frac{m\omega_0}{\pi} \right)^{\frac{3}{2}} \exp \left[-\frac{m\omega_0}{2} (\boldsymbol{\rho}^2 + \boldsymbol{\lambda}^2) \right]. \quad (4.1.38)$$

Combined with $SU(6)$ for spin and flavor, where the quarks are assigned to the fundamental $\mathbf{6}$ representation ($u \uparrow, d \uparrow, s \uparrow, u \downarrow, d \downarrow, s \downarrow$), the harmonic oscillator potential results in the band structure discussed earlier around Table 3.5.

A pure oscillator spectrum with $E_N \propto N$ does not describe the baryon spectrum particularly well, since all states in Table 3.5 with the same N would be mass-degenerate. However, it provides a useful calculational basis for further refinements. For example, one can solve the Schrödinger equation in terms of the oscillator potentials and evaluate anharmonic parts of the potential perturbatively in the oscillator basis.

The prototype of a nonrelativistic quark model is the one by **de Rujula, Georgi and Glashow** from 1975. It employs the Breit-Fermi interaction for one-gluon exchange, which breaks $SU(3)_f$ symmetry due to the different light and strange-quark masses and $SU(2)$ spin symmetry due to its spin-dependent interactions. For ground states the important part is the spin-spin contact interaction, which leads to mass formulas of the form

$$M = \sum_i m_i + \frac{2\alpha_s}{3} \frac{8\pi}{3} \left\langle \delta^3(\mathbf{r}) \sum_{i < j} \frac{\mathbf{s}_i \cdot \mathbf{s}_j}{m_i m_j} \right\rangle. \quad (4.1.39)$$

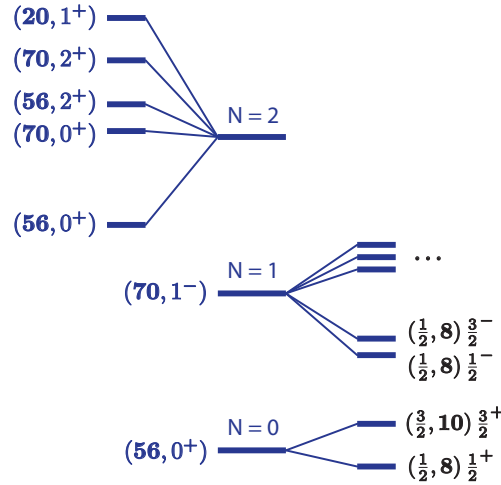


FIG. 4.2: Baryon mass splittings in the quark model (same notation as in Table 3.5).

In the limit of $SU(3)_f$ symmetry, the mass splitting from the spin-spin interactions can be determined like in Eq. (4.1.16) from $\mathbf{S} = \mathbf{s}_1 + \mathbf{s}_2 + \mathbf{s}_3$ and therefore

$$\sum_{i < j} \langle \mathbf{s}_i \cdot \mathbf{s}_j \rangle = \frac{S(S+1) - \frac{9}{4}}{2} = \pm \frac{3}{4}. \quad (4.1.40)$$

In this way, the ground-state octet baryons satisfy

$$\begin{aligned} M_N &= M_0 + 3m_n, \\ M_\Lambda &= M_\Sigma = M_0 + 2m_n + m_s, \\ M_\Xi &= M_0 + m_n + 2m_s \end{aligned} \quad (4.1.41)$$

where $m_{n,s}$ are the light and strange quark masses and the decuplet masses only differ by the spin splitting. This reproduces phenomenological **Gell-Mann-Okubo relations** such as $M_{\Sigma^*} - M_\Sigma = M_{\Xi^*} - M_\Xi$, whereas for $m_s \neq m_n$ they pick up corrections.

Another influential quark model has been the **Isgur-Karl model**, which implements a harmonic oscillator potential together with an anharmonic perturbation, a confinement part and a hyperfine interaction. The spin-orbit interactions are neglected since their inclusion would spoil the agreement with the spectrum (the resulting mass splittings tend to be too large). This leads to the pattern in Fig. 4.2, where the splittings between the $N = 0$ states is due to the spin-spin contact term, the splittings in the $N = 1$ band come from the spin-spin contact and tensor terms, and the splittings between the $SU(6)$ multiplets in the $N = 2$ band are due to the anharmonic perturbation. The Isgur-Karl model provides a good description of the light and strange baryon spectrum but also predicts more states than observed; however, it also predicts that most of those unobserved states are weakly coupled to the πN channel.

Many quark potential models have been constructed following up on the early developments, including relativized models, flux-tube and instanton-induced models, Goldstone-boson exchange models, diquark-based models and more.¹

¹For a review, see Capstick and Roberts, Prog. Part. Nucl. Phys. 45 (2000) 241, [nuc1-th/0008028](https://arxiv.org/abs/nuc1-th/0008028).

4.2 Spontaneous chiral symmetry breaking

In the quark model, the ‘constituent-quark masses’ enter as input parameters which cannot be further explained. How do they come about in QCD? This ties into the question of **mass generation**: if the light up and down quarks in the QCD Lagrangian have masses of a few MeV, how is it possible that the masses of the proton and other hadrons are of the order of 1 GeV? In fact, we could even set $m_u = m_d = 0$ and we would still get a proton mass not far from its physical value, so the overwhelming contribution to its mass must be generated in QCD.

Earlier we have seen that regularization introduces a scale. Without a scale in the theory, from a massless Lagrangian we would expect all hadrons to be massless as well, so the anomalous breaking of scale invariance is a necessary component. The other component is **spontaneous chiral symmetry breaking (S χ SB)**. We will see that this mechanism plays a quite important role in the light hadron spectrum: it is not only responsible for the Goldstone nature of the pions, but also the origin of the constituent-quark masses which produce the typical hadronic scales of ~ 1 GeV.

Spontaneous symmetry breaking. Let us go back to the beginning of Sec. (3.1) and start with some general considerations. Suppose ϕ_i are a set of (potentially composite) fields which transform nontrivially under some continuous global symmetry group G :

$$\phi'_i = D_{ij}(\varepsilon) \phi_j = \left(e^{i \sum_a \varepsilon_a \mathbf{t}_a} \right)_{ij} \phi_j = \phi_i + \delta \phi_i, \quad \delta \phi_i = i \sum_a \varepsilon_a (\mathbf{t}_a)_{ij} \phi_j, \quad (4.2.1)$$

where ε_a are the group parameters, the \mathbf{t}_a are the generators of the Lie algebra of G in the representation to which the ϕ_i belong, and $D(\varepsilon)$ are the representation matrices. The quantum-field theoretical version of this relation is

$$e^{i \sum_a \varepsilon_a Q_a} \phi_i e^{-i \sum_a \varepsilon_a Q_a} = D_{ij}^{-1}(\varepsilon) \phi_j. \quad (4.2.2)$$

where the charge operators Q_a form a representation of the algebra on the state space. Expanding the exponentials on both sides, we obtain for each ε_a :

$$[Q_a, \phi_i] = -(\mathbf{t}_a)_{ij} \phi_j. \quad (4.2.3)$$

We have encountered examples of this relation earlier:

- Eq. (3.1.71) for the quark field operators under a vector transformation;
- Eq. (3.1.68) for the collection of *composite* fields $\{S(x), S_a(x), P_a(x)\}$ under axial transformations (that this is a manifestation of the same relation will become clear in the discussion of the sigma model in Sec. 4.4.1).

If the symmetry group leaves the vacuum invariant, $e^{i\varepsilon_a Q_a} |0\rangle = |0\rangle$, then all generators Q_a must annihilate the vacuum: $Q_a |0\rangle = 0$. Hence, when we take the vacuum expectation value (VEV) of Eq. (4.2.2) we get

$$\langle 0 | \phi_i | 0 \rangle = D_{ij}^{-1}(\varepsilon) \langle 0 | \phi_j | 0 \rangle. \quad (4.2.4)$$

If the ϕ_i had been invariant under G to begin with, this relation would be trivially satisfied. Because they transform nontrivially, $D_{ij}^{-1}(\varepsilon)$ is not the identity matrix for all ε_a and so these vacuum expectation values must vanish:

$$Q_a |0\rangle = 0 \quad \Rightarrow \quad \langle 0 | \phi_i | 0 \rangle = 0. \quad (4.2.5)$$

This is the '**Wigner-Weyl**' realization of a symmetry, which simply means that the symmetry is unbroken.

On the other hand, if an operator that is not invariant under G develops a nonzero vacuum expectation value $\langle 0 | \phi_i | 0 \rangle \neq 0$, then the symmetry G is spontaneously broken. This is the '**Nambu-Goldstone** realization' of the symmetry, in which case we find

$$\langle 0 | [Q_a, \phi_i] | 0 \rangle = -(\mathbf{t}_a)_{ij} \langle 0 | \phi_j | 0 \rangle \neq 0. \quad (4.2.6)$$

Then we would conclude that the charges do not annihilate the vacuum: $Q_a |0\rangle \neq 0$. Since the symmetry is classically realized, they still commute with the Hamiltonian and we have found another energy-degenerate vacuum:

$$Q_a |0\rangle = |\eta\rangle \neq 0, \quad H|0\rangle = 0 \quad \Rightarrow \quad H|\eta\rangle = HQ_a|0\rangle = Q_a H|0\rangle = 0. \quad (4.2.7)$$

Unfortunately we have to be careful with these statements because in the case of spontaneous symmetry breaking the charges are not well defined. $|\eta\rangle$ is not a normalizable state, which we can see from using the definition of the charge (3.1.5) together with translation invariance:

$$\langle \eta | \eta \rangle = \langle 0 | Q_a^2 | 0 \rangle = \int d^3x \int d^3y \langle 0 | j_a^0(x) j_a^0(y) | 0 \rangle = \infty. \quad (4.2.8)$$

Fortunately, *commutators* involving the charges are still well-defined, so when discussing spontaneous symmetry breaking we should start from Eq. (4.2.6). To prove the **Goldstone theorem**, we insert the completeness relation (2.2.5) in that equation and follow the same steps as when deriving the spectral representation:

$$\begin{aligned} \langle 0 | [Q_a(x_0), \phi(0)] | 0 \rangle &= \int d^3x \langle 0 | [j_a^0(x), \phi(0)] | 0 \rangle \\ &= \sum_{\lambda} \int \frac{d^3p}{2E_p} \frac{i}{(2\pi)^3} \int d^3x (R_{a\lambda}(\mathbf{p}) e^{-ipx} + R_{a\lambda}^*(\mathbf{p}) e^{ipx}) \\ &= \sum_{\lambda} \frac{i}{2m_{\lambda}} (R_{a\lambda}(\mathbf{0}) e^{-im_{\lambda}x_0} + R_{a\lambda}^*(\mathbf{0}) e^{im_{\lambda}x_0}) \\ &= \sum_{\lambda} \frac{i}{m_{\lambda}} \operatorname{Re} \{ R_{a\lambda}(\mathbf{0}) e^{-im_{\lambda}x_0} \} \stackrel{!}{=} \text{const.} \end{aligned} \quad (4.2.9)$$

In going from the first to the second row we used translation invariance (2.2.11) to factor out the phases $e^{\pm ipx}$, and we defined

$$\langle 0 | j_a^0(0) | \lambda \rangle \langle \lambda | \phi(0) | 0 \rangle = iR_{a\lambda}(\mathbf{p}). \quad (4.2.10)$$

The integral over d^3x produces $\delta^3(\mathbf{p})$, so that $p_0 = E_p = (\mathbf{p}^2 + m_{\lambda}^2)^{1/2}$ becomes m_{λ} . By translation invariance, the VEV $\langle 0 | \phi_j(x) | 0 \rangle = \langle 0 | \phi_j(0) | 0 \rangle$ on the right-hand side of

Eq. (4.2.9) must also be independent of x_0 , whereas the left-hand side still contains x_0 in the exponential. Thus, if the VEV is nonzero, the above requirement can only be met if for each charge Q_a there is a mode $|\lambda\rangle$ with

$$m_\lambda = 0 \quad \text{and} \quad \frac{R_{a\lambda}(\mathbf{0})}{m_\lambda} \neq 0. \quad (4.2.11)$$

Thus, for each generator that does not leave the vacuum invariant there is a **massless Goldstone boson**, which has a non-zero vacuum overlap $\langle 0|j_a^0(0)|\lambda\rangle$ and $\langle 0|\phi(0)|\lambda\rangle$. The other modes with $m_\lambda \neq 0$ (excited states) must have $R_{a\lambda}(\mathbf{0}) = 0$.

S χ SB in QCD and chiral condensate. How does spontaneous breaking of chiral symmetry come about in QCD? The Goldstone theorem does not tell us *why* a non-zero VEV appears, it only says that *if* there is a non-zero VEV, we must have massless particles in the spectrum. Therefore, we must first identify potential candidates for vacuum condensates that break chiral symmetry. From Eq. (4.2.10) we already see that the ‘field’ $\phi(0)$ will have to be a composite operator, since only those produce overlaps with hadronic states.

Let us go back to the quark propagator,

$$S_{\alpha\beta}(x-y) = \langle 0|\mathrm{T}\psi_\alpha(x)\bar{\psi}_\beta(y)|0\rangle, \quad (4.2.12)$$

and contract it with either of the Dirac matrices $\Gamma \in \{\gamma^\mu, \gamma^\mu\gamma_5, \mathbb{1}, i\gamma_5\}$ and flavor matrices $\{\mathbf{t}_a, \mathbb{1}\}$. This gives us the vacuum expectation values of either of the currents in Eq. (3.1.23):

$$-\Gamma_{\beta\alpha}\mathbf{t}_a S_{\alpha\beta}(0) = \langle 0|j_a^\Gamma(0)|0\rangle. \quad (4.2.13)$$

Because of translation invariance, they cannot depend on x and must be (dimensionful) constants. Due to Lorentz and parity invariance these must all be zero, with the only possible exception of the **scalar condensates** which carry the quantum numbers of the vacuum (0^{++}):

$$\begin{aligned} \langle 0|\tilde{S}_a(0)|0\rangle &= \langle 0|\bar{\psi}(0)\mathbf{t}_a\psi(0)|0\rangle, \\ \langle 0|\tilde{S}(0)|0\rangle &= \langle 0|\bar{\psi}(0)\psi(0)|0\rangle =: \langle\bar{\psi}\psi\rangle. \end{aligned} \quad (4.2.14)$$

Here we put a tilde on the scalar densities S and S_a to avoid confusion with the quark propagator. Actually, if $SU(N_f)$ were unbroken, all flavor non-singlet scalar condensates would vanish as well. From Eq. (3.1.57) one can derive

$$[Q_a^V, \tilde{S}_b(x)] = if_{abc}\tilde{S}_c(x), \quad (4.2.15)$$

and since unbroken $SU(N_f)_V$ implies $Q_a^V|0\rangle = 0$, the VEV of this relation vanishes. The singlet condensate is then identical for all flavors:

$$\langle 0|\tilde{S}_a(0)|0\rangle = 0 \quad \Rightarrow \quad \langle\bar{u}u\rangle - \langle\bar{d}d\rangle = 0, \quad \langle\bar{u}u\rangle + \langle\bar{d}d\rangle - 2\langle\bar{s}s\rangle = 0, \quad (4.2.16)$$

and therefore $\langle\bar{u}u\rangle = \langle\bar{d}d\rangle = \langle\bar{s}s\rangle = \langle\bar{\psi}\psi\rangle/3$.

Finally, in the discussion below Eq. (3.1.47) we saw that a scalar bilinear of quarks breaks chiral symmetry, i.e., it breaks $SU(N_f)_A \times U(1)_A$. Thus we have a potential candidate for a condensate that breaks chiral symmetry. In a chirally symmetric theory of massless quarks, this quantity should vanish — but does it?

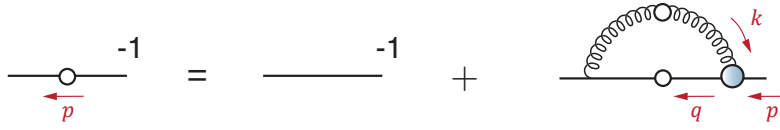


FIG. 4.3: Quark DSE.

Quark mass function. Since the quark condensate is the trace of the quark propagator, let us have a closer look at the propagator itself. For the following discussion we temporarily switch to **Euclidean conventions** to avoid cumbersome factors of $i\epsilon$. The transcription rules between Minkowski and Euclidean space can be found in Appendix C, but all we need to remember in the following is $p^2 = -p_E^2$ and the quark propagator in Euclidean conventions (we drop the subscript E):

$$S(p) = \frac{1}{A(p^2)} \frac{-i\not{p} + M(p^2)}{p^2 + M(p^2)^2}. \quad (4.2.17)$$

Recall the quark Dyson-Schwinger equation (DSE) in Fig. 4.3,

$$S^{-1}(p) = A(p^2) (i\not{p} + M(p^2)) = Z_2(i\not{p} + Z_m m) + \Sigma(p), \quad (4.2.18)$$

where $M(p^2)$ is the **quark mass function** and $\Sigma(p)$ the self-energy incorporating the quantum effects, which in one-loop perturbation theory reduces to Eq. (2.3.46). To obtain the quark condensate for a particular flavor, we need to take the Dirac and color trace of the quark propagator, which singles out the term with $M(p^2)$ and gives a factor $4N_c$. In addition, setting $x - y = 0$ corresponds to an integration over $d^4p/(2\pi)^4$ in momentum space, which from Eq. (C.0.32) entails

$$\int \frac{d^4p}{(2\pi)^4} f(p^2) = \frac{1}{(4\pi)^2} \int dp^2 p^2 f(p^2). \quad (4.2.19)$$

Thus we arrive at²

$$-\langle \bar{u}u \rangle = N_c \int \frac{d^4p}{(2\pi)^4} \text{Tr} S(p) = \frac{N_c}{(2\pi)^2} \int dp^2 \frac{p^2}{A(p^2)} \frac{M(p^2)}{p^2 + M(p^2)^2}. \quad (4.2.20)$$

The functions $M(p^2)$ and $A(p^2)$ should be positive for spacelike momenta $p^2 \geq 0$. Since for a chirally symmetric Lagrangian ($m = 0$) we expect the condensate to vanish, and because the condensate is proportional to the integrated quark mass function, this means that the quark mass function should be zero for all p^2 . The resulting quark propagator is then chirally symmetric: $\{\gamma_5, S(p)\} = 0$.

Indeed, this is what happens when we evaluate the self-energy order by order in perturbation theory, see Fig. 4.4. In the massless theory, the tree-level propagator is proportional to \not{p} and the tree-level vertex is proportional to γ^μ , so they both contain one γ matrix. However, *every* possible perturbative diagram has an odd number of γ matrices whose Dirac trace vanishes. In this way, we can never generate a mass function and $M(p^2) = 0$ to all orders in perturbation theory!

²Strictly speaking we should also attach a factor $Z_2 Z_m$, since the condensate renormalizes like the mass term in the Lagrangian and the product $m\langle \bar{\psi}\psi \rangle$ is renormalization-point independent.

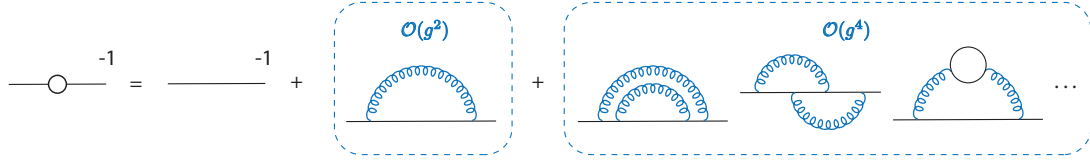


FIG. 4.4: Perturbative expansion of the inverse quark propagator. In massless QCD, each Feynman diagram contains an odd number of γ matrices whose trace vanishes.

On the other hand, we can generate a non-zero mass function **nonperturbatively**, which can already be illustrated in simple DSE models. Apart from renormalization constants, the exact expression for the self-energy is

$$\Sigma(p) = g^2 C_F \int \frac{d^4 k}{(2\pi)^4} \gamma^\mu S(q) \Gamma^\nu(q, p) D^{\mu\nu}(k), \quad (4.2.21)$$

which depends on the full gluon propagator and quark-gluon vertex. Let us assume that the quark-gluon vertex remains at tree-level, so that only the internal quark and gluon propagators are dressed (‘rainbow truncation’). In Feynman gauge the gluon is diagonal in its Lorentz indices, so we can write the self-energy as

$$\Sigma(p) = \int d^4 k \gamma^\mu S(p+k) \gamma^\mu D(k), \quad (4.2.22)$$

where $D(k)$ is proportional to the gluon propagator and absorbs all prefactors. Thus, if we can find a good ansatz for $D(k)$, we can solve the Dyson-Schwinger equation $S^{-1}(p) = i\not{p} + m + \Sigma(p)$ for the quark propagator. $D(k)$ must be a scalar function of the gluon momentum k^2 with mass dimension -2 . At large k^2 it should be proportional to QCD’s running coupling, $D(k^2) \propto \alpha_s(k^2)/k^2$, because this is where quarks and gluons become asymptotically free. In the following we employ two rather crude models: one where the gluon propagator is localized in momentum space and another one where it is localized in coordinate space.

Munczek-Nemirovsky model. In this case the gluon propagator is just a δ -function peaked at the origin, equipped with some mass scale Λ :

$$D(k) = \Lambda^2 \delta^4(k). \quad (4.2.23)$$

Here the self-energy can be integrated analytically, so the model is UV-finite and instead of imposing renormalization conditions we can set all renormalization constants to 1 (as we already did above). The result is

$$\Sigma(p) = \Lambda^2 \gamma^\mu S(p) \gamma^\mu = \Lambda^2 \frac{\gamma^\mu (-i\not{p} + M) \gamma^\mu}{(p^2 + M^2) A} = 2\Lambda^2 \frac{i\not{p} + 2M}{(p^2 + M^2) A}, \quad (4.2.24)$$

where we suppressed the momentum dependencies of $A(p^2)$ and $M(p^2)$ to avoid clutter. Putting this back into the DSE leads to selfconsistent algebraic equations for the two quark dressing functions:

$$A = 1 + \frac{2\Lambda^2}{(p^2 + M^2) A}, \quad AM = m + 2M \frac{2\Lambda^2}{(p^2 + M^2) A}. \quad (4.2.25)$$

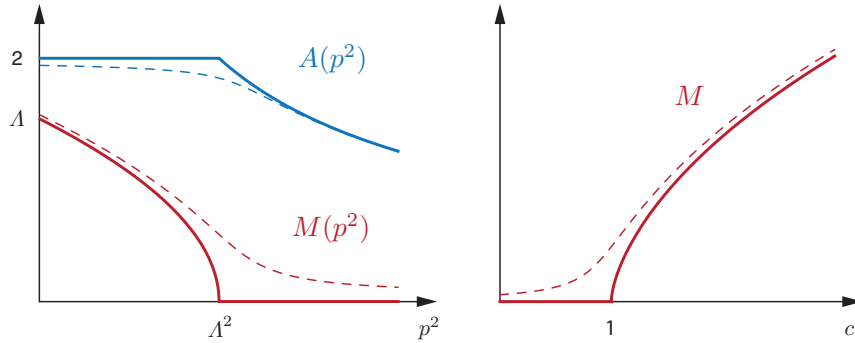


FIG. 4.5: Quark propagator in the Munczek-Nemirovsky model (left) and NJL model (right). The solid lines are the results in the chiral limit and the dashed lines exemplify the solutions for $m \neq 0$.

In the chiral limit ($m = 0$), we see from the second equation that the trivial solution $M = 0$ is always possible. It leads to a quadratic equation for A whose result is

$$M(p^2) = 0, \quad A(p^2) = \frac{1}{2} \left(1 + \sqrt{1 + 8 \Lambda^2 / p^2} \right). \quad (4.2.26)$$

It has the correct perturbative behavior for $p^2 \rightarrow \infty$, namely $M = 0$ and $A \rightarrow 1$, so it reverts the quark propagator back to its tree-level form and preserves chiral symmetry. On the other hand, $A(p^2)$ diverges for $p^2 \rightarrow 0$, so this cannot be the whole story. Indeed there is another solution with $M \neq 0$:

$$M(p^2) = \sqrt{\Lambda^2 - p^2}, \quad A(p^2) = 2. \quad (4.2.27)$$

It breaks chiral symmetry and is finite in the infrared. Both solutions are connected at the point $p^2 = \Lambda^2$, see Fig. 4.5. This is the typical shape of an order parameter of a spontaneously broken symmetry, like the magnetization in a ferromagnet when plotted over temperature. If we switch on a quark mass $m \neq 0$, the curves become smooth (in the ferromagnet this corresponds to a background magnetic field).

Despite the simplicity of the model, these results already capture the essence of more realistic DSE calculations. At large momenta, $M(p^2)$ is the renormalized current-quark mass in the Lagrangian. When lowering the momentum, the onset of the non-symmetric phase sets in at some typical hadronic scale Λ , below which a mass is spontaneously generated. The mass function in the infrared defines the quark mass at low momenta that is relevant for hadrons, so it can be viewed as a ‘constituent-quark’ mass scale. Thus, the quark mass function encodes the transition from a current quark at large momenta to a constituent quark in the infrared, and this effect cannot be described in QCD perturbation theory.

If we insert the combined solution in Eq. (4.2.20), the resulting quark condensate in the chiral limit becomes

$$-\langle \bar{u}u \rangle = \frac{N_c}{(2\pi)^2} \int_0^{\Lambda^2} dp^2 p^2 \frac{\sqrt{\Lambda^2 - p^2}}{2\Lambda^2} = \frac{2}{15} \frac{N_c}{(2\pi)^2} \Lambda^3. \quad (4.2.28)$$

With $\Lambda = 1$ GeV we even get a reasonable numerical value: $-\langle \bar{u}u \rangle \sim (220 \text{ MeV})^3$.

NJL model/contact interaction. The shortcoming of the Munczek-Nemirovsky model is that it does not have a *critical* coupling: a non-trivial solution for the quark mass function and thus a chiral condensate exist for any $\Lambda > 0$. The gluon propagator in Eq. (4.2.23) is localized in momentum space because of the δ -function. We could take the extreme opposite and localize it in coordinate space, which results in an effective four-fermi contact interaction between two quarks where the gluon shrinks to a point and is integrated out. This is the **NJL model** (Nambu, Jona-Lasinio), where the momentum dependence of the gluon is simply a constant:

$$D(k) = \frac{1}{(2\pi)^2} \frac{c}{\Lambda^2}. \quad (4.2.29)$$

In this case it is more convenient to integrate over the quark momentum $q = p - k$ instead of k in (4.2.22). However, now the self-energy integral must be regulated because it is divergent. We could impose a sharp cutoff at $q^2 = \Lambda^2$, so that the gluon propagator is a constant up to some scale Λ and vanishes above. As a consequence, the integrand no longer depends on the external momentum p ,

$$\Sigma(p) = \frac{1}{(2\pi)^2} \frac{c}{\Lambda^2} \int d^4q \gamma^\mu S(q) \gamma^\mu = \frac{1}{(2\pi)^2} \frac{c}{\Lambda^2} \int d^4q \frac{2}{A(q^2)} \frac{i\not{q} + 2M(q^2)}{q^2 + M(q^2)^2}, \quad (4.2.30)$$

which means that $\Sigma(p)$ is constant and therefore A and M will be constants as well. The integral over q , which is the self-energy contribution to A , vanishes and we get $A = 1$. The equation for M becomes:

$$M = m + cM \int_0^1 dy \frac{y}{y+a} = m + cM \left[1 - a \ln\left(1 + \frac{1}{a}\right)\right] = m + cM f(a), \quad (4.2.31)$$

where we set $y = q^2/\Lambda^2$ and $a = M^2/\Lambda^2$. The function $f(a)$ satisfies $f(a) \leq 1$ and $f(0) = 1$. In the chiral limit we obtain the algebraic equation

$$M = cM f(a), \quad (4.2.32)$$

which returns again the trivial solution $M = 0$, but also a nontrivial solution where M as a function of c is determined from the equation $f(a) = 1/c$. Because $f(a) \leq 1$, this solution only occurs above a critical value $c \geq 1$.

The result is shown in Fig. 4.5: In contrast to the previous case, the dynamical quark mass M is no longer a mass *function* that depends on the momentum but just a constant; however, it depends on the coupling strength c and vanishes for $c < 1$. Above that value, chiral symmetry is spontaneously broken. If we plug the result into the chiral condensate (4.2.20) using the same cutoff, we obtain the same form as in Eq. (4.2.28) except that the prefactor $2/15$ is replaced by $M(c)/(c\Lambda)$, which also vanishes for $c < 1$.

In general, the gluon propagator is neither a δ -function nor a constant, and the spontaneous breaking of chiral symmetry will not only generate a mass term for the quark propagator but also chirally asymmetric terms for other correlation functions with quark and antiquark legs such as the quark-gluon vertex. Nevertheless, both models encode general features:

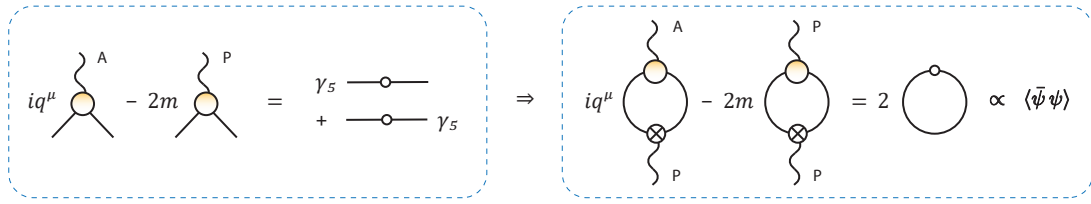


FIG. 4.6: Axialvector WTI for the three-point functions (left) and current correlators (right).

- Implementing a scale Λ was necessary to make them work. If we replace Λ^2 by k^2 in Eq. (4.2.23), the self-energy vanishes. In the NJL model, Λ is the regulator which cannot be removed. The quark mass function and other dimensionful quantities such as the chiral condensate, and eventually the masses of hadrons, are then proportional to this scale, so that S_χ SB can be viewed as the **mass generation** mechanism in the fermion sector of QCD.
- S_χ SB is a *critical* phenomenon: if the combined strength from the gluon propagator and quark-gluon vertex (the 'effective' running coupling) exceeds a critical value, a quark mass is generated dynamically; otherwise we remain with the chirally symmetric solution.
- In contrast to effective theories of QCD, where the terms that trigger S_χ SB already appear in the Lagrangian, the QCD Lagrangian tells us nothing about whether chiral symmetry is preserved at the quantum level or not. Its spontaneous breaking is a purely dynamical effect induced by the strong gluonic interactions, hence the name **dynamical chiral symmetry breaking (DCSB)**.

Gell-Mann-Oakes-Renner relation. Now let us return to the Goldstone theorem. We have explored the origin of S_χ SB and identified its order parameters: the scalar quark condensate or, equivalently, the quark mass function. Hence, any other quantity that depends on the mass function (and vanishes if the mass function does) will break chiral symmetry as well. In Eq. (3.1.143) we found that, as a simple consequence of the PCAC relation, either a pseudoscalar meson's mass or its electroweak decay constant must vanish in the chiral limit:

$$f_\lambda m_\lambda^2 = 2m r_\lambda \xrightarrow{m=0} 0. \quad (4.2.33)$$

Therefore, if we can show that the pion decay constant f_π is also proportional to the mass function and comes about by S_χ SB, we must have massless pions.

The right place to look for such a relation is the axialvector WTI in (3.1.81), which is pictorially shown in Fig. 4.6. On its l.h.s. we have the difference of the G_A and G_P three-point functions; the r.h.s. is the sum of quark propagators multiplied with γ_5 . If we multiply again with γ_5 and take the trace, we get a difference of AP and PP current correlators on the left and the quark condensate on the right. When inserting the completeness relation, both terms contain pseudoscalar poles *only*, where the residues depend on f_λ and r_λ as given in Eq. (3.1.144). Moreover, the hadronic poles must cancel out between G_A and G_P because the quark propagator does not have such poles. In this way we should be able to establish a relation between f_π and $\langle \bar{\psi}\psi \rangle$.

Let us start directly from the WTI (3.1.72) for the AP current correlator:

$$\partial_\mu^x \langle 0 | T A_a^\mu(x) P_b(0) | 0 \rangle - 2m \langle 0 | T P_a(x) P_b(0) | 0 \rangle = \delta(x^0) \langle 0 | [A_a^0(x), P_b(0)] | 0 \rangle. \quad (4.2.34)$$

We already inserted the PCAC relation for the PP term. If we integrate over d^4x on the r.h.s., we obtain the vacuum expectation value of the commutator that we derived earlier in Eq. (3.1.68),

$$\langle 0 | [Q_a^A, P_b(0)] | 0 \rangle = -i \langle 0 | \left[\frac{\delta_{ab}}{N_f} S(0) + d_{abc} S_c(0) \right] | 0 \rangle = -i \frac{\delta_{ab}}{N_f} \langle \bar{\psi} \psi \rangle, \quad (4.2.35)$$

where only the singlet condensate survives in the limit of exact $SU(N_f)_V$. This is the representative of the generic equation (4.2.6): since the condensate which is not invariant under axial symmetries is the scalar condensate and the respective charges are the axial charges, the corresponding field φ_i must be the pseudoscalar density. For the l.h.s. in Eq. (4.2.34), we insert the spectral decompositions of the AP and PP current correlators from (3.1.144) and (3.1.145) and integrate over x . This means taking the limit $p \rightarrow 0$:

$$\lim_{p \rightarrow 0} \sum_\lambda \frac{p^2 f_\lambda - 2m r_\lambda}{p^2 - m_\lambda^2 + i\varepsilon} i r_\lambda \delta_{ab} = \sum_\lambda i r_\lambda f_\lambda \delta_{ab} \stackrel{!}{=} -i \frac{\delta_{ab}}{N_f} \langle \bar{\psi} \psi \rangle, \quad (4.2.36)$$

where we used the relation $f_\lambda m_\lambda^2 = 2m r_\lambda$ in the second equality. The poles cancel indeed, and we arrive at the result that if chiral symmetry is realized and the quark condensate vanishes, all combinations $r_\lambda f_\lambda$ must vanish as well; if it is spontaneously broken, there is at least one mode where both r_λ and f_λ are nonzero. Since $f_\lambda \neq 0$ in that case, we must have $m_\lambda \rightarrow 0$, i.e., a massless Goldstone boson.

Each $|\lambda\rangle$ corresponds to one of the generators, so there is a massless Goldstone boson for each generator \mathfrak{t}_a (for three flavors with $SU(3)_A \times U(1)_A$ this means a pseudoscalar octet and a singlet). In turn, the decay constants f_λ must vanish for the remaining excited states with $m_\lambda \neq 0$, so we can remove the sum in the equation above and write

$$r_{\lambda_0} f_{\lambda_0} = -\frac{\langle \bar{\psi} \psi \rangle}{N_f}, \quad (4.2.37)$$

where $|\lambda_0\rangle$ is the ground state in each channel. If we substitute r_{λ_0} by the condensate and insert it in Eq. (4.2.33), we obtain the **Gell-Mann-Oakes-Renner (GMOR)** relation,

$$f_{\lambda_0}^2 m_{\lambda_0}^2 = -2m \frac{\langle \bar{\psi} \psi \rangle}{N_f}, \quad (4.2.38)$$

which is valid for each member of the lowest-lying pseudoscalar octet and singlet. (In the singlet case it only holds if we ignore the anomaly.)

All in all, S_χ SB has important consequences for the light hadron spectrum: It generates a large dynamical quark mass function, which translates to a large mass contribution for hadrons made of quarks and antiquarks even in the chiral limit. The pseudoscalar meson masses, on the other hand, behave like $m_{\text{PS}}^2 \propto m_q$ and vanish for $m_q \rightarrow 0$ as shown in Fig. 4.7.

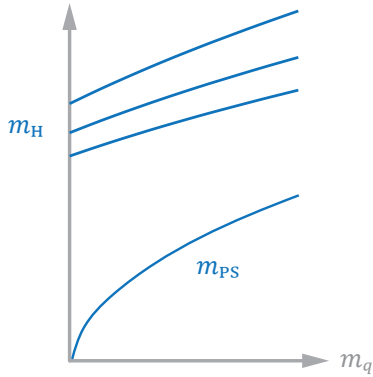


FIG. 4.7: Generic dependence of hadron masses on the current-quark mass

$SU(N_f)_V$ breaking. So far we have assumed that all quark masses are equal, $m_u = m_d = m_s$. In the case of $SU(3)_V$ breaking, we have to go back to the general PCAC relation (3.1.38) and evaluate the anticommutators, and also keep the d_{abc} terms in Eq. (4.2.35). In this case the form of the GMOR relation remains the same for each generator with index a if we replace the quark mass m by

$$\begin{aligned}
 a = 1, 2, 3: & \quad \frac{1}{2} (m_u + m_d), \\
 a = 4, 5: & \quad \frac{1}{2} (m_u + m_s), \\
 a = 6, 7: & \quad \frac{1}{2} (m_d + m_s), \\
 a = 8: & \quad \frac{1}{6} (m_u + m_d + 4m_s), \\
 a = 0: & \quad \frac{1}{3} (m_u + m_d + m_s),
 \end{aligned} \tag{4.2.39}$$

and the condensate accordingly:

$$\frac{\langle \bar{\psi}\psi \rangle}{3} \longrightarrow \frac{\langle \bar{u}u + \bar{d}d \rangle}{2} \quad (a = 1, 2, 3), \quad \frac{\langle \bar{u}u + \bar{s}s \rangle}{2} \quad (a = 4, 5), \quad \text{etc.} \tag{4.2.40}$$

Then we get for the pions and kaons:

$$f_\pi^2 m_\pi^2 = -\frac{m_u + m_d}{2} \langle \bar{u}u + \bar{d}d \rangle, \quad f_K^2 m_K^2 = -\frac{m_u + m_s}{2} \langle \bar{u}u + \bar{s}s \rangle. \tag{4.2.41}$$

Inserting the experimental values³ $f_\pi \approx 92$ MeV, $m_\pi \approx 140$ MeV and assuming an average quark mass $m_u = m_d = 3.5$ MeV yields $\langle \bar{u}u \rangle = \langle \bar{d}d \rangle \approx -(280 \text{ MeV})^3$. The same estimate for kaons ($f_K \approx 110$ MeV, $m_K \approx 494$ MeV, $m_s \approx 120$ MeV) gives us $\langle \bar{s}s \rangle \approx -(290 \text{ MeV})^3$. The renormalized quark masses and condensates are renormalization-point and -scheme dependent; the values quoted here are consistent with lattice QCD results⁴ obtained in an $\overline{\text{MS}}$ scheme at $\mu = 2$ GeV.

Strictly speaking, the GMOR relation as it stands is only valid in the chiral limit because the quark condensate is only well-defined for $m = 0$. We can see this from its definition (4.2.20) as the momentum integral of the quark mass function: In the chiral limit, $M(p^2 \rightarrow \infty)$ vanishes like $1/p^2$, so the integral only diverges logarithmically and is renormalized by $Z_2 Z_m$. For $m \neq 0$, the one-loop result in Eq. (2.3.88) entails that the mass function vanishes logarithmically and therefore the integral diverges quadratically. In this case, $f_\lambda m_\lambda^2 = 2mr_\lambda$ can be viewed as a generalized GMOR relation since the quantities f_λ and r_λ are well-defined for all quark masses. In principle, they can be used to define the quark condensate from a pseudoscalar meson's Bethe-Salpeter wave function, namely as the chiral limit of the combination $r_{\lambda_0} f_{\lambda_0}$ via Eq. (4.2.37).

³The decay constants are sometimes defined with a factor $\sqrt{2}$, in which case $f_\pi \approx 130$ MeV.

⁴McNeile et al., Phys. Rev. D87 (2013), 034503. [arXiv:1211.6577](https://arxiv.org/abs/1211.6577).

4.3 $U(1)_A$ anomaly

We have seen that spontaneous chiral symmetry breaking should affect all axial symmetries including the flavor-singlet $U(1)_A$. The fact that there is no good candidate for a flavor-singlet (pseudo-) Goldstone boson in the spectrum is related to the anomalous $U(1)_A$ breaking. **Anomalies** are symmetries of classical Lagrangians that are broken at the quantum level. They arise when regularization destroys a symmetry and there is no regulator choice that can preserve it. Since the symmetry is lost, there is no Goldstone boson because the quantum corrections generate a mass for that mode.

Anomalies are again a typical feature of **axial symmetries**. In contrast to spontaneous symmetry breaking, where the symmetry is lost due to dynamical effects, anomalies have their origin in short-distance singularities of the currents $A_a^\mu = \bar{\psi} \gamma^\mu \gamma_5 \mathbf{t}_a \psi$ and $A^\mu = \bar{\psi} \gamma^\mu \gamma_5 \psi$. These are composite operators at the same space-time point which are potentially divergent and have to be regularized. In principle, the problem would also affect vector currents, but in that case it is possible to find appropriate regularization prescriptions that leave their symmetry intact. Vector symmetries are related to conserved charges (color charge, electromagnetic charge, flavor charges, etc.). If they were broken at the quantum level, we would not only lose charge conservation but also gauge symmetry, and the theory would become nonrenormalizable and inconsistent. In this sense, global axial symmetries are ‘less important’ and the fact that they produce anomalies is not a serious problem for the theory. (Except when they are also promoted to gauge symmetries: if a gauge symmetry is broken anomalously, then one needs anomaly cancellations between different sectors of the theory.)

In the following we will see that

- QCD only leads to an anomalous $U(1)_A$ breaking, which has observable consequences for the η and η' masses, whereas
- QED also induces an anomalous $SU(N_f)_A$ breaking, which can be observed in the $\pi^0 \rightarrow \gamma\gamma$ decay.

We already wrote down the basic relations that characterize the anomalous $U(1)_A$ breaking in QCD. We have anticipated in Eq. (3.1.54) that the divergence of the axialvector singlet current picks up an anomalous contribution

$$\partial_\mu A^\mu = 2i \bar{\psi} \mathbf{M} \gamma_5 \psi + N_f \mathcal{Q}(x), \quad (4.3.1)$$

where $\mathcal{Q}(x)$ is the topological charge density that we encountered in Section 2.1:

$$\mathcal{Q}(x) = \frac{g^2}{8\pi^2} \text{Tr} \{ \tilde{F}_{\mu\nu} F^{\mu\nu} \}, \quad \tilde{F}^{\mu\nu} = \frac{1}{2} \varepsilon^{\mu\nu\alpha\beta} F_{\alpha\beta}. \quad (4.3.2)$$

The derived relation (3.2.42) entails that the mass of the η_0 does not vanish in the chiral limit, so there is no flavor-singlet Goldstone boson:

$$f_{\eta_0} m_{\eta_0}^2 = 2 \frac{m_u + m_d + m_s}{3} r_{\eta_0} + \frac{g^2 N_f}{(4\pi)^2} \langle 0 | \tilde{F}_a^{\mu\nu}(0) F_{\mu\nu}^a(0) | \eta_0 \rangle. \quad (4.3.3)$$

Anomalies from the path integral. To see how the anomalous term comes about, suppose we start from an action $S[\psi, \bar{\psi}]$ that is invariant under global $U(1)_A$ transformations, e.g. the fermionic part of the Lagrangian for massless quarks:

$$\mathcal{L} = \bar{\psi} i \not{\partial} \psi, \quad \psi' = e^{i\varepsilon\gamma_5} \psi, \quad \bar{\psi}' = \bar{\psi} e^{i\varepsilon\gamma_5}. \quad (4.3.4)$$

To derive WTIs for global flavor symmetries from the path integral, we need to employ the **background-field method** discussed below Eq. (3.1.108): we add a source term to the action with a background field B^μ , so that the total action that enters in the partition function is *locally* invariant by construction:

$$Z[B] = \int \mathcal{D}[\psi, \bar{\psi}] e^{i(S[\psi, \bar{\psi}] + \tilde{S}[\psi, \bar{\psi}, B])}. \quad (4.3.5)$$

This means we need to impose a $U(1)_A$ transformation behavior for the B^μ field with a covariant derivative:

$$B'_\mu = B_\mu + \frac{1}{g} \partial_\mu \varepsilon, \quad D_\mu = \partial_\mu - igB_\mu \gamma_5. \quad (4.3.6)$$

The resulting Lagrangian

$$\bar{\psi} i \not{D} \psi = \bar{\psi} (i \not{\partial} + g \not{B} \gamma_5) \psi = \mathcal{L} + g A_\mu B^\mu \quad (4.3.7)$$

is locally invariant as desired. As before, A_μ is the $U(1)_A$ axialvector current and not the gluon field (in the following we denote the gluon fields by \mathbf{A}^μ to avoid confusion) and the extra source term in the action is

$$\tilde{S}[\psi, \bar{\psi}, B] = g \int d^4x A_\mu B^\mu, \quad A^\mu = \bar{\psi} \gamma^\mu \gamma_5 \psi. \quad (4.3.8)$$

Because all terms in the path integral are locally gauge invariant, a gauge transformation $\{\psi, \bar{\psi}, B\} \rightarrow \{\psi', \bar{\psi}', B'\}$ does not change the partition function: $Z[B] = Z[B']$. If we then relabel the quark fields back to unprimed ones and work out the transformation of B only, we find

$$Z[B'] = \int \mathcal{D}[\psi, \bar{\psi}] e^{i(S[\psi, \bar{\psi}] + \tilde{S}[\psi, \bar{\psi}, B] + \delta\tilde{S})} = Z[B] \langle e^{i\delta\tilde{S}} \rangle_B \quad (4.3.9)$$

and therefore $\langle \delta\tilde{S} \rangle_B = 0$. Then, with

$$\delta\tilde{S} = \int d^4x A_\mu(x) \partial^\mu \varepsilon(x) = - \int d^4x \varepsilon(x) \partial_\mu A^\mu(x) \quad (4.3.10)$$

we arrive at the usual PCAC relation for the flavor-singlet case:

$$\langle \partial_\mu A^\mu \rangle_B = 0. \quad (4.3.11)$$

This means that current conservation holds inside the vacuum expectation value in the presence of the background field. Note that without it the relation would be trivial: $\langle \partial_\mu A^\mu \rangle = \partial_\mu \langle A^\mu \rangle = 0$ because $\langle A^\mu \rangle = 0$. If we had also included source terms $\eta, \bar{\eta}$ for the quarks, we would have obtained the usual WTIs for the n -point functions like in Eq. (3.1.111).

But where is the *anomalous* term? As always we assumed that the **path integral measure** remains invariant under the transformation. However, for axial transformations this is not necessarily the case. The origin of this behavior is the transformation of the Dirac spinors

$$\psi'(x) = e^{+i\varepsilon\gamma_5} \psi(x), \quad \bar{\psi}'(x) = \bar{\psi}(x) e^{+i\varepsilon\gamma_5}, \quad (4.3.12)$$

which leads to a Jacobian determinant of the transformation:

$$\mathcal{D}[\psi', \bar{\psi}'] = (\det C)^{-2} \mathcal{D}[\psi, \bar{\psi}]. \quad (4.3.13)$$

It turns out that this determinant is ill-defined ($0 \cdot \infty$) and requires regularization, which in turn breaks the $U(1)_A$ symmetry. The final result is just the anomalous term:

$$(\det C)^{-2} = \exp\left(-i \int d^4x \varepsilon(x) N_f \mathcal{Q}(x)\right). \quad (4.3.14)$$

As a consequence, $Z[B'] \neq Z[B]$ under a gauge transformation but instead

$$Z[B'] = Z[B] \left\langle \exp\left(-i \int d^4x \varepsilon(x) N_f \mathcal{Q}(x)\right) \right\rangle_B, \quad (4.3.15)$$

and comparison with Eq. (4.3.9) gives the anomalous correction to the PCAC relation:

$$\langle \partial_\mu A^\mu - N_f \mathcal{Q} \rangle_B = 0. \quad (4.3.16)$$

Fujikawa's method. In order to prove Eq. (4.3.14), let us expand the functional determinant into eigenfunctions of the **Dirac operator** $\mathcal{D} = \not{\partial} - ig\not{A}$. This is now again the usual covariant derivative with the gluon field and not the quantity in Eq. (4.3.6), which we no longer need. Assume that the Dirac operator \mathcal{D} is hermitian, so that it has real eigenvalues λ_n and a set of orthonormal, complete eigenfunctions:

$$\begin{aligned} \mathcal{D} \varphi_n(x) &= \lambda_n \varphi_n(x), & \int d^4x \varphi_{m,i}^\dagger(x) \varphi_{n,j}(x) &= \delta_{mn} \delta_{ij}, \\ \sum_n \varphi_{n,i}(x) \varphi_{n,j}^\dagger(y) &= \delta^4(x-y) \delta_{ij}, \end{aligned} \quad (4.3.17)$$

where i, j collect the Dirac, color and flavor indices. To ensure the (anti-) hermiticity of the Dirac operator, we should really do this in Euclidean space, but let us ignore this subtlety in what follows.

We can expand the spinors $\psi, \bar{\psi}$ into these eigenfunctions, where the coefficients a_n and \bar{b}_n are independent Grassmann variables, and write down the path integral measure:

$$\psi(x) = \sum_n a_n \varphi_n(x), \quad \bar{\psi}(x) = \sum_n \varphi_n^\dagger(x) \bar{b}_n, \quad \mathcal{D}[\psi, \bar{\psi}] = \prod_n da_n \prod_m d\bar{b}_m. \quad (4.3.18)$$

As a side remark, the fermionic path integral can be written as the determinant of the Dirac operator (which is useful in lattice calculations):

$$\det \mathcal{D} = \int \mathcal{D}[\psi, \bar{\psi}] e^{i \int d^4x \bar{\psi} i \not{\mathcal{D}} \psi} = \int \prod_n da_n d\bar{b}_n e^{-\sum_n \bar{b}_n \lambda_n a_n} = \prod_n \lambda_n. \quad (4.3.19)$$

Now, if we use the orthogonality relation to project out the coefficients, an axial transformation changes a_n and \bar{b}_n to

$$a'_n = \int d^4x \varphi_n^\dagger(x) \psi'(x) = \sum_m \underbrace{\int d^4x \varphi_n^\dagger(x) e^{i\varepsilon(x)\gamma_5} \varphi_m(x)}_{=:C_{nm}} a_m \quad (4.3.20)$$

so that we have

$$a'_n = \sum_m C_{nm} a_m, \quad \bar{b}'_m = \sum_n C_{nm} \bar{b}_n. \quad (4.3.21)$$

Note that because we are dealing with axial transformations, both a_n and \bar{b}_m transform with the same C_{mn} ,

$$C_{mn} = \delta_{mn} + i \int d^4x \varepsilon(x) \varphi_n^\dagger(x) \gamma_5 \varphi_m(x) + \dots, \quad (4.3.22)$$

and because the Grassmann measure transforms with the inverse determinant we arrive at Eq. (4.3.13). Using $\det C = e^{\text{Tr} \ln C}$ and expanding the logarithm, we obtain

$$(\det C)^{-2} = \exp \left(-2i \int d^4x \varepsilon(x) \sum_n \varphi_n^\dagger(x) \gamma_5 \varphi_n(x) \right), \quad (4.3.23)$$

which involves the ‘functional trace’ over γ_5 . With the completeness relation in (4.3.17), the sum becomes

$$\sum_n \varphi_n^\dagger(x) \gamma_5 \varphi_n(x) = \lim_{y \rightarrow x} \sum_n \varphi_{n,i}^\dagger(y) (\gamma_5)_{ij} \varphi_{n,j}(x) = \lim_{y \rightarrow x} \text{Tr} \{ \gamma_5 \} \delta^4(x - y), \quad (4.3.24)$$

where the trace goes over Dirac, color and flavor indices. The color-flavor trace gives a factor $N_f N_c$, whereas the Dirac trace vanishes but the δ -function diverges. Thus we have a $0 \cdot \infty$ situation: this expression is possibly finite, but it is not well-defined and must be regulated.

Fujikawa suggested to regulate it in a gauge-invariant way by damping the contribution from the large eigenvalues by a Gaussian cutoff, with a regulator mass M that is taken to infinity in the end:

$$\begin{aligned} & \lim_{M \rightarrow \infty} \sum_n \varphi_n^\dagger(x) \gamma_5 e^{-(\lambda_n/M)^2} \varphi_n(x) \\ &= \lim_{M \rightarrow \infty} \sum_n \varphi_n^\dagger(x) \gamma_5 e^{-(\not{D}/M)^2} \varphi_n(x) \\ &= \lim_{\substack{M \rightarrow \infty \\ y \rightarrow x}} \text{Tr} \left\{ \gamma_5 e^{-(\not{D}/M)^2} \right\} \delta^4(x - y) \\ &= \lim_{M \rightarrow \infty} \int \frac{d^4k}{(2\pi)^4} e^{-ikx} \text{Tr} \left\{ \gamma_5 e^{-(\not{D}/M)^2} \right\} e^{ikx}. \end{aligned} \quad (4.3.25)$$

This regularization is gauge-invariant because the covariant derivative appears in it; hence, it preserves the vector gauge symmetry. To proceed, we express \not{D}^2 by

$$\begin{aligned} \not{D}^2 &= \gamma^\mu \gamma^\nu D_\mu D_\nu = \frac{1}{2} \{ \gamma^\mu, \gamma^\nu \} D_\mu D_\nu + \frac{1}{2} [\gamma^\mu, \gamma^\nu] D_\mu D_\nu \\ &= D^2 + \frac{1}{4} [\gamma^\mu, \gamma^\nu] [D_\mu, D_\nu] = D^2 - \frac{ig}{4} [\gamma^\mu, \gamma^\nu] F_{\mu\nu} \end{aligned} \quad (4.3.26)$$

and exploit the relation (2.2.47),

$$e^{-ikx} f\left(\frac{\partial}{\partial x}\right) e^{ikx} = f\left(\frac{\partial}{\partial x} + ik\right), \quad (4.3.27)$$

where unsaturated derivatives vanish in the end. Eq. (4.3.25) then becomes

$$\dots = \lim_{M \rightarrow \infty} \int \frac{d^4 k}{(2\pi)^4} \text{Tr} \left\{ \gamma_5 \exp \left(-\frac{(D + ik)^2}{M^2} + \frac{ig}{4M^2} [\gamma^\mu, \gamma^\nu] F_{\mu\nu} \right) \right\}. \quad (4.3.28)$$

When expanding the exponential, only terms with at least four γ matrices can survive the trace with γ_5 , and only those $\propto 1/M^4$ which produce a dimensionless quantity after integration will survive the limit $M \rightarrow \infty$. These terms can only appear at quadratic order and produce

$$\frac{i}{4} \text{Tr} \{ \gamma_5 \gamma^\mu \gamma^\nu \gamma^\alpha \gamma^\beta \} = \varepsilon^{\mu\nu\alpha\beta}. \quad (4.3.29)$$

The resulting expression has the form

$$\dots = \lim_{M \rightarrow \infty} \int \frac{d^4 k}{(2\pi)^4} \left[e^{-\frac{k^2}{M^2}} \frac{g^2}{M^4} N_f \text{Tr} \{ \tilde{F}_{\mu\nu} F^{\mu\nu} \} + \dots \right]. \quad (4.3.30)$$

Then, after integrating out the momentum k and sending $M \rightarrow \infty$, the final result becomes

$$\sum_n \varphi_n^\dagger(x) \gamma_5 \varphi_n(x) = \lim_{y \rightarrow x} \text{Tr} \{ \gamma_5 \} \delta^4(x - y) = \frac{g^2 N_f}{16\pi^2} \text{Tr} \{ \tilde{F}_{\mu\nu} F^{\mu\nu} \}, \quad (4.3.31)$$

where only the color trace over the $SU(3)_C$ generators remains. Inserted in the determinant (4.3.23), we arrive at Eq. (4.3.14).

A few **remarks** are in order:

- Note that we did not perform an ‘additional renormalization’ because the theory was already renormalized before. Renormalization means that the regulator remains in the theory, but it is hidden in the renormalization constants which must cancel each other in observables, together with the regulator dependence. Here we have merely cured a $0 \cdot \infty$ situation by introducing a cutoff M that we sent to infinity at the end. However, the resulting finite expression has the property that it breaks the $U(1)_A$ symmetry. While we used exponential damping, one can show that this result is indeed independent of the chosen regularization as long as it is gauge invariant.

- Since the topological charge is essentially the trace over γ_5 , one can ask why only $U(1)_A$ and not the non-Abelian global $SU(N_f)_A$ transformations lead to anomalies. Repeating the analysis with $\varepsilon \rightarrow \sum_a \varepsilon_a \mathbf{t}_a$ yields

$$\partial_\mu A_a^\mu = \frac{g^2}{(4\pi)^2} \varepsilon^{\alpha\beta\mu\nu} F_{\alpha\beta}^b F_{\mu\nu}^c \text{Tr}_F \{ \mathbf{t}_a \} \text{Tr}_C \{ \mathbf{t}_b \mathbf{t}_c \}, \quad (4.3.32)$$

which vanishes in the flavor-octet case because $\text{Tr} \{ \mathbf{t}_a \} = 0$. In other words, gluons couple only to flavor-singlet currents, and the anomaly signals the breakdown of the $U(1)_A$ symmetry in the presence of gluons.

■ The topological charge density can be written as the divergence of a current, the **Chern-Simons current**:

$$\mathcal{Q}(x) = \partial_\mu K^\mu, \quad K^\mu = \frac{g^2}{8\pi^2} \varepsilon^{\mu\nu\alpha\beta} \text{Tr} \left\{ F_{\alpha\beta} \mathbf{A}_\nu + \frac{2ig}{3} \mathbf{A}_\alpha \mathbf{A}_\beta \mathbf{A}_\nu \right\}. \quad (4.3.33)$$

One could then conclude that the flavor-singlet PCAC relation (in the chiral limit) still induces a conserved current $\partial_\mu(A^\mu - N_f K^\mu) = 0$, which leads back to the argument that there should be a flavor-singlet Goldstone boson. However, K^μ and its corresponding charge $\int d^3x K^0$ are not gauge invariant, so they cannot couple to physical states and hence there is no conserved axial charge.

Triangle diagrams. The axial anomaly will show up (and was originally derived) in the calculation of correlation functions involving axialvector currents, e.g.

$$\langle 0 | T A^\mu(x) V^\alpha(y) V^\beta(z) | 0 \rangle, \quad \langle 0 | T A^\mu(x) A^\alpha(y) A^\beta(z) | 0 \rangle, \quad \text{etc.} \quad (4.3.34)$$

Take for example the WTI for an AVV correlator:

$$\begin{aligned} \partial_\mu^x \langle A^\mu V^\alpha V^\beta \rangle &= \langle (\partial_\mu A^\mu) V^\alpha V^\beta \rangle + \delta(x^0 - y^0) \langle [A^0, V^\alpha] V^\beta \rangle \\ &\quad + \delta(x^0 - z^0) \langle V^\alpha [A^0, V^\beta] \rangle = 0. \end{aligned} \quad (4.3.35)$$

The last two terms on the right-hand side are zero because the commutators of the singlet currents vanish, as one can infer from Eq. (3.1.57). The first term produces the pseudoscalar density via the PCAC relation. Repeating this for derivatives with respect to y and z , we arrive at

$$\partial_\mu^x \langle A^\mu V^\alpha V^\beta \rangle = 2m \langle P V^\alpha V^\beta \rangle, \quad \partial_\alpha^y \langle A^\mu V^\alpha V^\beta \rangle = 0, \quad \partial_\beta^z \langle A^\mu V^\alpha V^\beta \rangle = 0, \quad (4.3.36)$$

without taking into account the anomaly.

The problem is that these diagrams are linearly divergent and therefore not translationally invariant. If one calculates them explicitly to 1-loop order, shifting integration variables by a different momentum routing will produce results that differ by surface terms. The freedom in distributing these surface terms can be used in the regularization procedure when getting rid of all infinite pieces. It turns out that the relations (4.3.36) cannot be satisfied simultaneously, and in order to preserve the vector symmetries the axialvector WTI must pick up the additional anomalous term.

A theorem by **Adler and Bardeen** states that the full structure of the anomaly is already contained in the perturbative one-loop fermion diagrams. Higher-loop corrections do not renormalize the anomaly except for replacing the fields and coupling constants by their renormalized values. For anomaly considerations it is therefore enough to calculate the triangle and rectangle diagrams in Fig. 4.8. These are the superficially divergent ones (in fact, pentagon diagrams should be included as well although they are convergent), and they include an odd number of axial currents and thus an odd number of γ_5 matrices. For example, the anomalous contribution to the η_0 mass in a current correlator arises from quark-disconnected diagrams like the one on the right in Fig. 4.8, which contains intermediate gluon exchanges in the flavor-singlet channel.

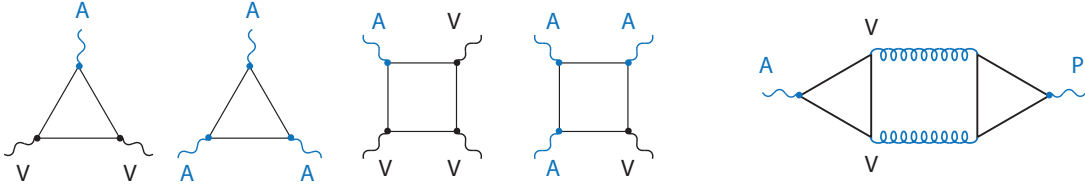


FIG. 4.8: Anomalous 1-loop fermion diagrams.

QED anomaly and $\pi^0 \rightarrow \gamma\gamma$ decay. Anomalies have observable consequences. The prime example are the η and η' masses, but in this case the anomalous contribution is also difficult to quantify due to the explicit breaking of chiral symmetry and mixing effects. A much cleaner system is the decay of the π^0 into two photons, which is almost exclusively caused by the axial anomaly from QED effects.

So far we have considered the axial anomaly in QCD (the ‘gluon anomaly’) which is the relevant one for the $\eta - \eta'$ problem. Quarks couple to gluons, and the quark’s flavor-singlet axialvector current A^μ picks up an anomalous term containing the gluonic field-strength tensor. On the other hand, quarks can also couple to photons, which will also produce an anomaly although the related effects are much weaker ($\alpha_{\text{QED}} \ll \alpha_{\text{QCD}}$). If we repeat the derivation for the QED Lagrangian, replace $F^{\mu\nu}$ by the electromagnetic field-strength tensor and the coupling g with e , we obtain the electromagnetic ‘photon anomaly’ (**Adler-Bell-Jackiw** or **ABJ anomaly**):

$$\partial_\mu A_a^\mu = \frac{e^2}{(4\pi)^2} \varepsilon^{\alpha\beta\mu\nu} F_{\alpha\beta} F_{\mu\nu} \text{Tr}_F \{t_a \mathbf{Q}^2\} \text{Tr}_C \{1\}, \quad \mathbf{Q} = \frac{1}{3} \begin{pmatrix} 2 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \quad (4.3.37)$$

stated here without the fermion mass term and for $N_f = 3$. The generator t_a comes from the axial transformation and the quark charge matrix \mathbf{Q} from the covariant derivative that enters quadratically in the regulator. Since fermions with different flavors have different charges (expressed by \mathbf{Q}), photons can also couple to flavor-nonsinglet currents. Therefore, the electromagnetic anomaly produces additional terms for the divergences of the axial currents A^μ and A_a^μ , i.e., for both $U(1)_A$ and $SU(N_f)_A$.

For the $\pi^0 \rightarrow \gamma\gamma$ decay, consider the three-point function of an axialvector current and two electromagnetic vector currents:

$$\langle 0 | T A_a^\mu(x) V_{\text{em}}^\alpha(x_1) V_{\text{em}}^\beta(x_2) | 0 \rangle. \quad (4.3.38)$$

The electromagnetic current is proportional to the quark charges and given by

$$V_{\text{em}}^\mu(x) = \bar{\psi}(x) \gamma^\mu \mathbf{Q} \psi(x) = V_3^\mu(x) + \frac{1}{\sqrt{3}} V_8^\mu(x). \quad (4.3.39)$$

To lowest order perturbation theory, Eq. (4.3.38) is the AVV triangle diagram in Fig. 4.8 which diverges linearly. However, it also has a spectral representation in terms of pseudoscalar poles, which we can derive in analogy to Eqs. (4.2.34–4.2.36). First, we write down its WTI by acting with the derivative on the index μ :

$$\partial_\mu^x \langle 0 | T A_a^\mu(x) V_{\text{em}}^\alpha(\frac{z}{2}) V_{\text{em}}^\beta(-\frac{z}{2}) | 0 \rangle - 2m \langle 0 | T P_a(x) V_{\text{em}}^\alpha(\frac{z}{2}) V_{\text{em}}^\beta(-\frac{z}{2}) | 0 \rangle = \dots \quad (4.3.40)$$

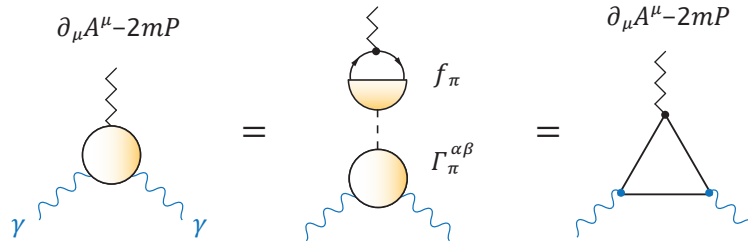


FIG. 4.9: $\pi^0 \rightarrow \gamma\gamma$ decay in the chiral limit.

We are interested in the π^0 with $a = 3$; in that case the commutators on the right-hand side obtained from (3.1.74) vanish, because they contain the structure constants $f_{338} = 0$, etc. Instead we have the contribution from the anomaly:

$$\dots = \frac{e^2 D}{(4\pi)^2} \varepsilon^{\mu\nu\rho\sigma} \langle 0 | T F_{\mu\nu}(x) F_{\rho\sigma}(x) V_{\text{em}}^\alpha(\frac{z}{2}) V_{\text{em}}^\beta(-\frac{z}{2}) | 0 \rangle, \quad (4.3.41)$$

where the factor $D = N_c/6$ comes from the flavor and color traces.

If we work out the time orderings on the left-hand side and insert the completeness relation, we can again isolate the Feynman propagator. The pole residues are the two decay constants from Eq. (3.1.142) and the $\pi^0 \rightarrow \gamma\gamma$ decay amplitude, defined via

$$\Gamma_\lambda^{\alpha\beta}(z, p) = i \langle \lambda | T V_{\text{em}}^\alpha(\frac{z}{2}) V_{\text{em}}^\beta(-\frac{z}{2}) | 0 \rangle = \int \frac{d^4 q}{(2\pi)^4} e^{-iqz} \Gamma_\lambda(q, p) \varepsilon^{\alpha\beta\rho\sigma} q_\rho p_\sigma. \quad (4.3.42)$$

Its structure in momentum space is due to Lorentz and parity invariance: p is the pion momentum, q the relative momentum between the photons, and the only possible Lorentz tensor is $\varepsilon^{\alpha\beta\rho\sigma} q_\rho p_\sigma$. Integrating (4.3.40) over x and z , the poles drop out again and the analogue of Eq. (4.2.36) becomes

$$\lim_{\substack{p \rightarrow 0 \\ q \rightarrow 0}} \sum_\lambda f_\lambda \Gamma_\lambda^{\alpha\beta}(q, p) = \lim_{\substack{p \rightarrow 0 \\ q \rightarrow 0}} f_\pi \Gamma_\pi^{\alpha\beta}(q, p) = 0, \quad (4.3.43)$$

as long as we discard the anomaly on the right-hand side. We have again removed the sum over λ because the decay constants are zero for all excited states with $m_\lambda \neq 0$. Since the transition matrix elements are defined at $p^2 = m_\pi^2 = 0$, this is a chiral-limit relation. Hence, the decay amplitude should be zero, which is known as the **Sutherland-Veltman theorem**.

In order to take the anomaly into account, we would have to work out the right-hand side of Eq. (4.3.41). However, since the anomaly is already produced in the lowest order perturbation theory, it is sufficient to start again from Eq. (4.3.40) and work out its perturbative 1-loop contributions, the *AVV* and *PVV* triangle diagrams. The ambiguity in shifting integration variables produces just the anomalous term. The result has the same structure in momentum space $\sim \varepsilon^{\alpha\beta\rho\sigma} q_\rho p_\sigma$, and the resulting decay amplitude becomes $\Gamma_\pi(0, 0) = e^2 D / (2\pi^2 f_\pi)$. The calculated $\pi \rightarrow \gamma\gamma$ decay width using this result is 7.862 eV; the experimental value is 7.8 ± 0.9 eV. Therefore, the neutral pion decay does not probe the nonperturbative structure of QCD at all — it is completely determined by the axial anomaly.

4.4 Chiral effective field theories

In the discussion so far we have seen that the information on hadrons that can be easily and directly extracted from the QCD Lagrangian is limited: some exact statements are possible, but in practice one needs numerical calculations and/or models to describe the dynamics of the theory. On the other hand, analytic calculations are still possible if we exploit the symmetries of QCD. In particular, the near chiral symmetry of the QCD Lagrangian and its spontaneous breaking can be used to construct low-energy effective theories of QCD, which are not formulated in terms of quarks and gluons but rather with hadrons as effective degrees of freedom. The fact that the pion mass is so much smaller than all other hadronic energy scales makes a perturbative expansion in powers of momenta and pion masses possible. The resulting field theory is called **chiral perturbation theory (ChPT)** and allows one to make rigorous statements as long as the momenta and pion masses are small.

4.4.1 Sigma model

Linear sigma model. We start with the **linear sigma model**, which is the prototype of an effective field theory that implements spontaneous chiral symmetry breaking. In its basic version it describes the interaction of nucleons with pions and scalar mesons:

- The nucleon is represented by spinor fields $\psi(x)$, $\bar{\psi}(x)$ which are isospin doublets, i.e., they transform under the fundamental representation of $SU(2)_f$.
- The three pions correspond to an isospin triplet $\pi_a(x)$ of pseudoscalar fields.
- The scalar meson $\sigma(x)$ is an isoscalar and identified with the $\sigma/f_0(500)$.

One could extend this by including more meson fields such as the ρ meson or other baryon fields, and various **quark-meson models** have been constructed by interpreting the spinors not as nucleons but as quarks.

We combine the pions and the scalar meson into a **meson matrix** ϕ , which is a matrix in Dirac and flavor space and depends linearly on π_a and σ :

$$\phi = \sigma + i\gamma_5 \boldsymbol{\tau} \cdot \boldsymbol{\pi}. \quad (4.4.1)$$

Here, τ_a are the Pauli matrices which are related to the $SU(2)_f$ generators by $\mathbf{t}_a = \tau_a/2$. We defined the ‘length’ of ϕ that will enter in the mass term by

$$|\phi|^2 := \frac{1}{2} \text{Tr} \{ \phi^\dagger \phi \} = \frac{1}{2} \text{Tr} \{ (\sigma - i\gamma_5 \boldsymbol{\tau} \cdot \boldsymbol{\pi})(\sigma + i\gamma_5 \boldsymbol{\tau} \cdot \boldsymbol{\pi}) \} = \sigma^2 + \boldsymbol{\pi}^2, \quad (4.4.2)$$

where we used the identities (same indices are summed over)

$$(\boldsymbol{\tau} \cdot \boldsymbol{\pi})^2 = \pi_a \pi_b \left(\underbrace{\frac{1}{2} [\tau_a, \tau_b]}_{if_{abc} \tau_c} + \underbrace{\frac{1}{2} \{ \tau_a, \tau_b \}}_{\delta_{ab}} \right) = \boldsymbol{\pi}^2, \quad \text{Tr} \{ \tau_a \tau_b \} = 2\delta_{ab} \quad (4.4.3)$$

with $f_{abc} = \varepsilon_{abc}$ in $SU(2)$. Likewise, for the kinetic term we have

$$|\partial_\mu \phi|^2 = \frac{1}{2} \text{Tr} \{ \partial_\mu \phi^\dagger \partial^\mu \phi \} = (\partial_\mu \sigma)^2 + (\partial_\mu \boldsymbol{\pi})^2. \quad (4.4.4)$$

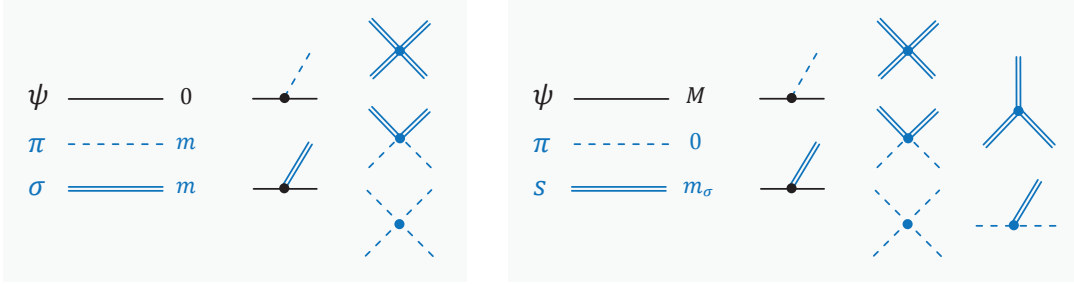


FIG. 4.10: Field content of the sigma model before (left) and after (right) spontaneous chiral symmetry breaking.

The Lagrangian then reads as follows:

$$\mathcal{L} = \bar{\psi} (i\cancel{\partial} - g\phi) \psi + \frac{1}{2} (|\partial_\mu\phi|^2 - m^2|\phi|^2) - V(|\phi|^2), \quad (4.4.5)$$

where the meson matrix couples to the spinors through a Yukawa interaction $\bar{\psi}\phi\psi$. We have assigned the same mass m to each meson, whereas the nucleon at this point is massless. The potential V depends on powers of the meson matrix and we will specify it below. For four-point interactions, the field content of this theory is shown in the left panel of Fig. 4.10.

Let us impose **chiral symmetry** $SU(2)_V \times SU(2)_A$ on the Lagrangian, where the fermions transform under Eqs. (3.1.18–3.1.19):

$$V: \quad U = \exp\left(i\varepsilon_a \frac{\tau_a}{2}\right) \quad \Rightarrow \quad U' = U\psi, \quad \bar{\psi}' = \bar{\psi}U^\dagger, \quad (4.4.6)$$

$$A: \quad U = \exp\left(i\gamma_5 \varepsilon_a \frac{\tau_a}{2}\right) \quad \Rightarrow \quad U' = U\psi, \quad \bar{\psi}' = \bar{\psi}U. \quad (4.4.7)$$

The fermion kinetic term is invariant under both operations, i.e., chirally symmetric:

$$(\bar{\psi}i\cancel{\partial}\psi)' = \left\{ \begin{array}{l} \bar{\psi}U^\dagger i\cancel{\partial}U\psi \quad \dots V \\ \bar{\psi}U i\cancel{\partial}U\psi \quad \dots A \end{array} \right\} = \bar{\psi}i\cancel{\partial}U^\dagger U\psi = \bar{\psi}i\cancel{\partial}\psi. \quad (4.4.8)$$

We have not yet defined how the meson fields transform under chiral symmetry. To do so, we impose invariance of the meson-fermion coupling term $\bar{\psi}\phi\psi$:

$$V: \quad (\bar{\psi}\phi\psi)' = \bar{\psi}U^\dagger\phi'U\psi \stackrel{!}{=} \bar{\psi}\phi\psi \quad \Rightarrow \quad \phi' = U\phi U^\dagger, \quad (4.4.9)$$

$$A: \quad (\bar{\psi}\phi\psi)' = \bar{\psi}U\phi'U\psi \stackrel{!}{=} \bar{\psi}\phi\psi \quad \Rightarrow \quad \phi' = U^\dagger\phi U^\dagger. \quad (4.4.10)$$

The infinitesimal transformations for the π_a and σ fields then become

$$V: \quad \sigma' = \sigma, \quad \pi'_a = \pi_a - f_{abc}\varepsilon_b\pi_c, \quad (4.4.11)$$

$$A: \quad \sigma' = \sigma + \varepsilon_a\pi_a, \quad \pi'_a = \pi_a - \varepsilon_a\sigma. \quad (4.4.12)$$

Observe that the $SU(2)_A$ transformation mixes the σ with the pion fields! This is why they belong together and we needed *both* of them in constructing a chirally invariant Lagrangian.

From the transformation behavior of the meson matrix ϕ it is clear that the remaining terms in the Lagrangian $|\phi|^2 = \frac{1}{2} \text{Tr} \{ \phi^\dagger \phi \}$, $|\partial_\mu \phi|^2$ and $V(|\phi|^2)$ are also chirally invariant. For the individual fields this entails

$$V : \quad \sigma'^2 = \sigma^2, \quad \boldsymbol{\pi}'^2 = \boldsymbol{\pi}^2 + 2f_{abc} \pi_a \pi_b \varepsilon_c = \boldsymbol{\pi}^2, \quad (4.4.13)$$

$$A : \quad \sigma'^2 = \sigma^2 + 2\sigma \pi_a \varepsilon_a, \quad \boldsymbol{\pi}'^2 = \boldsymbol{\pi}^2 - 2\sigma \pi_a \varepsilon_a. \quad (4.4.14)$$

While $SU(2)_V$ leaves both σ^2 and $\boldsymbol{\pi}^2$ invariant, $SU(2)_A$ only preserves their combination $\sigma^2 + \boldsymbol{\pi}^2$. Moreover, renormalizability entails that the possible self-interactions in the potential $V(|\phi|^2)$ can be of order four at most, since the couplings for higher interactions would have a negative mass dimension. A $|\phi|^4$ interaction then leads to the quartic interaction vertices shown in Fig. 4.10.

In this initial Lagrangian, chiral symmetry demands that both mesons must have the same mass m and coupling strength g . Recalling Eq. (3.1.49), we deliberately did not include a mass term for the nucleon since it would break chiral symmetry. Below we will generate a nucleon mass and eliminate the pion mass by means of spontaneous chiral symmetry breaking.

The vector and axialvector **currents** corresponding to the $SU(2)_V$ and $SU(2)_A$ symmetries can be derived from their definition in (3.1.2). They pick up additional terms from the meson fields σ and π_a :

$$V_a^\mu = \bar{\psi} \gamma^\mu \mathbf{t}_a \psi + f_{abc} \pi_b \partial^\mu \pi_c, \quad A_a^\mu = \bar{\psi} \gamma^\mu \gamma_5 \mathbf{t}_a \psi + \sigma \partial^\mu \pi_a - \pi_a \partial^\mu \sigma. \quad (4.4.15)$$

These currents are conserved because the Lagrangian is chirally invariant. The classical **equations of motion** of the linear sigma model are

$$\begin{aligned} \not{\partial} \psi &= -ig \phi \psi, & (\square + m^2) \sigma &= -g \bar{\psi} \psi, \\ \bar{\psi} \overleftarrow{\not{\partial}} &= ig \bar{\psi} \phi, & (\square + m^2) \pi_a &= -2ig \bar{\psi} \gamma_5 \mathbf{t}_a \psi, \end{aligned} \quad (4.4.16)$$

up to terms coming from the potential $V(|\phi|^2)$.

We note that one could rewrite the linear sigma model in terms of a meson matrix Σ which is a matrix in flavor space only:

$$\Sigma := \sigma + i\boldsymbol{\tau} \cdot \boldsymbol{\pi}. \quad (4.4.17)$$

Employing the chiral projectors $P_\pm = (\mathbb{1} \pm \gamma_5)/2$ from Eq. (3.1.42), we have

$$\begin{aligned} \phi &= \sigma + i\gamma_5 \boldsymbol{\tau} \cdot \boldsymbol{\pi} = (P_+ + P_-) \sigma + (P_+ - P_-) i\boldsymbol{\tau} \cdot \boldsymbol{\pi} \\ &= P_+ (\sigma + i\boldsymbol{\tau} \cdot \boldsymbol{\pi}) + P_- (\sigma - i\boldsymbol{\tau} \cdot \boldsymbol{\pi}) \\ &= P_+ \Sigma + P_- \Sigma^\dagger = P_+ \Sigma P_+ + P_- \Sigma^\dagger P_-. \end{aligned} \quad (4.4.18)$$

With the definition (3.1.43) of the right- and left-handed spinors, $\psi_\omega = P_\omega \psi$ and $\bar{\psi}_\omega = \bar{\psi} P_{-\omega}$, the Yukawa coupling becomes

$$\bar{\psi} \phi \psi = \bar{\psi}_- \Sigma \psi_+ + \bar{\psi}_+ \Sigma^\dagger \psi_-. \quad (4.4.19)$$

The remaining terms, defined via (4.4.2), have the same form as before:

$$|\phi|^2 = |\Sigma|^2, \quad |\partial_\mu \phi|^2 = |\partial_\mu \Sigma|^2. \quad (4.4.20)$$

With the transformation of the chiral spinors in Eq. (3.1.46), $\psi'_\omega = U_\omega \psi_\omega$ and $\bar{\psi}'_\omega = \bar{\psi}_\omega U_\omega^\dagger$, chiral symmetry demands

$$\bar{\psi}'_- \Sigma' \psi'_+ + \bar{\psi}'_+ \Sigma'^\dagger \psi'_- = \bar{\psi}_- U_-^\dagger \Sigma' U_+ \psi_+ + \bar{\psi}_+ U_+^\dagger \Sigma'^\dagger U_- \psi_- \stackrel{!}{=} \bar{\psi}_- \Sigma \psi_+ + \bar{\psi}_+ \Sigma^\dagger \psi_-, \quad (4.4.21)$$

hence the matrix Σ must transform under $SU(2)_L \times SU(2)_R$ as $\Sigma' = U_- \Sigma U_+^\dagger$.

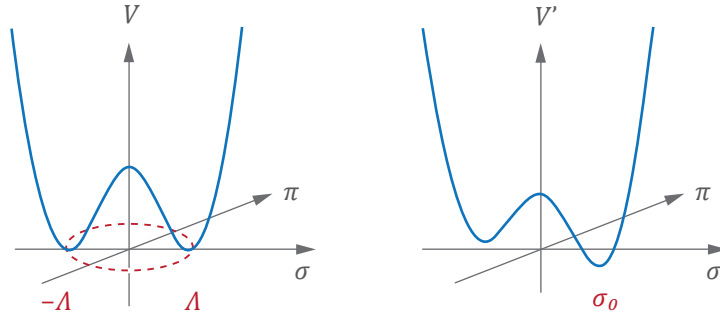


FIG. 4.11: Mexican hat potential of Eq. (4.4.22), with minima along the chiral circle. The right figure includes the explicit symmetry-breaking term of Eq. (4.4.29).

Spontaneous chiral symmetry breaking. Next, we want to generate a mass for the fermions and also get rid of the pion mass. To this end we drop the identification of m with the masses of σ and π . Instead we interpret it as a scale Λ via $-m^2 =: \lambda\Lambda^2$ that we absorb into the potential:

$$V(|\phi|^2) = \frac{\lambda}{4} |\phi|^4 - \frac{\lambda\Lambda^2}{2} |\phi|^2 = \frac{\lambda}{4} (|\phi|^2 - \Lambda^2)^2 - \frac{\lambda}{4} \Lambda^4. \quad (4.4.22)$$

Constant terms can always be dropped from the Lagrangian. The remainder is the **mexican hat** potential shown in Fig. 4.11, which has minima along the ‘chiral circle’ $|\phi|^2 = \sigma^2 + \pi^2 = \Lambda^2$. Note that this is still a chirally symmetric condition and the Lagrangian is invariant under chiral symmetry as before.

However, in this way we have prepared the groundwork that triggers a spontaneous symmetry breaking (SSB) in the quantum field theory. Recall the discussion of the quantum effective action $\Gamma[\varphi]$ and the classical field $\varphi(x) = \langle\phi(x)\rangle_J$ around Eq. (2.2.42). If we set the sources $J = 0$, then also $\varphi(x) = 0$. The first derivative of $\Gamma[\varphi]$ vanishes and the higher derivatives are the 1PI correlation functions for the field $\phi(x)$:

$$\left. \frac{\delta\Gamma[\varphi]}{\delta\varphi(x)} \right|_{\varphi=0} = 0, \quad \left. \frac{\delta^n\Gamma[\varphi]}{\delta\varphi(x_1)\cdots\delta\varphi(x_n)} \right|_{\varphi=0} = \Gamma_\phi^{(n)}(x_1, \dots, x_n). \quad (4.4.23)$$

In the presence of a non-zero vacuum expectation value, setting $J = 0$ entails $\varphi(x) = v$ and these relations are modified as follows:

$$\left. \frac{\delta\Gamma[\varphi]}{\delta\varphi(x)} \right|_{\varphi=v} = 0, \quad \left. \frac{\delta^n\Gamma[\varphi]}{\delta\varphi(x_1)\cdots\delta\varphi(x_n)} \right|_{\varphi=v} = \Gamma_{\phi-v}^{(n)}(x_1, \dots, x_n). \quad (4.4.24)$$

The higher derivatives are the 1PI correlation functions for the field $\phi(x) - v$ and the ‘one-point function’ still vanishes for $\varphi(x) = v$, which therefore extremizes the effective action. Since the classical potential gives the tree-level contribution to $\Gamma[\varphi]$, its minimum determines the leading-order result for the VEV.

Then again, the minimum of the mexican hat is still a chirally symmetric condition. What actually breaks the chiral symmetry of the vacuum is parity invariance, which entails $\langle 0|\pi_a|0\rangle = 0$ and leaves only $\sigma_0 = \langle 0|\sigma|0\rangle = \pm\Lambda$, i.e., it singles out two points on the chiral circle. To determine the true ground state, one must introduce an explicit symmetry-breaking term that tilts the potential towards one absolute minimum.

The next step is to expand the σ field around its minimum by introducing a new fluctuating field s . There are different ways to do so; one possible choice is $\sigma = \Lambda + s$. Since Λ is a constant, we have $\partial_\mu \sigma = \partial_\mu s$ and the form of the kinetic term for the mesons remains unchanged:

$$\frac{1}{2} |\partial_\mu \phi|^2 = \frac{1}{2} ((\partial_\mu \sigma)^2 + (\partial_\mu \boldsymbol{\pi})^2) \cong -\frac{1}{2} (s \square s + \boldsymbol{\pi} \square \boldsymbol{\pi}). \quad (4.4.25)$$

Instead, the potential becomes

$$\begin{aligned} V(|\phi|^2) &= \frac{\lambda}{4} (|\phi|^2 - \Lambda^2)^2 = \frac{\lambda}{4} ((\Lambda + s)^2 + \boldsymbol{\pi}^2 - \Lambda^2)^2 = \frac{\lambda}{4} (s^2 + \boldsymbol{\pi}^2 + 2\Lambda s)^2 \\ &= \lambda \left[\frac{1}{4} (s^2 + \boldsymbol{\pi}^2)^2 + \Lambda s (s^2 + \boldsymbol{\pi}^2) + \Lambda^2 s^2 \right]. \end{aligned} \quad (4.4.26)$$

Expressed in terms of s and $\boldsymbol{\pi}$, the Lagrangian (4.4.5) reads explicitly:

$$\begin{aligned} \mathcal{L} &= \bar{\psi} (i\not{\partial} - g\Lambda) \psi - g\bar{\psi} (s + i\gamma_5 \boldsymbol{\tau} \cdot \boldsymbol{\pi}) \psi \\ &\quad - \frac{1}{2} s (\square + 2\lambda\Lambda^2) s - \frac{1}{2} \boldsymbol{\pi} \square \boldsymbol{\pi} - \lambda\Lambda (s^3 + s\boldsymbol{\pi}^2) - \frac{\lambda}{4} (s^4 + 2s^2\boldsymbol{\pi}^2 + \boldsymbol{\pi}^4). \end{aligned} \quad (4.4.27)$$

In this way we have generated a nucleon mass M , a scalar mass m_σ , and two new cubic interaction vertices $\sim s^3$ and $\sim s\boldsymbol{\pi}^2$ (see right panel of Fig. 4.10):

$$M = g\Lambda, \quad m_\sigma = \sqrt{2\lambda}\Lambda, \quad g_{sss} = g_{\pi\pi s} = \lambda\Lambda. \quad (4.4.28)$$

The pions remain massless, hence they are the three Goldstone bosons of the spontaneously broken $SU(2)_A$. Observe that since we only redefined the fields, the Lagrangian is still the same as before and therefore chirally invariant (despite the fermion mass term!). The symmetry is merely ‘hidden’. However, the ground state is not invariant and thus chiral symmetry is spontaneously broken in the QFT.

Explicit chiral symmetry breaking. Since the pions in nature have a mass, we can add a term to the Lagrangian that breaks chiral symmetry explicitly,

$$V' = V - m_\pi^2 \Lambda \sigma \quad \Leftrightarrow \quad \mathcal{L}' = \mathcal{L} + m_\pi^2 \Lambda \sigma, \quad (4.4.29)$$

where we already named the coefficient accordingly. The potential is now tilted, and the absolute minimum appears at

$$\left. \frac{\partial V'}{\partial \sigma} \right|_{\substack{\sigma=\sigma_0, \\ \pi_a=0}} \stackrel{!}{=} 0 \quad \Rightarrow \quad \lambda(\sigma_0^2 - \Lambda^2) = m_\pi^2 \frac{\Lambda}{\sigma_0} \approx m_\pi^2 \quad \Rightarrow \quad \sigma_0 = +\sqrt{\Lambda^2 + \frac{m_\pi^2}{\lambda}}. \quad (4.4.30)$$

If we expand around the new minimum and insert $\sigma = \sigma_0 + s$ into the potential (4.4.22), we generate a mass term $\sim -\frac{1}{2} m_\pi^2 \boldsymbol{\pi}^2$ for the pion. The remaining Lagrangian has the same form as in Eq. (4.4.27) at first order in m_π^2 , but instead of the relations (4.4.28) we find:

$$M = g \sqrt{\Lambda^2 + \frac{m_\pi^2}{\lambda}}, \quad m_\sigma = \sqrt{2\lambda} \sqrt{\Lambda^2 + \frac{3m_\pi^2}{2\lambda}}, \quad g_{sss} = g_{\pi\pi s} = \lambda \sqrt{\Lambda^2 + \frac{m_\pi^2}{\lambda}}. \quad (4.4.31)$$

The resulting evolution of the nucleon mass with m_π^2 already resembles the outcome of realistic calculations in QCD, which we sketched earlier in Fig. 4.7. In the linear sigma model, the nucleon mass in the chiral limit is $g\Lambda$.

Non-linear representations. The linear sigma model needs both pions and a scalar field to respect chiral symmetry. This is not very satisfactory because the actual $\sigma/f_0(500)$ is a broad resonance and it seems unnatural that it would play the fundamental role suggested by the linear sigma model. While we cannot simply set $s = 0$ in the way we introduced it above ($\sigma = \Lambda + s$) without breaking the chiral symmetry of the Lagrangian, we can eliminate the σ meson by allowing the meson matrix to be **nonlinear** in the pion fields.

Let us introduce a new scalar field s and new pion fields φ_a by

$$\phi = (\Lambda + s)\Omega, \quad \Omega = \exp\left(i\gamma_5 \boldsymbol{\tau} \cdot \boldsymbol{\varphi} \frac{\alpha(z)}{\Lambda z}\right) = \cos \alpha(z) + i\gamma_5 \boldsymbol{\tau} \cdot \boldsymbol{\varphi} \frac{\sin \alpha(z)}{\Lambda z}. \quad (4.4.32)$$

Here we defined

$$(\boldsymbol{\tau} \cdot \boldsymbol{\varphi})^2 = \boldsymbol{\varphi}^2 = \Lambda^2 z^2, \quad (4.4.33)$$

where z is the dimensionless ‘length’ of the pion field, $\alpha(z)$ is some function of z , and we used $e^{iA\alpha} = \cos \alpha + iA \sin \alpha$ for $A^2 = 1$. As a consequence, the original fields σ and π_a are related to the new ones by

$$\sigma = (\Lambda + s) \cos \alpha(z), \quad \pi_a = (\Lambda + s) \frac{\varphi_a}{\Lambda z} \sin \alpha(z). \quad (4.4.34)$$

The advantage of doing this is that Ω depends only on the new pion fields (but on *all* powers of them), and because of $|\phi|^2 = (\Lambda + s)^2$ the potential depends only on the scalar field:

$$V(|\phi|^2) = \frac{\lambda}{4} (|\phi|^2 - \Lambda^2)^2 = \lambda \left(\frac{s^4}{4} + \Lambda s^3 + \Lambda^2 s^2 \right). \quad (4.4.35)$$

Because $|\phi|^2$ is chirally symmetric, in this way we have achieved a chirally symmetric separation of the scalar and pion fields. In turn, the kinetic term for the mesons becomes more complicated and also encodes the pion’s self-interactions via derivative couplings. With $\partial_\mu \phi = \partial_\mu s \Omega + (\Lambda + s) \partial_\mu \Omega$ we find

$$\begin{aligned} |\partial_\mu \phi|^2 &= \frac{1}{2} \text{Tr} \{ \partial_\mu \phi^\dagger \partial^\mu \phi \} \\ &= \frac{1}{2} \text{Tr} \left\{ \partial_\mu s \partial^\mu s + (\Lambda + s) \partial^\mu s \underbrace{\left(\partial_\mu \Omega^\dagger \Omega + \Omega^\dagger \partial_\mu \Omega \right)}_{\partial_\mu (\Omega^\dagger \Omega) = 0} + (\Lambda + s)^2 \partial_\mu \Omega^\dagger \partial^\mu \Omega \right\} \\ &= (\partial_\mu s)^2 + (\Lambda + s)^2 |\partial_\mu \Omega|^2, \end{aligned}$$

where $|\partial_\mu \Omega|^2$ is a complicated function of the pion fields. The explicit calculation yields

$$|\partial_\mu \Omega|^2 = \frac{1}{z^2} \left[\frac{(\partial_\mu \boldsymbol{\varphi})^2}{\Lambda^2} \sin^2 \alpha + \frac{(\boldsymbol{\varphi} \cdot \partial_\mu \boldsymbol{\varphi})^2}{\Lambda^4} \left(\alpha'(z)^2 - \frac{\sin^2 \alpha}{z^2} \right) \right]. \quad (4.4.36)$$

Depending on the function $\alpha(z)$, we could work with the

- **exponential representation:** $\alpha(z) = z \Rightarrow \Omega = \exp\left(i\gamma_5 \frac{\boldsymbol{\tau} \cdot \boldsymbol{\varphi}}{\Lambda}\right)$,
- **square-root representation:** $\sin \alpha(z) = z \Rightarrow \Omega = \sqrt{1 - z^2} + i\gamma_5 \frac{\boldsymbol{\tau} \cdot \boldsymbol{\varphi}}{\Lambda}$.

In any case, the Lagrangian in terms of the new fields becomes

$$\mathcal{L} = \bar{\psi} \left(i\cancel{\partial} - g(\Lambda + s)\Omega \right) \psi - \frac{1}{2} s (\square + 2\lambda\Lambda^2) s - \lambda \left(\frac{s^4}{4} + \Lambda s^3 \right) + \frac{1}{2} (\Lambda + s)^2 |\partial_\mu \Omega|^2.$$

Diagrammatically it still contains the Feynman rules from the right panel of Fig. 4.10, with the identifications in Eq. (4.4.28), but due to the appearance of Ω there are new vertices with pion legs: in fact, once we expand the exponential, there are infinitely many of them! While this looks very different from the Lagrangians (4.4.5) or (4.4.27), in principle it is still the same theory since all we have done is renaming the fields. Indeed one can show that onshell scattering amplitudes obtained from either of these representations are identical.

Non-linear sigma model. The main advantage of arranging the fields in this way is the following: because we separated the fields s and φ_a in a chirally invariant manner, setting $s = 0$ does no longer break chiral symmetry and we can safely eliminate it from the theory. The resulting Lagrangian is

$$\mathcal{L} = \mathcal{L}_N + \mathcal{L}_\pi = \bar{\psi} \left(i\cancel{\partial} - g\Lambda\Omega \right) \psi + \frac{\Lambda^2}{2} |\partial_\mu \Omega|^2 \quad (4.4.37)$$

and contains nucleons and pions only. The pionic part \mathcal{L}_π , where the self-interactions of the pions enter via derivative couplings, is called **nonlinear sigma model**. Note that we could have also obtained it by setting $\sigma^2 + \boldsymbol{\pi}^2 = \Lambda^2$ from the beginning, i.e., by restricting the fields to the chiral circle and thereby eliminating the σ field as an independent degree of freedom.

Unfortunately, the chirally invariant separation of scalar and pion fields comes at a price, namely the loss of renormalizability. The reason is that Ω and $|\partial_\mu \Omega|^2$ contain all powers of the pion field φ_a . Suppose we work with the exponential representation $\alpha(z) = z$:

$$|\partial_\mu \Omega|^2 = \frac{1}{z^2} \left[\frac{(\partial_\mu \boldsymbol{\varphi})^2}{\Lambda^2} \sin^2 z + \frac{(\boldsymbol{\varphi} \cdot \partial_\mu \boldsymbol{\varphi})^2}{\Lambda^4} \left(1 - \frac{\sin^2 z}{z^2} \right) \right]. \quad (4.4.38)$$

The Taylor expansion of $(\sin z/z)^2$ gives

$$\frac{\sin^2 z}{z^2} = 1 - \frac{z^2}{3} + \frac{2z^4}{45} + \dots = 1 - z^2 \sum_{r=0}^{\infty} c_r z^{2r}, \quad (4.4.39)$$

so that \mathcal{L}_π becomes

$$\mathcal{L}_\pi = \frac{1}{2} (\partial_\mu \boldsymbol{\varphi})^2 + \frac{1}{2} \left((\boldsymbol{\varphi} \cdot \partial_\mu \boldsymbol{\varphi})^2 - \boldsymbol{\varphi}^2 (\partial_\mu \boldsymbol{\varphi})^2 \right) \sum_{r=0}^{\infty} \frac{c_r}{(\Lambda^2)^{r+1}} (\boldsymbol{\varphi}^2)^r. \quad (4.4.40)$$

The first term is the inverse tree-level propagator. The second contains an infinite number of tree-level vertices with even numbers of pion legs and derivative couplings (Fig. 4.12): The term with $r = 0$ returns a four-point vertex with coupling constant c_0/Λ^2 , the term with $r = 1$ a six-point vertex with coupling constant c_1/Λ^4 , and so on. As a result, the perturbative expansion of an n -point function not only contains infinitely many loop diagrams but also depends on infinitely many *vertices*.

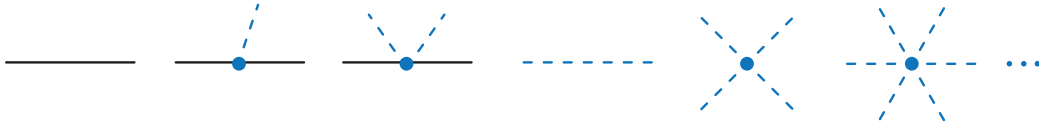


FIG. 4.12: Tree-level diagrams contained in the Lagrangians (4.4.37) and (4.4.64).

Even worse, the couplings carry negative mass dimensions and therefore these interactions are non-renormalizable. Because we deleted the field s , the non-linear sigma model is no longer equivalent to the original Lagrangian, which was renormalizable even though in the end this was no longer obvious (still, we could have transformed back to the original fields σ and π_a). In practice this means that each n -point function produces new divergences, so we would also need infinitely many renormalization conditions and the theory loses its predictivity.

While this would make practical applications hopeless, the fact that all couplings contain **derivatives** opens up a new interpretation: derivatives become momenta in momentum space, and higher powers of momenta are suppressed at low energies. If we can show that higher-loop diagrams also correlate with higher momentum powers, then we can stop the expansion at some given order and fix the necessary renormalization constants at that order by outside information, e.g., from experimental data. The convergence radius is then limited to low momenta and low energies; hence, we can interpret the model as a **low-energy effective theory**. In fact, the nonlinear sigma model \mathcal{L}_π constitutes the lowest-order term in chiral perturbation theory.

Weinberg's power counting. For illustration, let us go back to a ϕ^p theory, where we add up the ϕ^p couplings in the Lagrangian:

$$\mathcal{L} = \mathcal{L}_{\text{kin}} + \lambda_4 \phi^4 + \frac{\lambda_6}{\Lambda^2} \phi^6 + \frac{\lambda_8}{\Lambda^4} \phi^8 \dots \quad (4.4.41)$$

Because the non-renormalizable couplings for $p > 4$ carry negative mass dimensions, we pulled out powers of a scale Λ so that the λ_p are dimensionless. In this case the formula (2.3.62) which we established earlier generalizes to

$$[\Gamma_n] = 4L - 2I + \sum_p (4 - p) V_p, \quad \sum_p V_p = V. \quad (4.4.42)$$

Here, $[\Gamma_n]$ is the mass dimension of a given n -point function. For some perturbative loop diagram that contributes to Γ_n , L is the number of loops, I the number of propagators and V_p the number of ϕ^p vertices, with V the total number of vertices in the diagram. $D = 4L - 2I$ is the degree of divergence of the diagram.

Now observe that in order to preserve the mass dimension of the n -point function, internal momentum powers in a loop must translate to powers of the external momenta and masses. This is easiest to see in dimensional regularization from the loop formulas (2.3.40–2.3.41):

$$I_{nm}^{(4)} = \int \frac{d^4 l_E}{(2\pi)^4} \frac{(l_E^2)^m}{(l_E^2 + \Delta)^n} \propto \Delta^{2+m-n}, \quad (4.4.43)$$










	$V_4 = 0$	$V_4 = 1$	$V_4 = 2$	$V_4 = 3$
$V_6 = 0$	 D=0 D=0	 D=0 D=2	 D=0 D=4	 D=0 D=6
$V_6 = 1$	 D=2 D=4	 D=2 D=6	... D=2 D=8	... D=2 D=10
$V_6 = 2$	 D=4 D=8	 D=4 D=10	... D=4 D=12	
$V_6 = 3$	 D=6 D=12	... D=6 D=14		

FIG. 4.13: Loop diagrams contributing to a 1PI four-point function with four- and six-point interactions. The degree of divergence D for ordinary couplings with $d_\chi = 0$ is given in bold black font and the one for derivative couplings with $d_\chi = 2$ in red, cf. Eq. (4.4.48).

where the quantity Δ defined in Eq. (2.3.32) has mass dimension two and depends on the external momenta and the masses in the loop. It is attached to an expression that contains divergent $1/\varepsilon$ terms and finite parts, where the former drop out after renormalization. The renormalization scale M (or μ) only enters through logarithms. Likewise, had we employed a cutoff regulator, the divergent terms would scale with powers of the cutoff and drop out after renormalization, whereas the finite pieces scale with powers of the external momenta and the masses.

If a diagram contains higher ϕ^p vertices, its degree of divergence D raises according to Eq. (4.4.42):

$$D = [\Gamma_n] + \sum_p (p-4) V_p. \quad (4.4.44)$$

The reason is that those vertices come with higher powers of Λ^2 in the denominators, which must be compensated by momentum powers in the numerators to preserve the mass dimension in the n -point function; these make the diagrams more divergent. However, in this way $D = 4L - 2I$ not only counts the degree of divergence, but also the **powers in the external momenta and masses**. As long as the momenta and masses are small, diagrams with higher D (i.e., higher loop diagrams) will be more and more suppressed. If we supply enough renormalization conditions at a given order D to remove the infinities, we can stop the perturbative expansion after a few terms. Thus, non-renormalizable theories can be viewed as low-energy effective field theories.

The situation is illustrated in Fig. 4.13, which collects the lowest perturbative loop diagrams contributing to a four-point function with ϕ^4 and ϕ^6 vertices. In the horizontal direction we increase V_4 , the number of vertices corresponding to the renormalizable ϕ^4 interaction, which does not increase D . In the vertical direction we increase the number of vertices V_6 of the non-renormalizable ϕ^6 interaction. Each subsequent row increases D by two, so its diagrams are suppressed by a power of two at small momenta and masses compared to those in the previous row.

On the other hand, there is no suppression in the horizontal direction where D does not change; here we would need to rely on the smallness of λ_4 to stop the series. This is where **derivative couplings** come in. The nonlinear sigma model in (4.4.40) contains φ^p interactions ($p = 4 + 2r$) with derivatives, which become powers of momenta l^μ in momentum space. The generic structure of the φ^p couplings is

$$p = 4 : \frac{l^2}{\Lambda^2}, \quad p = 6 : \frac{l^2}{\Lambda^4}, \quad p = 8 : \frac{l^2}{\Lambda^6}, \quad \text{etc.} \quad (4.4.45)$$

Each vertex in the nonlinear sigma model has two powers of derivatives, which we denote by $d_\chi = 2$. In principle we could construct theories with higher powers of derivatives, e.g. $d_\chi = 4$; in that case, the couplings would have the form

$$p = 4 : \frac{l^4}{\Lambda^4}, \quad p = 6 : \frac{l^4}{\Lambda^6}, \quad p = 8 : \frac{l^4}{\Lambda^8}, \quad \text{etc.} \quad (4.4.46)$$

Eq. (4.4.42) still holds for derivative couplings since the mass dimensions of the φ^p couplings are still $4 - p$. However, they are now the differences stemming from the numerators with momentum powers d_χ and denominators with powers of Λ^2 , where the former contribute to the degree of divergence D . If we split $4 - p = d_\chi - (d_\chi + p - 4)$, then for a theory with fixed d_χ we have

$$\sum_p (4 - p) V_p = \sum_p [d_\chi - (\dots)] V_p = d_\chi V - (\dots), \quad (4.4.47)$$

where the rest does not contribute to D but only to the mass dimension $[\Gamma_n]$. Therefore we arrive at

$$D = 4L - 2I + d_\chi V, \quad (4.4.48)$$

which counts the powers in the external momenta and masses.

If we now go back to Fig. 4.13 and interpret the ϕ^4 and ϕ^6 interactions as derivative couplings with $d_\chi = 2$ like in the nonlinear sigma model, we see that D not only increases vertically but also horizontally. Then up to $D = 4$, for instance, we only need to keep a small number of diagrams. The same diagrams in a theory with $d_\chi = 4$ would carry even larger D . For a general theory containing couplings with any possible d_χ , Eq. (4.4.48) generalizes to

$$D = 4L - 2I + \sum_{d_\chi} d_\chi V^{(d_\chi)}, \quad \sum_{d_\chi} V^{(d_\chi)} = V, \quad (4.4.49)$$

where $V^{(d_\chi)}$ is the number of vertices in a diagram from a given order in d_χ . In this way, D tells us where to stop the perturbative expansion: diagrams with higher D become less and less important at low momenta and small masses. The assumption of ‘small masses’ is justified since the propagators in the loops are pions and their masses are indeed small.

We finally note that the three quantities L , I and V are not independent but related by $L + V = I + 1$. Thus we could substitute I and write Eq. (4.4.48) as

$$D = 2 + 2L + (d_\chi - 2)V, \quad (4.4.50)$$

which shows directly that D grows with the number of loops and vertices.

Fermion Lagrangian. We have not yet addressed the fermionic part

$$\mathcal{L}_N = \bar{\psi} \left(i \not{\partial} - g \Lambda \Omega \right) \psi \quad (4.4.51)$$

of the Lagrangian (4.4.37), which contains the nucleon mass term through a complicated dependence on the pion fields encoded in Ω . In analogy to Eq. (4.4.18), we rewrite the Dirac-flavor matrix Ω in terms of

$$\Sigma = \exp \left(i \frac{\boldsymbol{\tau} \cdot \boldsymbol{\varphi}}{\Lambda} \right) = \cos z + i \boldsymbol{\tau} \cdot \boldsymbol{\varphi} \frac{\sin z}{\Lambda z}, \quad (4.4.52)$$

which is a matrix in $SU(2)$ flavor space only. Here we used the exponential representation $\alpha(z) = z$. If we use the chiral projectors P_ω and the right- and left-handed spinors $\psi_\omega, \bar{\psi}_\omega$, with $\omega = \pm$, and follow the same steps as in (4.4.18) and below, we find

$$\Omega = P_+ \Sigma P_+ + P_- \Sigma^\dagger P_- \quad \Rightarrow \quad \bar{\psi} \Omega \psi = \bar{\psi}_- \Sigma \psi_+ + \bar{\psi}_+ \Sigma^\dagger \psi_- \quad (4.4.53)$$

as well as

$$|\partial_\mu \Omega|^2 = |\partial_\mu \Sigma|^2 = \frac{1}{2} \text{Tr} \left\{ \partial_\mu \Sigma^\dagger \partial^\mu \Sigma \right\}. \quad (4.4.54)$$

The chiral invariance of \mathcal{L}_N implies the following transformation behavior under the group $SU(2)_L \times SU(2)_R$, where U_- and U_+ are left- and right-handed transformation matrices with independent group parameters:

$$\Sigma' = U_- \Sigma U_+^\dagger. \quad (4.4.55)$$

Note that in the literature it is common to write

$$\psi_- = \psi_L, \quad \psi_+ = \psi_R, \quad U_- = L, \quad U_+ = R, \quad \Sigma = U. \quad (4.4.56)$$

Next, we redefine the fermion fields such that $\bar{\psi} \Omega \psi$ becomes a simple mass term. To do so, we introduce the $SU(2)$ matrices

$$\xi_\pm(x) = \exp \left(\pm i \frac{\boldsymbol{\tau} \cdot \boldsymbol{\varphi}}{2\Lambda} \right) \quad \Rightarrow \quad \xi_\omega^\dagger = \xi_{-\omega}, \quad \Sigma = \xi_+ \xi_+, \quad \Sigma^\dagger = \xi_- \xi_-, \quad (4.4.57)$$

and insertion in Eq. (4.4.53) yields

$$\Omega = \sum_\omega P_\omega \xi_\omega \xi_\omega P_\omega \quad \Rightarrow \quad \bar{\psi} \Omega \psi = \sum_\omega \bar{\psi}_{-\omega} \xi_\omega \xi_\omega \psi_\omega. \quad (4.4.58)$$

Defining the spinors

$$\begin{aligned} \Psi_\omega &= \xi_\omega \psi_\omega = \xi_\omega P_\omega \psi, & \Psi &= \sum_\omega \Psi_\omega \quad \Rightarrow \quad \Psi_\omega = P_\omega \Psi, \\ \bar{\Psi}_\omega &= \bar{\psi}_\omega \xi_{-\omega} = \bar{\psi} P_{-\omega} \xi_{-\omega}, & \bar{\Psi} &= \sum_\omega \bar{\Psi}_\omega \quad \Rightarrow \quad \bar{\Psi}_\omega = \bar{\Psi} P_{-\omega}, \end{aligned} \quad (4.4.59)$$

we arrive at

$$\bar{\psi} \Omega \psi = \sum_\omega \bar{\Psi}_{-\omega} \Psi_\omega = \bar{\Psi} \sum_\omega P_\omega^2 \Psi = \bar{\Psi} \sum_\omega P_\omega \Psi = \bar{\Psi} \Psi. \quad (4.4.60)$$

In turn, the kinetic term $\bar{\psi} i \not{\partial} \psi$ becomes more complicated:

$$\begin{aligned} \bar{\psi} i \not{\partial} \psi &= \sum_{\omega} \bar{\psi}_{\omega} i \not{\partial} \psi_{\omega} = \sum_{\omega} \bar{\Psi}_{\omega} \xi_{\omega} i \not{\partial} \xi_{-\omega} \Psi_{\omega} = \bar{\Psi} \left(\sum_{\omega} \xi_{\omega} i \not{\partial} \xi_{-\omega} P_{\omega} \right) \Psi \\ &= \bar{\Psi} \left(i \not{\partial} + \sum_{\omega} i \xi_{\omega} (\not{\partial} \xi_{-\omega}) P_{\omega} \right) \Psi = \bar{\Psi} (i \not{\partial} + \not{\phi} + \not{\phi} \gamma_5) \Psi. \end{aligned} \quad (4.4.61)$$

In the last step we introduced the vector and axialvector fields

$$\begin{aligned} v^{\mu} &= \frac{i}{2} [\xi_{+}, \partial^{\mu} \xi_{-}] = \frac{i}{2} (\xi_{+} \partial^{\mu} \xi_{-} + \xi_{-} \partial^{\mu} \xi_{+}), \\ a^{\mu} &= \frac{i}{2} \{\xi_{+}, \partial^{\mu} \xi_{-}\} = \frac{i}{2} (\xi_{+} \partial^{\mu} \xi_{-} - \xi_{-} \partial^{\mu} \xi_{+}), \end{aligned} \quad (4.4.62)$$

where we used $\partial_{\mu} (\xi_{-} \xi_{+}) = 0$. Their combination gives

$$\not{\phi} + \not{\phi} \gamma_5 = i (\xi_{+} \not{\partial} \xi_{-} P_{+} + \xi_{-} \not{\partial} \xi_{+} P_{-}), \quad (4.4.63)$$

which is the combination that appears in Eq. (4.4.61).

Putting everything together, the Lagrangian (4.4.37) takes the form

$$\mathcal{L} = \mathcal{L}_N + \mathcal{L}_{\pi} = \bar{\Psi} (i \not{\partial} + \not{\phi} + \not{\phi} \gamma_5 - M) \Psi + \frac{\Lambda^2}{4} \text{Tr} \left\{ \partial_{\mu} \Sigma^{\dagger} \partial^{\mu} \Sigma \right\}, \quad (4.4.64)$$

where the nucleon mass is $M = g\Lambda$ and the original Yukawa couplings between the nucleon and the pion have turned into vector and axialvector couplings. Expanding v^{μ} and a^{μ} in the lowest powers of the pion fields, we obtain

$$v_{\mu} = -\frac{1}{4\Lambda^2} \boldsymbol{\tau} \cdot (\boldsymbol{\varphi} \times \partial_{\mu} \boldsymbol{\varphi}) + \dots, \quad a_{\mu} = \frac{1}{2\Lambda} \boldsymbol{\tau} \cdot \partial_{\mu} \boldsymbol{\varphi} + \dots \quad (4.4.65)$$

The **axialvector coupling** of the pion to the nucleon induced by a^{μ} corresponds to the second diagram in Fig. 4.12. The third diagram is the **Weinberg-Tomozawa term** stemming from v^{μ} , a seagull-like contact interaction between two pions and the nucleon which gives the dominant tree-level contribution to $N\pi$ scattering.

Moreover, with $\Sigma = \xi_{+} \xi_{+}$ it follows that the transformation behavior $\Sigma' = U_{-} \Sigma U_{+}^{\dagger}$ is satisfied if ξ_{+} transforms like

$$\xi'_{+} = U_{-} \xi_{+} K^{\dagger} \equiv K \xi_{+} U_{+}^{\dagger} \quad \Rightarrow \quad \xi'_{-} = U_{+} \xi_{-} K^{\dagger} = K \xi_{-} U_{-}^{\dagger}, \quad (4.4.66)$$

where K is a unitary $SU(2)$ matrix which depends on U_{+} , U_{-} but also on the pion fields φ_a themselves which carry a dependence on x . As a result, we find

$$\begin{aligned} \Psi'_{\omega} &= K \Psi_{\omega}, & \Psi' &= K \Psi, & v'_{\mu} &= K v_{\mu} K^{\dagger} + i K (\partial_{\mu} K^{\dagger}), \\ & & & & a'_{\mu} &= K a_{\mu} K^{\dagger}. \end{aligned} \quad (4.4.67)$$

If we define the **chiral covariant derivative** by $D_{\mu} = \partial_{\mu} - i v_{\mu}$, the comparison with Eq. (2.1.5) shows that the transformation of v_{μ} is that of a vector field under a *local* symmetry. In other words, instead of a global invariance with respect to U_{+} and U_{-} , chiral symmetry has turned into a local invariance under a transformation with K .

Explicit symmetry breaking. In order to make contact with QCD, we add the following term to the Lagrangian of the nonlinear sigma model:

$$\mathcal{L} = \mathcal{L}_N + \mathcal{L}_\pi + \mathcal{L}_{\text{sb}}, \quad \mathcal{L}_{\text{sb}} = \frac{b\Lambda^3}{2} \text{Tr} \left(\mathbf{M} \Sigma^\dagger + \Sigma \mathbf{M}^\dagger \right). \quad (4.4.68)$$

It depends on the two-flavor quark mass matrix \mathbf{M} from Eq. (3.1.15), and b is a dimensionless parameter. We can compare this to $\mathcal{L}_{QCD}|_{\text{massless}} - \bar{\psi} \mathbf{M} \psi$ based on the following arguments:

- \mathcal{L}_{sb} has mass dimension four and is linear in the quark mass matrix.
- \mathcal{L}_{sb} breaks chiral symmetry explicitly. To see this, consider equal quark masses $\mathbf{M} = m \mathbf{1}$; then from $\Sigma' = U_- \Sigma U_+^\dagger$ we have

$$\text{Tr} \left(\Sigma'^\dagger + \Sigma' \right) = \text{Tr} \left(U_+ \Sigma U_-^\dagger + U_- \Sigma U_+^\dagger \right). \quad (4.4.69)$$

Recall from Eq. (3.1.44) that a $SU(2)_V$ transformation implies $\varepsilon_+ = \varepsilon_-$ and thus $U_+ = U_-$, whereas a $SU(2)_A$ transformation implies $\varepsilon_+ = -\varepsilon_-$ and $U_+ = U_-^\dagger$. Therefore, \mathcal{L}_{sb} is still invariant under isospin symmetry $SU(2)_V$ but it breaks $SU(2)_A$, as does the mass term in the QCD Lagrangian.

- Eq. (4.4.52) tells us that

$$\Sigma + \Sigma^\dagger = 2 \cos z = 2 \cos \frac{\varphi^2}{\Lambda^2} = 2 \left[1 - \frac{\varphi^2}{2\Lambda^2} + \dots \right], \quad (4.4.70)$$

and because $\mathbf{M} = \mathbf{M}^\dagger$ we find

$$\mathcal{L}_{\text{sb}} = b\Lambda^3 (m_u + m_d) \left[1 - \frac{\varphi^2}{2\Lambda^2} + \dots \right] = -\frac{1}{2} b\Lambda (m_u + m_d) \varphi^2 + \dots \quad (4.4.71)$$

Hence we can identify the **pion mass** from $m_\pi^2 = b\Lambda (m_u + m_d)$.

- From the mass term in the QCD Lagrangian we can infer the VEV of the Hamiltonian density

$$\langle \mathcal{H}_{QCD} \rangle = \langle \bar{\psi} \mathbf{M} \psi \rangle = m_u \langle \bar{u}u \rangle + m_d \langle \bar{d}d \rangle, \quad (4.4.72)$$

from where we obtain the quark condensate by taking the derivative of $\langle \mathcal{H}_{QCD} \rangle$ with respect to any of the quark masses and setting $m_q = 0$. Comparison with the effective Hamiltonian at vanishing meson fields ($\Sigma = 1$) yields

$$\langle \mathcal{H}_{\text{sb}} \rangle = -b\Lambda^3 (m_u + m_d) \quad \Rightarrow \quad -b\Lambda^3 = \langle \bar{u}u \rangle = \langle \bar{d}d \rangle, \quad (4.4.73)$$

which suggests the identification of b with the dimensionless **quark condensate**. Comparing this with m_π^2 from above, we arrive at

$$\Lambda^2 m_\pi^2 = -\frac{m_u + m_d}{2} \langle \bar{u}u + \bar{d}d \rangle, \quad (4.4.74)$$

which is identical to the GMOR relation (4.2.41) if we identify the scale Λ with the chiral-limit **pion decay constant** f_π .

We can establish more analogies if we derive the **vector and axialvector currents** and their divergences. Usually we would do this via Eq. (3.1.2) by taking the derivative of the Lagrangian with respect to the derivative of the fields. However, this can become cumbersome if \mathcal{L} depends on the fields in a complicated way. A simpler method is to consider the variation of the (globally invariant) action under a *local* gauge transformation, and evaluate it for the solutions of the classical equations of motion:

$$\delta S = - \int d^4x \partial_\mu \sum_a \varepsilon_a j_a^\mu = - \int d^4x \sum_a (\partial_\mu \varepsilon_a j_a^\mu + \varepsilon_a \partial_\mu j_a^\mu). \quad (4.4.75)$$

The variation is then no longer zero because the Lagrangian is not locally invariant, i.e., the surface term is nonvanishing. On the other hand, in this way we can read off both the currents and their divergences (which vanish if the global symmetry is intact) as the coefficients of $\partial_\mu \varepsilon_a$ and ε_a .

To compute the variation of the Lagrangian $\mathcal{L}_\pi = \frac{\Lambda^2}{4} \text{Tr} \{ \partial_\mu \Sigma^\dagger \partial^\mu \Sigma \}$, we work out the infinitesimal transformation of the meson matrix Σ ,

$$\begin{aligned} \Sigma' &= U_- \Sigma U_+^\dagger = (1 + i\varepsilon_-) \Sigma (1 - i\varepsilon_+) = 1 + i\varepsilon_- \Sigma - i\Sigma \varepsilon_+ \\ &= 1 + i[\varepsilon_V, \Sigma] - i\{\varepsilon_A, \Sigma\} = 1 + \delta\Sigma, \end{aligned} \quad (4.4.76)$$

where we used the infinitesimal relation $\varepsilon_\pm = \varepsilon_V \pm \varepsilon_A$ from Eq. (3.1.44). After some algebra, the currents can be read off from the coefficients of $\partial_\mu \varepsilon_V^a$ and $\partial_\mu \varepsilon_A^a$ in δS :

$$\begin{aligned} V_a^\mu &= -i \frac{\Lambda^2}{4} \text{Tr} \left(\tau_a \left[\Sigma, \partial^\mu \Sigma^\dagger \right] \right) = \varepsilon_{abc} \varphi_b \partial^\mu \varphi_c + \dots, \\ A_a^\mu &= i \frac{\Lambda^2}{4} \text{Tr} \left(\tau_a \left\{ \Sigma, \partial^\mu \Sigma^\dagger \right\} \right) = \Lambda \partial^\mu \varphi_a + \dots \end{aligned} \quad (4.4.77)$$

Vice versa, the coefficients of ε_V^a and ε_A^a vanish and therefore the currents are conserved: $\partial_\mu V_a^\mu = \partial_\mu A_a^\mu = 0$.

On the other hand, the variation of the symmetry-breaking mass term \mathcal{L}_{sb} is

$$\begin{aligned} \delta\Sigma + \delta\Sigma^\dagger &= i[\varepsilon_V, \Sigma + \Sigma^\dagger] - i\{\varepsilon_A, \Sigma - \Sigma^\dagger\} = \frac{2}{\Lambda} \sum_a \varepsilon_A^a \varphi_a + \dots \\ \Rightarrow \delta\mathcal{L}_{\text{sb}} &= \sum_a \varepsilon_A^a b \Lambda^2 (m_u + m_d) \varphi_a, \end{aligned} \quad (4.4.78)$$

in which case the divergences of the currents become

$$\partial_\mu V_a^\mu = 0, \quad \partial_\mu A_a^\mu = -b \Lambda^2 (m_u + m_d) \varphi_a = -\Lambda m_\pi^2 \varphi_a. \quad (4.4.79)$$

This is the analogue of the PCAC relation (3.1.39). If we take the divergence of the current in Eq. (4.4.77), we get back the classical equation of motion for the pion field, the Klein-Gordon equation:

$$\partial_\mu A_a^\mu = \Lambda \square \varphi_a = -\Lambda m_\pi^2 \varphi_a \quad \Rightarrow \quad (\square + m_\pi^2) \varphi_a = 0. \quad (4.4.80)$$

4.4.2 Chiral perturbation theory

The approach developed so far looks promising, but it is also not quite satisfactory: We have eliminated the scalar meson in the linear sigma model at the price of a non-renormalizable low-energy effective theory. How is this better than the original approach? After all, it is still just a model that contains certain chosen interactions.

The idea of **chiral perturbation theory (ChPT)**, formulated by Weinberg and then applied by Gasser and Leutwyler, is the following: we do not know the underlying microscopic interactions that constitute hadronic n -point functions, so instead of providing a specific model for them (like the linear or nonlinear sigma model) we write down a systematic expansion in *all* possible terms that are compatible with the symmetries of QCD. The resulting theory is an effective theory formulated in terms of nucleons and pions, and eventually also other $SU(3)$ multiplet members, but since it contains all possible interactions it is a low-energy expansion of QCD.

This theory is non-renormalizable and the resulting Lagrangian will contain infinitely many terms with infinitely many free parameters. However, this is not a serious issue as long as we stick to the lowest orders in the expansion in derivatives and pion masses (i.e., we work at small momenta and close to the chiral limit). If we can fix a small number of unknown parameters — the **low-energy constants (LECs)** — from experiment or lattice QCD, we should be able to make a range of predictions already at tree level or at a low loop orders, e.g. for $\pi\pi$ or $N\pi$ scattering amplitudes, electromagnetic and weak interaction processes, etc.

With the meson matrix $\Sigma = e^{i\tau\cdot\varphi/\Lambda}$, the quark mass matrix M and the scale Λ (identified with the chiral-limit pion decay constant f_π) as building blocks, we can organize the infinitely many possible terms in the Lagrangian by their number of derivatives and powers of pion masses. Because of Lorentz invariance, each term in the Lagrangian must contain an even number $d_\chi = 2, 4, 6, \dots$ of derivatives, and there can be no derivative-free term because $\text{Tr}\{\Sigma^\dagger\Sigma\}$ is a constant. We can then write

$$\mathcal{L} = \sum_{d_\chi} \mathcal{L}^{(d_\chi)} = \mathcal{L}^{(2)} + \mathcal{L}^{(4)} + \mathcal{L}^{(6)} + \dots \quad (4.4.81)$$

Because each derivative translates to a factor of momentum when taking matrix elements, low orders in derivatives correspond to small momenta and larger-derivative terms only have a small effect. On the other hand, each instance of the mass matrix M becomes a factor $\propto m_\pi^2$, as we saw for the lowest-order term in Eq. (4.4.71), and therefore also enters at order $d_\chi = 2$. The lowest-order terms are then given by

$$\begin{aligned} &\bullet \mathcal{L}^{(2)} : \text{Tr}\{\partial_\mu\Sigma^\dagger\partial^\mu\Sigma\}, \text{Tr}\{M\Sigma^\dagger + \Sigma M^\dagger\} \\ &\bullet \mathcal{L}^{(4)} : (\text{Tr}\{\partial_\mu\Sigma^\dagger\partial^\mu\Sigma\})^2, \text{Tr}\{\partial_\mu\Sigma^\dagger\partial_\nu\Sigma\}\text{Tr}\{\partial^\mu\Sigma^\dagger\partial^\nu\Sigma\}, \\ &\quad \text{Tr}\{\partial_\mu\Sigma^\dagger\partial_\nu\Sigma\}\text{Tr}\{M\Sigma^\dagger + \Sigma M^\dagger\}, (\text{Tr}\{M\Sigma^\dagger \pm \Sigma M^\dagger\})^2, \\ &\quad \text{Tr}\{M\Sigma^\dagger M\Sigma^\dagger + M^\dagger\Sigma M^\dagger\Sigma\}, \text{Tr}\{M^\dagger M\} \\ &\bullet \mathcal{L}^{(6)} : \dots \end{aligned} \quad (4.4.82)$$

The lowest-order Lagrangian for $d_\chi = 2$ is just the nonlinear sigma model.

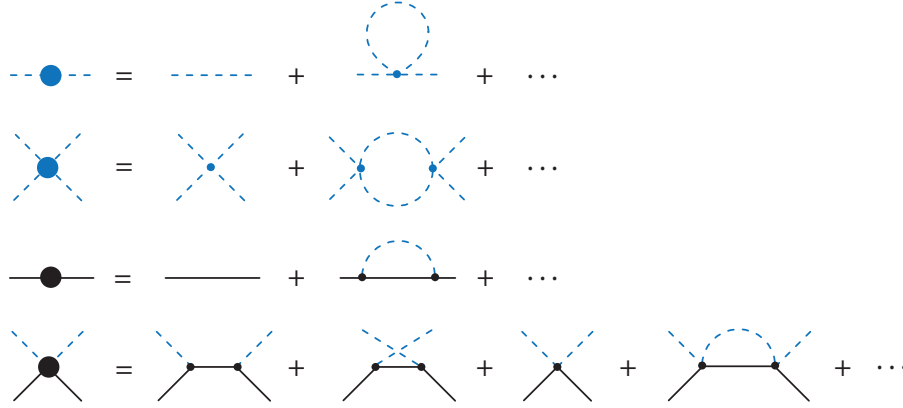


FIG. 4.14: Chiral expansion of π and N propagators, $\pi\pi$ and $N\pi$ scattering.

Unfortunately, the inclusion of nucleons disrupts the power counting because the nucleon mass is not a small scale, and loop diagrams may contribute at the same order as tree-level diagrams. This problem is addressed in **heavy-baryon ChPT**, where manifest covariance is traded for a systematic power counting.

One can then perform loop expansions⁵ for amplitudes like those in Fig. 4.14:

- π and N propagators, where the latter allow one to determine the nucleon mass as a function of m_π^2 .
- $\pi\pi$ scattering near pion-production threshold ($s = 4m_\pi^2$, $t = u = 0$), which is the onset of the physical region. At threshold, the scattering amplitude is expressed by the $\pi\pi$ **scattering lengths**, which vanish in the chiral limit.
- $N\pi$ scattering close to pion production threshold $s = (M + m_\pi)^2$, where at the threshold one can extract the $N\pi$ scattering lengths; etc.

Several ChPT extensions are possible:

- In the case of $SU(3)_f$, the meson matrix becomes $\Sigma = e^{i\lambda \cdot \varphi/\Lambda}$, where the λ_a are the Gell-Mann matrices:

$$\lambda \cdot \varphi = \sqrt{2} \begin{pmatrix} \frac{\pi_0}{\sqrt{2}} + \frac{\eta_8}{\sqrt{6}} & \pi^+ & K^+ \\ \pi^- & -\frac{\pi_0}{\sqrt{2}} + \frac{\eta_8}{\sqrt{6}} & K^0 \\ K^- & \bar{K}^0 & -2\frac{\eta_8}{\sqrt{6}} \end{pmatrix}. \quad (4.4.83)$$

- The effect of the axial anomaly can be implemented through a **Wess-Zumino-Witten (WZW)** term.
- Since ChPT is still a *perturbation theory* around small momenta and pion masses, it cannot generate resonance poles which are nonperturbative effects. This is addressed in **unitarized ChPT**, which amounts to solving self-consistent Bethe-Salpeter equations of the form (3.1.153) but for hadronic correlation functions.

⁵See e.g. S. Scherer, *Adv. Nucl. Phys.* **27** (2003) 277, [hep-ph/0210398](https://arxiv.org/abs/hep-ph/0210398) for explicit calculations.

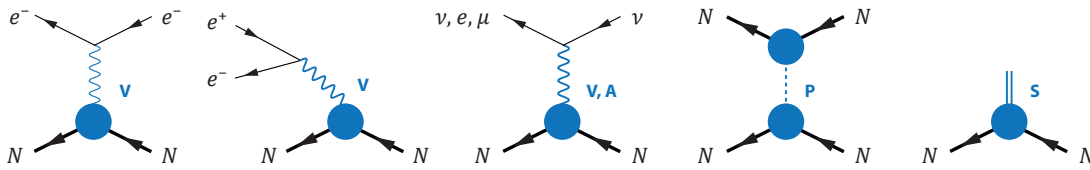


FIG. 4.15: Experimental processes involving the nucleon and different types of currents.

4.5 Hadron matrix elements

4.5.1 Scattering amplitudes

Since quarks and gluons are confined in hadrons, we cannot probe them directly in experiments. Instead, we can learn about their dynamics by probing hadrons with external currents, e.g. when scattering leptons off hadrons, or by scattering hadrons on other hadrons. Among the basic observables extracted from such reactions are the **form factors** of hadrons, which encode their momentum-dependent interactions with photons, W and Z bosons. Some of the relevant processes involving the nucleon are shown in Fig. 4.15:

- **e^-N scattering** has been essential for learning about the substructure of the proton. Due to the smallness of the electromagnetic coupling constant $\alpha_{\text{QED}} \approx 1/137$, the process is dominated by one-photon exchange. The pioneering experiments by Robert Hofstadter revealed that the proton and neutron are not pointlike; instead, their **electromagnetic form factors** provide information on their substructure in terms of electric charge and magnetization distributions (Nobel prize 1961). Even today, the nucleon's electromagnetic form factors are not fully understood, as evidenced by the proton radius puzzle and other open questions. The emission/absorption of virtual photons can also turn a nucleon into a resonance, and electromagnetic **transition form factors** provide insight on the internal structure of nucleon resonances (N^*). Moreover, crossing symmetry implies that the same form factors describing the interaction with a virtual photon also enter in crossed processes such as $e^+e^- \leftrightarrow N\bar{N}$ or $N^{(*)} \rightarrow Ne^+e^-$. The latter is an important tool to probe the initial stages of heavy-ion collisions when forming a quark-gluon plasma, since dileptons (e^+e^- or $\mu^+\mu^-$ pairs) escape the interaction zone mostly unharmed.

- The **axial form factors** of the nucleon can be probed by the weak interaction using W and Z bosons. Examples are neutrino scattering off the nucleon, or the neutron beta decay which is the process $n \rightarrow pe^-\bar{\nu}_e$. The nucleon's axial charge g_A is a basic ingredient in many low-energy relations.

- The **pseudoscalar form factor** of the nucleon is proportional to the $N\pi$ coupling, which enters in the NN interaction through pion exchange.

- The nucleon's **scalar form factor** is not directly measurable but related to the derivative dM/dm_q of the nucleon mass with respect to the current-quark mass (the so-called nucleon **sigma term**) by the **Feynman-Hellmann theorem**. In addition, the Higgs boson could couple to the nucleon through a top-quark loop.

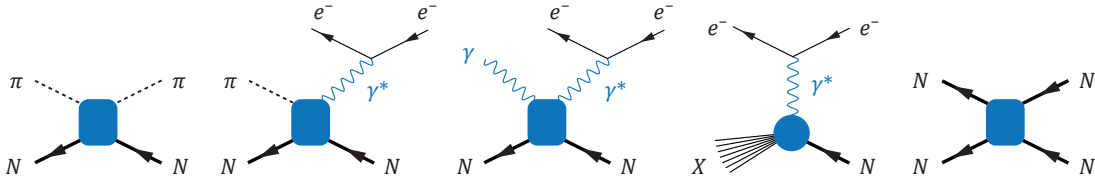


FIG. 4.16: Experimental processes involving nucleons, pions and photons.

On the one hand, measuring form factors gives us experimental information on the internal quark-gluon structure of the probed hadrons and their resonances. On the other hand, these basic quantities appear as building blocks in other scattering processes (Fig. 4.16), which are experimentally accessible but theoretically more complicated:

- **$N\pi$ scattering** has been the traditional tool for extracting nucleon resonances. The scattering amplitude has resonance poles, $N\pi \rightarrow N^* \rightarrow N\pi$, whose residues are the $NN^*\pi$ coupling strengths. Historically, the nucleon resonances have been named after the incoming partial wave $L_{2I,2J}$ in $N\pi$ scattering, with $L = S, P, D, F, \dots$: the Roper resonance is P_{11} , the $N(1535)$ is S_{11} , etc. Because many of the higher-lying excitations predicted by the quark model have not been seen in $N\pi$ scattering, a common assumption has been that they may not couple strongly to $N\pi$.

- **Meson electroproduction** is the process $N\gamma^* \rightarrow NM$, where the virtual photon is produced by the electron and M is the meson. In **photoproduction** the photon is real. These reactions also create nucleon resonances but involve their electromagnetic transition form factors. Thus, if some resonance couples weakly to $N\pi$ but has a large electromagnetic coupling, it should be easier to detect in this way. Combined with improved partial-wave analyses, photo- and electroproduction experiments have indeed found new baryon resonances in recent years. A typical question here concerns the separation of the resonance contributions (like $N\gamma^* \rightarrow N^* \rightarrow NM$) from the quantum-field theoretical ‘background’.

- **Compton scattering** is the process where two photons couple to the nucleon, each of which can be real or virtual. It encodes the nucleon’s polarizabilities, which describe the nucleon’s response to an external electromagnetic field, but also structure functions and generalized parton distributions (GPDs).

- In **deep inelastic scattering (DIS)** the nucleon is broken up by the highly virtual photon and one measures the inclusive cross section $eN \rightarrow eX$. DIS encodes the nucleon’s structure functions, which give us access to the partonic structure of the nucleon and its parton distribution functions (PDFs). The early DIS measurements in the 1960s/70s have provided first convincing evidence for the existence of quarks.

- Finally, **NN scattering** is the elementary reaction in nucleus-nucleus and heavy-ion collisions which are performed e.g. at the LHC and RHIC. The NN interaction is also the basic ingredient in nuclear physics, and in contrast to its long-range part which is mediated by pion exchange, the short-range nuclear force is still not well understood.

Even though this list only contains processes involving the nucleon, the discussion should make it clear that a good understanding of reactions with baryons and mesons, and the various couplings and form factors they contain, is essential for many questions in hadron physics, among them hadron spectroscopy.

A general matrix element can be written as

$$\langle p'_1 \dots p'_n | \mathbb{T} j^\Gamma(x) \dots | p_1 \dots p_n \rangle, \quad (4.5.1)$$

where $\{p_i\}$ and $\{p'_i\}$ are the onshell momenta of the incoming and outgoing particles. The legs associated with the currents (if there are any) are offshell, i.e., their squared momentum is not fixed but arbitrary. After splitting off the spinors u_α, \bar{u}_α for onshell spin- $\frac{1}{2}$ particles, or polarization vectors ε^μ for spin-1 particles etc., the remainder can be expanded in a tensor basis:

$$\mathcal{M}_{\alpha\beta\dots}^{\mu\nu\dots}(p'_1 \dots p_n) = \sum_{i=1}^N F_i(\dots) \tau_i(p'_1 \dots p_n)_{\alpha\beta\dots}^{\mu\nu\dots}, \quad (4.5.2)$$

where the Lorentz-invariant amplitudes or form factors $F_i(\dots)$ are analytic functions of the invariant momentum variables and carry the information on the process.

Kinematic variables. Let us work out the kinematics for a generic scattering process shown in Fig. 4.17,

$$A(p_i) + B(k_i) \rightarrow A'(p_f) + B'(k_f), \quad (4.5.3)$$

where the incoming and outgoing states are not necessarily on their mass shells. If all particles are onshell, then the process describes e.g. eN scattering, $N\pi$ scattering or NN scattering in Figs. (4.15–4.16). If k_i is offshell, it corresponds to meson electroproduction or virtual Compton scattering, and if also k_f is offshell, doubly-virtual Compton scattering. An example where only p_f is offshell is inelastic eN scattering, where p_f is the total momentum of the decay products.

Let us express the amplitude in terms of three independent momenta

$$p = \frac{p_i + p_f}{2}, \quad k = \frac{k_i + k_f}{2}, \quad q = p_f - p_i = k_i - k_f, \quad (4.5.4)$$

where p and k are the average momenta of A and B , respectively, and q is the momentum transfer, with the inverse relations

$$\begin{aligned} p_i &= p - \frac{q}{2}, & k_i &= k + \frac{q}{2}, \\ p_f &= p + \frac{q}{2}, & k_f &= k - \frac{q}{2}. \end{aligned} \quad (4.5.5)$$

The amplitudes $F_i(\dots)$ can then depend on six Lorentz invariants $p^2, k^2, q^2, p \cdot k, p \cdot q$ and $k \cdot q$. It is convenient to define the **Mandelstam variables** s, u and t ,

$$\begin{aligned} s &= (p_i + k_i)^2 = (p_f + k_f)^2 = (p + k)^2, \\ u &= (p_i - k_f)^2 = (p_f - k_i)^2 = (p - k)^2, \\ t &= (p_f - p_i)^2 = (k_i - k_f)^2 = q^2, \end{aligned} \quad (4.5.6)$$

whose sum is

$$s + t + u = 2p^2 + 2k^2 + q^2 = p_i^2 + p_f^2 + k_i^2 + k_f^2. \quad (4.5.7)$$

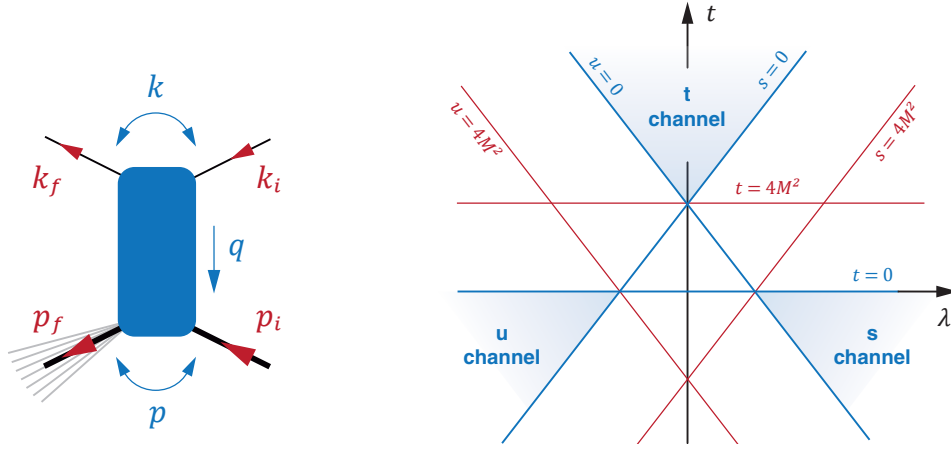


FIG. 4.17: Kinematics in $2 \rightarrow 2$ scattering and Mandelstam plane for identical masses.

If we write $p_i^2 = M^2$, $k_i^2 = m^2$, $p_f^2 = M'^2$ and $k_f^2 = m'^2$, then the six independent variables are given by

$$t = q^2, \quad \lambda = \frac{p \cdot k}{M^2} = \frac{s - u}{4M^2}, \quad (4.5.8)$$

where we defined the **crossing variable** λ , and

$$\begin{aligned} \frac{p^2 + k^2}{2} &= \frac{M^2 + M'^2 + m^2 + m'^2 - q^2}{4} = \frac{s + u}{4}, & \omega &= \frac{p \cdot q}{2M^2} = \frac{M'^2 - M^2}{4M^2}, \\ \frac{p^2 - k^2}{2} &= \frac{M^2 + M'^2 - m^2 - m'^2}{4}, & \omega' &= \frac{k \cdot q}{2M^2} = \frac{m^2 - m'^2}{4M^2}. \end{aligned}$$

For one-particle exchanges in the t channel like in Fig. 4.15, t is the squared momentum of the exchange particle and it is common to define the momentum transfer variable

$$\tau = -\frac{q^2}{4M^2} = \frac{Q^2}{4M^2}. \quad (4.5.9)$$

In a $2 \rightarrow 2$ scattering process where all particles are onshell and their masses are fixed, only t and λ remain independent. They define the **Mandelstam plane**, where the physical regions of the process and its singularity structure can be visualized. The s , t and u -channel processes correspond to

$$\begin{aligned} s \text{ channel} : & \quad 1 + 2 \rightarrow 3 + 4, \\ t \text{ channel} : & \quad 1 + \bar{3} \rightarrow \bar{2} + 4, \\ u \text{ channel} : & \quad 1 + \bar{4} \rightarrow \bar{2} + 3, \end{aligned} \quad (4.5.10)$$

where e.g. $\bar{3}$ is the antiparticle of 3 with opposite momentum. Crossing symmetry implies that all these processes are described by the same amplitudes $F_i(s, t, u)$ but with different physical domains on the Mandelstam plane. In the s -channel reaction, $s > (M + m)^2$ is the square of the total energy in the center-of-mass (CM) frame and t

is the momentum transfer. In the t and u -channel processes, these roles are exchanged. For example, in $N\pi$ scattering the three processes are

$$\begin{aligned} s \text{ channel : } & N(p_i) + \pi^+(k_i) \rightarrow N(p_f) + \pi^+(k_f), \\ t \text{ channel : } & N(p_i) + \bar{N}(-p_f) \rightarrow \pi^(-k_i) + \pi^+(k_f), \\ u \text{ channel : } & N(p_i) + \pi^(-k_f) \rightarrow \pi^(-k_i) + N(p_f), \end{aligned} \quad (4.5.11)$$

where the s , u channels correspond to $N\pi$ scattering and the t channel to $N\bar{N} \leftrightarrow \pi^+\pi^-$ annihilation. In this case the amplitude is symmetric under $s \leftrightarrow u$ crossing and therefore the F_i can only depend on t and λ^2 .

In general, the physical regions in the Mandelstam plane are determined by the **Kibble function**

$$\Phi = stu - \frac{as + bt + cu}{M^2 + M'^2 + m^2 + m'^2} \geq 0, \quad (4.5.12)$$

where

$$\begin{aligned} a &= (M^2 m^2 - M'^2 m'^2)(M^2 + m^2 - M'^2 - m'^2), \\ b &= (M^2 M'^2 - m^2 m'^2)(M^2 - m^2 + M'^2 - m'^2), \\ c &= (M^2 m'^2 - M'^2 m^2)(M^2 - m^2 - M'^2 + m'^2). \end{aligned} \quad (4.5.13)$$

For an elastic scattering process with $M = M'$ and $m = m'$, this reduces to $a = c = 0$ and therefore

$$\Phi = t [su - (M^2 - m^2)^2] \geq 0. \quad (4.5.14)$$

The simplest case where all masses are equal and therefore $stu \geq 0$ is shown in Fig. 4.17. Examples for such processes are $\pi\pi$ or NN scattering. The natural variables to describe the physical s -channel region ($s \geq 0$, $t \leq 0$, $u \leq 0$) are then the Mandelstam variable $s \geq 4M^2$ and the angular variable $z = \cos \theta_{\text{CM}}$ with $-1 \leq z \leq 1$, where θ_{CM} is the scattering angle in the CM frame, see Eq. (4.5.26). Forward scattering ($\theta_{\text{CM}} = 0$) corresponds to $t = 0$ and backward scattering ($\theta_{\text{CM}} = \pi$) to $u = 0$. Employing Legendre polynomials $P_l(z)$, one can perform a **partial-wave expansion** of the amplitude:

$$F(s, z) = \sum_{l=0}^{\infty} (2l+1) F_l(s) P_l(z). \quad (4.5.15)$$

The Mandelstam plane is particularly useful for studying the analytic structure of the amplitudes. Analyticity means that the scattering amplitudes are analytic functions of s , t and u regarded as complex variables. Their only singularities are those imposed by unitarity, which are simple poles due to the exchange of physical particles and branch cuts due to intermediate multiparticle states. For example, if the s -channel process in Fig. 4.17 generates a bound state below the two-particle threshold, it will show up as a pole at some constant $s < 4M^2$ which is a line in the Mandelstam plane. A resonance above threshold will form another line of constant s , however with a pole on a higher Riemann sheet that can at best produce a bump in the partial wave $F_i(s)$. If the amplitude is crossing-symmetric in $s \leftrightarrow u$, the same singularity structure in the u channel appears through lines of constant u . If the t -channel process produces bound states and resonances, they will form lines of constant t . Thus, the angular dependence of the amplitude at fixed s will be influenced by singularities in crossed channels.

In **deep inelastic scattering** $eN \rightarrow eX$, $m = m'$ is the electron mass and M the nucleon mass, whereas $W = M'$ is the total invariant mass of the particles in the final state X , which is not fixed (also called the ‘missing mass’). The process is described by the three variables $\{s, t, u\}$ or, equivalently, τ, λ and any of the variables

$$\omega = \frac{W^2 - M^2}{4M^2}, \quad \nu = \frac{p_i \cdot q}{M} = 2M(\tau + \omega), \quad x = -\frac{q^2}{2M\nu} = \frac{\tau}{\tau + \omega}. \quad (4.5.16)$$

With $W \geq M$, the **Bjorken variable** $0 < x \leq 1$ reduces to $x = 1$ for elastic scattering. In Sec. 5.1 we will see that for a one-photon exchange interaction the hadronic part $\gamma^*N \rightarrow X$ does not depend on λ but only on τ and ω . It is parametrized by the nucleon’s **structure functions**, which are usually expressed in terms of $\{\tau, x\}$ or $\{\nu, x\}$.

Finally, in processes such as Compton scattering or photo- and electroproduction, the onshell nucleon has mass $M = M'$ whereas m^2 or m'^2 (or both of them) can be virtual. The amplitude then depends on four variables, for example $\{s, t, u, \omega'\}$ or $\{\tau, \lambda, m^2, m'^2\}$, where ω' is related to the ‘skewness’ variable.

So far we have not specified any Lorentz frame because the above variables are Lorentz-invariant. Their interpretation in terms of energies and a scattering angle, however, depends on the reference frame. In the following we work out the kinematics in the center-of mass and laboratory frames.

■ In the s -channel **center-of-mass (CM) frame** the spatial component of the total momentum $p_i + k_i = p_f + k_f$ vanishes:

$$p_i = \begin{pmatrix} \varepsilon \\ \mathbf{k} \end{pmatrix}, \quad k_i = \begin{pmatrix} E \\ -\mathbf{k} \end{pmatrix}, \quad p_f = \begin{pmatrix} \varepsilon' \\ \mathbf{k}' \end{pmatrix}, \quad k_f = \begin{pmatrix} E' \\ -\mathbf{k}' \end{pmatrix}. \quad (4.5.17)$$

For fixed masses, only two of the variables in (4.5.17) are independent, for example the CM momentum $|\mathbf{k}|$ and the scattering angle defined by $\mathbf{k} \cdot \mathbf{k}' = |\mathbf{k}||\mathbf{k}'| \cos \theta_{\text{CM}}$, which can be related to the Mandelstam variables s and t . The variable $s = (p_i + k_i)^2 = (\varepsilon + E)^2 = (\varepsilon' + E')^2$ is the total CM energy squared. From the mass-shell conditions one can express all energies in terms of s ,

$$\varepsilon = \frac{s + M^2 - m^2}{2\sqrt{s}}, \quad E = \frac{s - M^2 + m^2}{2\sqrt{s}}, \quad \varepsilon' = \frac{s + M'^2 - m'^2}{2\sqrt{s}}, \quad E' = \frac{s - M'^2 + m'^2}{2\sqrt{s}} \quad (4.5.18)$$

as well as the three-momenta:

$$\mathbf{k}^2 = \varepsilon^2 - M^2 = E^2 - m^2 = \frac{[s - (M + m)^2][s - (M - m)^2]}{4s} = \frac{\bar{\lambda}(s, M^2, m^2)}{4s}, \quad (4.5.19)$$

$$\mathbf{k}'^2 = \varepsilon'^2 - M'^2 = E'^2 - m'^2 = \frac{[s - (M' + m')^2][s - (M' - m')^2]}{4s} = \frac{\bar{\lambda}(s, M'^2, m'^2)}{4s}.$$

The **triangle function** $\bar{\lambda}(x, y, z)$ is defined as

$$\bar{\lambda}(x, y, z) = x^2 + y^2 + z^2 - 2xy - 2yz - 2xz \quad (4.5.20)$$

and invariant under permutation of its arguments. The variable t is given by

$$t = (p_i - p_f)^2 = (\varepsilon - \varepsilon')^2 - (\mathbf{k} - \mathbf{k}')^2 = M^2 + M'^2 - 2\varepsilon\varepsilon' + 2|\mathbf{k}||\mathbf{k}'| \cos \theta_{\text{CM}}, \quad (4.5.21)$$

from where one can relate the scattering angle to s and t :

$$\cos \theta_{\text{CM}} = \frac{s^2 + 2st - s(M^2 + M'^2 + m^2 + m'^2) + (M^2 - m^2)(M'^2 - m'^2)}{\sqrt{\bar{\lambda}(s, M^2, m^2)\bar{\lambda}(s, M'^2, m'^2)}}. \quad (4.5.22)$$

For **elastic scattering** with $M = M'$, $m = m'$ these relations simplify to $\varepsilon = \varepsilon'$, $E = E'$, $\mathbf{k}^2 = \mathbf{k}'^2$,

$$\cos \theta_{\text{CM}} = 1 + \frac{2st}{\bar{\lambda}(s, M^2, m^2)}, \quad \bar{\lambda}(s, M^2, m^2) = (s - M^2 - m^2)^2 - 4M^2m^2 \quad (4.5.23)$$

and the Mandelstam variables in terms of \mathbf{k}^2 and $\cos\theta_{\text{CM}}$ become

$$\begin{aligned} s &= \left(\sqrt{\mathbf{k}^2 + M^2} + \sqrt{\mathbf{k}^2 + m^2} \right)^2, \\ t &= -2\mathbf{k}^2 (1 - \cos\theta_{\text{CM}}), \\ u &= -2\mathbf{k}^2 (1 + \cos\theta_{\text{CM}}) + \left(\sqrt{\mathbf{k}^2 + M^2} - \sqrt{\mathbf{k}^2 + m^2} \right)^2. \end{aligned} \quad (4.5.24)$$

If all masses are equal ($m = M$), we obtain

$$\varepsilon = E = \frac{\sqrt{s}}{2}, \quad \mathbf{k}^2 = \frac{s - 4M^2}{4}, \quad \cos\theta_{\text{CM}} = 1 + \frac{2t}{s - 4M^2}, \quad \bar{\lambda}(s, M^2, M^2) = s(s - 4M^2) \quad (4.5.25)$$

and the Mandelstam variables become

$$s = 4(\mathbf{k}^2 + M^2), \quad t = -2\mathbf{k}^2(1 - \cos\theta_{\text{CM}}), \quad u = -2\mathbf{k}^2(1 + \cos\theta_{\text{CM}}). \quad (4.5.26)$$

Here, forward scattering ($\theta_{\text{CM}} = 0$) corresponds to $t = 0$ and backward scattering ($\theta_{\text{CM}} = \pi$) to $u = 0$.

■ Next, we consider the **lab frame** where p_i is at rest:

$$p_i = \begin{pmatrix} M \\ \mathbf{0} \end{pmatrix}, \quad k_i = \begin{pmatrix} E \\ \mathbf{k} \end{pmatrix}, \quad p_f = \begin{pmatrix} \varepsilon' \\ \mathbf{p}' \end{pmatrix}, \quad k_f = \begin{pmatrix} E' \\ \mathbf{k}' \end{pmatrix}. \quad (4.5.27)$$

We use the same symbols E , E' , ε , \mathbf{k} and \mathbf{k}' as before, but keep in mind that these quantities are not the same as in Eq. (4.5.17) since they are the energies and three-momenta in the lab frame. The scattering angle in the lab frame is defined by $\mathbf{k} \cdot \mathbf{k}' = |\mathbf{k}||\mathbf{k}'| \cos\theta$. In this case we have

$$\begin{aligned} s &= (p_i + k_i)^2 = (M + E)^2 - \mathbf{k}^2 = M^2 + m^2 + 2ME, \\ t &= (p_f - p_i)^2 = (\varepsilon' - M)^2 - \mathbf{p}'^2 = M^2 + M'^2 - 2M\varepsilon', \\ u &= (p_i - k_f)^2 = (M - E')^2 - \mathbf{k}'^2 = M^2 + m'^2 - 2ME', \end{aligned} \quad (4.5.28)$$

from where we can relate the energies to the Mandelstam variables:

$$E = \frac{s - M^2 - m^2}{2M}, \quad E' = \frac{M^2 + m'^2 - u}{2M}, \quad \varepsilon' = \frac{M^2 + M'^2 - t}{2M}. \quad (4.5.29)$$

The three-momenta are then given by

$$\begin{aligned} \mathbf{k}^2 &= E^2 - m^2 = \frac{\bar{\lambda}(s, M^2, m^2)}{4M^2}, \\ \mathbf{k}'^2 &= E'^2 - m'^2 = \frac{\bar{\lambda}(u, M^2, m'^2)}{4M^2}, \\ \mathbf{p}'^2 &= \varepsilon'^2 - M'^2 = \frac{\bar{\lambda}(t, M^2, M'^2)}{4M^2}. \end{aligned} \quad (4.5.30)$$

The Mandelstam variables are related by $s + t + u = M^2 + M'^2 + m^2 + m'^2$. The scattering angle in the lab frame can be worked out using

$$t = (k_i - k_f)^2 = (E - E')^2 - (\mathbf{k} - \mathbf{k}')^2 = m^2 + m'^2 - 2EE' + 2|\mathbf{k}||\mathbf{k}'| \cos\theta, \quad (4.5.31)$$

from where we obtain

$$\cos\theta = \frac{2M^2(t - m^2 - m'^2) - (s - M^2 - m^2)(u - M^2 - m'^2)}{\sqrt{\bar{\lambda}(s, M^2, m^2)\bar{\lambda}(u, M^2, m'^2)}}. \quad (4.5.32)$$

■ A practical example for the case $m = m' = 0$ is (elastic or inelastic) **eN scattering** with a nucleon mass M and electron mass $m = m' \ll M$. The lab frame is the natural frame for fixed-target experiments, where the experimental control parameters are the initial and final lepton energies E , E' and the scattering angle θ . In this case the above relations become

$$E = |\mathbf{k}| = \frac{s - M^2}{2M}, \quad E' = |\mathbf{k}'| = \frac{M^2 - u}{2M}, \quad \varepsilon' = \frac{M^2 + M'^2 - t}{2M}, \quad \mathbf{p}'^2 = \frac{\lambda(t, M^2, M'^2)}{4M^2}, \quad (4.5.33)$$

and the scattering angle is given by

$$\sin^2 \frac{\theta}{2} = \frac{1 - \cos \theta}{2} = \frac{M^2 t}{(s - M^2)(u - M^2)}. \quad (4.5.34)$$

Expressing the Mandelstam variables through τ , λ and ω defined in Eqs. (4.5.8–4.5.9),

$$\begin{aligned} s &= M^2 [1 + 2(\tau + \omega + \lambda)], & t &= -4M^2 \tau \\ u &= M^2 [1 + 2(\tau + \omega - \lambda)], \end{aligned} \quad (4.5.35)$$

we find

$$\begin{aligned} E &= M(\lambda + \tau + \omega), & \sin^2 \frac{\theta}{2} &= \frac{\tau}{\lambda^2 - (\tau + \omega)^2}. \\ E' &= M(\lambda - \tau - \omega), \end{aligned} \quad (4.5.36)$$

The condition (4.5.12) for the physical region becomes

$$\Phi = t(su - M^2 M'^2) = -16M^4 \tau [\tau + (\tau + \omega)^2 - \lambda^2] \geq 0. \quad (4.5.37)$$

From the inverse relations

$$\lambda = \frac{E + E'}{2M}, \quad \tau = \frac{EE'}{M^2} \sin^2 \frac{\theta}{2}, \quad \omega + \tau = \frac{E - E'}{2M} \quad (4.5.38)$$

we see that the crossing variable is proportional to the average lepton energy. The variable ν from Eq. (4.5.16) plays the role of the energy transfer from the electron to the proton, $\nu = E - E'$, and another commonly used variable is

$$y = \frac{p_i \cdot q}{p_i \cdot k_i} = \frac{2(\tau + \omega)}{\lambda + \tau + \omega - \omega'} = 1 - \frac{E'}{E}, \quad (4.5.39)$$

which becomes the lepton energy loss ($0 \leq y \leq 1$) in the lab frame. In the case of elastic scattering ($\omega = 0$) there are only two independent variables and we have the additional constraint

$$E' = \frac{E}{1 + \frac{2E}{M} \sin^2 \frac{\theta}{2}}. \quad (4.5.40)$$

■ We also work out the **cross section** for eN scattering. The general form of the cross section for $2 \rightarrow n$ -particle scattering has the form

$$d\sigma = \frac{|\mathcal{M}|^2 d\Phi}{4\sqrt{(p_i \cdot k_i)^2 - M^2 m^2}}, \quad (4.5.41)$$

where $|\mathcal{M}|^2$ is the invariant amplitude, $d\Phi$ is the phase space element and the denominator is the incoming flux factor. For two particles in the final state, the phase space is given by

$$d\Phi = \frac{d^3 p_f}{(2\pi)^3 2\varepsilon'} \frac{d^3 k_f}{(2\pi)^3 2E'} (2\pi)^4 \delta^4(p_i + k_i - k_f - p_f), \quad (4.5.42)$$

where p_f and k_f are the outgoing momenta and ε' and E' their energies in the lab frame, cf. (4.5.27). Integration over $d^3 p_f$ removes the three-dimensional δ -function for three-momentum conservation. For vanishing electron masses, inserting $d^3 k_f = dE' E'^2 d\Omega$ yields

$$d\Phi = \frac{d\Omega}{(4\pi)^2} \frac{E'}{\varepsilon'} dE' \delta(M + E - E' - \varepsilon'). \quad (4.5.43)$$

We can express the final-state energy by $\varepsilon' = \sqrt{\mathbf{q}^2 + W^2} = \sqrt{\mathbf{q}^2 + M^2 + 4M^2\omega}$, where for elastic scattering the energy-conservation constraint is satisfied for $\omega = 0$. Hence, we can rewrite the δ -function in the variable ω :

$$\delta(M + E - E' - \varepsilon') = \frac{\varepsilon'}{2M^2} \delta(\omega) \quad \Rightarrow \quad d\Phi = \frac{d\Omega}{(4\pi)^2} \frac{E'}{2M^2} dE' \delta(\omega). \quad (4.5.44)$$

On the other hand, we have

$$\int dE' f(E') \frac{\delta(\omega)}{2M} = \frac{f(E')}{2M \left| \frac{d\omega}{dE'} \right|_{\omega=0}} \stackrel{(4.5.38)}{=} \frac{f(E')}{1 + \frac{2E}{M} \sin^2 \frac{\theta}{2}} \Big|_{\omega=0} \stackrel{(4.5.40)}{=} f(E') \frac{E'}{E}. \quad (4.5.45)$$

Combining this with the flux factor $4p_i \cdot k_i = 4ME$, we arrive at

$$\frac{d\sigma}{d\Omega} = \frac{1}{4M^2} \frac{|\mathcal{M}|^2}{16\pi^2} \frac{E'^2}{E^2}. \quad (4.5.46)$$

4.5.2 Form factors

We already motivated the basic ideas behind form factors from an experimental point of view in Sec. 4.5.1. When resolving the structure of a hadron using electromagnetic and weak interactions, the elementary structure observables are form factors. Viewed from a distance, the proton looks like a point fermion that only carries a charge and a magnetic moment, but when probed with short-wavelength photons (or other currents) it reveals more and more of its composite nature which is encoded in the momentum dependence of its form factors.

Form factors are encoded in the **current matrix elements**

$$\langle \lambda'(p_f) | j_a^\Gamma(0) | \lambda(p_i) \rangle, \quad (4.5.47)$$

where $|\lambda(p_i)\rangle$ is a one-particle state with onshell momentum p_i and $|\lambda'(p_f)\rangle$ one with momentum p_f . The currents $j_a^\Gamma(x)$ can be any of the quark bilinears in Eq. (3.1.23) such as vector, axialvector, scalar or pseudoscalar currents. On theoretical grounds, a current matrix element can be motivated in several ways:

- In Eq. (3.1.147) we saw that current matrix elements arise from elementary correlation functions at the pole positions, i.e., they are the residues at the hadronic double pole. This also gives us an intuitive way to understand **elastic** and **transition form factors**: in the second case, the initial and final hadrons can be different and correspond to different poles, provided that the symmetries in the process (e.g. baryon number conservation) are preserved.

- In Eq. (3.1.141) we found that the contraction of the Bethe-Salpeter wave function with a Dirac-flavor matrix $\Gamma \mathbf{t}_a$ gives rise to decay constants $\langle 0 | j^\Gamma(0) | \lambda(p) \rangle$, which are gauge invariant and depend on the onshell momentum p . In the same way, a current matrix element arises from the contraction of the object $\langle \lambda'(p_f) | \mathbb{T} \psi_\alpha(x) \bar{\psi}_\beta(y) | \lambda(p_i) \rangle$ with open quark and antiquark legs:

$$\begin{aligned} & - (\mathbf{t}_a)_{ji} \Gamma_{\beta\alpha} \langle \lambda'(p_f) | \mathbb{T} \psi_{\alpha i}(x) \bar{\psi}_{\beta j}(x) | \lambda(p_i) \rangle \\ & = \langle \lambda'(p_f) | j_a^\Gamma(x) | \lambda(p_i) \rangle = \langle \lambda'(p_f) | j_a^\Gamma(0) | \lambda(p_i) \rangle e^{iq \cdot x}, \end{aligned} \quad (4.5.48)$$

where $q = p_f - p_i$. Thus, the current couples to the quarks inside the hadrons as shown in Fig. 4.18. The resulting matrix element is again gauge invariant and depends on the two onshell momenta p_f and p_i .

The three-point function depends on two independent momenta, e.g. the incoming and outgoing momenta p_i and p_f , or their combinations $p = (p_i + p_f)/2$ and the momentum transfer $q = p_f - p_i$. For elastic form factors we have $p_i^2 = p_f^2 = M^2$ and therefore

$$\begin{aligned} p_f = p + \frac{q}{2} & \Rightarrow p_f^2 = p^2 + \frac{q^2}{4} + p \cdot q \stackrel{!}{=} M^2 & \Rightarrow p \cdot q = 0 \\ p_i = p - \frac{q}{2} & \Rightarrow p_i^2 = p^2 + \frac{q^2}{4} - p \cdot q \stackrel{!}{=} M^2 & \Rightarrow p^2 = M^2 - \frac{q^2}{4}. \end{aligned} \quad (4.5.49)$$

Thus, the only independent variable is the squared momentum transfer q^2 . Because $q^2 < 0$ is spacelike in s -channel processes such as eN scattering, we work with the spacelike momentum transfer $Q^2 = -q^2$ or equivalently $\tau = Q^2/(4M^2)$.

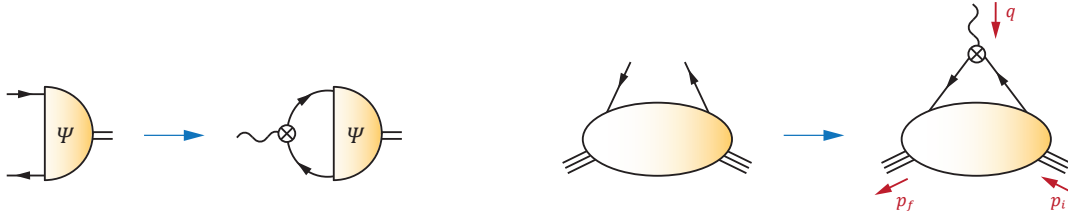


FIG. 4.18: Current matrix element of a hadron, viewed as the Dirac-flavor contraction of the four-point function in Eq. (4.5.48) in analogy to the Bethe-Salpeter wave function.

Form factors. As a specific case of Eq. (4.5.2), form factors are the Lorentz-invariant coefficients of current matrix elements. As an example, consider a spin- $\frac{1}{2}$ particle and a vector current $V^\mu = \bar{\psi}\gamma^\mu\psi$. In this case, the general decomposition involves the **Dirac** and **Pauli form factors** $F_1(q^2)$ and $F_2(q^2)$:

$$\langle p_f, s' | V^\mu(0) | p_i, s \rangle = \bar{u}_{s'}(p_f) \left[\gamma^\mu F_1(q^2) + \sigma^{\mu\nu} \frac{iq_\nu}{2M} F_2(q^2) \right] u_s(p_i). \quad (4.5.50)$$

Here $\sigma^{\mu\nu} = \frac{i}{2} [\gamma^\mu, \gamma^\nu]$ and $u_s(p_i)$, $\bar{u}_{s'}(p_f)$ are the onshell Dirac spinors with normalization $\bar{u}_{s'}(p) u_s(p) = 2M \delta_{ss'}$. The dimensionless form factors depend only on q^2 . For two flavors, the **isoscalar** and **isovector** form factors correspond to the currents

$$\begin{aligned} V^\mu &= \bar{\psi}\gamma^\mu\psi = \bar{u}\gamma^\mu u + \bar{d}\gamma^\mu d, \\ V_3^\mu &= \bar{\psi}\gamma^\mu \mathbf{t}_3 \psi = \frac{1}{2} (\bar{u}\gamma^\mu u - \bar{d}\gamma^\mu d) \end{aligned} \quad (4.5.51)$$

and the **electromagnetic form factors**, which are linear combinations of them, to the electromagnetic current V_{em}^μ from Eq. (3.1.92) with the quark charge matrix

$$\mathbf{Q} = \begin{pmatrix} q_u & 0 \\ 0 & q_d \end{pmatrix} = \frac{\mathbb{1}}{6} + \frac{\tau_3}{2}. \quad (4.5.52)$$

Why are there just *two* form factors in Eq. (4.5.50)? Consider the most generic form of a vector-spinor three-point function $\Omega^\mu(p, q)$. Poincaré covariance and parity invariance in principle allows for 12 tensor structures, for example

$$\{\gamma^\mu, p^\mu, q^\mu\} \times \{\mathbb{1}, \not{p}, \not{q}, [\not{p}, \not{q}]\} \quad (4.5.53)$$

or linear combinations of those. After sandwiching between the onshell nucleon spinors $\bar{u}(p_f)$ and $u(p_i)$, we can use the Dirac equation $(\not{p} - M) u(p) = 0$ to eliminate all slashes:

$$\begin{aligned} \bar{u}(p_f) \not{q} u(p_i) &= \bar{u}(p_f) (\not{p}_f - \not{p}_i) u(p_i) = 0, \\ \bar{u}(p_f) \not{p} u(p_i) &= \bar{u}(p_f) \frac{\not{p}_f + \not{p}_i}{2} u(p_i) = M \bar{u}(p_f) u(p_i), \\ \bar{u}(p_f) [\gamma^\mu, \not{q}] u(p_i) &= 4 \bar{u}(p_f) (p^\mu - M\gamma^\mu) u(p_i), \end{aligned} \quad (4.5.54)$$

where the last relation is the **Gordon identity**. As a result, we are left with γ^μ , p^μ and q^μ only.

In addition, charge conjugation imposes the condition

$$C \Omega^\mu(-p, q)^T C^T \stackrel{!}{=} -\Omega^\mu(p, q), \quad (4.5.55)$$

where $C = i\gamma^2\gamma^0$ is the charge-conjugation matrix. This is satisfied for γ^μ and p^μ but not for q^μ , which has opposite C -parity. To restore it, we would need to attach a factor $p \cdot q$ so that $(p \cdot q) q^\mu$ becomes the third basis element, but $p \cdot q = 0$ because the nucleon is onshell, cf. Eq. (4.5.49). Hence, a vector current matrix element can only depend on γ^μ and p^μ . Finally, we use the Gordon identity to express p^μ in terms of γ^μ and $\sigma^{\mu\nu}q_\nu = \frac{i}{2}[\gamma^\mu, \not{q}]$, which leads to the form in (4.5.50).

The same principles can be used to establish the matrix elements of an axialvector current $A^\mu(0)$, a pseudoscalar density $P(0)$ and a scalar density $S(0)$. In these cases, the bracket in (4.5.50) must be replaced with

$$\gamma^\mu \gamma_5 G_A(q^2) + \gamma_5 \frac{q^\mu}{2M} G_P(q^2), \quad G_5(q^2) i\gamma_5, \quad G_S(q^2), \quad (4.5.56)$$

respectively. $G_A(q^2)$ is the **axial form factor** and $G_P(q^2)$ the ‘induced’ pseudoscalar form factor of a spin-1/2 baryon. In the limit $q^2 \rightarrow 0$, the axial form factor becomes the **axial charge** $g_A = G_A(0)$, whose experimental value $g_A \approx 1.27$ for the nucleon is known from neutron beta decay. $G_5(q^2)$ and $G_S(q^2)$ are the pseudoscalar and scalar form factors.

One can also write down current matrix elements for baryons with higher spin (e.g. for $J = 3/2$ the Dirac spinors must be replaced by **Rarita-Schwinger spinors**), which produces more tensors and thus more form factors, or transition matrix elements between baryons with different spins, or meson form factors, etc.

Current conservation. Next, we want to work out the implications of current conservation for the matrix elements. Vector current conservation $\partial_\mu V^\mu = 0$ implies

$$\partial_\mu \langle \lambda' | V^\mu(x) | \lambda \rangle = \langle \lambda' | V^\mu(0) | \lambda \rangle \partial_\mu e^{iq \cdot x} = iq_\mu \langle \lambda' | V^\mu(0) | \lambda \rangle e^{iq \cdot x} \stackrel{!}{=} 0, \quad (4.5.57)$$

which means that the vector current matrix element must be transverse with respect to the momentum transfer q_μ . Eq. (4.5.50) already satisfies that constraint because $\sigma^{\mu\nu}q_\mu q_\nu = 0$ and $\bar{u}(p_f) \not{q} u(p_i) = 0$, so this does not impose any constraints on the Dirac and Pauli form factors.

In the axialvector case, the PCAC relation $\partial_\mu A^\mu = 2mP$ tells us that

$$\partial_\mu \langle \lambda' | A^\mu(x) | \lambda \rangle = iq_\mu \langle \lambda' | A^\mu(0) | \lambda \rangle e^{iq \cdot x} \stackrel{!}{=} 2m \langle \lambda' | P(0) | \lambda \rangle e^{iq \cdot x}. \quad (4.5.58)$$

Inserted into the matrix elements and using Eq. (4.5.56), this entails

$$\cdots \left[i\not{q}\gamma_5 G_A(q^2) + i\gamma_5 \frac{q^2}{2M} G_P(q^2) \right] \cdots = \cdots [2m G_5(q^2) i\gamma_5] \cdots \quad (4.5.59)$$

and with $\not{q}\gamma_5 = \not{p}_f \gamma_5 + \gamma_5 \not{p}_i \cong 2M\gamma_5$ using the Dirac equation, we find that the axial and pseudoscalar form factors are related:

$$G_A(q^2) + \frac{q^2}{4M^2} G_P(q^2) \stackrel{!}{=} \frac{m}{M} G_5(q^2). \quad (4.5.60)$$

It appears that g_A should vanish in the chiral limit $m \rightarrow 0$, but this is not the case because the pseudoscalar form factor contains pion poles. The four-point function in Fig. 4.18 must develop meson poles in the t channel for $q^2 = m_\lambda^2$, since $N\bar{N}$ and $q\bar{q}$ are compatible with meson quantum numbers. According to Eq. (3.1.121), the residue involving the $q\bar{q}$ pair is the Bethe-Salpeter wave function of the respective meson and the residue on the $N\bar{N}$ pair is the nucleon-meson coupling constant. Contracted with γ_5 , this only leaves pseudoscalar-meson poles whose residues r_λ we defined in Eq. (3.1.142). Thus, at the pion pole $G_5(q^2)$ must have the form

$$G_5(q^2) = -\frac{2r_\pi}{q^2 - m_\pi^2} G_{\pi NN}(q^2) \stackrel{(3.1.143)}{=} -\frac{m_\pi^2}{q^2 - m_\pi^2} \frac{f_\pi}{m} G_{\pi NN}(q^2), \quad (4.5.61)$$

which is illustrated in Fig. 4.19. Here we defined an effective pion-nucleon form factor $G_{\pi NN}(q^2)$, which absorbs all further pseudoscalar pole contributions and non-resonant terms and reduces to the pion-nucleon coupling constant $G_{\pi NN}(q^2 = m_\pi^2) = g_{\pi NN}$ at the pion pole. The factor 2 accounts for $G_5 \tau_a \sim 2 G_{\pi NN} \tau_a = G_{\pi NN} \tau_a$ and the minus sign makes $G_5(q^2 < 0)$ positive. Combined with Eq. (4.5.60), we arrive at the **Goldberger-Treiman relation**

$$g_A = \frac{f_\pi}{M} G_{\pi NN}(0) \xrightarrow{\text{chiral limit}} \frac{f_\pi}{M} g_{\pi NN}, \quad (4.5.62)$$

which connects the nucleon's axial charge with the pion-nucleon coupling. $G_{\pi NN}(0)$ is not measurable in contrast to $g_{\pi NN} \approx 13.2$, which is the residue at the physical pion mass. Together with the experimental values for $f_\pi \approx 92$ MeV, $M \approx 940$ MeV and $g_A \approx 1.27$, the Goldberger-Treiman relation is well realized in nature.

Meson resonances. The appearance of t -channel meson poles in form factors has far-reaching consequences for their analytic structure. Timelike poles appear not only in the pseudoscalar form factor $G_5(q^2)$, but also in

- the Dirac and Pauli vector form factors $F_1(q^2)$ and $F_2(q^2)$, which have 1^{--} vector-meson poles e.g. at $q^2 = m_\rho^2$ or m_ω^2 (depending on the isovector/isoscalar channel) and whose residues are the products of the ρ/ω -nucleon couplings combined with the ρ/ω -meson decay constants;
- the axial form factor $G_A(q^2)$ which has axialvector 1^{++} poles,
- the scalar form factor with scalar poles 0^{++} , etc.

The singularity structure is independent of the hadron because microscopically it originates from the vertex that describes the coupling of the current to the quarks, like the quark-photon vertex discussed in Eq. (3.1.146) in the vector case.

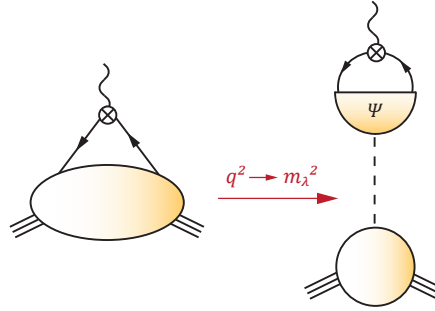


FIG. 4.19: Form factors have meson poles in the t channel.

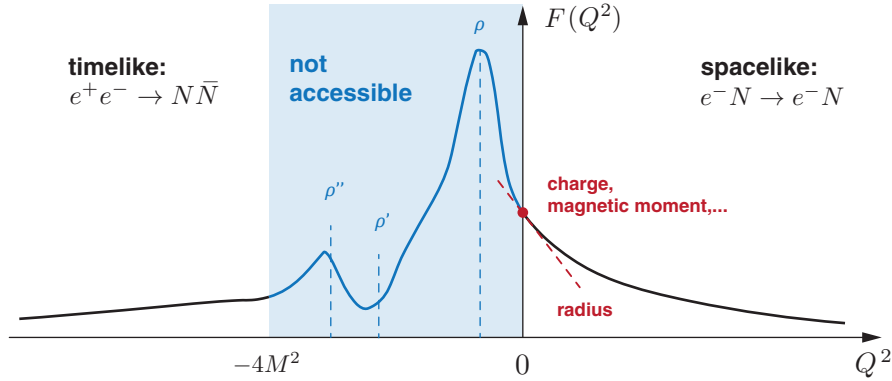


FIG. 4.20: Sketch of a nucleon electromagnetic form factor containing ρ -meson bumps. The fictitious curve in the unphysical window is based on the pole structure in the (measured) pion electromagnetic form factor.

In fact, since only the pion is stable with respect to the strong interaction, all other mesons have non-zero hadronic decay widths. Their poles must then move into the complex q^2 plane onto higher Riemann sheets and only produce bumps on the timelike q^2 axis. The respective branch cuts are generated by intermediate multiparticle states containing two pions ($\rho \rightarrow \pi\pi$), three pions ($\omega \rightarrow \pi\pi\pi$), $K\bar{K}$, etc.

The situation is illustrated in Fig. 4.20 for a generic elastic nucleon vector-isovector form factor with ρ -meson bumps. A similar picture with appropriate J^{PC} poles would arise for other types of form factors as well. The form factor's momentum dependence in the spacelike domain ($Q^2 = -q^2 > 0$) can be extracted from elastic electron-nucleon scattering as long as the one-photon exchange process is dominant (more on that below). The timelike region above $p\bar{p}$ production threshold ($q^2 > 4M^2$) can be accessed in e^+e^- annihilation. However, meson resonances should be most pronounced in the window $q^2 \sim 0 \dots 4 \text{ GeV}^2$ which is experimentally not accessible; in the deep timelike region the resonance peaks are already washed out. Fortunately, precise data are available for the *pion* electromagnetic form factor which should display a similar resonance structure as in the nucleon case. Here the unphysical window is much smaller ($q^2 = 0 \dots 4m_\pi^2 \approx 0.08 \text{ GeV}^2$) and the resonance peaks are indeed directly visible in the data, with a similar shape as in Fig. 4.20.

The timelike resonance structure can be connected with the spacelike behavior of the form factors through **dispersion relations**. Like physical scattering amplitudes, form factors must be analytic everywhere in the complex Q^2 plane except for branch-point singularities starting at $q^2 = 4m_\pi^2$ and extending to infinity, which are due to intermediate two-pion and multiparticle states. The Cauchy formula then tells us that the form factor in the domain of analyticity can be inferred from knowledge of its value on a closed contour, which can be deformed to encompass only the branch cut (see Fig. 4.21). Since the form factor is analytic everywhere else, the difference above and below the branch cut is proportional to its imaginary part, i.e., the discontinuity along the branch cut:

$$F(z_0) = \frac{1}{2\pi i} \oint dz \frac{F(z)}{z - z_0} = \frac{1}{\pi} \int_{4m_\pi^2}^{\infty} dz \frac{\text{Im } F(z)}{z - z_0}. \quad (4.5.63)$$

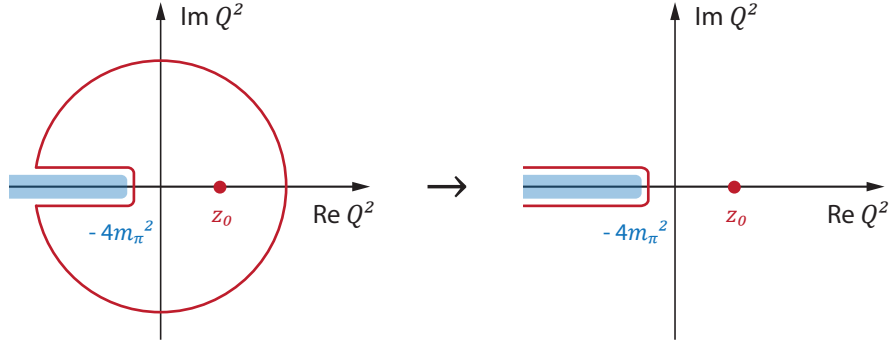


FIG. 4.21: Analytic structure of the form factor $F(Q^2)$ in the complex Q^2 plane and deformation of the integration contour.

Hence, knowledge of the spectral function $\text{Im} F(z)$ along the cut is sufficient to determine the spacelike form factor as well. On the other hand, since the experimental knowledge is limited to $q^2 > 4M^2 \sim 4 \text{ GeV}^2$, one usually has to make assumptions about the timelike behavior to extract such information.

Cross section for elastic eN scattering. The nucleon's electromagnetic form factors in the spacelike region $Q^2 \geq 0$ are experimentally extracted from elastic eN scattering. Since the process is reasonably well described by one-photon exchange, we start from the amplitude (4.1.3) for scattering leptons from a point-like Dirac particle through one-photon exchange (Born approximation):

$$\mathcal{M}_{\sigma\sigma'\lambda\lambda'}(q, p, k) = \frac{e^2}{q^2} \bar{u}_{\sigma'}(p_f) \gamma^\mu u_\sigma(p_i) \bar{u}_{\lambda'}(k_f) \gamma_\mu u_\lambda(k_i). \quad (4.5.64)$$

We worked out the kinematic variables in Eqs. (4.5.4–4.5.9); in particular, since the nucleon and electron scatter elastically, we have $M = M'$, $m = m'$ and therefore

$$p^2 = M^2 - \frac{q^2}{4}, \quad p \cdot q = 0, \quad k^2 = m^2 - \frac{q^2}{4}, \quad k \cdot q = 0 \quad (4.5.65)$$

so that only $\tau = Q^2/(4M^2)$ and $\lambda = p \cdot k/M^2$ remain independent variables. For unpolarized scattering, we take the spin average

$$|\mathcal{M}|^2 = \frac{1}{4} \sum_{\text{spins}} |\mathcal{M}|^2 = \frac{e^4}{q^4} L^{\mu\nu} W_{\mu\nu} \quad (4.5.66)$$

which factorizes into a leptonic and a hadronic part. The lepton tensor has the form

$$\begin{aligned} L^{\mu\nu} &= \frac{1}{2} \sum_{\lambda\lambda'} \bar{u}_{\lambda'}(k_f) \gamma^\mu u_\lambda(k_i) \bar{u}_\lambda(k_i) \gamma^\nu u_{\lambda'}(k_f) = \\ &= \frac{1}{2} \text{Tr}[(\not{k}_f + m) \gamma^\mu (\not{k}_i + m) \gamma^\nu] \\ &= 2 \left(k_f^\mu k_i^\nu + k_i^\mu k_f^\nu - k_i \cdot k_f g^{\mu\nu} + m^2 g^{\mu\nu} \right) \\ &= 4k^\mu k^\nu - q^\mu q^\nu + 2 \left(\frac{q^2}{4} - k^2 + m^2 \right) g^{\mu\nu} = 4 \left(k^\mu k^\nu + \frac{q^2}{4} T_q^{\mu\nu} \right), \end{aligned} \quad (4.5.67)$$

where in the final step we used the transverse projector

$$T_q^{\mu\nu} = g^{\mu\nu} - \frac{q^\mu q^\nu}{q^2}. \quad (4.5.68)$$

Because $k \cdot q = 0$, the lepton tensor is transverse with respect to the photon momentum in both Lorentz indices, which reflects the conservation of the leptonic vector current. For elastic scattering on the hadron side, the hadronic tensor for a structureless fermion has the analogous form

$$\begin{aligned} W^{\mu\nu} &= \frac{1}{2} \sum_{\sigma\sigma'} \bar{u}_{\sigma'}(p_f) \gamma^\mu u_\sigma(p_i) \bar{u}_\sigma(p_i) \gamma^\nu u_{\sigma'}(p_f) = \\ &= \frac{1}{2} \text{Tr} [(\not{p}_f + M) \gamma^\mu (\not{p}_i + M) \gamma^\nu] \\ &= 2 \left(p_f^\mu p_i^\nu + p_i^\mu p_f^\nu - p_i \cdot p_f g^{\mu\nu} + M^2 g^{\mu\nu} \right) \\ &= 4p^\mu p^\nu - q^\mu q^\nu + 2 \left(\frac{q^2}{4} - p^2 + M^2 \right) g^{\mu\nu} = 4 \left(p^\mu p^\nu + \frac{q^2}{4} T_q^{\mu\nu} \right), \end{aligned} \quad (4.5.69)$$

which is again transverse in both Lorentz indices. Using $T_q^{\mu\nu} T_{\mu\nu,q} = 3$ and neglecting the small electron mass, their combination becomes

$$L^{\mu\nu} W_{\mu\nu} = 16 \left[(p \cdot k)^2 + \frac{q^2}{4} (k^2 + p^2) + 3 \frac{q^4}{16} \right] = 16M^4 (\lambda^2 + \tau^2 - \tau), \quad (4.5.70)$$

and with $e^2 = 4\pi\alpha$ the result for the invariant squared amplitude is

$$|\mathcal{M}|^2 = \frac{e^4}{q^4} L^{\mu\nu} W_{\mu\nu} = \frac{16\pi^2\alpha^2}{\tau^2} (\lambda^2 + \tau^2 - \tau). \quad (4.5.71)$$

We already worked out the cross section for elastic eN scattering in Eq. (4.5.46),

$$\frac{d\sigma}{d\Omega} = \frac{1}{4M^2} \frac{|\mathcal{M}|^2}{16\pi^2} \frac{E'^2}{E^2}, \quad (4.5.72)$$

which is expressed through the initial and final lepton energies E, E' in the lab frame. Plugging in the result for $|\mathcal{M}|^2$, the differential cross section becomes

$$\frac{d\sigma}{d\Omega} = \frac{\alpha^2}{4M^2\tau^2} \frac{E'^2}{E^2} (\lambda^2 + \tau^2 - \tau) = \underbrace{\frac{\alpha^2 \cos^2 \frac{\theta}{2}}{4E^2 \sin^4 \frac{\theta}{2}} \frac{E'}{E}}_{\text{Mott}} \left(1 + 2\tau \tan^2 \frac{\theta}{2} \right), \quad (4.5.73)$$

where we exploited the relations (4.5.36–4.5.38) with $\omega = 0$ to arrive at the second form. The **Mott cross section** describes lepton scattering off a pointlike scalar particle in Born approximation. The parenthesis reflects the nucleon's nature as a spin- $\frac{1}{2}$ particle, which at this point carries no internal structure.

To take the composite nature of the nucleon into account, we must replace the pointlike Dirac current with the general current matrix element

$$\bar{u}(p_f) \gamma^\mu u(p_i) \longrightarrow \bar{u}(p_f) \left(\gamma^\mu F_1(q^2) + \sigma^{\mu\nu} \frac{iq_\nu}{2M} F_2(q^2) \right) u(p_i) \quad (4.5.74)$$

with Pauli and Dirac form factors F_1 and F_2 . Here it is more convenient to work with the **Sachs electric and magnetic form factors**

$$G_E(q^2) = F_1(q^2) - \tau F_2(q^2), \quad G_M(q^2) = F_1(q^2) + F_2(q^2) \quad (4.5.75)$$

since they do not produce interference terms $\propto F_1 F_2$ in the cross section. The invariant amplitude then becomes

$$|\mathcal{M}|^2 = \frac{16 \alpha^2 \pi^2}{\tau^2} \left[\frac{G_E^2 + \tau G_M^2}{1 + \tau} (\lambda^2 - \tau^2 - \tau) + 2\tau^2 G_M^2 \right], \quad (4.5.76)$$

and the resulting cross section is the **Rosenbluth cross section**:

$$\frac{d\sigma}{d\Omega} = \left(\frac{d\sigma}{d\Omega} \right)_{\text{Mott}} \left(\frac{G_E^2 + \tau G_M^2}{1 + \tau} + 2\tau G_M^2 \tan^2 \frac{\theta}{2} \right). \quad (4.5.77)$$

For a structureless fermion ($F_1 = 1$, $F_2 = 0$ or $G_E = G_M = 1$) these formulas reduce to the previous forms (4.5.71) and (4.5.73).

The Rosenbluth cross section allows one to extract the nucleon's electromagnetic form factors under the assumption of one-photon exchange. If we define the kinematic variable

$$\varepsilon = \frac{\lambda^2 - \tau(1 + \tau)}{\lambda^2 + \tau(1 + \tau)} = \left(1 + 2(1 + \tau) \tan^2 \frac{\theta}{2} \right)^{-1}, \quad (4.5.78)$$

where $\varepsilon = 1$ corresponds to forward scattering and $\varepsilon = 0$ to backward scattering, the cross section takes the form

$$\frac{d\sigma}{d\Omega} = \left(\frac{d\sigma}{d\Omega} \right)_{\text{Mott}} \frac{\varepsilon G_E^2 + \tau G_M^2}{\varepsilon(1 + \tau)}. \quad (4.5.79)$$

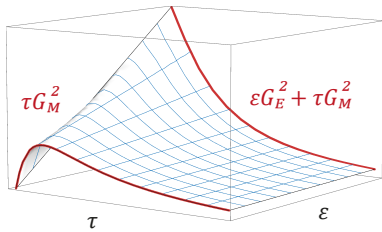


FIG. 4.22: Sketch of the numerator in the Rosenbluth cross section.

Because the form factors only depend on τ , at fixed τ the dependence of the numerator on ε is linear, which allows one to extract the magnetic form factor from the intercept at $\varepsilon = 0$ and the electric form factor from the slope in ε , see Fig. 4.22. This is known as the **Rosenbluth method**. In turn, at large τ (large photon virtualities Q^2) one is less sensitive to G_E and therefore G_E is not so well known at large Q^2 . The traditional Rosenbluth results yielded $G_E/G_M \approx \text{const.}$ for the proton at

large Q^2 , which was in agreement with perturbative scaling arguments. However, more recent polarization transfer experiments at Jefferson Lab measured the ratio G_E/G_M directly and found a falloff with Q^2 , which even points towards a zero crossing. A likely explanation is that G_E indeed falls off and that the discrepancy is due to **two-photon exchange** effects: although the corresponding diagrams enter with α_{QED}^2 in the cross section, they are large enough to interfere with the extraction of G_E at large Q^2 .

The form factors of the proton are directly accessible in $ep \rightarrow ep$ scattering, whereas those of the neutron are extracted from scattering on deuterium since there is no free neutron target in nature.

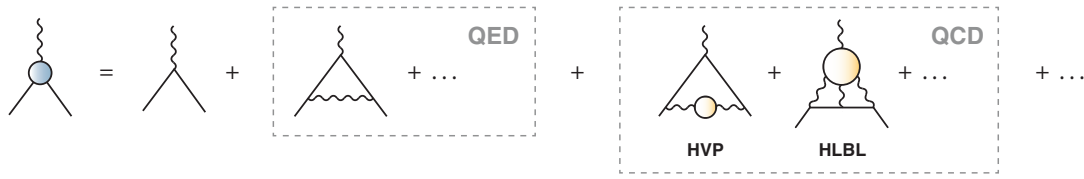


FIG. 4.23: Vertex corrections to the muon-photon vertex which contribute to the muon anomalous magnetic moment. The leading QCD contributions are the hadronic vacuum polarization and hadronic light-by-light scattering diagrams.

Form factor phenomenology. How can we interpret electromagnetic form factors? The Dirac and Pauli form factors at vanishing photon momentum encode the nucleons' charges and their **anomalous magnetic moments**:

$$F_1^p(0) = 1, \quad F_1^n(0) = 0, \quad F_2^p(0) = \kappa_p \approx 1.79, \quad F_2^n(0) = \kappa_n \approx -1.91.$$

For the Sachs form factors G_E and G_M this implies

$$G_M^p(0) = \mu_p = 1 + \kappa_p = 2.79, \quad G_M^n(0) = \mu_n = \kappa_n = -1.91.$$

The fact that the anomalous magnetic moments differ from zero means that the nucleon is not a pointlike Dirac particle but carries structure. In the analogous case of leptons, the coupling of the photon to an electron or muon has the same form as in Eq. (4.5.50). For pointlike Dirac particles $F_2(0)$ is zero, but due to QED corrections one finds

$$F_2(0) = \frac{\alpha_{\text{QED}}}{2\pi} + \dots \approx 1\%_0. \quad (4.5.80)$$

The leading diagram is the one-loop vertex dressing in Fig. 4.23, followed by higher-order QED corrections. The fact that $F_2(0)$ is much larger for the proton and neutron implies that they are far from pointlike. The **muon** anomalous magnetic moment ('muon $g - 2$ ') is particularly interesting: it has been measured to great precision but there is a current $\sim 4\sigma$ discrepancy between experiment and the Standard Model prediction, which could point towards new physics. Also QCD contributes to this process through the hadronic vacuum polarization (the diagram in Fig. 3.5) and the much smaller hadronic light-by-light scattering diagram. Both of these contributions are tiny compared to the QED effects, but they are almost alone responsible for the theory uncertainty of the Standard Model prediction.

The slopes of the Dirac and Pauli form factors at $Q^2 = 0$ define the Dirac and Pauli **charge radii**:

$$F_1(Q^2) = F_1(0) - \frac{r_1^2}{6} Q^2 + \dots, \quad F_2(Q^2) = F_2(0) \left[1 - \frac{r_2^2}{6} Q^2 + \dots \right]. \quad (4.5.81)$$

The electric and magnetic charge radii are defined accordingly from G_E and G_M . Also here there has been a surprise in the form of the **proton radius puzzle**: The electric charge radius of the proton measured in muonic hydrogen was found to be significantly smaller ($r_E^p \approx 0.84$ fm) than the previously established CODATA value inferred from ep scattering and hydrogen spectroscopy ($r_E^p \approx 0.88$ fm). Possible explanations include again new physics or two-photon effects, although several new measurements (including ep scattering) tend to agree with the lower radius as well.

Empirically, it turns out that the Sachs form factors can be reasonably well described by a dipole shape over a wide Q^2 range (except for G_E^n which vanishes at the origin). The 'dipole mass' Λ can then be used to estimate the charge radii:

$$G_i(Q^2) \approx \frac{G_i(0)}{(1 + Q^2/\Lambda^2)^2}, \quad \Lambda \approx 0.84 \text{ GeV} \quad \Rightarrow \quad r_i \approx \hbar c \frac{\sqrt{12}}{\Lambda} \approx 0.8 \text{ fm}, \quad (4.5.82)$$

with $\hbar c = 0.197 \text{ GeV fm}$. Such a dipole behavior for the Sachs form factors agrees with perturbative QCD predictions but has been challenged by measurements of G_E^p/G_M^p at larger Q^2 as mentioned above.

Non-relativistically, form factors can be interpreted as Fourier transforms of charge distributions. Consider the scattering of an electron from a static, spinless source generated by a charge distribution $\rho(\mathbf{x})$ that generates the vector potential $A^\mu(\mathbf{x})$:

$$\square A^\mu = j^\mu, \quad A^\mu = \begin{pmatrix} A^0 \\ \mathbf{0} \end{pmatrix}, \quad j^\mu = \begin{pmatrix} e\rho \\ \mathbf{0} \end{pmatrix}. \quad (4.5.83)$$

The invariant matrix element is given by

$$\mathcal{M} = ie \bar{u}(k_f) \gamma^\mu u(k_i) \underbrace{\int d^4x e^{-iq \cdot x} A_\mu(x)}_{(2\pi) \delta(E - E') \frac{e}{\mathbf{q}^2} F(\mathbf{q}) \delta_\mu^0} \quad (4.5.84)$$

which comes about as follows: Because $A^\mu(\mathbf{x})$ is time-independent, its Fourier transform in time produces a δ -function $\delta(q^0)$, which enforces $E = E'$ for the lab energies of the incoming and outgoing electron, cf. Eq. (4.5.27). The Maxwell equation $\square A^\mu = j^\mu$ then reduces to $\Delta A^0 = -e\rho$, and a partial integration yields

$$\int d^3x e^{iq \cdot x} A^0(\mathbf{x}) = \frac{e}{\mathbf{q}^2} \int d^3x e^{iq \cdot x} \rho(\mathbf{x}) =: \frac{e}{\mathbf{q}^2} F(\mathbf{q}), \quad (4.5.85)$$

where we defined the form factor $F(\mathbf{q})$ as the Fourier transformation of the charge density. Therefore, it measures the deviation from the pointlike nature of the source. For a spherically symmetric charge distribution $\rho(\mathbf{x}) = \rho(|\mathbf{x}|) = \rho(r)$ normalized to $\int d^3x \rho(\mathbf{x}) = 1$, the form factor at small $|\mathbf{q}|$ can be expanded in

$$F(\mathbf{q}) = \int d^3x \rho(\mathbf{x}) \left(1 + i\mathbf{q} \cdot \mathbf{x} - \frac{(\mathbf{q} \cdot \mathbf{x})^2}{2} + \dots \right) = 1 - \frac{|\mathbf{q}|^2}{6} \underbrace{4\pi \int dr \rho(r) r^4}_{\langle r^2 \rangle} + \dots$$

The coefficient of the quadratic term is the mean-square radius of the 'charge cloud', which motivates the definition of the charge radius in Eq. (4.5.81).

Examples for charge distributions and their corresponding form factors are shown in Fig. 4.24: A pointlike charge corresponds to a constant form factor, an exponential charge distribution to a dipole form factor,

$$\rho(r) = \frac{\Lambda^3}{8\pi} e^{-\Lambda r} \quad \Leftrightarrow \quad F(\mathbf{q}) = \int d^3x e^{iq \cdot x} \rho(\mathbf{x}) = \frac{1}{(1 + |\mathbf{q}|^2/\Lambda^2)^2}, \quad (4.5.86)$$

a Gaussian to a Gaussian and a homogeneous sphere to an oscillating form factor.

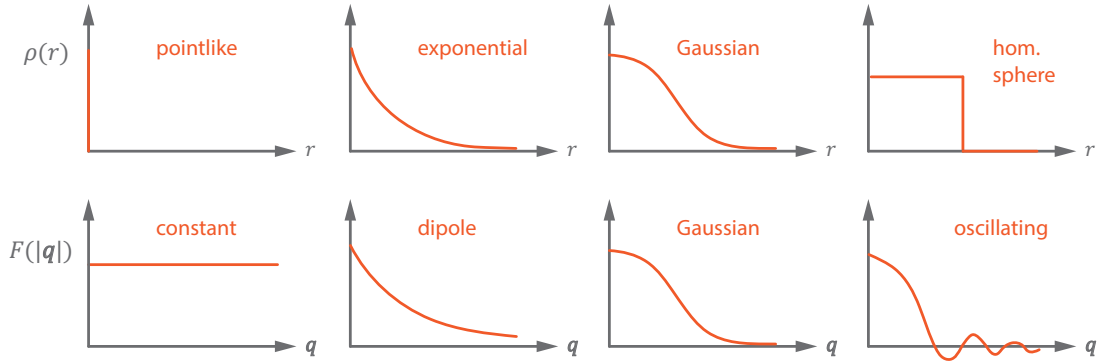


FIG. 4.24: Charge distributions and form factors.

For relativistic nucleons, the interpretation of form factors being Fourier transforms of charge and magnetization distributions has to be taken with a grain of salt. The formulas still look similar to the nonrelativistic case in the **Breit frame**, where the incoming and outgoing proton have opposite momenta ($\mathbf{p}_f = -\mathbf{p}_i = \mathbf{q}/2$) and hence the same energies, so that the photon transfers no energy and $E' = E$; this also implies $Q^2 = -q^2 = |\mathbf{q}|^2$. Furthermore, the vector current matrix element in the Breit frame reduces to the form

$$\langle p_f, \sigma' | V^0 | p_i, \sigma \rangle = 2M G_E \delta_{\sigma'\sigma}, \quad \langle p_f, \sigma' | \mathbf{V} | p_i, \sigma \rangle = G_M \chi_{\sigma'}^\dagger i\boldsymbol{\tau} \times \mathbf{q} \chi_\sigma,$$

hence the name ‘electric’ and ‘magnetic’ form factors. The charge densities extracted from the experimentally measured G_E^p and G_E^n have shapes shown in Fig. 4.25, which has led to the picture of a neutron behaving like a proton with a positively charged core and a negatively charged **pion cloud**. However, since there is a different Breit frame for each value of Q^2 , the relation to charge densities in the lab frame (the rest frame of the nucleon) will suffer from relativistic boost corrections and hence the interpretation of the radii as actual charge and magnetization radii is not directly applicable. In general, while the Lorentz-invariant form factors uniquely specify the electromagnetic structure of a hadron, their physical interpretation in terms of spatial densities depends on the reference frame.

Magnetic moments in the quark model. Current matrix elements encode the complicated nonperturbative substructure of hadrons and have become amenable to first-principle calculations only in recent years. Nevertheless, we can infer simple relations already from the nonrelativistic quark model. We saw in Eq. (3.2.79) that the spin-flavor wave functions for ground-state baryon octet states can be written as the combination of a flavor and a spin doublet:

$$|\lambda\sigma\rangle = \mathcal{D}^\lambda \cdot \mathcal{D}^\sigma = \sum_{m=1}^2 \mathcal{D}_m^\lambda \mathcal{D}_m^\sigma, \quad \lambda \in \{p, n, \Sigma^+, \dots\}, \quad \sigma \in \{\uparrow, \downarrow\}, \quad (4.5.87)$$

combined with a symmetric spatial wave function and the antisymmetric color part. The flavor doublets \mathcal{D}^λ are the flavor wave functions in Table 3.4, and the $SU(2)$ spin doublets \mathcal{D}^σ follow if we replace u by \uparrow and d by \downarrow . The index m denotes the doublet entries.

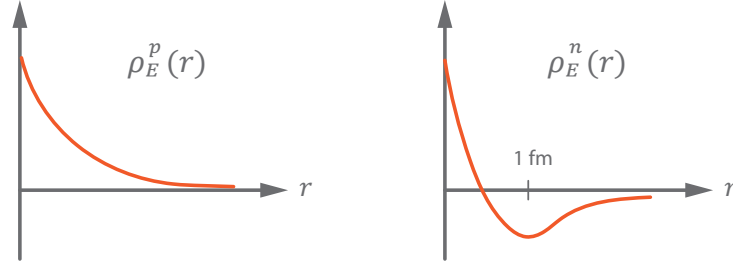


FIG. 4.25: Sketch of the electric charge distributions for proton and neutron in the Breit frame extracted from the measured form factors $G_E^p(Q^2)$ and $G_E^n(Q^2)$.

In the following we are only interested in the spin-flavor part. Its unit normalization is ensured via

$$\langle \lambda' \sigma' | \lambda \sigma \rangle = \frac{1}{\mathcal{N}} \sum_{m'm} (\mathcal{D}_{m'}^{\lambda'})^\dagger (\mathcal{D}_m^\lambda) (\mathcal{D}_{m'}^{\sigma'})^\dagger (\mathcal{D}_m^\sigma) \stackrel{!}{=} \delta_{\lambda'\lambda} \delta_{\sigma'\sigma}, \quad (4.5.88)$$

from where the factor \mathcal{N} has to be determined. From Table 3.4 one can verify

$$(\mathcal{D}_{m'}^p)^\dagger (\mathcal{D}_m^p) = (\mathcal{D}_{m'}^n)^\dagger (\mathcal{D}_m^n) = (\mathcal{D}_{m'}^\uparrow)^\dagger (\mathcal{D}_m^\uparrow) = \delta_{m'm}, \quad (4.5.89)$$

e.g. with $u^\dagger u = d^\dagger d = 1$, $u^\dagger d = d^\dagger u = 0$:

$$(\mathcal{D}_1^p)^\dagger (\mathcal{D}_1^p) = \frac{1}{2} (u^\dagger d^\dagger u^\dagger - d^\dagger u^\dagger u^\dagger) (udu - duu) = 1, \quad \text{etc.} \quad (4.5.90)$$

Inserting this in (4.5.88) yields

$$\langle p^\uparrow | p^\uparrow \rangle = \langle n^\uparrow | n^\uparrow \rangle = \frac{1}{\mathcal{N}} \text{Tr} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \frac{2}{\mathcal{N}} \quad \Rightarrow \quad \mathcal{N} = 2. \quad (4.5.91)$$

The expectation value of a generic flavor (F) and spin (Γ) operator is then

$$\langle \lambda' \sigma' | F \Gamma | \lambda \sigma \rangle = \frac{3}{\mathcal{N}} \sum_{m'm} \underbrace{(\mathcal{D}_{m'}^{\lambda'})^\dagger F (\mathcal{D}_m^\lambda)}_{=: F_{m'm}^{\lambda'\lambda}} \underbrace{(\mathcal{D}_{m'}^{\sigma'})^\dagger \Gamma (\mathcal{D}_m^\sigma)}_{=: \Gamma_{m'm}^{\sigma'\sigma}} = \frac{3}{2} \text{Tr} \left\{ F^{\lambda'\lambda T} \Gamma^{\sigma'\sigma} \right\}, \quad (4.5.92)$$

which is understood in the sense that F and Γ act on the flavor and spin indices of the third quark in each doublet \mathcal{D} , and the factor 3 counts the three possible permutations. The trace in the last equation goes over the doublet indices. It is useful to work out the flavor and spin matrix elements of the $SU(2)$ unit matrix and the Pauli matrix τ_3 for proton and neutron (use $\tau_3 u = u$, $\tau_3 d = -d$):

$$\mathbb{1}^{\uparrow\uparrow} = \mathbb{1}^{pp} = \mathbb{1}^{nn} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \tau_3^{\uparrow\uparrow} = \tau_3^{pp} = -\tau_3^{nn} = \begin{pmatrix} 1 & 0 \\ 0 & -\frac{1}{3} \end{pmatrix}. \quad (4.5.93)$$

The matrix elements of the unit matrix are just those in (4.5.89). Their combination yields the two-flavor quark charge matrix, cf. Eq. (4.5.52):

$$\mathbb{Q} = \begin{pmatrix} q_u & 0 \\ 0 & q_d \end{pmatrix} = \frac{1}{6} + \frac{\tau_3}{2} \quad \Rightarrow \quad \mathbb{Q}^{pp} = \frac{2}{3} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad \mathbb{Q}^{nn} = \frac{1}{3} \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \quad (4.5.94)$$

from where one obtains the charges of proton and neutron:

$$\langle p \uparrow | Q | p \uparrow \rangle = \frac{3}{2} \text{Tr } Q^{pp} = 1, \quad \langle n \uparrow | Q | n \uparrow \rangle = \frac{3}{2} \text{Tr } Q^{nn} = 0 \quad (4.5.95)$$

as well as their magnetic moments:

$$\begin{aligned} \langle p \uparrow | Q \tau_3 | p \uparrow \rangle &= \frac{3}{2} \text{Tr} \left\{ Q^{pp} \tau_3^{\uparrow\uparrow} \right\} = 1, \\ \langle n \uparrow | Q \tau_3 | n \uparrow \rangle &= \frac{3}{2} \text{Tr} \left\{ Q^{nn} \tau_3^{\uparrow\uparrow} \right\} = -\frac{2}{3}, \end{aligned} \quad (4.5.96)$$

apart from the remaining spatial integral. However, since the spatial part is taken to be identical for proton and neutron, the last relation yields the quark-model relation $\mu_n/\mu_p = -\frac{2}{3}$ which is quite close to the experimental value -0.685 . Similarly, one can also work out the magnetic moments for the other ground-state octet members:

$$\mu_{\Sigma^+} = 1, \quad \mu_{\Sigma^0} = \frac{1}{3}, \quad \mu_{\Sigma^-} = \mu_{\Xi^-} = \mu_{\Lambda} = -\frac{1}{3}, \quad \mu_{\Xi^0} = -\frac{2}{3}, \quad (4.5.97)$$

and in principle also those of the decuplet baryons.

References and further reading

Quark models:

- J. F. Donoghue, E. Golowich, B. R. Holstein, *Dynamics of the Standard Model*. Cambridge University Press, 1992
- A. W. Thomas, W. Weise, *The Structure of the Nucleon*. Wiley-VCH, 2001
- A. Hosaka, H. Toki, *Quarks, Baryons and Chiral Symmetry*. World Scientific, 2001.
- F. Jegerlehner, *Quantum Chromodynamics and strong interaction physics*. Lecture notes, 2009. <http://www-com.physik.hu-berlin.de/~fjeger/books.html>
- S. Capstick, W. Roberts, *Quark Models of Baryon Masses and Decays*, Prog. Part. Nucl. Phys. 45 (2000) 241, [arXiv:nuc1-th/0008028](https://arxiv.org/abs/nuc1-th/0008028)

Spontaneous chiral symmetry breaking:

- A. W. Thomas, W. Weise, *The Structure of the Nucleon*. Wiley-VCH, 2001
- S. Pokorski, *Gauge Field Theories*. Cambridge University Press, 1987
- V. P. Nair, *Quantum Field Theory: A Modern Perspective*. Springer, 2005
- C. D. Roberts, *Strong QCD and Dyson-Schwinger Equations*. Lecture notes, 2012. [arXiv:1203.5341](https://arxiv.org/abs/1203.5341) [[nucl-th](#)]

Axial anomaly:

- S. Pokorski, *Gauge Field Theories*. Cambridge University Press, 1987
- J. F. Donoghue, E. Golowich, B. R. Holstein, *Dynamics of the Standard Model*. Cambridge University Press, 1992
- V. P. Nair, *Quantum Field Theory: A Modern Perspective*. Springer, 2005
- R. A. Bertlmann, *Anomalies in Quantum Field Theory*. Oxford University Press, 1996
- M. Kaku, *Quantum Field Theory: A Modern Introduction*. Oxford University Press, 1993

Chiral effective field theories:

- J. F. Donoghue, E. Golowich, B. R. Holstein, *Dynamics of the Standard Model*. Cambridge University Press, 1992
- A. W. Thomas, W. Weise, *The Structure of the Nucleon*. Wiley-VCH, 2001
- S. Pokorski, *Gauge Field Theories*. Cambridge University Press, 1987
- S. Scherer, *Introduction to Chiral Perturbation Theory*. Adv. Nucl. Phys. **27** (2003) 277. [arXiv:hep-ph/0210398](https://arxiv.org/abs/hep-ph/0210398)
- M. Birse and J. McGovern, *Chiral perturbation theory*. Appears in: F. Close, S. Donnachie and G. Shaw (ed.), *Electromagnetic interactions and hadronic structure*, Cambridge University Press, 2007
- J. D. Walecka, *Theoretical nuclear and subnuclear physics*. World Scientific, 1995

Hadron matrix elements:

- V. Barone, E. Predazzi, *High-Energy Particle Diffraction*. Springer, 2002
- V. Gribov, *Strong Interactions of Hadrons at High Energies*. Cambridge University Press, 2009
- F. Halzen, A. D. Martin, *Quarks and Leptons: An Introductory Course in Modern Particle Physics*. Wiley, 1984.
- A. W. Thomas, W. Weise, *The Structure of the Nucleon*. Wiley-VCH, 2001
- J. F. Donoghue, E. Golowich, B. R. Holstein, *Dynamics of the Standard Model*. Cambridge University Press, 1992

Chapter 5

High-energy phenomenology

5.1 Deep inelastic scattering

From elastic eN scattering one can extract the nucleon's electromagnetic form factors. In a general *inelastic* scattering process the nucleon does not stay intact; instead it breaks up and produces hadronic final states. Depending on the invariant mass of the hadronic end product, the inelastic cross section then contains nucleon resonance peaks and nucleon-meson continua. Moreover, **deep inelastic scattering (DIS)** has given us first convincing evidence for the existence of quarks since it probes the composite nature of the nucleon. In DIS the transferred momentum of the photon is so large that it strikes the perturbative ‘partons’ inside the nucleon, which allows us to describe the longitudinal momentum distributions of the quarks and gluons through parton distribution functions.

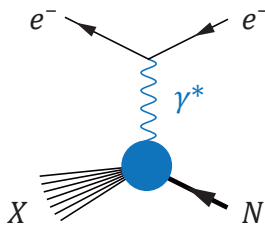


FIG. 5.1: DIS

Phase space. Employing the variables (4.5.8–4.5.9) with massless electrons ($m = m' = 0$), there are three independent Lorentz invariants: the spacelike momentum transfer $\tau = Q^2/(4M^2) \geq 0$, the crossing variable λ , and the invariant mass $W = M'$ of the hadrons in the final state which is no longer fixed but also a variable. The kinematic phase space of the process, which in elastic scattering was described by the two-dimensional Mandelstam plane, thus becomes three-dimensional. Instead of W , we could work with either of the variables ω , ν or the Bjorken variable x defined in Eq. (4.5.16):

$$W = M\sqrt{1 + 4\omega}, \quad \omega + \tau = \frac{\nu}{2M} = \frac{\tau}{x}. \quad (5.1.1)$$

Below we will see that in the one-photon exchange approximation the cross section factorizes again into a leptonic and a hadronic part, where the hadronic subprocess only depends on τ and ω .

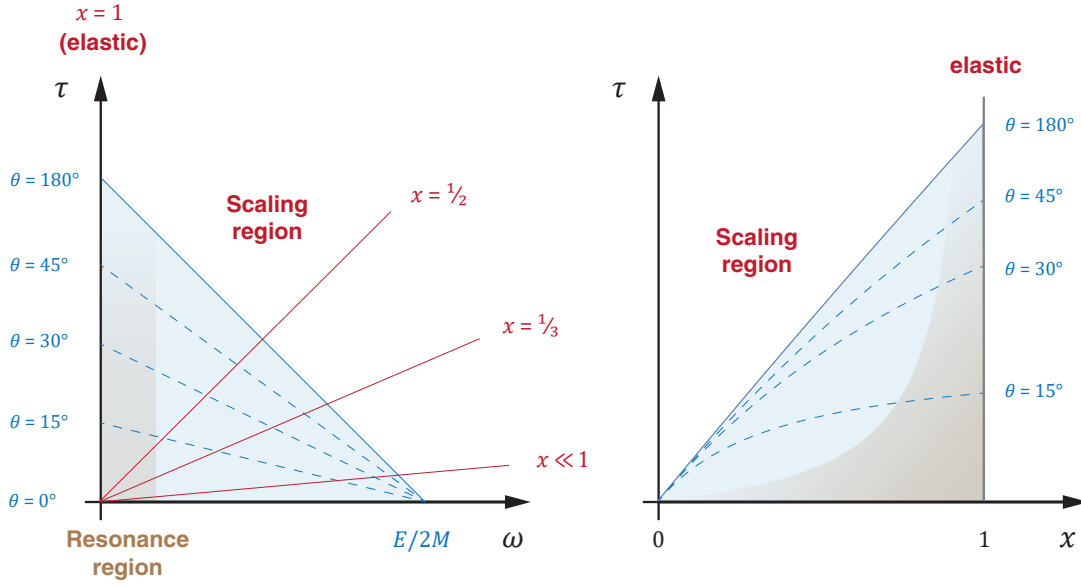


FIG. 5.2: Kinematic phase space in inelastic eN scattering at fixed lepton energy E in the variables (ω, τ) , left, and (x, τ) , right. The physical region is shown in blue and the resonance region in brown. Lines of constant scattering angle θ and Bjorken- x are also included.

To relate the Lorentz invariants to the incoming lepton energy E and the scattering angle θ in the lab frame, we infer from the relations (4.5.38):

$$\tau = (\varepsilon - \omega) \frac{4\varepsilon \sin^2 \frac{\theta}{2}}{1 + 4\varepsilon \sin^2 \frac{\theta}{2}} = \varepsilon x \frac{4\varepsilon \sin^2 \frac{\theta}{2}}{x + 4\varepsilon \sin^2 \frac{\theta}{2}}, \quad \varepsilon := \frac{E}{2M}. \quad (5.1.2)$$

The resulting phase space in the (ω, τ) and (x, τ) planes is sketched in Fig. 5.2. For fixed lepton energy E , the physically allowed region is bounded by $\omega = 0 \Leftrightarrow x = 1$ (elastic scattering), $\tau = 0$ (forward angles $\theta = 0$), and backward angles $\theta = \pi$ which for large energies implies $\tau \approx \varepsilon - \omega \approx \varepsilon x$. This is the blue area in the plot, whose size is characterized by the external control parameter E : if we increase the energy of the lepton beam, we can reach higher τ and ω values. We can locate different regions in these plots:

- At the elastic threshold $\omega = 0 \Leftrightarrow x = 1$, the invariant mass is $W = M$. The region $W \gtrsim M$ (or $\omega \lesssim 1$) is the **resonance region** where nucleon resonances appear in the cross section, starting with the $\Delta(1232)$ peak as shown in Fig. 5.3. Above $W \sim 2$ GeV, there is no visible resonance structure left.
- The limit $\tau + \omega \rightarrow \infty$ and $\omega/\tau = \text{const.}$ defines the **Bjorken limit**: this is the DIS region where scaling occurs. From Eq. (5.1.1) the Bjorken limit corresponds to $\nu \rightarrow \infty$ and constant Bjorken- x .
- The region of small x is interesting for several reasons and assumed to give experimental access to the properties of gluons.

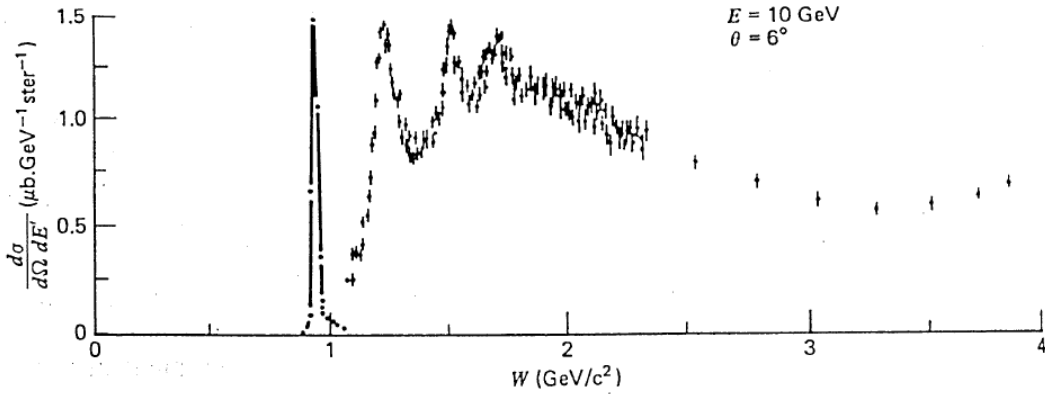


FIG. 5.3: Double-differential inelastic eN cross section from Eq. (5.1.11) at fixed lepton energy E and scattering angle θ . At large invariant masses, the resonance peaks are washed out. (*Halzen and Martin, Quarks and Leptons: An Introductory Course in Modern Particle Physics, Wiley, 1984.*)

Cross section and structure functions. Let us work out the cross section for inelastic eN scattering. In an inclusive measurement only the outgoing electron is detected but not the remnants of the proton. The cross section in the one-photon approximation still has the generic form of Eqs. (4.5.41–4.5.42) with the same leptonic tensor (4.5.67). However, the hadronic contribution to the invariant matrix element $|\mathcal{M}|^2$ and to the phase space factor now sums over all possible final states,

$$d\sigma = \frac{1}{4ME} \frac{d^3k_f}{(2\pi)^3} \frac{e^4}{2E'} \frac{1}{q^4} L_{\mu\nu} 4\pi M W^{\mu\nu},$$

$$4\pi M W^{\mu\nu} = \sum_X \frac{d^3p_f}{(2\pi)^3} \frac{1}{2E_X} \langle N(p_i) | V_{\text{em}}^\mu(0) | X(p_f) \rangle \langle X(p_f) | V_{\text{em}}^\nu(0) | N(p_i) \rangle \quad (5.1.3)$$

$$\times (2\pi)^4 \delta^4(q + p_i - p_f).$$

Here we absorbed the integral over d^3p_f and the δ -function for energy-momentum conservation into a hadronic tensor $W^{\mu\nu}$, and V_{em}^μ is the electromagnetic current operator from Eq. (3.1.92) that enters in the electromagnetic transition from the nucleon to all possible final states X .

Observe that the hadronic tensor comprises the completeness relation (2.2.5). If we write the δ -function in momentum space as

$$(2\pi)^4 \delta^4(q + p_i - p_f) = \int d^4z e^{i(q+p_i-p_f)z}, \quad (5.1.4)$$

use translation invariance (2.2.10–2.2.11) to shuffle the z -dependence in the phase factor $e^{i(p_i-p_f)z}$ into the current operators, and sum over the complete set of states X , we obtain:

$$4\pi M W^{\mu\nu}(p, q) = \int d^4z e^{iqz} \langle N(p_i) | V_{\text{em}}^\mu\left(\frac{z}{2}\right) V_{\text{em}}^\nu\left(-\frac{z}{2}\right) | N(p_i) \rangle$$

$$= \int d^4z e^{iqz} \langle N(p_i) | [V_{\text{em}}^\mu\left(\frac{z}{2}\right), V_{\text{em}}^\nu\left(-\frac{z}{2}\right)] | N(p_i) \rangle. \quad (5.1.5)$$

In the second line we replaced the product of the currents by their commutator because the matrix element of $V_{\text{em}}^\nu(-\frac{z}{2})V_{\text{em}}^\mu(\frac{z}{2})$ is zero: it gives rise to a δ -function $\delta(q-p_i+p_f)$ which cannot be saturated by any intermediate state. Energy conservation would require $E_X = M - E + E' = M - \nu \leq M$, but the nucleon is the lightest ground-state baryon. In this way, the hadronic tensor is the matrix element of the current commutator, which is analogous to the commutators in Eq. (3.1.58) and vanishes outside the light cone. We will return to this expression later.

For now, let us work out the general form of the hadronic tensor $W^{\mu\nu}(p, q)$ in momentum space. For unpolarized scattering, it can only depend on the Lorentz tensors

$$T_q^{\mu\nu}, \quad p_T^\mu p_T^\nu, \quad q^\mu q^\nu, \quad p_T^\mu q^\nu \pm q^\mu p_T^\nu, \quad (5.1.6)$$

where $T_q^{\mu\nu} = g^{\mu\nu} - q^\mu q^\nu / q^2$ is the transverse projector and $p_T^\mu = T_q^{\mu\nu} p_\nu$ the momentum transverse to q^μ . Current conservation still holds because the sum of the outgoing charges must equal the nucleon charge, so $W^{\mu\nu}$ must be transverse in its Lorentz indices: $q_\mu W^{\mu\nu} = W^{\mu\nu} q_\nu = 0$. The most general transverse tensor according to these constraints is given by

$$W^{\mu\nu} = -W_1(\tau, \omega) T_q^{\mu\nu} + \frac{W_2(\tau, \omega)}{M^2} p_T^\mu p_T^\nu, \quad (5.1.7)$$

where the response functions W_1 and W_2 depend on the Lorentz invariants τ and ω . From these one defines the dimensionless **nucleon structure functions** as

$$F_1(\tau, \omega) = M W_1(\tau, \omega), \quad F_2(\tau, \omega) = \nu W_2(\tau, \omega). \quad (5.1.8)$$

(For polarized scattering, there are two further spin-dependent structure functions g_1, g_2 and there is also another term in the lepton tensor.)

Combining this with the leptonic tensor yields

$$\begin{aligned} L^{\mu\nu} W_{\mu\nu} &= 4 \left[\frac{W_2}{M^2} \left((p \cdot k)^2 + \frac{q^2}{4} p_T^2 \right) - W_1 \left(k^2 + \frac{3}{4} q^2 \right) \right] \\ &= 4M^2 [W_2 (\lambda^2 - (\tau + \omega)^2 - \tau) + 2W_1 \tau] \\ &= 4EE' \cos^2 \frac{\theta}{2} \left[W_2 + 2W_1 \tan^2 \frac{\theta}{2} \right]. \end{aligned} \quad (5.1.9)$$

In going from the first to the second line we used $k_T^\mu = k^\mu$, $p \cdot k = M^2 \lambda$, $k^2 = M^2 \tau$ and

$$p_T^2 = p^2 - \frac{(p \cdot q)^2}{q^2} = M^2 \left(1 + 2\omega + \tau + \frac{\omega^2}{\tau} \right) = \frac{M^2}{\tau} (\tau + (\tau + \omega)^2), \quad (5.1.10)$$

and to obtain the third line we exploited Eqs. (4.5.36) and (4.5.38). The resulting cross section, which is shown in Fig. 5.3, is

$$\frac{d^2\sigma}{d\Omega dE'} = \frac{\alpha^2}{q^4} \frac{E'}{E} L_{\mu\nu} W^{\mu\nu} = \frac{\alpha^2 \cos^2 \frac{\theta}{2}}{4E^2 \sin^4 \frac{\theta}{2}} \left[W_2 + 2W_1 \tan^2 \frac{\theta}{2} \right]. \quad (5.1.11)$$

How does this compare to the limit of elastic scattering? From (4.5.44) and (4.5.71) we can write down the double-differential cross section for a pointlike fermion in the elastic case:

$$\begin{aligned} \frac{d^2\sigma}{d\Omega dE'} &= \frac{|\mathcal{M}|^2}{4ME} \frac{1}{(4\pi)^2} \frac{E' \delta(\omega)}{2M^2} = \frac{\alpha^2}{4M^2\tau^2} \frac{E'}{E} \frac{\delta(\omega)}{2M} (\lambda^2 + \tau^2 - \tau) \\ &= \frac{\alpha^2 \cos^2 \frac{\theta}{2}}{4E^2 \sin^4 \frac{\theta}{2}} \frac{\delta(\omega)}{2M} \left(1 + 2\tau \tan^2 \frac{\theta}{2}\right). \end{aligned} \quad (5.1.12)$$

Hence, in the elastic limit the response functions reduce to

$$W_1(\tau, \omega) = \tau \frac{\delta(\omega)}{2M}, \quad W_2(\tau, \omega) = \frac{\delta(\omega)}{2M}. \quad (5.1.13)$$

We can trade the dependence on ω by a dependence on the Bjorken variable x using the relations

$$\tau = \frac{\nu}{2M} x, \quad \omega = \frac{\nu}{2M} (1 - x). \quad (5.1.14)$$

As a consequence, when expressed in terms of τ and x , the structure functions defined in Eq. (5.1.14) become

$$\begin{aligned} F_1(\tau, x) &= M W_1(\tau, x) = \frac{1}{2} \delta(1 - x), \\ F_2(\tau, x) &= \nu W_2(\tau, x) = \delta(1 - x). \end{aligned} \quad (5.1.15)$$

We see that for elastic scattering on a pointlike particle, the dimensionless structure functions $F_1(\tau, x)$ and $F_2(\tau, x)$ are functions of x only; in addition, the δ -function enforces $x = 1$ in the elastic limit.

For scattering on a composite nucleon, the expressions (5.1.13) in the elastic limit must be multiplied with the Sachs form factor combinations in the Rosenbluth cross section (4.5.77). Here one should remember not to confuse the structure functions F_1 and F_2 with the equally named Dirac and Pauli form factors. In fact, even their physical meanings are reversed: By comparing the two cross sections, one can see that W_1 encodes the spin of the target and vanishes for a spinless particle. Thus, the structure function F_1 carries the spin dependence, whereas in the form factor case it is rather the Pauli form factor (or the magnetic form factor G_M) that contains the nucleon spin.

Bjorken scaling and the parton model. One might expect that for inelastic scattering processes ($x \neq 1$), away from the nucleon resonance region, the structure functions F_1 and F_2 are complicated functions of τ and x . However, it turns out that in the DIS region they are almost independent of τ and only functions of x :

$$F_{1,2}(\tau, x) \approx F_{1,2}(x). \quad (5.1.16)$$

This is visible in the left of Fig. 5.4 and called **Bjorken scaling**. Another observation is the **Callan-Gross relation**, which implies that F_1 and F_2 are not independent:

$$F_2(x) = 2xF_1(x). \quad (5.1.17)$$

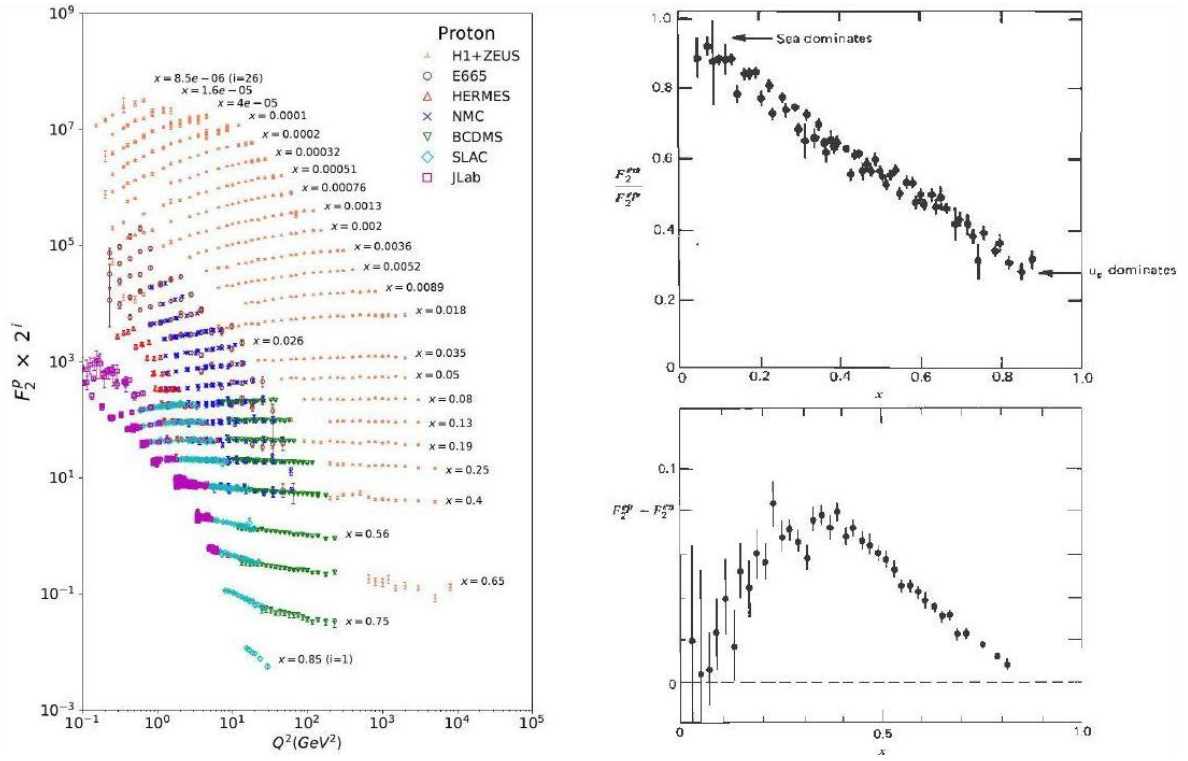


FIG. 5.4: *Left:* scaling behavior in the structure function $F_2(Q^2, x)$, PDG 2020, P. A. Zyla et al., Prog. Theor. Exp. Phys. **2020**, 083C01 (2020). *Right:* experimental data for the ratio and difference of proton and neutron structure functions in Eq. (5.1.29). Source: Halzen and Martin (see Fig. 5.3).

The origin of scaling can be understood from dimensional arguments, which follow from the near scale invariance of massless perturbative QCD (up to logarithmic corrections). A dimensionless function can only depend on dimensionless variables. τ and ω are only dimensionless because we scaled the momenta with the nucleon mass M , which requires the presence of a nonperturbative nucleon mass to begin with. If we scatter instead on (nearly) massless quarks, no such scale is available and therefore the dimensionless structure functions cannot depend on τ and ω individually but only on their dimensionless combination $\tau/\omega \sim q^2/p \cdot q$. Hence, the observation of scaling is an indication for the composite nature of the nucleon in terms of pointlike, essentially massless quarks and gluons.

The experimental observation of Bjorken scaling has led to the development of the **parton model**. Here the proton is viewed as a collection of ‘partons’, namely valence quarks, sea quarks and gluons. The incoming momentum p_i of the proton (mass M) is the sum of the parton momenta, $p_i = \sum_k p_k$, where p_k is the four-momentum of a single onshell parton with mass m_k . The basic assumption we need in the following is **collinearity**: $p_k = \xi_k p_i$, which can be justified in the **infinite momentum frame**. If we write

$$p_i = \begin{pmatrix} \sqrt{\mathbf{p}^2 + M^2} \\ \mathbf{p} \end{pmatrix}, \quad p_k = \begin{pmatrix} \sqrt{\xi_k^2 \mathbf{p}^2 + (\mathbf{p}_k^\perp)^2 + m_k^2} \\ \xi_k \mathbf{p} + \mathbf{p}_k^\perp \end{pmatrix}, \quad (5.1.18)$$

then ξ_k defines the longitudinal momentum fraction of parton k in the direction of the proton's three-momentum \mathbf{p} . In the infinite-momentum frame ($|\mathbf{p}| \rightarrow \infty$) we can neglect the transverse components and masses:

$$|\mathbf{p}_k^\perp| \ll |\mathbf{p}|, \quad m_k \ll |\mathbf{p}|, \quad \Rightarrow \quad p_k \approx \xi_k p_i, \quad \sum_k \xi_k = 1. \quad (5.1.19)$$

The collinearity assumption allows for a simple interpretation of the Bjorken scaling variable. We know from Eq. (4.5.16) that elastic scattering on the nucleon corresponds to $x = -q^2/(2p_i \cdot q) = 1$. In the inelastic process ($x \neq 1$), elastic scattering on a single parton k then entails that

$$x_k := -\frac{q^2}{2p_k \cdot q} = -\frac{q^2}{2p_i \cdot q} \frac{1}{\xi_k} = \frac{x}{\xi_k} = 1 \quad \Rightarrow \quad \xi_k = x. \quad (5.1.20)$$

In this way, the Bjorken variable x assumes the meaning of the parton's **longitudinal momentum fraction** in the infinite-momentum frame. The photon only couples to those partons whose momentum fraction is $\xi_k = x$, hence a measurement of the structure function $F_2(x)$ allows us to 'see' how the parton momenta are distributed inside the proton. In elastic scattering we have $x = 1$ and the photon couples to the whole proton since it carries the full momentum. Note that if we want to guarantee $p_i^2 = M^2$, we should set \mathbf{p}_k^\perp and $m_k = \xi_k M$. Although this last relation is a bit nonsensical as it would imply that the 'mass' of a parton changes with its momentum fraction, we need it for consistency of the naive parton model.

Let us define the **parton distribution function** or **PDF** as the momentum distribution $f_k(\xi)$ of a parton in the hadron, so that $f_k(\xi) d\xi$ is the probability density that a parton carries a momentum fraction between ξ and $\xi + d\xi$. Momentum conservation implies

$$\sum_k \int_0^1 d\xi \xi f_k(\xi) = 1. \quad (5.1.21)$$

Now suppose we scatter on spin- $\frac{1}{2}$ quarks. Using the relations (5.1.14) with $x_k = x/\xi_k$ and $m_k = \xi_k M$, the structure functions $F_j^{(k)}$ with $j = 1, 2$ for the parton k are

$$\begin{aligned} 2F_1^{(k)} &= 2MW_1^{(k)} = \frac{M}{m_k} 2m_k W_1^{(k)} = \frac{M}{m_k} x_k \delta(1 - x_k) = \delta(\xi_k - x), \\ F_2^{(k)} &= \nu W_2^{(k)} = \delta(1 - x_k) = \frac{\xi_k^2}{x} \delta(\xi_k - x) = x \delta(\xi_k - x), \end{aligned} \quad (5.1.22)$$

and integrating over all partons yields

$$F_j(x) = \sum_k e_k^2 \int d\xi f_k(\xi) F_j^{(k)}(\xi, x) \quad \Rightarrow \quad \begin{aligned} F_1(x) &= \frac{1}{2} \sum_k e_k^2 f_k(x), \\ F_2(x) &= x \sum_k e_k^2 f_k(x). \end{aligned} \quad (5.1.23)$$

Hence we have shown that in the parton model F_1 and F_2 are indeed only functions of x , and we can confirm the Callan-Gross relation (5.1.16). The latter is also an experimental indication for the spin- $\frac{1}{2}$ nature of the quarks: if quarks had spin zero, $F_1(x)$ would vanish.

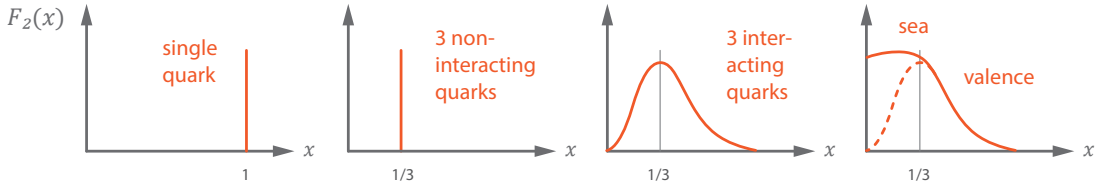


FIG. 5.5: Structure functions for different compositions of the proton.

Due to the Callan-Gross relation only the structure function $F_2(x)$ will be relevant in what follows. What does it look like? If the proton consisted of a single ‘quark’ that carried all of its momentum, the structure function would have a single peak at $x = 1$ (see Fig. 5.5). If it consisted of three non-interacting quarks, the quarks would all carry the same momentum fraction and $F_2(x)$ would have a peak at $x = \frac{1}{3}$. If the three quarks interact with each other, they can exchange momentum and hence the momentum fraction carried by each quark will fluctuate; the resulting structure function is a smooth distribution peaked near $x = \frac{1}{3}$. Finally, the presence of sea quarks will lead to an enhancement at small x because sea quarks are created in Bremsstrahlung-like processes which are typically enhanced at small momenta and lead to $xf(x) \xrightarrow{x \rightarrow 0} \text{const.}$ Note that gluons will also contribute to the momentum sum rule (5.1.21) whereas the structure function only probes electrically charged partons (quarks).

Parton distribution functions. Now let’s see how much information on the PDFs we can gather from experimental data on $F_2(x)$. There is no sensible way to distinguish two identical partons within a proton, but we can still group them according to the various quark and antiquark flavors: $f_k(x) = u(x), \bar{u}(x), d(x), \bar{d}(x)$, etc., so that we have

$$\frac{F_2^p(x)}{x} = q_u^2 ((u(x) + \bar{u}(x)) + q_d^2 (d(x) + \bar{d}(x)) + q_s^2 (s(x) + \bar{s}(x)) + \dots \quad (5.1.24)$$

It is usually sufficient to stop at the strange quark because the probability for finding charm in the proton is very small. $u(x)$ is the probability distribution for up quarks in the proton, $\bar{u}(x)$ that of anti-up quarks, and so on. One can also measure the structure function F_2^n of the neutron via electron-deuteron scattering. Charge symmetry entails that the d distribution in the neutron is identical to the u distribution in the proton: $u = u^p = d^n$, $d = d^p = u^n$, $s = s^p = s^n$, and analogously for the antiquark PDFs:

$$\frac{F_2^n(x)}{x} = q_d^2 ((u(x) + \bar{u}(x)) + q_u^2 (d(x) + \bar{d}(x)) + q_s^2 (s(x) + \bar{s}(x)) + \dots \quad (5.1.25)$$

In the following it will be more convenient to work with valence- and sea-quark distributions, defined via

$$\begin{aligned} u &= u_v + u_s, & \bar{u} &= \bar{u}_s, & s &= s_s, \\ d &= d_v + d_s, & \bar{d} &= \bar{d}_s, & \bar{s} &= \bar{s}_s, \end{aligned} \quad (5.1.26)$$

because antiquarks and strange quarks can only appear in the sea. Now, since the PDFs are number densities defined on the momentum fraction x , the integrals over

this range are just the total flavor numbers of each quark type:

$$\int_0^1 dx u_v(x) = 2, \quad \int_0^1 dx d_v(x) = 1, \quad \int_0^1 dx [f_s(x) - \bar{f}_s(x)] = 0. \quad (5.1.27)$$

The third relation expresses fermion number conservation for each flavor $f = u, d, s$: by summing over all individual partons, we must recover charge 1, baryon number 1 and strangeness 0 of the proton.

Can we extract the valence and sea distributions from the data for $F_2^{p,n}(x)$? We have two measured quantities but too many unknowns. Let's make the further simplifying assumption that all sea-quark distributions are identical: $f_s(x) = \bar{f}_s(x) =: S(x)$. Then the structure functions for the proton and neutron become

$$\begin{aligned} \frac{F_2^p(x)}{x} &= q_u^2 u_v(x) + q_d^2 d_v(x) + (q_u^2 + q_d^2 + q_s^2) 2S(x), \\ \frac{F_2^n(x)}{x} &= q_d^2 u_v(x) + q_u^2 d_v(x) + (q_u^2 + q_d^2 + q_s^2) 2S(x), \end{aligned} \quad (5.1.28)$$

from where we can form their ratio and their difference:

$$R = \frac{F_2^n}{F_2^p} = \frac{u_v + 4d_v + 12S}{4u_v + d_v + 12S}, \quad F_2^p - F_2^n = \frac{x}{3}(u_v - d_v). \quad (5.1.29)$$

The ratio satisfies the **Nachtman inequality** $\frac{1}{4} \leq R(x) \leq 4$: in a region of x where the up (down) quarks dominate, we have $R = \frac{1}{4}$ ($R = 4$); if the sea quarks dominate we will find $R = 1$. The ratio is plotted in Fig. 5.4 and reveals that the sea quarks are indeed dominant at small x whereas valence up quarks are important at large x . The difference in (5.1.29) is also plotted: it measures only the valence-quark contribution and shows a peak around $x = 1/3$, as we had expected. Finally, the sum

$$\frac{9}{5}(F_2^p + F_2^n) = x \left(u_v + d_v + \frac{24}{5} S \right) \quad (5.1.30)$$

can be plugged into the momentum sum rule (5.1.21) which now takes the form

$$\int dx x (u_v + d_v + 6S) + \varepsilon = 1, \quad (5.1.31)$$

where ε is the gluon contribution to the proton's longitudinal momentum. From the experimental data we can roughly estimate

$$\frac{9}{5} \int dx (F_2^p + F_2^n) \approx 0.54 \approx 1 - \varepsilon, \quad (5.1.32)$$

which entails that the gluons carry almost half of the proton's momentum. In fact, the gluon PDFs dominate at small values of x , see Fig. 5.6.

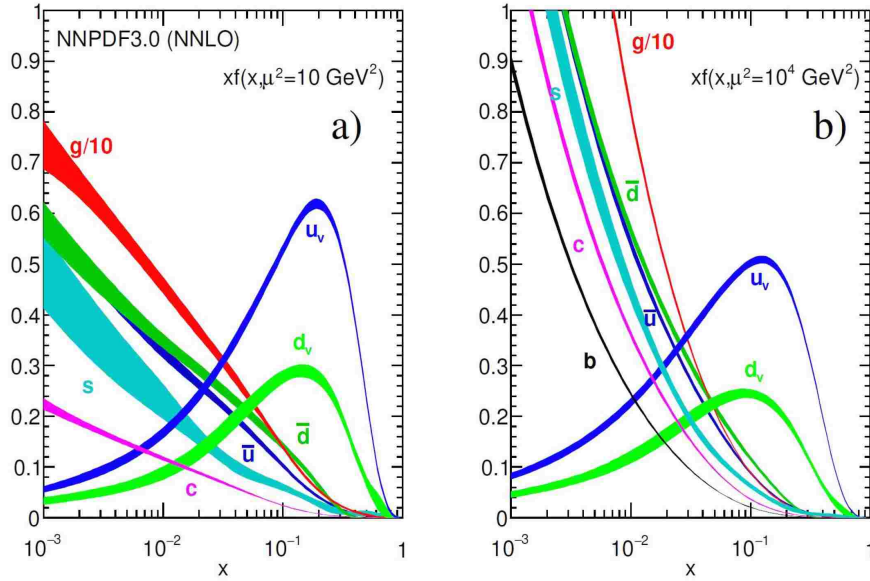


FIG. 5.6: Valence, sea-quark and gluon PDFs shown at two different resolution scales (PDG, same reference as in Fig. 5.4).

How good is the assumption that all sea-quark distributions are identical? If we go back to the original equations (5.1.24) and (5.1.25), take their difference and integrate over x , we have

$$\int dx \frac{F_2^p - F_2^n}{x} = \frac{1}{3} \int dx (u_v - d_v + u_s + \bar{u}_s - d_s - \bar{d}_s) \stackrel{(5.1.27)}{=} \frac{1}{3} + \frac{2}{3} \int dx (\bar{u}_s - \bar{d}_s)$$

which should equal $\frac{1}{3}$ if $\bar{u}_s = \bar{d}_s = S$ (this is the **Gottfried sum rule**). Instead, the experimental value is $\sim 0.23 \Rightarrow \int dx (\bar{d}_s - \bar{u}_s) \sim 0.15$, which entails that the light quark sea is indeed flavor-asymmetric.

Scaling violations. The left plot in Fig. 5.4 demonstrates that scaling is not exact because the structure functions show a Q^2 dependence, which is most pronounced at small and large values of x . In terms of the PDFs, this implies that their x -dependence is not completely independent of the resolution scale Q^2 but also evolves with Q^2 , which can be seen in Fig. 5.6. We can intuitively understand this as follows: a photon with intermediate Q^2 does not resolve the full spatial structure of the proton and mainly sees three interacting quarks, together with parts of the sea. In contrast, a high- Q^2 photon can resolve small distances and will reveal more and more of the quark sea which contains short-distance processes such as gluon emission from a quark or gluon splitting into $q\bar{q}$ pairs. As a result, the sea-quark contributions will be more prominent at higher Q^2 . On the other hand, since the photon can resolve more partons, momentum conservation implies that each parton now carries a smaller fraction of the total momentum, and hence the PDFs will be shifted to smaller x . The resulting structure function $F_2(x)$ that sums up the individual quark PDFs will rise with higher Q^2 at small x and fall with higher Q^2 at large x .

The short-distance dynamics depend on the resolution scale through the coupling $\alpha_s(Q^2)$. As a consequence, the individual quark structure functions F_i^k will no longer be mere δ -functions as in Eq. (5.1.22) but also inherit a Q^2 dependence from the coupling. Since the coupling is dimensionless, it also introduces a scale μ (the **factorization scale**), so that Eq. (5.1.23) becomes

$$F_j(x, Q^2) = \sum_k e_k^2 \int d\xi f_k(\xi, \mu) F_j^{(k)}(\xi, x, \frac{Q^2}{\mu^2}). \quad (5.1.33)$$

The $F_j^{(k)}$ encode the short-distance splitting processes and are calculable in perturbative QCD. The PDFs f_k , which now also depend on μ , are inherently nonperturbative and have to be fitted to experimental data or calculated with nonperturbative methods.

Because the nucleon structure function must be independent of the factorization scale μ , its total derivative with respect to μ must vanish. Similarly to the Callan-Symanzik equation (2.3.66), one then derives the **DGLAP equations** (Dokshitzer, Gribov, Lipatov, Altarelli, Parisi) $dF_j/d\mu = 0$. They relate PDFs at different μ with each other and thereby allow one to calculate the scaling violations using QCD perturbation theory.

Compton amplitude and PDFs. How can PDFs be calculated nonperturbatively? Let us return to the hadronic tensor $W^{\mu\nu}(q)$ from Eq. (5.1.5) which enters in the inelastic eN cross section. By means of the optical theorem, it can be written as the imaginary part of the nucleon's **forward Compton scattering amplitude**: $4\pi M W^{\mu\nu}(p, q) = 2 \text{Im} T^{\mu\nu}(p, q)$. The forward Compton amplitude $N\gamma^* \rightarrow N\gamma^*$ is given by

$$T^{\mu\nu}(p, q) = i \int d^4z e^{iqz} \langle N(p_i) | T V_{\text{em}}^\mu(\frac{z}{2}) V_{\text{em}}^\nu(-\frac{z}{2}) | N(p_i) \rangle. \quad (5.1.34)$$

If we apply the kinematics in Eqs. (4.5.8–4.5.9), then in the forward limit (vanishing momentum transfer) we have $p = p_i = p_f$ and the photon momentum is $k = k_i = k_f$, so that k^2 and the crossing variable λ are the independent Lorentz-invariants. The variables τ and x defined in DIS are related to these by

$$\tau = -\frac{k^2}{4M^2}, \quad x = -\frac{k^2}{2p \cdot k} = \frac{2\tau}{\lambda}. \quad (5.1.35)$$

Thus, the structure functions, which depend on τ and x , can be expressed through the Lorentz-invariant form factors of the Compton amplitude in the forward limit, which depend on those same variables. The Mandelstam variables s and u in Compton scattering are given by

$$\left\{ \begin{matrix} s \\ u \end{matrix} \right\} = (p \pm k)^2 = M^2 (1 - 4\tau \pm 2\lambda) = M^2 \left(1 - 4\tau \pm \frac{4\tau}{x} \right), \quad (5.1.36)$$

and the resulting Mandelstam plane is shown in Fig. 5.7. For real or virtual photons we have $\tau \geq 0$, and the physical region for $s \geq M^2$ corresponds to $0 \leq x \leq 1$. The Compton amplitude has non-analyticities arising from intermediate baryon resonances and baryon-meson continua. Hence, a theoretical handle on nucleon Compton scattering allows us to compute the nucleon's structure functions in DIS.

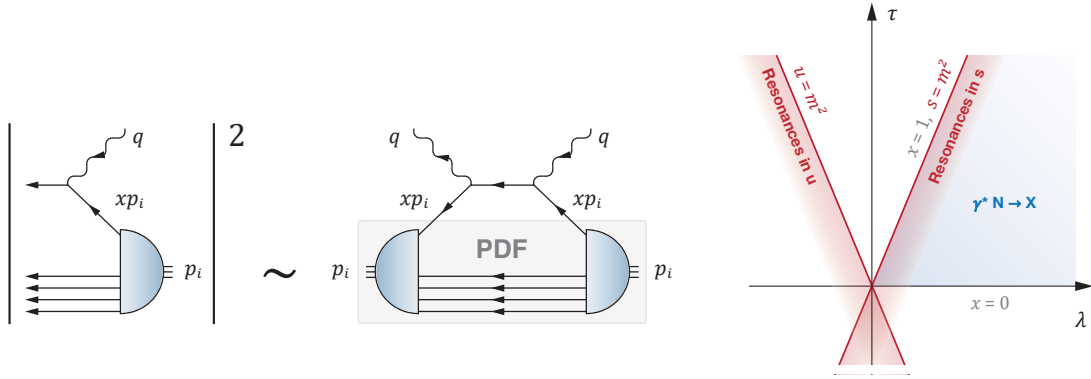


FIG. 5.7: Left: Hadronic tensor $W^{\mu\nu}$ in the parton model, and its relation with the forward Compton scattering amplitude and its factorized handbag structure. Right: Mandelstam plane in forward Compton scattering.

What about the PDFs? To begin with, it is important to realize that in the Bjorken limit the Fourier transform in Eq. (5.1.5) is dominated by the behavior close to the light cone $z^2 \rightarrow 0$, i.e., where the two interaction points are separated by a lightlike distance. This is easiest seen using **light-cone variables**:

$$a_{\pm} := \frac{1}{\sqrt{2}}(a^0 \pm a^3), \quad \mathbf{a}_{\perp} = (a^1, a^2) \quad \Rightarrow \quad a \cdot b = a_+ b_- + a_- b_+ - \mathbf{a}_{\perp} \cdot \mathbf{b}_{\perp}. \quad (5.1.37)$$

Then the integral (5.1.5) becomes schematically:

$$W(p, q) = \int dz_- e^{iq_+ z_-} \int dz_+ e^{iq_- z_+} \int_{z_{\perp}^2 < 2z_+ z_-} d^2 z_{\perp} e^{-iq_{\perp} \cdot z_{\perp}} W(p, z). \quad (5.1.38)$$

The domain of the z_{\perp} integration is restricted since the current commutator vanishes outside the light cone ($z^2 = 2z_+ z_- - z_{\perp}^2 < 0$) due to causality. In light-cone variables, the Bjorken limit $\nu \rightarrow \infty$, $x = \text{const.}$ corresponds to $q_+ \rightarrow \infty$ and $q_- = \text{const.}$:

$$\sqrt{2} q_{\pm} = q_0 \pm q^3 \stackrel{(4.5.27)}{=} \nu \pm \sqrt{\nu^2 - q^2} = \nu \left(1 \pm \sqrt{1 + \frac{2Mx}{\nu}} \right) \approx \begin{cases} 2\nu + Mx + \dots \\ -Mx + \dots \end{cases}$$

For $q_+ \rightarrow \infty$ and $q_- = \text{const.}$, the integral (5.1.38) is determined by the behavior of the integrand for $z_- \rightarrow 0$ and z_+ finite; this is the area with the least oscillations according to the Riemann-Lebesgue lemma. The condition $z_{\perp}^2 < 2z_+ z_-$ then implies $z^2 \rightarrow 0^+$ but $z^{\mu} \neq 0$, which is the light cone.

To proceed, we need to work out the current commutator in Eq. (5.1.5). We derived equal-time current commutators earlier in Eq. (3.1.57) using the anticommutation relations for the quark fields. For free fields one can generalize that formula to unequal times $x_0 \neq y_0$ with the generalized anticommutation relations

$$\{\psi(x), \bar{\psi}(y)\} = S(x - y), \quad \{\psi(x), \psi(y)\} = \{\bar{\psi}(x), \bar{\psi}(y)\} = 0, \quad (5.1.39)$$

where $S(z) := (i\rlap{/}\partial + m)\Delta(z)$, and $\Delta(z)$ is the **causal propagator** which vanishes outside the light cone, i.e., for spacelike distances $z^2 < 0$:

$$\Delta(z) := \int \frac{d^3p}{2E_p} \frac{e^{-ipz} - e^{ipz}}{(2\pi)^3} \Big|_{p^0=E_p} = \int \frac{d^4p}{(2\pi)^3} e^{-ipz} \varepsilon(p^0) \delta(p^2 - m^2), \quad (5.1.40)$$

and $\varepsilon(a) = a/|a| = \Theta(a) - \Theta(-a)$ is the sign function. At equal times $z_0 = 0$, the causal propagators reduce to $\Delta(z) = 0$, $\partial_0\Delta(z) = -i\delta^3(\mathbf{z})$ and $S(z) = \gamma_0\delta^3(\mathbf{z})$ which reproduces the equal-time (anti-)commutation relations for scalar and fermion fields. (In contrast to the Feynman propagator (2.2.14), the causal propagator sums up the positive- and negative-energy pole residues of a free scalar propagator.)

Rederiving the current commutator relation in this case gives the result¹

$$\left[j_a^\Gamma(x), j_b^{\Gamma'}(y) \right] = if_{abc} j_c^+(x, y) + d_{abc} j_c^-(x, y) + \frac{\delta_{ab}}{N} j^-(x, y), \quad (5.1.41)$$

which depends on the bilocal currents

$$j_a^\pm(x, y) := \frac{1}{2} \left(\bar{\psi}(x) \Gamma S(x-y) \Gamma' \mathbf{t}_a \psi(y) \pm \bar{\psi}(y) \Gamma' S(y-x) \Gamma \mathbf{t}_a \psi(x) \right). \quad (5.1.42)$$

Here we recognize the ‘**handbag**’ structure from Fig. 5.7 when putting the result back in the hadronic tensor $W^{\mu\nu}$; for the electromagnetic current commutator we have $\Gamma = \gamma^\mu$ and $\Gamma' = \gamma^\nu$. The light-cone singularities come from the free propagator $S(z)$ which for a massless fermion reduces to

$$S(z) \xrightarrow{m=0} \frac{1}{2\pi} \rlap{/}\partial (\varepsilon(z_0) \delta(z^2)). \quad (5.1.43)$$

It represents the hard part of the process, namely the scattering of the photon on a single perturbative quark which was the underlying assumption of the parton model.

The soft part is expressed through the remaining matrix element of bilocal quark-antiquark currents which is closely related to the quantity in Eq. (4.5.48). One can work out the Dirac structures for $\Gamma S(z)\Gamma'$ and $\Gamma' S(-z)\Gamma$ and expand the resulting currents in Taylor series about $z = 0$. This leads to the **operator product expansion (OPE)**, schematically written as

$$j\left(\frac{z}{2}, -\frac{z}{2}\right) = \sum_i c_i(z) \mathcal{O}_i(0), \quad (5.1.44)$$

where the $\mathcal{O}_i(0)$ are local operators and the $c_i(z)$ are the **Wilson coefficients**. The operators which are most important at high Q^2 are those for which the $c_i(z)$ are most singular as $z^2 \rightarrow 0$. This allows for a rigorous definition of PDFs that enter in Eq. (5.1.33) and makes them accessible for nonperturbative calculations.

Finally, the relation with the Compton amplitude also allows one to define non-forward **generalized parton distributions (GPDs)**. They encode the transverse structure of the proton, which is related to the orbital momentum carried by the quarks and gluons. In contrast to PDFs, they are no longer connected with DIS because a nonvanishing momentum transfer implies $p_f \neq p_i$. Hence, they have to be extracted directly from deeply virtual Compton scattering (DVCS) or related processes.

¹Extra care should be taken with regard to Schwinger terms, which include derivatives of the δ -function and do not show up in commutators of zero components of currents.

Appendix A

SU(N)

The group $SU(N)$ is probably the most important symmetry group in particle physics. $SU(2)$ encodes spin and isospin, $SU(3)$ describes both color and the physics of three light quark flavors, $SU(2) \times SU(2)$ is the universal cover of the (Euclidean) Lorentz group, etc. In the following we will collect some basic facts and useful formulas.

A.1 Basic properties of $SU(N)$

The group $SU(N)$ is the special unitary Lie group ($U^\dagger = U^{-1}$, $\det U = 1$) with $N^2 - 1$ real group parameters. The group element in a given representation can be written as

$$U = \exp\left(i \sum_{a=1}^{N^2-1} \varepsilon_a \mathbf{t}_a\right) = \exp(i\varepsilon), \quad (\text{A.1.1})$$

where the $N^2 - 1$ generators \mathbf{t}_a are hermitian and traceless. They form the basis of a Lie algebra with commutator relations

$$[\mathbf{t}_a, \mathbf{t}_b] = i f_{abc} \mathbf{t}_c, \quad (\text{A.1.2})$$

where f_{abc} are the totally antisymmetric and real structure constants of $SU(N)$. For $SU(2)$ one has $f_{abc} = \epsilon_{abc}$; the structure constants of $SU(3)$ are given in Table A.1. The Jacobi identity for the generators,

$$[\mathbf{t}_a, [\mathbf{t}_b, \mathbf{t}_c]] + [\mathbf{t}_b, [\mathbf{t}_c, \mathbf{t}_a]] + [\mathbf{t}_c, [\mathbf{t}_a, \mathbf{t}_b]] = 0, \quad (\text{A.1.3})$$

implies for the structure constants the relation $f_{abe} f_{cde} + f_{bce} f_{ade} + f_{cae} f_{bde} = 0$.

In the fundamental representation, the generators are $N \times N$ matrices. In the case of $SU(2)$, they are proportional to the Pauli matrices: $\mathbf{t}_a = \tau_a/2$, with

$$\tau_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \tau_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \tau_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (\text{A.1.4})$$

f_{abc}	abc	d_{abc}	abc
1	123	$\frac{1}{\sqrt{3}}$	118, 228, 338
$\frac{1}{2}$	147, 246, 257, 345	$\frac{1}{2}$	146, 157, 256, 344, 355
$-\frac{1}{2}$	156, 367	$-\frac{1}{2}$	247, 366, 377
$\frac{\sqrt{3}}{2}$	458, 678	$-\frac{1}{2\sqrt{3}}$	448, 558, 668, 778
		$-\frac{1}{\sqrt{3}}$	888

TABLE A.1: Antisymmetric structure constants f_{abc} and symmetric symbols d_{abc} for the group $SU(3)$. The values for the remaining indices are obtained via permutation.

For $SU(3)$ they are given by the Gell-Mann matrices, $\mathfrak{t}_a = \lambda_a/2$, with

$$\begin{aligned}
\lambda_1 &= \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, & \lambda_2 &= \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, & \lambda_3 &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \\
\lambda_4 &= \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, & \lambda_5 &= \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix}, & & \\
\lambda_6 &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, & \lambda_7 &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix}, & \lambda_8 &= \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}.
\end{aligned} \tag{A.1.5}$$

In the adjoint representation of $SU(N)$, the generators are given by $(\mathfrak{t}_a)_{bc} = -if_{abc}$, so they are $(N^2 - 1) \times (N^2 - 1)$ matrices. The generators satisfy

$$\text{Tr}(\mathfrak{t}_a \mathfrak{t}_b) = T(R) \delta_{ab} \quad \text{and} \quad \left(\sum_a \mathfrak{t}_a^2 \right)_{ij} = C(R) \delta_{ij}, \tag{A.1.6}$$

where $T(R)$ is the Dynkin index and $C(R)$ the Casimir in the representation R :

- fundamental representation: $T(R) = \frac{1}{2}$, $C(R) = (N^2 - 1)/(2N)$,
- adjoint representation: $T(R) = C(R) = N$.

From Eq. (A.1.6) it follows that $T(R)D(A) = C(R)D(R)$, where $D(R)$ is the dimension of the representation R and $D(A)$ is the dimension of the adjoint (which defines the dimension of the group).

In the fundamental representation, one has the anticommutation relation

$$\{\mathfrak{t}_a, \mathfrak{t}_b\} = \frac{1}{N} \delta_{ab} + d_{abc} \mathfrak{t}_c, \tag{A.1.7}$$

where the totally symmetric d_{abc} are collected in Table A.1 for the case of $SU(3)$; for $SU(2)$, they are zero. The structure constants are then obtained by taking traces of the generators in the fundamental representation, as follows from (A.1.2), (A.1.6) and (A.1.7):

$$f_{abc} = -2i \text{Tr}([\mathfrak{t}_a, \mathfrak{t}_b] \mathfrak{t}_c), \quad d_{abc} = 2 \text{Tr}(\{\mathfrak{t}_a, \mathfrak{t}_b\} \mathfrak{t}_c). \tag{A.1.8}$$

A.2 $SU(N)$ representations

The group $SU(N)$ has $N^2 - 1$ generators \mathfrak{t}_a which form the basis of a Lie algebra defined by the commutator relations in (A.1.2). The group has rank $N - 1$, so there are $N - 1$ Casimir operators which commute with all generators and label the irreducible representations of $SU(N)$. A given irreducible representation is then defined by $N - 1$ numbers. It has dimension D and defines a D -dimensional invariant subspace that can be visualized by collecting its basis states in a multiplet. Rank $N - 1$ also entails that there are at most $N - 1$ generators that commute with each other; they form the Cartan subalgebra (the maximal Abelian subalgebra) and label the states within the multiplet. The multiplets are therefore geometric structures in $N - 1$ dimensions. The remaining generators are ladder operators that connect the states with each other. Here are examples for $N = 2, 3$ and 4:

SU(2): The group $SU(2)$ describes angular momentum (spin, isospin, etc) and has three generators \mathfrak{t}_a (often called J_a). It has rank one, so there is one Casimir operator ($\mathfrak{t}_a \mathfrak{t}_a = \mathbf{J}^2$) whose eigenvalues $j(j + 1)$ label the irreducible representations; the spin j can take integer and half-integer values. The Cartan subalgebra consists of only one generator (\mathfrak{t}_3) whose eigenvalues cover the interval $-j \dots j$ and label the states within each multiplet. Anticipating the notation for $SU(3)$, we denote the irreducible representations of $SU(2)$ by D^p , where $p = 2j = 0, 1, 2, \dots$. Their dimension (which we also call D^p for brevity) is then $D^p = p + 1$:

$$D^0 = \mathbf{1}, \quad D^1 = \mathbf{2}, \quad D^2 = \mathbf{3}, \quad D^3 = \mathbf{4}, \quad \dots \quad (\text{A.2.1})$$

In the two-dimensional fundamental representation D^1 , the generators are the Pauli matrices ($\mathfrak{t}_a = \tau_a/2$). Because $SU(2)$ has three generators, the adjoint representation is the three-dimensional one, D^2 .

SU(3): The group $SU(3)$ has eight generators \mathfrak{t}_a . It has rank two and therefore there are two Casimir operators (namely $\mathfrak{t}_a \mathfrak{t}_a$ and $d_{abc} \mathfrak{t}_a \mathfrak{t}_b \mathfrak{t}_c$) which label its irreducible representations. We call these representations D^{pq} ; they depend on two quantum numbers $p, q = 0, 1, 2, \dots$ and their dimension is

$$D^{pq} = \frac{1}{2} (p + 1)(q + 1)(p + q + 2). \quad (\text{A.2.2})$$

The lowest-dimensional irreducible representations are:

$$D^{00} = \mathbf{1}, \quad D^{10} = \mathbf{3}, \quad D^{20} = \mathbf{6}, \quad D^{11} = \mathbf{8}, \quad D^{30} = \mathbf{10}, \quad \dots \quad (\text{A.2.3})$$

$$D^{01} = \bar{\mathbf{3}}, \quad D^{02} = \bar{\mathbf{6}}, \quad D^{03} = \bar{\mathbf{10}}, \quad \dots$$

The fundamental triplet and antitriplet representations are D^{10} and D^{01} ; the adjoint representation is the octet D^{11} ; D^{00} is the singlet and D^{30} the decuplet. The multiplets can be constructed graphically as shown in Fig. A.1. The generators \mathfrak{t}_3 and $(2/\sqrt{3})\mathfrak{t}_8$ commute with each other and form the Cartan subalgebra, so their eigenvalues I_3 and Y label the states within the multiplets which are therefore planar objects. The remaining generators are ladder operators and connect these states with each other:

$$\mathfrak{t}_\pm = \mathfrak{t}_1 \pm i\mathfrak{t}_2, \quad \mathfrak{u}_\pm = \mathfrak{t}_6 \pm i\mathfrak{t}_7, \quad \mathfrak{v}_\pm = \mathfrak{t}_4 \pm i\mathfrak{t}_5. \quad (\text{A.2.4})$$

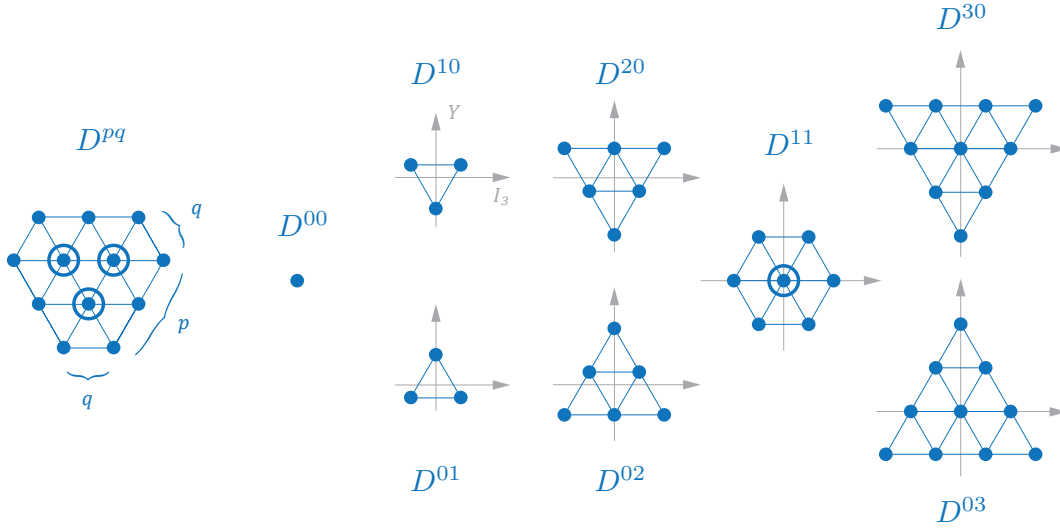


FIG. A.1: $SU(3)$ multiplets. The left diagram shows the generic construction and the remaining ones correspond to the lowest-dimensional representations from Eq. (B.2.7).

A generic multiplet is a hexagon in the (I_3, Y) plane; for $p = 0$ or $q = 0$ the hexagons degenerate to triangles. Each hexagon includes further degenerate states that are obtained by lowering p and q by one unit each.

SU(4): The group $SU(4)$ has rank three and therefore we have three quantum numbers $p, q, r = 0, 1, 2, \dots$ to label the representations D^{prq} . Their dimensions are

$$D^{prq} = \frac{1}{12} (p+1)(q+1)(r+1)(p+r+2)(q+r+2)(p+q+r+3). \quad (\text{A.2.5})$$

The lowest-dimensional irreducible representations are

$$D^{000} = \mathbf{1}, \quad \begin{matrix} D^{100} = \mathbf{4} \\ D^{001} = \bar{\mathbf{4}} \end{matrix}, \quad D^{010} = \mathbf{6}, \quad \begin{matrix} D^{200} = \mathbf{10} \\ D^{002} = \bar{\mathbf{10}} \end{matrix}, \quad D^{101} = \mathbf{15}, \quad \dots \quad (\text{A.2.6})$$

The fundamental representations are $\mathbf{4}$ and $\bar{\mathbf{4}}$, and $\mathbf{15}$ is the adjoint.

A.3 Product representations

Vectors that transform under the N -dimensional fundamental or antifundamental representations of $SU(N)$ satisfy the transformation law

$$\psi' = U\psi \quad \Leftrightarrow \quad \psi'_i = U_{ij} \psi_j, \quad \psi'^{\dagger} = \psi^{\dagger} U^{\dagger} \quad \Leftrightarrow \quad \psi'^{\star}_i = U^{\star}_{ij} \psi^{\star}_j, \quad (\text{A.3.1})$$

where $U \in \{D^1, D^{10}, D^{100}, \dots\}$ and $U^{\star} \in \{D^1, D^{01}, D^{001}, \dots\}$. What happens if we take tensor products of ψ and ψ^{\star} ? Higher-rank tensors are defined as those quantities that have the same transformation properties as the direct product of vectors. To

keep track of the (anti-)fundamental nature of the representations where they originate from, it is helpful to introduce upper and lower indices and write:

$$\psi_i \rightarrow \psi^i, \quad \psi_i^* \rightarrow \psi_i, \quad U_{ij} \rightarrow U^i_j, \quad U_{ij}^* \rightarrow U_i^j, \quad (\text{A.3.2})$$

together with the Einstein summation convention (for example, $UU^\dagger = 1$ becomes $U^i_j U_k^j = \delta_k^i$). The coefficients of a generic $SU(N)$ tensor of rank (n, m) then transform under the product representation

$$\psi'^{i_1 \dots i_n}_{j_1 \dots j_m} = (U^{i_1}_{k_1} \dots U^{i_n}_{k_n}) (U_{j_1}^{l_1} \dots U_{j_m}^{l_m}) \psi^{k_1 \dots k_n}_{l_1 \dots l_m} \quad (\text{A.3.3})$$

which is, however, not irreducible. To see this, permute the indices for example of a rank $(2, 0)$ tensor — it commutes with the $SU(N)$ transformation:

$$\psi'^{ij} = U^i_k U^j_l \psi^{kl}, \quad \psi'^{ji} = U^j_k U^i_l \psi^{kl} = U^i_k U^j_l \psi^{lk}. \quad (\text{A.3.4})$$

Therefore, the symmetric and antisymmetric combinations

$$S^{ij} = \frac{1}{2} (\psi^{ij} + \psi^{ji}), \quad A^{ij} = \frac{1}{2} (\psi^{ij} - \psi^{ji}) \quad (\text{A.3.5})$$

do not mix under $SU(N)$ and form irreducible subspaces. A 2×2 matrix can be decomposed into one antisymmetric and three symmetric components, a 3×3 matrix has three antisymmetric and six symmetric components; so we can write¹

$$SU(2): \quad \mathbf{2} \otimes \mathbf{2} = \mathbf{1}_A \oplus \mathbf{3}_S, \quad SU(3): \quad \mathbf{3} \otimes \mathbf{3} = \bar{\mathbf{3}}_A \oplus \mathbf{6}_S. \quad (\text{A.3.6})$$

These components transform now again under irreducible representations of $SU(N)$:

$$D^1 \otimes D^1 = D^0 \oplus D^2, \quad D^{10} \otimes D^{10} = D^{01} \oplus D^{20}. \quad (\text{A.3.7})$$

That is, if we arrange the components of the 2×2 tensor ψ^{ij} into a four-dimensional vector, the reducible representation matrix $D^1 \otimes D^1$ becomes block-diagonal:

$$\begin{pmatrix} \bullet \\ \bullet \\ \bullet \\ \bullet \end{pmatrix}' = \left(\begin{array}{c|c} D^0 & \\ \hline & D^2 \end{array} \right) \begin{pmatrix} \bullet \\ \bullet \\ \bullet \\ \bullet \end{pmatrix}, \quad (\text{A.3.8})$$

and similarly for the 3×3 matrix ψ^{ijk} .

While the symmetry argument is not directly applicable for tensors of mixed rank, the trace ψ^i_i is invariant under $SU(N)$ and can be factored out. For example for a tensor of rank $(1, 1)$:

$$\psi'^i_i = U^i_k U_i^l \psi^k_l = \delta_k^l \psi^k_l = \psi^i_i, \quad (\text{A.3.9})$$

so we have $\mathbf{2} \otimes \bar{\mathbf{2}} = \mathbf{1} \oplus \mathbf{3}$ in $SU(2)$, $\mathbf{3} \otimes \bar{\mathbf{3}} = \mathbf{1} \oplus \mathbf{8}$ in $SU(3)$, etc. Generally, in order to study product representations of $SU(N)$ one must work out the simultaneous irreducible representations of $SU(N)$ and the permutation group. The latter can be most easily obtained with the help of Young diagrams.

¹It will become clear from the discussion of Young diagrams why $\bar{\mathbf{3}}$ instead of $\mathbf{3}$ appears here.

Young diagrams. Let's forget again about $SU(N)$ for the moment. Consider a quantity $f_{a_1 \dots a_n}$ that carries n indices (where each can run from $1 \dots N$), or suppose we want to form the n -fold tensor product of N -dimensional vectors: $\psi_{a_1} \otimes \dots \otimes \psi_{a_n}$. Without specifying the index range N , there are $n!$ possible permutations which fall into irreducible subspaces of the permutation group S_n . These can be visualized by Young diagrams: take n boxes and stick them together in all possible ways so that the number of boxes in each consecutive row (from top to bottom) and each consecutive column (from left to right) does not increase. For example:

$$S_2: \quad \square, \begin{array}{|c|} \hline \square \\ \hline \end{array} \quad S_3: \quad \square\square\square, \begin{array}{|c|c|} \hline \square & \square \\ \hline \end{array}, \begin{array}{|c|} \hline \square \\ \hline \square \\ \hline \end{array} \quad S_4: \quad \square\square\square\square, \begin{array}{|c|c|c|} \hline \square & \square & \square \\ \hline \end{array}, \begin{array}{|c|c|} \hline \square & \square \\ \hline \square & \square \\ \hline \end{array}, \begin{array}{|c|} \hline \square \\ \hline \square \\ \hline \square \\ \hline \end{array}, \begin{array}{|c|} \hline \square \\ \hline \square \\ \hline \square \\ \hline \end{array}.$$

Eventually we will fill these boxes with the N possible indices (or the N possible vector components), but let's see first how far we can get without doing that. A row denotes symmetrization, a column antisymmetrization. For S_3 , $\square\square\square$ is totally symmetric in all indices, $\begin{array}{|c|c|} \hline \square & \square \\ \hline \end{array}$ has mixed symmetry, and the remaining diagram is totally antisymmetric. Each Young diagram corresponds to an irreducible representation of S_n .

The dimension d of a representation can be inferred from the number of 'standard tableaux' that it permits. A standard tableau is filled with (distinct) numbers from $1 \dots n$ which increase in each row and each column. For example:

$$\underbrace{\begin{array}{|c|c|c|} \hline 1 & 2 & 3 \\ \hline \end{array}}_{d=1}, \quad \underbrace{\begin{array}{|c|c|} \hline 1 & 2 \\ \hline 3 & \end{array}, \begin{array}{|c|c|} \hline 1 & 3 \\ \hline 2 & \end{array}}_{d=2}, \quad \underbrace{\begin{array}{|c|c|} \hline 1 & 2 \\ \hline 3 & 4 \\ \hline \end{array}, \begin{array}{|c|c|} \hline 1 & 3 \\ \hline 2 & 4 \\ \hline \end{array}}_{d=2}, \quad \text{etc.} \quad (\text{A.3.10})$$

For larger Young tableaux this exercise can become tedious. Fortunately, the dimension of a Young diagram can be also determined from the hook factor h which is the product of all hook lengths in a diagram. The hook length of a box counts the number of boxes directly below and to its right, plus counting the box itself. In the following Young tableau the hook lengths are given for each box:

$$\begin{array}{|c|c|c|} \hline 5 & 3 & 1 \\ \hline 3 & 1 & \\ \hline 1 & & \\ \hline \end{array} \Rightarrow h = 5 \cdot 3 \cdot 3 \cdot 1 \cdot 1 \cdot 1 = 45. \quad (\text{A.3.11})$$

The dimension of a Young diagram is then given by $d = n!/h$ (in this example, we would have $d = 6!/45 = 16$).

The dimension d is also the multiplicity of each diagram in the $n!$ -dimensional reducible representation of S_n . For example in S_3 , the 6 possible permutations of a function of three indices can be arranged in a symmetric singlet, an antisymmetric singlet, and two doublets:

$$\square\square\square + 2 \cdot \begin{array}{|c|c|} \hline \square & \square \\ \hline \end{array} + \begin{array}{|c|} \hline \square \\ \hline \square \\ \hline \end{array} \Rightarrow n! = 6 = 1 + 2 \cdot 2 + 1 = \sum_i d_i^2. \quad (\text{A.3.12})$$

For S_2 , we have $2 = 1 + 1$ ($\square\square$ and $\begin{array}{|c|} \hline \square \\ \hline \end{array}$) and for S_4 : $4! = 1 + 3 \cdot 3 + 2 \cdot 2 + 3 \cdot 3 + 1$. Phrased differently: in S_2 , we can arrange the 2 possible permutations of f_{ab} into a singlet and an antisinglet; in S_3 , we can distribute the 6 permutations of f_{abc} into a singlet, an antisinglet and two doublets; and in S_4 , we can arrange the 24 permutations of f_{abcd} into a singlet, an antisinglet, two doublets, three triplets and three antitriplets.

Conjugate representation. In terms of Young diagrams, the conjugate representations can be obtained in the following way. For each column, replace the j boxes in the column by $N - j$ boxes, and flip the diagram around the vertical axis. For example in $SU(4)$:

$$\overline{\begin{array}{|c|} \hline \square \\ \hline \square \\ \hline \square \\ \hline \end{array}} = \begin{array}{|c|} \hline \square \\ \hline \square \\ \hline \square \\ \hline \end{array}, \quad \overline{\begin{array}{|c|} \hline \square \\ \hline \square \\ \hline \end{array}} = \begin{array}{|c|} \hline \square \\ \hline \square \\ \hline \end{array}, \quad \overline{\begin{array}{|c|c|} \hline \square & \square \\ \hline \square & \square \\ \hline \square & \square \\ \hline \end{array}} = \begin{array}{|c|c|} \hline \square & \square \\ \hline \square & \square \\ \hline \square & \square \\ \hline \end{array}, \quad \overline{\begin{array}{|c|c|} \hline \square & \square \\ \hline \square & \square \\ \hline \end{array}} = \begin{array}{|c|c|} \hline \square & \square \\ \hline \square & \square \\ \hline \end{array}. \quad (\text{A.3.16})$$

This entails that representations which only differ by columns of length N attached to the left are equivalent, for example in $SU(3)$:

$$\overline{\begin{array}{|c|c|} \hline \square & \square \\ \hline \square & \square \\ \hline \square & \square \\ \hline \end{array}} = \begin{array}{|c|} \hline \square \\ \hline \square \\ \hline \square \\ \hline \end{array} = \overline{\begin{array}{|c|} \hline \square \\ \hline \square \\ \hline \square \\ \hline \end{array}} \Rightarrow \begin{array}{|c|c|} \hline \square & \square \\ \hline \square & \square \\ \hline \square & \square \\ \hline \end{array} = \begin{array}{|c|} \hline \square \\ \hline \square \\ \hline \square \\ \hline \end{array}, \quad (\text{A.3.17})$$

or in $SU(2)$:

$$\overline{\begin{array}{|c|c|} \hline \square & \square \\ \hline \square & \square \\ \hline \end{array}} = \begin{array}{|c|} \hline \square \\ \hline \square \\ \hline \end{array}, \quad \overline{\begin{array}{|c|c|c|} \hline \square & \square & \square \\ \hline \square & \square & \square \\ \hline \end{array}} = \begin{array}{|c|c|} \hline \square & \square \\ \hline \square & \square \\ \hline \end{array}, \quad \overline{\begin{array}{|c|c|c|} \hline \square & \square & \square \\ \hline \square & \square & \square \\ \hline \square & \square & \square \\ \hline \end{array}} = \begin{array}{|c|c|} \hline \square & \square \\ \hline \square & \square \\ \hline \square & \square \\ \hline \end{array} = \begin{array}{|c|} \hline \square \\ \hline \square \\ \hline \square \\ \hline \end{array}. \quad (\text{A.3.18})$$

It also implies that in $SU(2)$ each conjugate representation is identical to the representation itself:

$$\overline{\begin{array}{|c|} \hline \square \\ \hline \square \\ \hline \square \\ \hline \end{array}} = \begin{array}{|c|} \hline \square \\ \hline \square \\ \hline \square \\ \hline \end{array}, \quad \overline{\begin{array}{|c|c|} \hline \square & \square \\ \hline \square & \square \\ \hline \end{array}} = \begin{array}{|c|c|} \hline \square & \square \\ \hline \square & \square \\ \hline \end{array}, \quad \overline{\begin{array}{|c|c|c|} \hline \square & \square & \square \\ \hline \square & \square & \square \\ \hline \square & \square & \square \\ \hline \end{array}} = \begin{array}{|c|c|c|} \hline \square & \square & \square \\ \hline \square & \square & \square \\ \hline \square & \square & \square \\ \hline \end{array}. \quad (\text{A.3.19})$$

$SU(N)$ representations as Young diagrams. If we put everything together we arrive at Table A.2, which states the dimension D for the irreducible representations of $SU(N)$, together with their Young diagrams which carry dimension d . In this way we can identify each irreducible representation of $SU(N)$ directly with a Young diagram, for example for $SU(3)$:

$$D^{00} = \begin{array}{|c|} \hline \square \\ \hline \square \\ \hline \square \\ \hline \end{array}, \quad D^{10} = \begin{array}{|c|} \hline \square \\ \hline \square \\ \hline \end{array}, \quad D^{20} = \begin{array}{|c|c|} \hline \square & \square \\ \hline \square & \square \\ \hline \square & \square \\ \hline \end{array}, \quad D^{11} = \begin{array}{|c|c|} \hline \square & \square \\ \hline \square & \square \\ \hline \end{array}, \quad D^{30} = \begin{array}{|c|c|} \hline \square & \square \\ \hline \square & \square \\ \hline \square & \square \\ \hline \end{array}, \quad D^{03} = \begin{array}{|c|c|c|} \hline \square & \square & \square \\ \hline \square & \square & \square \\ \hline \square & \square & \square \\ \hline \end{array}, \quad \dots$$

In general, the correspondence is given by:

- $SU(2) : D^p = \overbrace{\begin{array}{|c|c|c|} \hline \square & \square & \square \\ \hline \square & \square & \square \\ \hline \end{array}}^p$
- $SU(3) : D^{pq} = \overbrace{\begin{array}{|c|c|c|} \hline \square & \square & \square \\ \hline \square & \square & \square \\ \hline \square & \square & \square \\ \hline \end{array}}^q \overbrace{\begin{array}{|c|c|c|} \hline \square & \square & \square \\ \hline \square & \square & \square \\ \hline \end{array}}^p$
- $SU(4) : D^{prq} = \overbrace{\begin{array}{|c|c|c|} \hline \square & \square & \square \\ \hline \square & \square & \square \\ \hline \square & \square & \square \\ \hline \end{array}}^q \overbrace{\begin{array}{|c|c|c|} \hline \square & \square & \square \\ \hline \square & \square & \square \\ \hline \square & \square & \square \\ \hline \end{array}}^r \overbrace{\begin{array}{|c|c|c|} \hline \square & \square & \square \\ \hline \square & \square & \square \\ \hline \end{array}}^p$

	d	h	D	$SU(2)$	$SU(3)$	$SU(4)$
	1	1	N	2	3	4
} $N-1$	1	$(N-1)!$	N	2	$\bar{3}$	$\bar{4}$
} N	1	$N!$	1	1	1	1
} $N-1$	$N-1$	$\frac{N!}{N-1}$	N^2-1	3	8	15
	1	2	$\frac{N(N+1)}{2}$	3	6	10
	1	2	$\frac{N(N-1)}{2}$	1	$\bar{3}$	6
	1	6	$\frac{N(N+1)(N+2)}{6}$	4	10	20
	2	3	$\frac{N(N+1)(N-1)}{3}$	2	8	20
	1	6	$\frac{N(N-1)(N-2)}{6}$	-	1	$\bar{4}$
	1	24	$\frac{N(N+1)(N+2)(N+3)}{24}$	5	15	35
	3	8	$\frac{N(N+1)(N+2)(N-1)}{8}$	3	15	45
	2	12	$\frac{N^2(N+1)(N-1)}{12}$	1	$\bar{6}$	20
	3	8	$\frac{N(N+1)(N-1)(N-2)}{8}$	-	3	15
	1	24	$\frac{N(N-1)(N-2)(N-3)}{24}$	-	-	1

TABLE A.2: Identification of irreducible $SU(N)$ representations with Young diagrams for S_2 , S_3 and S_4 . The second and third columns state the dimension d of the diagram and the hook factor h . The remaining columns show the dimension D of the $SU(N)$ representation, with examples for $SU(2)$, $SU(3)$ and $SU(4)$. The first four rows collect the fundamental, antifundamental, singlet and (highlighted in red) adjoint representations.

In addition, arbitrarily many columns of length N can be attached from the left because this produces an equivalent representation.

Product representations. Now let's return to the construction of product representations. So far we have only looked at products of fundamental representations, e.g.

$$\square \otimes \square \otimes \square = \left(\square\square \oplus \begin{array}{|c|} \hline \square \\ \hline \end{array} \right) \otimes \square = \square\square\square \oplus \begin{array}{|c|c|} \hline \square & \square \\ \hline \end{array} \oplus \begin{array}{|c|c|} \hline \square & \square \\ \hline \end{array} \oplus \begin{array}{|c|} \hline \square \\ \hline \square \\ \hline \end{array}, \quad (\text{A.3.20})$$

which becomes $\mathbf{3} \otimes \mathbf{3} \otimes \mathbf{3} = \mathbf{10} \oplus \mathbf{8} \oplus \mathbf{8} \oplus \mathbf{1}$ in $SU(3)$. Now suppose we want to evaluate a tensor product such as this:

$$\begin{array}{|c|c|} \hline \square & \square \\ \hline \square & \square \\ \hline \square & \square \\ \hline \end{array} \otimes \begin{array}{|c|c|} \hline \square & \square \\ \hline \end{array}. \quad (\text{A.3.21})$$

How does the general rule work? Let's denote the left diagram by X and the right one by Y . The prescription goes like this: start by filling the boxes in the top row of Y with labels 'a', the boxes in the second row with labels 'b', etc. Take the 'a' boxes from the top row and attach them to X in all possible ways, so that the number of boxes in each consecutive row (from top to down) and each consecutive column (from left to right) does not increase. Then,

- If you end up with more than one a in a column, delete the diagram (because a column means antisymmetrization).
- If a diagram contains a column with N boxes, delete just that column (because it yields an equivalent representation). If it contains a column with more than N boxes, delete the diagram (because we cannot antisymmetrize more flavors than we have).

All identical diagrams count just once. Repeat these steps for the second row in Y with the b 's, the third row with the c 's and so on until you're done. One final step:

- In the resulting diagrams, go from right to left in the first row, then in the second row, etc. At any point along that path, the number of b boxes you have picked up must be smaller than the number of a boxes. If this is not the case, delete the diagram. (For example, if the top right box contains 'b', you can delete the diagram because the number of a 's at this point is zero.) Apply the same logic for $\#c < \#b$, etc.

Here's an example:

$$\begin{aligned} \square \otimes \begin{array}{|c|} \hline a \\ \hline b \\ \hline \end{array} &= \left(\square a \oplus \begin{array}{|c|} \hline \square \\ \hline a \\ \hline \end{array} \right) \otimes \begin{array}{|c|} \hline b \\ \hline \end{array} = \square a b \oplus \begin{array}{|c|c|} \hline \square & a \\ \hline b & \end{array} \oplus \begin{array}{|c|c|} \hline \square & b \\ \hline a & \end{array} \oplus \begin{array}{|c|} \hline \square \\ \hline a \\ \hline b \\ \hline \end{array} \\ &= \begin{array}{|c|c|} \hline \square & a \\ \hline b & \end{array} \oplus \begin{array}{|c|} \hline \square \\ \hline a \\ \hline b \\ \hline \end{array}. \end{aligned} \quad (\text{A.3.22})$$

In the case of $SU(3)$, this becomes $\mathbf{3} \otimes \bar{\mathbf{3}} = \mathbf{8} \oplus \mathbf{1}$. (For $SU(2)$, the same relation entails $\mathbf{2} \otimes \mathbf{1} = \mathbf{2}$ because the second diagram vanishes, and for $SU(4)$ we get $\mathbf{4} \otimes \bar{\mathbf{6}} = \mathbf{20} \oplus \bar{\mathbf{4}}$.) Of course we would have obtained the same result faster if we had started from $\bar{\mathbf{3}} \otimes \mathbf{3}$, but with this strategy it is also straightforward to verify more complicated tensor products.

We collect some useful results:

- $SU(2)$: here the tensor representations can also be inferred from the angular momentum addition rules:

$$(2j+1) \otimes (2j'+1) = \bigoplus_{J=|j-j'|}^{j+j'} (2J+1) \Rightarrow \begin{aligned} \mathbf{2} \otimes \mathbf{2} &= \mathbf{1} \oplus \mathbf{3}, \\ \mathbf{3} \otimes \mathbf{3} &= \mathbf{1} \oplus \mathbf{3} \oplus \mathbf{5}, \\ \mathbf{4} \otimes \mathbf{4} &= \mathbf{1} \oplus \mathbf{3} \oplus \mathbf{5} \oplus \mathbf{7}, \\ \mathbf{2} \otimes \mathbf{2} \otimes \mathbf{2} &= \mathbf{2} \oplus \mathbf{2} \oplus \mathbf{4}. \end{aligned} \quad (\text{A.3.23})$$

- $SU(3)$: conjugate representations are here no longer equivalent. Some frequently used decompositions are

$$\begin{aligned} \mathbf{3} \otimes \bar{\mathbf{3}} &= \mathbf{1} \oplus \mathbf{8}, & \mathbf{6} \otimes \bar{\mathbf{6}} &= \mathbf{1} \oplus \mathbf{8} \oplus \mathbf{27}, \\ \mathbf{3} \otimes \mathbf{3} &= \bar{\mathbf{3}} \oplus \mathbf{6}, & \mathbf{8} \otimes \mathbf{8} &= \mathbf{1} \oplus \mathbf{8} \oplus \mathbf{8} \oplus \mathbf{10} \oplus \bar{\mathbf{10}} \oplus \mathbf{27}, \\ \mathbf{3} \otimes \mathbf{6} &= \mathbf{8} \oplus \mathbf{10}, & \mathbf{3} \otimes \mathbf{3} \otimes \mathbf{3} &= \mathbf{1} \oplus \mathbf{8} \oplus \mathbf{8} \oplus \mathbf{10}. \end{aligned} \quad (\text{A.3.24})$$

- $SU(6)$: $\mathbf{6} \otimes \mathbf{6} \otimes \mathbf{6} = \mathbf{20} \oplus \mathbf{70} \oplus \mathbf{70} \oplus \mathbf{56}$.

Appendix B

Poincaré group

Poincaré invariance is the fundamental symmetry in particle physics. A relativistic quantum field theory must have a Poincaré-invariant action. This means that its fields must transform under representations of the Poincaré group and Poincaré invariance must be implemented unitarily on the state space. Here we will collect some properties of the Lorentz and Poincaré groups together with their representation theory.

B.1 Lorentz and Poincaré group

Lorentz group. We work in Minkowski space with the metric tensor $g = (g_{\mu\nu}) = \text{diag}(1, -1, -1, -1)$, where the scalar product is given by

$$x \cdot y := x^T g y = x^0 y^0 - \mathbf{x} \cdot \mathbf{y} = g_{\mu\nu} x^\mu y^\nu = x_\mu y^\mu. \quad (\text{B.1.1})$$

Instead of carrying around explicit instances of g , it is more convenient to use the index notation where upper and lower indices are summed over. **Lorentz transformations** are those transformations $x' = \Lambda x$ that leave the scalar product invariant:

$$(\Lambda x) \cdot (\Lambda y) = x \cdot y \quad \Rightarrow \quad x^T \Lambda^T g \Lambda y = x^T g y \quad \Rightarrow \quad \Lambda^T g \Lambda = g. \quad (\text{B.1.2})$$

Written in components, this condition takes the form

$$g_{\alpha\beta} = g_{\mu\nu} \Lambda^\mu_\alpha \Lambda^\nu_\beta. \quad (\text{B.1.3})$$

Since the metric tensor is symmetric, this gives 10 constraints; the Lorentz transformation Λ is a 4×4 matrix, so it depends on $16 - 10 = 6$ independent parameters. If we write an infinitesimal transformation as $\Lambda^\alpha_\beta = \delta^\alpha_\beta + \varepsilon^\alpha_\beta + \dots$, then it follows from Eq. (B.1.3) that $\varepsilon_{\alpha\beta} = -\varepsilon_{\beta\alpha}$ must be totally antisymmetric.

The transformations of a space with coordinates $\{y_1 \dots y_n, x_1 \dots x_m\}$ that leave the quadratic form $(y_1^2 + \dots + y_n^2) - (x_1^2 + \dots + x_m^2)$ invariant constitute the orthogonal group $O(m, n)$, so the Lorentz group is $O(3, 1)$. The group axioms are satisfied; there

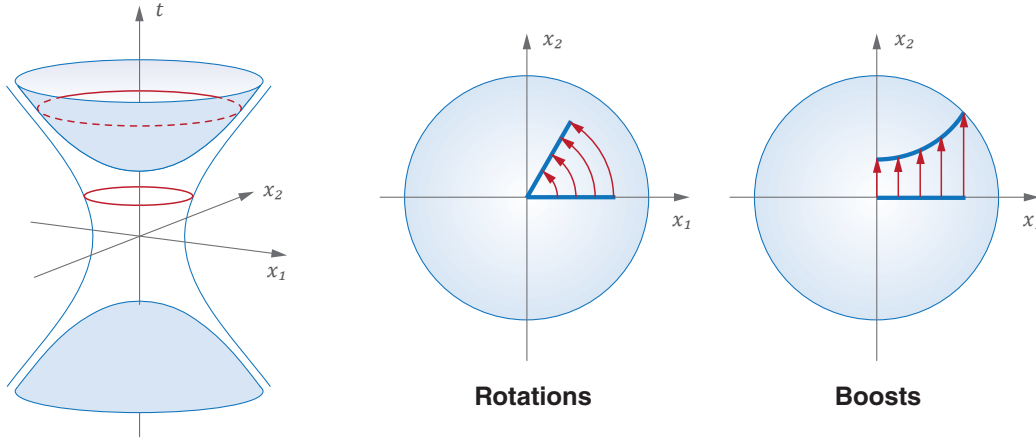


FIG. B.1: Invariant hyperboloids for the Lorentz group. Rotations go around circles and boosts in fixed directions \mathbf{n} along the surface.

is a unit element ($\Lambda = 1$), and each Λ has an inverse element because it is invertible: $\Lambda^T g \Lambda = g \Rightarrow (\det \Lambda)^2 = 1 \Rightarrow \det \Lambda = \pm 1$. Eq. (B.1.3) also entails

$$g_{\mu\nu} \Lambda^\mu{}_0 \Lambda^\nu{}_0 = (\Lambda_0^0)^2 - \sum_k (\Lambda^k{}_0)^2 = 1 \quad \Rightarrow \quad (\Lambda_0^0)^2 \geq 1. \quad (\text{B.1.4})$$

Depending of the signs of $\det \Lambda$ and Λ_0^0 , the Lorentz group has four disconnected components. The subgroup with $\det \Lambda = 1$ and $\Lambda_0^0 \geq 1$ is called the proper orthochronous Lorentz group $SO(3,1)^\uparrow$; it contains the identity matrix and preserves the direction of time and parity. The other three branches can be constructed from a given $\Lambda \in SO(3,1)^\uparrow$ combined with a space and/or time reflection:

- $SO(3,1)^\uparrow \times$ spatial reflections: $\Lambda_0^0 \geq 1$, $\det \Lambda = -1$
- $SO(3,1)^\uparrow \times$ time reversal: $\Lambda_0^0 \leq -1$, $\det \Lambda = -1$
- $SO(3,1)^\uparrow \times$ spacetime reflection: $\Lambda_0^0 \leq -1$, $\det \Lambda = 1$

Lorentz transformations preserve the norm $x^2 = x \cdot x$ in Minkowski space, which is positive for timelike four-vectors, negative for spacelike vectors, or zero for lightlike vectors. Therefore, they are transformations along the hypersurfaces of constant norm (Fig. B.1). For a four-momentum with positive norm $p^2 = m^2$ these are the forward and backward mass shells. For vanishing norm the hypersurface becomes the light cone, and for negative norm the hyperboloid lies outside of the light cone.

Each $\Lambda \in SO(3,1)^\uparrow$ can be reconstructed from a **Lorentz boost** with velocity $\beta = \frac{v}{c}$ in direction \mathbf{n} (with $|\beta| < 1$) together with a spatial **rotation** $R(\boldsymbol{\alpha}) \in SO(3)$:

$$\Lambda = \underbrace{\begin{pmatrix} \gamma & \gamma \beta \mathbf{n}^T \\ \gamma \beta \mathbf{n} & \mathbf{1} + (\gamma - 1) \mathbf{n} \mathbf{n}^T \end{pmatrix}}_{\mathbf{L}(\boldsymbol{\beta})} \underbrace{\begin{pmatrix} 1 & \mathbf{0}^T \\ \mathbf{0} & R(\boldsymbol{\alpha}) \end{pmatrix}}_{\mathbf{R}(\boldsymbol{\alpha})}, \quad \gamma = \frac{1}{\sqrt{1 - \beta^2}}. \quad (\text{B.1.5})$$

In the nonrelativistic limit $|\beta| \ll 1 \Rightarrow \gamma \approx 1$ this recovers the Galilei transformation.

The six group parameters can therefore be chosen as the three components of the velocity $\beta \mathbf{n}$ and the three rotation angles $\boldsymbol{\alpha}$. One can show that interchanging the order in Eq. (B.1.5) yields

$$\Lambda = L(\boldsymbol{\beta}) R(\boldsymbol{\alpha}) = R(\boldsymbol{\alpha}) L(R(\boldsymbol{\alpha})^{-1} \boldsymbol{\beta}). \quad (\text{B.1.6})$$

The rotation group $SO(3)$ forms a subgroup of the Lorentz group (two consecutive rotations form another one) whereas boosts do not: the product of two boosts generally also involves a rotation as in Eq. (B.1.6). There are two properties that will become important later in the context of representations: the Lorentz group is not compact because it contains boosts (hence all unitary representations are infinite-dimensional); and it is not simply connected because it contains rotations (so we need to study the representations of its universal covering group $SL(2, \mathbb{C})$).

Poincaré group. Actually, the fact that the Lorentz group leaves the norm x^2 of a vector invariant is not enough because on physical grounds we need the line element $(dx)^2 = g_{\mu\nu} dx^\mu dx^\nu = c^2(dt)^2 - (d\mathbf{x})^2$ to be invariant. This guarantees that the speed of light is the same in every inertial frame, and it allows us to add constant translations to the Lorentz transformation:

$$x' = T(\Lambda, a) x = \Lambda x + a. \quad (\text{B.1.7})$$

The resulting 10-parameter group which contains translations, rotations and boosts is the **Poincaré group** or inhomogeneous Lorentz group. We can check again that the group axioms are satisfied: two consecutive Poincaré transformations form another one,

$$T(\Lambda', a') T(\Lambda, a) = T(\Lambda' \Lambda, a' + \Lambda' a), \quad (\text{B.1.8})$$

the transformation is associative: $(T T') T'' = T (T' T'')$, the unit element is $T(1, 0)$, and by equating Eq. (B.1.8) with $T(1, 0)$ we can read off the inverse element:

$$T^{-1}(\Lambda, a) = T(\Lambda^{-1}, -\Lambda^{-1} a). \quad (\text{B.1.9})$$

In analogy to above, the component which contains the identity $T(1, 0)$ is called $ISO(3, 1)^\uparrow$, where I stands for *inhomogeneous*. This is the fundamental symmetry group of physics that transforms inertial frames into one another.

Poincaré algebra. Consider now the representations $U(\Lambda, a)$ of the Poincaré group on some vector space. They inherit the transformation properties from Eqs. (B.1.8–B.1.9), and we use the symbol U although they are not necessarily unitary. The Poincaré group $ISO(3, 1)^\uparrow$ is a Lie group and therefore its elements can be written as

$$U(\Lambda, a) = e^{\frac{i}{2} \varepsilon_{\mu\nu} M^{\mu\nu}} e^{i a_\mu P^\mu} = 1 + \frac{i}{2} \varepsilon_{\mu\nu} M^{\mu\nu} + i a_\mu P^\mu + \dots, \quad (\text{B.1.10})$$

where the explicit forms of $U(\Lambda, a)$ and the generators $M^{\mu\nu}$ and P^μ depend on the representation. Since $\varepsilon_{\mu\nu}$ is totally antisymmetric, $M^{\mu\nu}$ can also be chosen antisymmetric. It contains the six generators of the Lorentz group, whereas the momentum operator

P^μ is the generator of spacetime translations. $M^{\mu\nu}$ and P^μ form a Lie algebra whose commutator relations can be derived from

$$U(\Lambda, a)U(\Lambda', a')U^{-1}(\Lambda, a) = U(\Lambda\Lambda'\Lambda^{-1}, a + \Lambda a' - \Lambda\Lambda'\Lambda^{-1}a), \quad (\text{B.1.11})$$

which follows from the composition rules (B.1.8) and (B.1.9). Inserting infinitesimal transformations (B.1.10) for each $U(\Lambda = 1 + \varepsilon, a)$, with $U^{-1}(\Lambda, a) = U(1 - \varepsilon, -a)$, keeping only linear terms in all group parameters ε , ε' , a and a' , and comparing coefficients of the terms $\sim \varepsilon\varepsilon'$, $a\varepsilon'$, $\varepsilon a'$ and aa' leads to the identities

$$i[M^{\mu\nu}, M^{\rho\sigma}] = g^{\mu\sigma}M^{\nu\rho} + g^{\nu\rho}M^{\mu\sigma} - g^{\mu\rho}M^{\nu\sigma} - g^{\nu\sigma}M^{\mu\rho}, \quad (\text{B.1.12})$$

$$i[P^\mu, M^{\rho\sigma}] = g^{\mu\rho}P^\sigma - g^{\mu\sigma}P^\rho, \quad (\text{B.1.13})$$

$$[P^\mu, P^\nu] = 0 \quad (\text{B.1.14})$$

which define the **Poincaré algebra**. A shortcut to arrive at the Lorentz algebra relation (B.1.12) is to calculate the generator $M^{\mu\nu}$ directly in the four-dimensional representation, where $U(\Lambda, 0) = \Lambda$ is the Lorentz transformation itself:

$$U(\Lambda, 0)^\alpha{}_\beta = \delta^\alpha_\beta + \frac{i}{2}\varepsilon_{\mu\nu}(M^{\mu\nu})^\alpha{}_\beta + \dots = \Lambda^\alpha{}_\beta = \delta^\alpha_\beta + \varepsilon^\alpha{}_\beta + \dots \quad (\text{B.1.15})$$

This is solved by the tensor

$$(M^{\mu\nu})^\alpha{}_\beta = -i(g^{\mu\alpha}\delta^\nu{}_\beta - g^{\nu\alpha}\delta^\mu{}_\beta) \quad (\text{B.1.16})$$

which satisfies the commutator relation (B.1.12).

We can cast the Poincaré algebra relations in a less compact but more useful form. The antisymmetric matrix $\varepsilon_{\mu\nu}$ contains the six group parameters and the antisymmetric matrix $M^{\mu\nu}$ the six generators. If we define the generator of $SO(3)$ rotations \mathbf{J} (the angular momentum) and the generator of boosts \mathbf{K} via

$$M^{ij} = -\varepsilon_{ijk}J^k \quad \Leftrightarrow \quad J^i = -\frac{1}{2}\varepsilon_{ijk}M^{jk}, \quad M^{0i} = K^i, \quad (\text{B.1.17})$$

then the commutator relations take the form

$$\begin{aligned} [J^i, J^j] &= i\varepsilon_{ijk}J^k, & [J^i, P^j] &= i\varepsilon_{ijk}P^k, & [P^i, P^j] &= 0, \\ [J^i, K^j] &= i\varepsilon_{ijk}K^k, & [K^i, P^j] &= i\delta_{ij}P_0, & [J^i, P_0] &= 0, \\ [K^i, K^j] &= -i\varepsilon_{ijk}J^k, & [K^i, P_0] &= iP^i, & [P^i, P_0] &= 0. \end{aligned} \quad (\text{B.1.18})$$

If we similarly define $\varepsilon_{ij} = -\varepsilon_{ijk}\phi^k$ and $\varepsilon_{0i} = s^i$, we obtain

$$\frac{i}{2}\varepsilon_{\mu\nu}M^{\mu\nu} = i\boldsymbol{\phi} \cdot \mathbf{J} + i\mathbf{s} \cdot \mathbf{K}. \quad (\text{B.1.19})$$

\mathbf{J} is hermitian but, because the Lorentz group is not compact, \mathbf{K} is antihermitian for all finite-dimensional representations which prevents them from being unitary. From (B.1.18) we see that boosts and rotations generally do not commute unless the boost and rotation axes coincide. Moreover, P_0 (which becomes the Hamilton operator in the quantum theory) commutes with rotations and spatial translations but not with boosts and therefore the eigenvalues of \mathbf{K} cannot be used for labeling physical states.

Casimir operators. The Casimir operators of a Lie group are those that commute with all generators and therefore allow us to label the irreducible representations. The Lorentz group has two Casimirs which are given by

$$C_1 = \frac{1}{2} M^{\mu\nu} M_{\mu\nu} = \mathbf{J}^2 - \mathbf{K}^2, \quad C_2 = \frac{1}{2} \widetilde{M}^{\mu\nu} M_{\mu\nu} = 2\mathbf{J} \cdot \mathbf{K}. \quad (\text{B.1.20})$$

The 'dual' generator is defined in analogy to Eq. (2.1.32): $\widetilde{M}_{\mu\nu} = \frac{1}{2} \varepsilon_{\mu\nu\alpha\beta} M^{\alpha\beta}$. Using $[AB, C] = A[B, C] + [A, C]B$ it is straightforward to check that both operators commute with $M^{\mu\nu}$; they are Lorentz-*invariant*.

Unfortunately, when we turn to the full Poincaré group C_1 and C_2 do not commute with P^μ , so they are not Poincaré-invariant. In turn, $P^2 = P^\mu P_\mu$ is invariant; from Eqs. (B.1.13–B.1.14) it is easy to see that it commutes with all generators P^μ and $M^{\mu\nu}$ (for example, the contraction of (B.1.13) with P^μ gives zero). P^2 is therefore a Casimir operator of the Poincaré group. The second Casimir is the square $W^2 = W^\mu W_\mu$ of the **Pauli-Lubanski vector**

$$W_\mu = -\frac{1}{2} \varepsilon_{\mu\rho\sigma\lambda} M^{\rho\sigma} P^\lambda. \quad (\text{B.1.21})$$

Since W^μ is a four-vector, W^2 is Lorentz-invariant and must commute with $M^{\mu\nu}$. W^μ commutes with the momentum operator because of Eq. (B.1.13), $[P^\mu, W^\nu] = 0$, and therefore also $[P^\mu, W^2] = 0$. Hence, both P^2 and W^2 are not only Lorentz- but also Poincaré-invariant. Written in components, the Pauli-Lubanski vector has the form

$$W_0 = \mathbf{P} \cdot \mathbf{J}, \quad \mathbf{W} = P_0 \mathbf{J} + \mathbf{P} \times \mathbf{K}. \quad (\text{B.1.22})$$

Working out W^2 in generality is a bit cumbersome, but for $P^2 = m^2 > 0$ we can define a rest frame where $\mathbf{P} = 0$. In that frame one has $W_0 = 0$, $\mathbf{W} = m \mathbf{J}$ and $W^2 = -m^2 \mathbf{J}^2$. The eigenvalues of \mathbf{J}^2 in the rest frame are $j(j+1)$, but since W^2 is Poincaré-invariant, so must be j . Here lies the origin of *spin*: from the point of view of the Poincaré group, the **mass** m and **spin** j are the only Poincaré-invariant quantum numbers that we can assign to a physical state.

We can derive this in another way so that also the connection with the Casimir operators (B.1.20) of the Lorentz group becomes more transparent. Define the transverse projection of $M^{\mu\nu}$ with respect to P :

$$M_{\perp}^{\mu\nu} := T^{\mu\alpha} T^{\nu\beta} M_{\alpha\beta} \quad \text{with} \quad T^{\mu\nu} = g^{\mu\nu} - \frac{P^\mu P^\nu}{P^2}. \quad (\text{B.1.23})$$

Because the components P^μ commute among themselves and also with P^2 , they also commute with the transverse projector,

$$[P^\mu, M_{\perp}^{\rho\sigma}] = [P^\mu, T^{\rho\alpha} T^{\sigma\beta} M_{\alpha\beta}] = T^{\rho\alpha} T^{\sigma\beta} [P^\mu, M_{\alpha\beta}] \stackrel{(\text{B.1.13})}{=} 0, \quad (\text{B.1.24})$$

and the commutator relations (B.1.12–B.1.14) become

$$\begin{aligned} i[M_{\perp}^{\mu\nu}, M_{\perp}^{\rho\sigma}] &= T^{\mu\sigma} M_{\perp}^{\nu\rho} + T^{\nu\rho} M_{\perp}^{\mu\sigma} - T^{\mu\rho} M_{\perp}^{\nu\sigma} - T^{\nu\sigma} M_{\perp}^{\mu\rho}, \\ [P^\mu, M_{\perp}^{\rho\sigma}] &= 0, \\ [P^\mu, P^\nu] &= 0. \end{aligned} \quad (\text{B.1.25})$$

The square of $M_{\perp}^{\mu\nu}$ is now indeed Poincaré-invariant because it commutes not only with $M_{\perp}^{\mu\nu}$ but also with P^{μ} . To establish the relation with W^2 , one can derive¹

$$\begin{aligned} M_{\perp}^{\mu\nu} &= -\frac{1}{P^2} \varepsilon^{\mu\nu\alpha\beta} P_{\alpha} W_{\beta}, \\ \widetilde{M}_{\perp}^{\mu\nu} &= \frac{1}{2} \varepsilon^{\mu\nu\alpha\beta} (M_{\perp})_{\alpha\beta} = \frac{1}{P^2} (P^{\mu} W^{\nu} - P^{\nu} W^{\mu}), \end{aligned} \quad (\text{B.1.26})$$

from where it follows that

$$W^2 = -\frac{P^2}{2} M_{\perp}^{\mu\nu} (M_{\perp})_{\mu\nu}, \quad \widetilde{M}_{\perp}^{\mu\nu} (M_{\perp})_{\mu\nu} = 0. \quad (\text{B.1.27})$$

W^2 is therefore the analogue of C_1 from the Lorentz group whereas the remaining possible Casimir vanishes identically. Along the same lines one obtains the relation

$$[W^{\mu}, W^{\nu}] = -iP^2 M_{\perp}^{\mu\nu} = i\varepsilon^{\mu\nu\alpha\beta} P_{\alpha} W_{\beta} \quad (\text{B.1.28})$$

that will become useful later. From the $1/P^2$ factors in the denominators of these expressions we also see that the massless case $P^2 = 0$ will be special, cf. Sec. B.3.

B.2 Representations of the Lorentz group

Reducible vs. irreducible representations. Let's work out the irreducible representations of the Lorentz group. The discussion is similar to that in App. A for $SU(N)$ except for some additional complications due to the richer structure of the group. A Lorentz tensor of rank n is defined by the transformation law

$$(T')^{\mu\nu\dots\tau} = \underbrace{\Lambda^{\mu}_{\alpha} \Lambda^{\nu}_{\beta} \dots \Lambda^{\tau}_{\lambda}}_{n \text{ times}} T^{\alpha\beta\dots\lambda}, \quad (\text{B.2.1})$$

so we can always construct the representation matrices $\Lambda^{\mu}_{\alpha} \Lambda^{\nu}_{\beta} \dots$ of the Lorentz transformation as the outer product $\mathbf{4} \otimes \mathbf{4} \otimes \dots$ of the 4-dimensional defining representation Λ . However, these representations are not irreducible. Take for example the 4×4 tensor $T^{\mu\nu}$, which has in principle 16 components. Its trace, its antisymmetric component, and its symmetric and traceless part,

$$S = T^{\alpha}_{\alpha}, \quad A^{\mu\nu} = \frac{1}{2} (T^{\mu\nu} - T^{\nu\mu}), \quad S^{\mu\nu} = \frac{1}{2} (T^{\mu\nu} + T^{\nu\mu}) - \frac{1}{4} g^{\mu\nu} S, \quad (\text{B.2.2})$$

do not mix under Lorentz transformations: an (anti-)symmetric tensor is still (anti-)symmetric after the transformation, and the trace S is Lorentz-invariant. The trace is one-dimensional, the antisymmetric part defines a 6-dimensional subspace, and the symmetric and traceless part a 9-dimensional subspace, so we have the decomposition $\mathbf{4} \otimes \mathbf{4} = \mathbf{1} \oplus \mathbf{6} \oplus \mathbf{9}$.

¹Use the properties that $\varepsilon_{\mu\rho\sigma\lambda} M^{\rho\sigma} P^{\lambda} = \varepsilon_{\mu\rho\sigma\lambda} M_{\perp}^{\rho\sigma} P^{\lambda}$ in the definition of W^{μ} , that P^{λ} commutes with $M_{\perp}^{\rho\sigma}$ and W^{μ} , and insert the identity $\varepsilon_{\mu\alpha\beta\lambda} \varepsilon^{\mu}_{\rho\sigma\tau} P^{\lambda} P^{\tau} = -P^2 (T_{\alpha\rho} T_{\beta\sigma} - T_{\alpha\sigma} T_{\beta\rho})$. Note that the ε -tensor switches sign when lowering or raising spatial indices; $\varepsilon_{\mu\nu\alpha\beta} = 1$ and $\varepsilon^{\mu\nu\alpha\beta} = -1$ for an even permutation of the indices (0123).

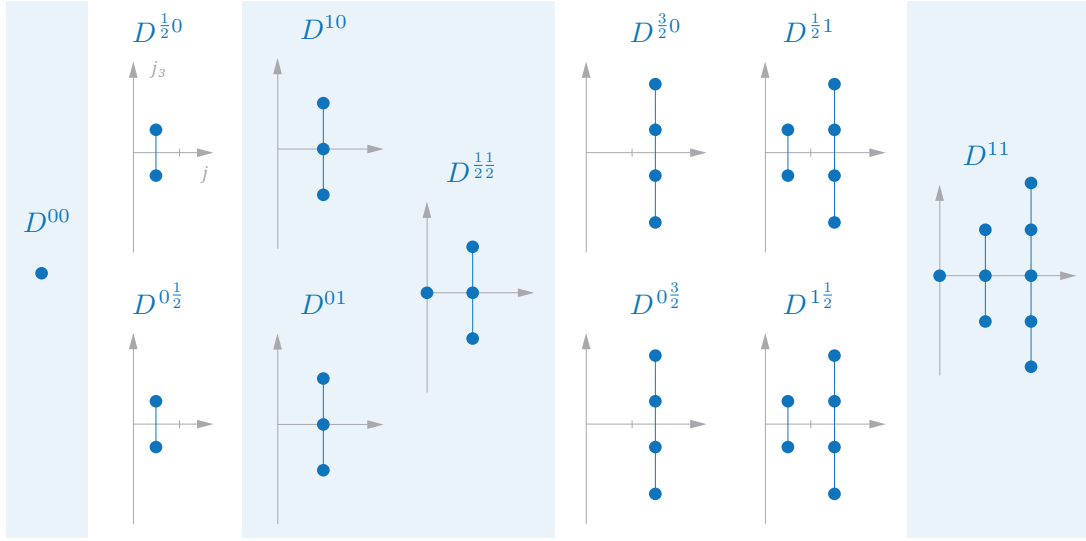


FIG. B.2: Multiplets of the Lorentz group: tensor (shaded) vs. spinor representations.

Is there a simple way to classify the **irreducible representations** of the Lorentz group? If we define

$$\mathbf{A} = \frac{1}{2}(\mathbf{J} - i\mathbf{K}), \quad \mathbf{B} = \frac{1}{2}(\mathbf{J} + i\mathbf{K}) \quad (\text{B.2.3})$$

and calculate their commutator relations using Eq. (B.1.18), we obtain two copies of an $SU(2)$ algebra with hermitian generators A_i and B_i :

$$[A_i, A_j] = i\varepsilon_{ijk} A_k, \quad [B_i, B_j] = i\varepsilon_{ijk} B_k, \quad [A_i, B_j] = 0. \quad (\text{B.2.4})$$

The two Casimir operators \mathbf{A}^2 and \mathbf{B}^2 are linear combinations of Eq. (B.1.20) with eigenvalues $a(a+1)$ and $b(b+1)$, hence there are two quantum numbers $a, b = 0, \frac{1}{2}, 1, \dots$ to label the multiplets. We will denote the irreducible representation matrices by

$$D(\Lambda) = e^{\frac{i}{2}\omega_{\mu\nu} M^{\mu\nu}} = e^{i\phi \cdot \mathbf{J} + i\mathbf{s} \cdot \mathbf{K}}, \quad M^{ij} = -\varepsilon_{ijk} J^k, \quad M^{0i} = K^i, \quad (\text{B.2.5})$$

where in an n -dimensional representation $D(\Lambda)$, $M^{\mu\nu}$, \mathbf{J} and \mathbf{K} are $n \times n$ matrices. The generators $M^{\mu\nu}$ are not hermitian because they contain the boost generators, and therefore the representation matrices are not unitary. Their dimension is

$$D^{ab} = (2a+1)(2b+1), \quad (\text{B.2.6})$$

which leads to

$$D^{00} = \mathbf{1}, \quad D^{\frac{1}{2}^0} = \mathbf{2}, \quad D^{10} = \mathbf{3}, \quad D^{\frac{1}{2}^{\frac{1}{2}}} = \mathbf{4}, \quad \dots \quad D^{11} = \mathbf{9}, \quad \dots \quad (\text{B.2.7})$$

$$D^{0^{\frac{1}{2}}} = \bar{\mathbf{2}}, \quad D^{01} = \bar{\mathbf{3}},$$

The generator of rotations is $\mathbf{J} = \mathbf{A} + \mathbf{B}$, so we can use the $SU(2)$ angular momentum addition rules to construct the states within each multiplet: the states come with all possible spins $j = |a-b| \dots a+b$, where j_3 goes from $-j$ to j , see Fig. B.2.

Tensor representations. Let's first discuss the 'tensor representations' where $a + b$ is integer (the shaded multiplets in Fig. B.2). These are the actual irreducible representations of the Lorentz group that can be constructed via Eq. (B.2.1):

- **Trivial representation** $D^{00} = \mathbf{1}$: here the generator is $M^{\mu\nu} = 0$ and the representation matrix is 1. This is how Lorentz scalars transform.
- **Antisymmetric representation**: the 6-dimensional antisymmetric part $A^{\mu\nu}$ of a 4×4 tensor belongs here. It is the adjoint representation because its dimension is the same as the number of generators. If $A^{\mu\nu}$ is real, it is also irreducible; if it is complex it can be further decomposed into a self-dual (D^{10}) and an anti-self-dual representation (D^{01}), depending on the sign of the condition $A^{\mu\nu} = \pm \frac{i}{2} \varepsilon^{\mu\nu\rho\sigma} A_{\rho\sigma}$. In Euclidean space $A^{\mu\nu}$ is always reducible and therefore the antisymmetric representation has the form $D^{10} \oplus D^{01}$.
- **Vector representation** $D^{\frac{1}{2}\frac{1}{2}} = \mathbf{4}$: The four-dimensional vector representation plays a special role because the transformation matrix is Λ itself, and it can be used to construct all further (reducible) tensor representations according to Eq. (B.2.1). The transformation matrices act on four-vectors, for example the space-time coordinate x^μ or the four-momentum p^μ , and they are irreducible because Λ mixes all components of the four-vector. The generator $M^{\mu\nu}$ has the form of Eq. (B.1.16).
- **Tensor representation** $D^{11} = \mathbf{9}$: This is where the 9-dimensional symmetric and traceless part $S^{\mu\nu}$ of a 4×4 tensor belongs.

The Lorentz group has two invariant tensors $g^{\mu\nu}$ and $\varepsilon^{\mu\nu\alpha\beta}$ which transform as

$$\begin{aligned} g'^{\mu\nu} &= \Lambda^\mu_\alpha \Lambda^\nu_\beta g^{\alpha\beta} = g^{\mu\nu}, \\ \varepsilon'^{\mu\nu\rho\sigma} &= \Lambda^\mu_\alpha \Lambda^\nu_\beta \Lambda^\rho_\gamma \Lambda^\sigma_\delta \varepsilon^{\alpha\beta\gamma\delta} = (\det \Lambda) \varepsilon^{\mu\nu\rho\sigma}. \end{aligned} \quad (\text{B.2.8})$$

$g^{\mu\nu}$ is a scalar and $\varepsilon^{\mu\nu\alpha\beta}$ is a pseudoscalar since it is odd under parity ($\det \Lambda = -1$). Their (anti-)symmetry can be exploited to construct the irreducible components of higher-rank tensors. For example, higher antisymmetric tensors in four dimensions become simple because we cannot antisymmetrize over more than four indices. $A^{\mu\nu\rho}$ has 4 components; they can be rearranged into a four-vector $\varepsilon_{\alpha\mu\nu\rho} A^{\mu\nu\rho}$ that transforms under the vector representation. $A^{\mu\nu\rho\sigma}$ has only one independent component A^{0123} that can be combined into the pseudoscalar $\varepsilon_{\mu\nu\rho\sigma} A^{\mu\nu\rho\sigma}$, and $A^{\mu\nu\rho\sigma\tau} = 0$.

Spinor representations. The analysis also produces spinor representations where $a + b$ is half-integer. These are not representations of the Lorentz group itself but rather *projective* representations, where instead of $D(\Lambda') D(\Lambda) = D(\Lambda' \Lambda)$ one has

$$D(\Lambda') D(\Lambda) = e^{i\varphi(\Lambda', \Lambda)} D(\Lambda' \Lambda), \quad (\text{B.2.9})$$

with a phase that depends on Λ and Λ' . In our case, $e^{i\varphi} = \pm 1$ and so the projective representations are double-valued: one can find two representation matrices $\pm D(\Lambda)$ that belong to the same Λ . However, both of them are physically equivalent and therefore the representations in Fig. B.2 are all relevant.

The origin of this behavior is that the Lorentz group, and in particular its subgroup $SO(3)$, is not simply connected. The projective representations of a group correspond to the representations of its universal covering group: it has the same Lie algebra, which reflects the property of the group close to the identity, but it is simply connected. In the same way as $SU(2)$ is the double cover of $SO(3)$, the double cover of $SO(3, 1)^\uparrow$ is the group $SL(2, \mathbb{C})$. It is the set of complex 2×2 matrices with unit determinant and, like the Lorentz group, it also depends on six real parameters. A double-valued projective representation of $SO(3, 1)^\uparrow$ corresponds to a single-valued representation of $SL(2, \mathbb{C})$. Similarly, the double cover of the *Euclidean* Lorentz group $SO(4)$ is $SU(2) \times SU(2)$; these are the representations that we actually derived in Fig. B.2. Hence we arrive at another type of chiral symmetry, labeled by the Casimir eigenvalues a (left-handed) and b (right-handed): representations with $a = 0$ or $b = 0$ have definite chirality, whereas those with $a = b$ are called non-chiral. Here are some of the lowest-dimensional irreducible spinor representations:

- **Fundamental representation:** $D^{\frac{1}{2}0}$ and $D^{0\frac{1}{2}}$ have both dimension two and carry spin $j = 1/2$. They are the (anti-)fundamental representations because all other representations can be built from them. The generators are $\mathbf{A} = \frac{\boldsymbol{\sigma}}{2}$ and $\mathbf{B} = 0$ for the left-handed representation and vice versa for the right-handed one, where σ_i are the Pauli matrices, and hence the spin and boost generators become

$$D^{\frac{1}{2}0} : \mathbf{J} = \frac{\boldsymbol{\sigma}}{2}, \quad \mathbf{K} = i\frac{\boldsymbol{\sigma}}{2}, \quad D^{0\frac{1}{2}} : \mathbf{J} = \frac{\boldsymbol{\sigma}}{2}, \quad \mathbf{K} = -i\frac{\boldsymbol{\sigma}}{2}. \quad (\text{B.2.10})$$

The representation matrices are complex 2×2 matrices $\in SL(2, \mathbb{C})$, and the corresponding spinors are left- and right-handed Weyl spinors ψ_L, ψ_R .

- **Dirac (bispinor) representation $D^{\frac{1}{2}0} \oplus D^{0\frac{1}{2}}$:** Under a parity transformation, the rotation generators are invariant whereas the boost generators change their sign: $\mathbf{J} \rightarrow \mathbf{J}, \mathbf{K} \rightarrow -\mathbf{K}$. Therefore, parity exchanges $\mathbf{A} \leftrightarrow \mathbf{B}$ and transforms the two fundamental representations into each other, and a theory that is invariant under parity must necessarily include both doublets. This is the reason why spin- $1/2$ fermions are treated as four-dimensional Dirac spinors ψ_α , which can be constructed as the direct sums of left- and right-handed Weyl spinors:

$$\mathbf{J} = \begin{pmatrix} \boldsymbol{\sigma}/2 & 0 \\ 0 & \boldsymbol{\sigma}/2 \end{pmatrix} = \frac{\boldsymbol{\Sigma}}{2}, \quad \mathbf{K} = \begin{pmatrix} i\boldsymbol{\sigma}/2 & 0 \\ 0 & -i\boldsymbol{\sigma}/2 \end{pmatrix}, \quad \psi = \begin{pmatrix} \psi_L \\ \psi_R \end{pmatrix}. \quad (\text{B.2.11})$$

The resulting generator $M^{\mu\nu} = -\frac{i}{4} [\gamma^\mu, \gamma^\nu]$ satisfies again the Lorentz algebra relation. The Dirac spinors transform under the four-dimensional representation matrices: $\psi' = D(\Lambda)\psi$, $\bar{\psi}' = \bar{\psi}D(\Lambda)^{-1}$. Therefore, a bilinear $\bar{\psi}\psi$ is Lorentz-invariant, $\bar{\psi}\gamma^\mu\psi$ transforms like a vector because $D(\Lambda)^{-1}\gamma^\mu D(\Lambda) = \Lambda^\mu{}_\nu\gamma^\nu$, etc.

- **Rarita-Schwinger representation:** The same point would in principle apply to spin- $\frac{3}{2}$ fermions in the (eight-dimensional) $D^{\frac{3}{2}0} \oplus D^{0\frac{3}{2}}$ representation, but it is more convenient to construct them as Rarita-Schwinger vector-spinors ψ_α^μ via

$$D^{\frac{1}{2}\frac{1}{2}} \otimes (D^{\frac{1}{2}0} \oplus D^{0\frac{1}{2}}) = (D^{\frac{1}{2}0} \oplus D^{\frac{1}{2}1}) \oplus (D^{0\frac{1}{2}} \oplus D^{1\frac{1}{2}}), \quad (\text{B.2.12})$$

which in turn requires additional constraints to single out the spin- $\frac{3}{2}$ subspace.

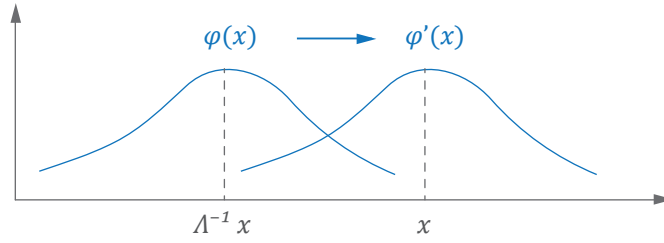


FIG. B.3: Visualization of $\varphi'(x) = \varphi(\Lambda^{-1}x)$. Compare this with quantum mechanics: if $\mathbf{x} \rightarrow R\mathbf{x}$ and $\varphi \rightarrow U\varphi$, then $\langle \mathbf{x} | U\varphi \rangle = \langle R^{-1}\mathbf{x} | \varphi \rangle$, or equivalently: $\varphi(\mathbf{x}) \rightarrow U\varphi(\mathbf{x}) = \varphi(R^{-1}\mathbf{x})$.

This last example may seem a bit contrived, but remember that from the perspective of the Poincaré group only the Casimirs P^2 and W^2 are relevant. For a massive particle the eigenvalues of W^2 in the rest frame coincide with j , but since W^2 is Poincaré-invariant, all properties associated with j hold in general. Therefore, the multiplet assignment D^{ab} in Fig. B.2 is strictly speaking meaningless because the only quantity that really matters is the spin content j : a particle with spin $j = \frac{1}{2}$ has two spin polarizations, a spin-1 particle three, and so on.

In the **nonrelativistic limit** where Lorentz transformations reduce to spatial rotations, the multiplets in Fig. B.2 are no longer irreducible but we can decompose them with respect to $SO(3)$ (or its universal cover $SU(2)$). For example, a four-vector $V^\mu = (V^0, \mathbf{V})$ defines an irreducible representation of the Lorentz group, but from the point of view of the $SO(3)$ subgroup it is reducible ($\mathbf{4} = \mathbf{1} \oplus \mathbf{3}$) because V^0 is invariant under spatial rotations (it has $j = 0$), whereas the three spatial components form an irreducible representation with $j = 1$. Similarly, the symmetric and traceless part of a 4×4 tensor is reducible: $\mathbf{9} = \mathbf{1} \oplus \mathbf{3} \oplus \mathbf{5}$.

B.3 Poincaré invariance in field theories

Field representations. So far we have only considered the Lorentz transformations of spacetime-independent quantities (scalars, vectors, spinors etc.). They transform generically as $\varphi'_i = D_{ij}(\Lambda) \varphi_j$, where i and j are the matrix indices in the given representation. When we consider fields $\varphi_i(x)$, the transformation $x' = \Lambda x$ must also act on the spacetime argument:

$$\varphi'_i(x) = D_{ij}(\Lambda) \varphi_j(\Lambda^{-1}x) \quad \Leftrightarrow \quad \varphi'_i(x') = D_{ij}(\Lambda) \varphi_j(x). \quad (\text{B.3.1})$$

The appearance of Λ^{-1} is consistent with the usual symmetry operations in quantum mechanics, cf. Fig. B.3. We can now define two types of infinitesimal transformations. The first is the same as before and expresses the ‘change in perspective’:

$$\delta\varphi_i = \varphi'_i(x') - \varphi_i(x) = \frac{i}{2} \varepsilon_{\mu\nu} (M_S^{\mu\nu})_{ij} \varphi_j(x), \quad (\text{B.3.2})$$

with the finite-dimensional matrix representation of the generator $M^{\mu\nu}$ (we added the subscript S for *spin* to distinguish it from what comes next). For example, a scalar

field $\varphi'(x') = \varphi(x)$ is Lorentz-invariant and has $\delta\varphi = 0$. On the other hand, when we want to measure how the *functional form* of the field changes at the position x (see again Fig. B.3), we have to work out

$$\delta_0\varphi_i = \varphi'_i(x) - \varphi_i(x) = \varphi'_i(x' - \delta x) - \varphi_i(x) = \delta\varphi_i - \delta x_\mu \partial^\mu \varphi_i. \quad (\text{B.3.3})$$

The infinitesimal Lorentz transformation has the form $\delta x_\mu = \varepsilon_{\mu\nu} x^\nu$, and therefore

$$-\delta x_\mu \partial^\mu \varphi_i = -\varepsilon_{\mu\nu} x^\nu \partial^\mu \varphi_i = \frac{i}{2} \varepsilon_{\mu\nu} \underbrace{[-i(x^\mu \partial^\nu - x^\nu \partial^\mu)]}_{=: M_L^{\mu\nu}} \varphi_i, \quad (\text{B.3.4})$$

where $M_L^{\mu\nu}$ contains the orbital angular momentum and satisfies again the Lorentz algebra relations. Before discussing it further, let's generalize this to Poincaré transformations right away. For pure translations each component of the field is a scalar:

$$\varphi'_i(x) = \varphi_i(x - a) \quad \Leftrightarrow \quad \varphi'_i(x') = \varphi_i(x), \quad (\text{B.3.5})$$

and hence $\delta\varphi_i = 0$ and $\delta_0\varphi_i = -a_\mu \partial^\mu \varphi_i = ia_\mu P^\mu \varphi_i$, with $P^\mu = i\partial^\mu$. The total change of the field is therefore

$$\varphi'_i(x) = \varphi_i(x) + \left[\frac{i}{2} \varepsilon_{\mu\nu} (M_S^{\mu\nu} + M_L^{\mu\nu}) + ia_\mu P^\mu \right]_{ij} \varphi_j(x). \quad (\text{B.3.6})$$

$M_L^{\mu\nu}$ and P^μ are differential operators that satisfy the Poincaré algebra relations when applied to $\varphi_i(x)$. They are diagonal in i, j whereas the spin matrix $M_S^{\mu\nu}$ depends on the representation of the field. In the same way as $M^{\mu\nu} = M_S^{\mu\nu} + M_L^{\mu\nu}$, the angular momentum and boost generators extracted from Eq. (B.1.17) are the sums of spin and orbital angular momentum parts: $\mathbf{J} = \mathbf{S} + \mathbf{L}$ and $\mathbf{K} = \mathbf{K}_S + \mathbf{K}_L$, with

$$\mathbf{L} = \mathbf{x} \times \mathbf{P}, \quad \mathbf{K}_L = \mathbf{x} P^0 - x^0 \mathbf{P}, \quad P^\mu = i\partial^\mu. \quad (\text{B.3.7})$$

Note that the boost generator is explicitly time-dependent.

Poincaré invariance of the action. The invariance of the classical action under Poincaré transformations has similar consequences as for global symmetry groups, cf. Sec. 3.1: there are conserved Noether currents, and after quantization the corresponding charges form a representation of the Poincaré algebra on the state space.

To derive the current we have to add variations of spacetime to Eq. (3.1.2):

$$\delta S = \underbrace{\int d^4x \delta_0 \mathcal{L}}_{\text{Eq. (3.1.2)}} + \int d^4x \partial_\mu \mathcal{L} \delta x^\mu + \int (\delta d^4x) \mathcal{L} = \int d^4x \left[\delta_0 \mathcal{L} + \partial_\mu (\mathcal{L} \delta x^\mu) \right]. \quad (\text{B.3.8})$$

The first term is the same as in Eq. (3.1.2) except for the replacement $\delta \rightarrow \delta_0$, because it contains only the variation in the functional form of the fields. To arrive at the last expression we used $\delta d^4x = d^4x \partial_\mu \delta x^\mu$. The new derivative term will contribute to the current, which becomes

$$-\delta j^\mu = \mathcal{L} \delta x^\mu + \sum_i \frac{\partial \mathcal{L}}{\partial (\partial_\mu \varphi_i)} \delta_0 \varphi_i. \quad (\text{B.3.9})$$

Inserting $\delta_0\varphi_i = \delta\varphi_i - \delta x_\alpha \partial^\alpha\varphi_i$ from Eq. (B.3.3), we can reexpress this in terms of $\delta\varphi_i$:

$$\delta j^\mu = \underbrace{\left[\sum_i \frac{\partial \mathcal{L}}{\partial(\partial_\mu\varphi_i)} \partial^\alpha\varphi_i - g^{\mu\alpha} \mathcal{L} \right]}_{=: T^{\mu\alpha}} \delta x_\alpha - \sum_i \frac{\partial \mathcal{L}}{\partial(\partial_\mu\varphi_i)} \delta\varphi_i. \quad (\text{B.3.10})$$

$T^{\mu\alpha}$ defines the **energy-momentum tensor** whose T^{00} component is the Hamiltonian density: $T^{00} = \pi_i \dot{\varphi}_i - \mathcal{L} = \mathcal{H}$. We can now derive two types of conserved currents that reflect the invariance under translations or Lorentz transformations:

- For pure translations $x \rightarrow x + a$ we have $\delta x_\alpha = a_\alpha$ and the fields are invariant, $\delta\varphi_i = 0$. Hence, the second term in (B.3.10) drops out and the translation current is just the energy-momentum tensor itself: $\delta j^\mu = a_\alpha T^{\mu\alpha}$. Translation invariance of the action entails that its divergence vanishes: $\partial_\mu T^{\mu\alpha} = 0$.
- For pure Lorentz transformations the group parameters are $\varepsilon_{\alpha\beta}$ and therefore

$$\delta x_\alpha = \varepsilon_{\alpha\beta} x^\beta, \quad \delta\varphi_i = \frac{i}{2} \varepsilon_{\alpha\beta} (M_S^{\alpha\beta})_{ij} \varphi_j. \quad (\text{B.3.11})$$

Inserting this into Eq. (B.3.10), writing $\delta j^\mu = \frac{1}{2} \varepsilon_{\alpha\beta} m^{\mu,\alpha\beta}$, and using the antisymmetry of $\varepsilon_{\alpha\beta}$ we find the conserved current

$$m^{\mu,\alpha\beta} = T^{\mu\alpha} x^\beta - T^{\mu\beta} x^\alpha + s^{\mu,\alpha\beta}, \quad s^{\mu,\alpha\beta} = -i \frac{\partial \mathcal{L}}{\partial(\partial_\mu\varphi_i)} (M_S^{\alpha\beta})_{ij} \varphi_j, \quad (\text{B.3.12})$$

with $\partial_\mu m^{\mu,\alpha\beta} = 0$. The first two terms encode the orbital angular momentum and the third term is the spin current.²

If we substitute the explicit form of the energy-momentum tensor into Eq. (B.3.12) together with $P^\alpha = i\partial^\alpha$ and $M^{\mu\nu} = M_S^{\mu\nu} + M_L^{\mu\nu}$, we can write the two currents as

$$\begin{aligned} T^{\mu\alpha} &= -i \frac{\partial \mathcal{L}}{\partial(\partial_\mu\varphi_i)} P^\alpha \varphi_i - g^{\mu\alpha} \mathcal{L}, \\ m^{\mu,\alpha\beta} &= -i \frac{\partial \mathcal{L}}{\partial(\partial_\mu\varphi_i)} M_{ij}^{\alpha\beta} \varphi_j + (x^\alpha g^{\mu\beta} - x^\beta g^{\mu\alpha}) \mathcal{L}. \end{aligned} \quad (\text{B.3.13})$$

The corresponding constants of motion, whose total time derivatives vanish, are the zero components of the currents $T^{\mu\alpha}$ and $m^{\mu,\alpha\beta}$ when integrated over d^3x :

$$\hat{P}^\alpha = \int d^3x T^{0\alpha}, \quad \hat{M}^{\alpha\beta} = \int d^3x m^{0,\alpha\beta}. \quad (\text{B.3.14})$$

In the quantum field theory they will form another representation of the Poincaré algebra that acts on the state space.

²An alternative form of the energy-momentum tensor is the Belinfante tensor, which is still conserved (and hence physically equivalent) but symmetric in α and β : $\Theta^{\alpha\beta} = T^{\alpha\beta} - \frac{1}{2} \partial_\mu (s^{\mu,\alpha\beta} + s^{\alpha,\beta\mu} - s^{\beta,\mu\alpha})$. To prove this, use the antisymmetry of $s^{\mu,\alpha\beta}$ in α, β and the conservation law $\partial_\mu m^{\mu,\alpha\beta} = 0$.

Dirac theory. As an example, consider a free Dirac Lagrangian $\mathcal{L} = \bar{\psi}(\not{P} - m)\psi$. The Poincaré transformation of the field is $\psi'(x') = D(\Lambda)\psi(x)$, where $D(\Lambda)$ has the form of Eq. (B.2.5) with $M_S^{\mu\nu} = -\frac{1}{2}\sigma^{\mu\nu} = -\frac{i}{4}[\gamma^\mu, \gamma^\nu]$. From Eq. (B.3.13) we have

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial(\partial_\mu \psi)} = \bar{\psi} i \gamma^\mu &\Rightarrow T^{00} = \psi^\dagger P^0 \psi - \mathcal{L}, & m^{0,ij} &= \psi^\dagger M^{ij} \psi, \\ \frac{\partial \mathcal{L}}{\partial(\partial_\mu \bar{\psi})} = 0 &\Rightarrow T^{0i} = \psi^\dagger P^i \psi, & m^{0,0i} &= \psi^\dagger M^{0i} \psi - x^i \mathcal{L}, \end{aligned} \quad (\text{B.3.15})$$

and we can read off the constants of motion ($\Sigma^i = \frac{1}{2}\varepsilon_{ijk}\sigma^{jk}$):

$$\hat{P}^0 = \int d^3x \bar{\psi}(\boldsymbol{\gamma} \cdot \mathbf{P} + m)\psi, \quad \hat{\mathbf{P}} = \int d^3x \psi^\dagger \mathbf{P} \psi, \quad \hat{\mathbf{J}} = \int d^3x \psi^\dagger \left[\mathbf{x} \times \mathbf{P} + \frac{\boldsymbol{\Sigma}}{2} \right] \psi.$$

In relativistic quantum mechanics the field $\psi(x)$ is interpreted as a particle's wave function that belongs to a Hilbert space, and a Lorentz-invariant scalar product for solutions of the Dirac equation $(\not{P} - m)\psi = 0$ is imposed:

$$\langle \psi | \psi \rangle := \int d\sigma_\mu \bar{\psi}(x) \gamma^\mu \psi(x) = \int d^3x \psi^\dagger(x) \psi(x). \quad (\text{B.3.16})$$

It has the same value on each spacelike hypersurface σ in Minkowski space, and choosing it to be a slice at fixed time yields the second form. For solutions of the classical equations of motion the terms proportional to the Dirac Lagrangian \mathcal{L} in (B.3.15) can be dropped and the conserved charges become the expectation values of the operators P^α and $M^{\alpha\beta}$:

$$\hat{P}^\alpha = \int d^3x T^{0\alpha} = \langle \psi | P^\alpha \psi \rangle, \quad \hat{M}^{\alpha\beta} = \int d^3x m^{0,\alpha\beta} = \langle \psi | M^{\alpha\beta} \psi \rangle. \quad (\text{B.3.17})$$

One can show that both operators P^α and $M^{\alpha\beta}$ are hermitian: $\langle \psi_1 | O \psi_2 \rangle = \langle O \psi_1 | \psi_2 \rangle$, and therefore the representation provided by Eq. (B.3.6) is unitary. This has become possible because, when applied to spacetime-dependent fields $\psi(x)$ that depend on a continuous and unbound variable x , the representations are now infinite-dimensional (they are differential operators). Specifically, the spin contribution to the boost generator $K_S^i = -\frac{1}{2}\sigma^{0i} = -\frac{i}{2}\gamma^0\gamma^i$ is still an antihermitian matrix, but its sum $\mathbf{K} = \mathbf{K}_S + \mathbf{K}_L$ with the differential operator $\mathbf{K}_L = \mathbf{x}P^0 - x^0\mathbf{P}$ is indeed hermitian. An analogous Lorentz-invariant scalar product for scalar fields $\phi(x)$ is

$$\langle \phi | \phi \rangle = \frac{i}{2} \int d\sigma^\mu \phi^*(x) \overleftrightarrow{\partial}_\mu \phi(x) = \frac{i}{2} \int d^3x \phi^*(x) \overleftrightarrow{\partial}_0 \phi(x), \quad \overleftrightarrow{\partial}_\mu = \overrightarrow{\partial}_\mu - \overleftarrow{\partial}_\mu. \quad (\text{B.3.18})$$

Unitary representations of the Poincaré group. Now what about the quantum field theory? A theorem by Wigner states that continuous symmetries must be implemented by unitary operators on the state space. The Lorentz group is not compact because it contains boosts, hence all unitary representations must be infinite-dimensional. This is realized in the quantum field theory: the fields $\varphi_i(x)$ become operators on the Fock space, and the constants of motion in Eq. (B.3.14) are hermitian operators that define a unitary representation of the Poincaré algebra on the state space:

$$U(\Lambda, a) = e^{\frac{i}{2}\varepsilon_{\mu\nu} \hat{M}^{\mu\nu}} e^{ia_\mu \hat{P}^\mu} = 1 + \frac{i}{2}\varepsilon_{\mu\nu} \hat{M}^{\mu\nu} + ia_\mu \hat{P}^\mu + \dots \quad (\text{B.3.19})$$

What is the irreducible state space? One of the axioms of quantum field theory is that the vacuum is the only Poincaré-invariant state: $U(\Lambda, a)|0\rangle = |0\rangle$.³ The Poincaré group has two Casimir operators P^2 and W^2 (we dropped the hats again). With $[P^\mu, W^\nu] = 0$ and Eq. (B.1.28) there are at most six operators that commute with each other and can be used to label the eigenstates: P^μ , W^2 , and one component of the Pauli-Lubanski vector W^μ . Considering one-particle states, this allows us to work with eigenstates of the momentum operator:

$$P^\mu |p, \dots\rangle = p^\mu |p, \dots\rangle \quad \Rightarrow \quad U(1, a) |p, \dots\rangle = e^{ia \cdot p} |p, \dots\rangle, \quad (\text{B.3.20})$$

where the dots are the remaining quantum numbers.

To construct the general form of the representation, let's start with a massive particle at rest. We denote the rest-frame momentum by $\hat{p} = (m, \mathbf{0})$. The group that leaves a given choice of momentum p^μ invariant is called the **little group**; its generators are the independent components of the Pauli-Lubanski vector. Since rotations leave the rest-frame momentum \hat{p}^μ invariant, the independent components are the generators J^i , cf. Eq. (B.1.22), and the little group is $SO(3)$ — or actually $SU(2)$ because we want to include spinor representations as well. Hence these operators take the form $P^2 = m^2$, $W^2 = -m^2 \mathbf{J}^2$ and $W^3 = mJ^3$, where J^3 has eigenvalue σ and the eigenvectors are

$$P^\mu |\hat{p}, j\sigma\rangle = \hat{p}^\mu |\hat{p}, j\sigma\rangle, \quad \mathbf{J}^2 |\hat{p}, j\sigma\rangle = j(j+1) |\hat{p}, j\sigma\rangle, \quad J^3 |\hat{p}, j\sigma\rangle = \sigma |\hat{p}, j\sigma\rangle. \quad (\text{B.3.21})$$

This is the standard angular momentum algebra, and therefore rotations R are represented by the unitary matrices $\mathcal{D}^j(R)$ with $\sigma \in [-j, j]$:

$$U(R, 0) |\hat{p}, j\sigma\rangle = \sum_{\sigma'} \mathcal{D}_{\sigma'\sigma}^j(R) |\hat{p}, j\sigma\rangle. \quad (\text{B.3.22})$$

On the other hand, a boost from \hat{p} to p , which we denote by $\mathbf{L}(p)$, will have the effect

$$U(\mathbf{L}(p), 0) |\hat{p}, j\sigma\rangle = |p, j\sigma\rangle. \quad (\text{B.3.23})$$

With that we have everything in place to apply a general Lorentz transformation $U(\Lambda, 0)$ to a state vector $|p, j\sigma\rangle$:

$$\begin{aligned} U(\Lambda, 0) |p, j\sigma\rangle &= U(\Lambda, 0) U(\mathbf{L}(p), 0) |\hat{p}, j\sigma\rangle \\ &= U(\mathbf{L}(\Lambda p) \underbrace{\mathbf{L}^{-1}(\Lambda p) \Lambda \mathbf{L}(p)}_{=: R_W}, 0) |\hat{p}, j\sigma\rangle. \end{aligned} \quad (\text{B.3.24})$$

The Wigner rotation $R_W(\Lambda, p)$ is a pure rotation that leaves the rest-frame vector invariant, because $\mathbf{L}(p)\hat{p} = p$ entails $R_W\hat{p} = \mathbf{L}^{-1}(\Lambda p)\Lambda p = \hat{p}$. Think of it as a journey along the mass shell that leads back to the starting point: $\hat{p} \rightarrow p \rightarrow \Lambda p \rightarrow \hat{p}$. This is extremely helpful because from Eq. (B.3.22) we know how rotations act on the state space, and in combination with Eqs. (B.3.23) and (B.3.20) we arrive at the final result:

$$U(\Lambda, a) |p, j\sigma\rangle = e^{ia \cdot (\Lambda p)} \sum_{\sigma'} \mathcal{D}_{\sigma'\sigma}^{(j)}(R_W) |\Lambda p, j\sigma'\rangle. \quad (\text{B.3.25})$$

³Actually, translation invariance and uniqueness of the vacuum is sufficient to prove this.

That the representation is unitary can be seen from the scalar product:

$$\langle p, j\sigma | U^\dagger(\Lambda, a) U(\Lambda, a) | p', j'\sigma' \rangle = \langle \Lambda p, j\sigma | \Lambda p', j'\sigma' \rangle = \langle p, j\sigma | p', j'\sigma' \rangle. \quad (\text{B.3.26})$$

In the first equality the representation matrices \mathcal{D}^j and the phases $e^{ia \cdot (\Lambda p)}$ cancel each other, and the second equality holds because $\langle \lambda | \lambda' \rangle = (2\pi)^3 2E_{\mathbf{p}} \delta(\mathbf{p} - \mathbf{p}') \delta_{\lambda\lambda'}$ is Lorentz-invariant. Hence, we have a unitary implementation of the Poincaré group in the quantum field theory, as required by Wigner's theorem.

Massless particles. Massless particles with $P^2 = 0$ do not have a rest frame, but the construction of the irreducible representations is very similar. Here we can choose $\hat{p} = \omega(1, \mathbf{n})$ to be some momentum on the light cone, and the little group $SO(2)$ (or equivalently $U(1)$) consists of the rotations around the momentum axis \mathbf{n} . The generator is the **helicity** $\mathbf{J} \cdot \mathbf{n}$, whose eigenvalue λ can be shown to be quantized: $\lambda = 0, \pm\frac{1}{2}, \pm 1$, etc. Hence, massless particles have no spin but only two components of the helicity that are measurable.⁴ The steps are the same as before, with the Wigner rotation R_W defined as in Eq. (B.3.24) except that $\mathcal{D}(R_W) = e^{i\lambda\theta(\Lambda, p)}$ is just a phase:

$$U(\Lambda, a) | p, \lambda \rangle = e^{ia \cdot (\Lambda p)} \mathcal{D}(R_W) | \Lambda p, \lambda \rangle. \quad (\text{B.3.27})$$

In principle this also implies that the helicity is Poincaré-invariant and $\pm\lambda$ corresponds to different species of particles. However, the same reasoning that required us earlier to implement spinors with both chiralities also applies here: $\mathbf{J} \cdot \mathbf{n}$ is a pseudoscalar and changes sign under parity, and a theory that conserves parity must treat both helicity states symmetrically. A combined representation of the Poincaré group and parity identifies $\pm\lambda$ with the two polarizations of the same particle (e.g. the photon in QED).

Transformation of field operators and n -point functions. Field operators transform in the same way as in Eq. (B.3.1) if we insert $\varphi'_i = U(\Lambda, a)^{-1} \varphi_i U(\Lambda, a)$. Shuffling things around between the left and right, it is more convenient to write

$$U(\Lambda, a) \varphi_i(x) U(\Lambda, a)^{-1} = D(\Lambda)_{ij}^{-1} \varphi_j(\Lambda x + a). \quad (\text{B.3.28})$$

As before, the field operator $\varphi_i(x)$ belongs to some finite-dimensional multiplet of the Lorentz group and $D(\Lambda)$ is the corresponding spin matrix of the Lorentz transformation. For example, we have $D(\Lambda) = 1$ for a scalar field, $D(\Lambda) = \Lambda$ for a vector field or $D(\Lambda) = \exp(-\frac{i}{4} \varepsilon_{\mu\nu} \sigma^{\mu\nu})$ for a Dirac spinor field.

Matrix elements are Lorentz-covariant and transform under these matrix representations. Take for example a scalar Bethe-Salpeter wave function of two scalar fields, $\chi(x_1, x_2, p) = \langle 0 | \mathbb{T} \varphi(x_1) \varphi(x_2) | p \rangle$. In that case Eqs. (B.3.25) and (B.3.28) simplify to

$$U_X = U(0, X) : \quad U_X | p \rangle = e^{ip \cdot X} | p \rangle, \quad U_X \varphi(x) U_X^{-1} = \varphi(x + X), \quad (\text{B.3.29})$$

$$U_\Lambda = U(\Lambda, 0) : \quad U_\Lambda | p \rangle = | \Lambda p \rangle, \quad U_\Lambda \varphi(x) U_\Lambda^{-1} = \varphi(\Lambda x). \quad (\text{B.3.30})$$

⁴In fact, the Pauli-Lubanski operator W^μ has *three* independent components in the massless case: the helicity $\mathbf{J} \cdot \mathbf{n}$ and two components perpendicular to \mathbf{n} . One can show, however, that the transverse components lead to representations with continuous spin $W^2 > 0$, which are not observed in nature and must be excluded. Evaluated on the helicity states, the spin is zero: $W^2 = 0$.

Translation invariance has the consequence that only the relative coordinate $x := x_1 - x_2$ is relevant because the dependence on the total position $X := \frac{x_1 + x_2}{2}$ can only enter through a phase:

$$\begin{aligned}\chi(x_1, x_2, p) &= \langle 0 | \mathbb{T} \varphi(x_1) \varphi(x_2) | p \rangle = \langle 0 | \mathbb{T} \varphi(X + \frac{x}{2}) \varphi(X - \frac{x}{2}) | p \rangle \\ &= \langle 0 | \mathbb{T} U_X \varphi(\frac{x}{2}) U_X^{-1} U_X \varphi(-\frac{x}{2}) U_X^{-1} | p \rangle \\ &= \langle 0 | \mathbb{T} \varphi(\frac{x}{2}) \varphi(-\frac{x}{2}) | p \rangle e^{-ip \cdot X} = \chi(x, p) e^{-ip \cdot X},\end{aligned}\tag{B.3.31}$$

where we used translation invariance of the vacuum. In turn, the wave function $\chi(x, p)$ is Lorentz-invariant:

$$\begin{aligned}\chi(x, p) &= \langle 0 | \mathbb{T} \varphi(\frac{x}{2}) \varphi(-\frac{x}{2}) | p \rangle \\ &= \langle 0 | \mathbb{T} U_\Lambda^{-1} U_\Lambda \varphi(\frac{x}{2}) U_\Lambda^{-1} U_\Lambda \varphi(-\frac{x}{2}) U_\Lambda^{-1} U_\Lambda | p \rangle \\ &= \langle 0 | \mathbb{T} \varphi(\frac{\Lambda x}{2}) \varphi(-\frac{\Lambda x}{2}) | \Lambda p \rangle = \chi(\Lambda x, \Lambda p).\end{aligned}\tag{B.3.32}$$

The time ordering commutes with the transformation because the sign of $(x_1 - x_2)^0$ is invariant under $ISO(3, 1)^\uparrow$. If we set $p = 0$ in the first equation we also see that translation invariance for the two-point function $\langle 0 | \mathbb{T} \varphi(x_1) \varphi(x_2) | 0 \rangle$ (and generally for any n -point function) means that the total coordinate drops out completely.

We can repeat the steps in Eq. (B.3.32) for matrix elements that contain fields in some general Lorentz representation. For example, for a $q\bar{q}$ vector Green function $G^\mu(x, x_1, x_2) = \langle 0 | \mathbb{T} j^\mu(x) \psi(x_1) \bar{\psi}(x_2) | 0 \rangle$ we obtain

$$G^\mu(x, x_1, x_2) = (\Lambda^{-1})^\mu{}_\nu D^{-1}(\Lambda) G^\nu(\Lambda x, \Lambda x_1, \Lambda x_2) D(\Lambda),\tag{B.3.33}$$

where $D(\Lambda)$ is again the transformation matrix for Dirac spinors coming from the quark fields. The analogous equation in momentum space,

$$G^\mu(p, q) = (\Lambda^{-1})^\mu{}_\nu D^{-1}(\Lambda) G^\nu(\Lambda p, \Lambda q) D(\Lambda),\tag{B.3.34}$$

can be immediately verified for the various tensor structures that contribute to the three-point function: γ^μ , p^μ , $p^\mu \not{p}$, $\gamma^\mu \not{p}$, etc. In covariant equations where these objects are combined in loop integrals (perturbation series, Dyson-Schwinger equations, etc.), all internal representation matrices cancel each other and only the overall factors of the diagrams remain, which can be factored out. It is then not necessary to perform explicit Lorentz transformations when changing the frame; one can simply evaluate the equation in a different frame and the result must be the same.

Literature:

- S. Weinberg, *The Quantum Theory of Fields. Vol. 1: Foundations*. Cambridge University Press, 1995.
- M. Maggiore, *A Modern Introduction to Quantum Field Theory*. Oxford University Press, 2005.
- W.-K. Tung, *Group Theory in Physics*. World Scientific, 1985.
- F. Scheck, *Quantum Physics*. Springer, 2007.
- B. Thaller, *The Dirac Equation*. Springer, 1992.

Appendix C

Euclidean conventions

When employing a Minkowski metric (which is what we use throughout the main text), one must be mindful of the $i\epsilon$ prescription that is necessary to make many relations in QFT well-defined. It arises from the imaginary-time boundary conditions (2.2.21), which lead to boundary conditions on d^4x and d^4p integrals. An alternative is to define $x^4 = ix_0$ and $p^4 = ip_0$ and perform a **Wick rotation** to write

$$\begin{aligned}\int d^4x &= \int d^3x \int_{-\infty(1-i\epsilon)}^{\infty(1-i\epsilon)} dx_0 = -i \int d^3x \int_{-\infty}^{\infty} dx_4, \\ \int d^4p &= \int d^3p \int_{-\infty(1+i\epsilon)}^{\infty(1+i\epsilon)} dp_0 = i \int d^3p \int_{-\infty}^{\infty} dp_4.\end{aligned}\tag{C.0.1}$$

Note that the integration paths in x_0 and p_0 rotate in opposite directions and thus

$$i \int d^4x = \int d^4x_E \quad \text{but} \quad \int d^4p = i \int d^4p_E.\tag{C.0.2}$$

Since $x^2 = x_0^2 - \mathbf{x}^2 = -x_4^2 - \mathbf{x}^2 = -x_E^2$, this amounts to using a **Euclidean metric** with signature $(+, +, +, +)$.

Euclidean conventions. In general, we define Euclidean vectors a_E^μ and tensors $T_E^{\mu\nu}$ such that their spatial parts agree with Minkowski space:

$$a_E^\mu = \begin{bmatrix} \mathbf{a} \\ ia_0 \end{bmatrix}, \quad T_E^{\mu\nu} = \begin{bmatrix} T^{ij} & iT^{i0} \\ iT^{0i} & -T^{00} \end{bmatrix},\tag{C.0.3}$$

where ‘ E ’ stands for Euclidean and no subscript refers to the Minkowski quantity. As a consequence, the Lorentz-invariant scalar product of any two four-vectors differs by a minus sign from its Minkowski counterpart:

$$a_E \cdot b_E = \sum_{k=1}^4 a_E^k b_E^k = -a \cdot b.\tag{C.0.4}$$

Therefore, a vector is spacelike if $a^2 > 0$ and timelike if $a^2 < 0$. Because the Euclidean metric is positive, we can drop the distinction between upper and lower indices.

To preserve the meaning of the slash $\not{a} = a^0 \gamma^0 - \mathbf{a} \cdot \boldsymbol{\gamma}$, we must also redefine the γ -matrices:

$$i\gamma_E^\mu = \begin{bmatrix} \boldsymbol{\gamma} \\ i\gamma_0 \end{bmatrix}, \quad \gamma_E^5 = \gamma^5 \quad \Rightarrow \quad \not{a}_E = a_E \cdot \boldsymbol{\gamma}_E = i\not{a}, \quad \{\gamma_E^\mu, \gamma_E^\nu\} = 2\delta^{\mu\nu}. \quad (\text{C.0.5})$$

Our sign convention for the Euclidean γ -matrices changes all signs in the Clifford algebra relation to be positive, and since this implies $(\gamma_E^i)^2 = 1$ for $i = 1 \dots 4$ we can choose them to be hermitian: $\gamma_E^\mu = (\gamma_E^\mu)^\dagger$. In the standard representation they read

$$\gamma_E^k = \begin{bmatrix} 0 & -i\tau_k \\ i\tau_k & 0 \end{bmatrix}, \quad \gamma_E^4 = \begin{bmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{bmatrix}, \quad \gamma_E^5 = \begin{bmatrix} 0 & \mathbb{1} \\ \mathbb{1} & 0 \end{bmatrix},$$

where the τ_k are the usual Pauli matrices from Eq. (A.1.4). Also the generators of the Clifford algebra are then hermitian, with $(\sigma_E^{\mu\nu})^\dagger = \sigma_E^{\mu\nu}$:

$$\sigma^{\mu\nu} = \frac{i}{2} [\gamma^\mu, \gamma^\nu] \quad \Rightarrow \quad \sigma_E^{\mu\nu} = -\frac{i}{2} [\gamma_E^\mu, \gamma_E^\nu]. \quad (\text{C.0.6})$$

Despite appearances, this does not alter the Lorentz transformation properties and the definition of the conjugate spinor as $\bar{\psi} = \psi^\dagger \gamma^4$ (which was necessary to make a bilinear $\bar{\psi}\psi$ Lorentz-invariant) remains intact. Denoting the representation matrix $\psi'(x') = D(\Lambda) \psi(x)$ of the Lorentz transformation by

$$D(\Lambda) = \exp \left[-\frac{i}{4} \omega_{\mu\nu} \sigma^{\mu\nu} \right] = \exp \left[-\frac{i}{4} \omega_E^{\mu\nu} \sigma_E^{\mu\nu} \right], \quad (\text{C.0.7})$$

then irrespective of $\gamma^4 (\sigma_E^{\mu\nu})^\dagger \gamma^4 \neq \sigma_E^{\mu\nu}$ the relation $\gamma^4 D(\Lambda)^\dagger \gamma^4 = D(\Lambda)^{-1}$ still holds, because the infinitesimal Lorentz transformation $\omega_E^{\mu\nu}$ which is related to its Minkowski counterpart via (C.0.3) is now complex. Hence

$$\bar{\psi}'(x') = \psi^\dagger(x) D(\Lambda)^\dagger \gamma^4 = \psi^\dagger(x) \gamma^4 D(\Lambda)^{-1} = \bar{\psi} D(\Lambda)^{-1}, \quad (\text{C.0.8})$$

and therefore $\bar{\psi}\psi$ is Lorentz-invariant, $\bar{\psi}\gamma_E^\mu\psi$ transforms like a Lorentz vector, etc.

For derivatives, Eq. (C.0.3) implies

$$\begin{aligned} \partial \cdot \mathbf{a} &= \partial_0 a^0 + \nabla \cdot \mathbf{a} = (\partial \cdot \mathbf{a})_E, \\ \partial_\mu^E &= \begin{bmatrix} \nabla \\ -i\partial_0 \end{bmatrix} \quad \Rightarrow \quad \not{\partial} = \gamma^0 \partial_0 + \boldsymbol{\gamma} \cdot \nabla = i\not{\partial}_E, \\ \square &= \partial_0^2 - \nabla^2 = -\square_E. \end{aligned} \quad (\text{C.0.9})$$

As a result, a fermionic action becomes

$$e^{iS} = \exp \left[i \int d^4x \bar{\psi} (i\not{\partial} - m) \psi \right] = \exp \left[- \int d^4x_E \bar{\psi} (\not{\partial}_E + m) \psi \right] = e^{-S_E}. \quad (\text{C.0.10})$$

In this way, the **Euclidean action** S_E is non-negative and the term e^{-S_E} defines a probability measure in the path integral formulation.

Another advantage of the Euclidean metric is that one can perform numerical calculations directly in a given frame (e.g. using *Mathematica*), with explicit γ -matrices and without the need for inserting the metric tensor in each summation. To transform an expression from Minkowski to Euclidean space, it is usually sufficient to employ the replacement rules collected in Table C.1 which can be read off from the spatial components.

Minkowski	Euclidean	Minkowski	Euclidean
$a \cdot b$	$-a \cdot b$	$[\gamma^\mu, \gamma^\nu]$	$-[\gamma^\mu, \gamma^\nu]$
a^μ	a^μ	$[\gamma^\mu, \not{a}]$	$[\gamma^\mu, \not{a}]$
γ^μ	$i\gamma^\mu$	$[\gamma^\mu, \gamma^\nu, \not{a}]$	$i[\gamma^\mu, \gamma^\nu, \not{a}]$
γ_5	γ_5	$[\not{a}, \not{b}]$	$-[\not{a}, \not{b}]$
\not{a}	$-i\not{a}$	$\varepsilon^{\mu\nu\rho\alpha} a_\alpha$	$i\varepsilon^{\mu\nu\rho\alpha} a^\alpha$
$g^{\mu\nu}$	$-\delta^{\mu\nu}$	$\varepsilon^{\mu\nu\alpha\beta} a_\alpha b_\beta$	$i\varepsilon^{\mu\nu\alpha\beta} a^\alpha b^\beta$
$a^\mu b^\nu$	$a^\mu b^\nu$	$\varepsilon^{\mu\alpha\beta\gamma} a_\alpha b_\beta c_\gamma$	$i\varepsilon^{\mu\alpha\beta\gamma} a^\alpha b^\beta c^\gamma$
∂_μ	∂^μ	$\varepsilon^{\mu\nu\alpha\beta} a_\alpha \gamma_\beta$	$-\varepsilon^{\mu\nu\alpha\beta} a^\alpha \gamma^\beta$

TABLE C.1: Replacement rules for some frequently occurring quantities. For expressions with Lorentz indices, the right columns define their Euclidean version in the sense of Eqs. (C.0.3) and (C.0.5). Each additional Minkowski summation over Lorentz indices leads to a minus sign in Euclidean conventions.

Expressions involving $\varepsilon^{\mu\nu\alpha\beta}$ work along the same lines: the spatial parts of Lorentz tensors are identical in Minkowski and Euclidean conventions, so this must also hold for $\varepsilon^{\mu\nu\alpha\beta} a_\alpha b_\beta$. In Euclidean space the ε -tensor is defined by $\varepsilon_{1234} = \varepsilon^{1234} = 1$, whereas in Minkowski conventions one has $\varepsilon_{0123} = -\varepsilon^{0123} = 1$, i.e., the spatial components switch sign when lowering or raising indices. Denoting spatial indices by i, j, k and summing over k , one has

$$\begin{aligned} \varepsilon^{ij\alpha\beta} a_\alpha b_\beta &= \varepsilon^{ijk0} (a_k b_0 - a_0 b_k) = -\varepsilon^{ijk0} (a^k b^0 - a^0 b^k) \\ &= i\varepsilon^{ijk4} (a^k b^4 - a^4 b^k)_E = (i\varepsilon^{ij\alpha\beta} a^\alpha b^\beta)_E, \end{aligned} \quad (\text{C.0.11})$$

because $\varepsilon^{1234} = 1 = \varepsilon^{1230}$ and $a^0 = -ia_E^4$. Repeating this for rank-1 and rank-3 tensors results in the identities in Table C.1 (which would also follow from Eq. (C.0.29) below).

Euclidean Feynman rules. We now drop the index ‘ E ’ and write all subsequent formulas in Euclidean space. The Euclidean action of QCD is

$$S = \int d^4x [\bar{\psi} (\not{D} + M) \psi + \frac{1}{4} F_{\mu\nu}^a F_a^{\mu\nu}], \quad (\text{C.0.12})$$

where $D_\mu = \partial_\mu + igA_\mu$ (this is consistent with Eq. (2.1.3) because the spatial component is $\mathbf{D} = \nabla + ig\mathbf{A}$). As a consequence,

$$F_{\mu\nu}(x) = -\partial_\mu A_\nu + \partial_\nu A_\mu - ig[A_\mu, A_\nu]. \quad (\text{C.0.13})$$

The Fourier transform is *defined* in Euclidean space,

$$F(x) = \int \frac{d^4p}{(2\pi)^4} e^{ip \cdot x} F(p), \quad F(p) = \int d^4x e^{-ip \cdot x} F(x), \quad (\text{C.0.14})$$

which would technically lead to i factors in front of the integrals in Minkowski space, but this can be compensated by removing factors of i from the propagators and vertices in momentum space.

The resulting **Feynman rules** are obtained by taking the Feynman rules in Minkowski space, transforming to Euclidean space, and splitting off a factor i from each 1PI quantity (vertices and inverse propagators). This yields:

- Quark propagator:

$$S_0^{-1}(p) = Z_\psi (i\not{p} + m_B), \quad S_0(p) = \frac{1}{Z_\psi} \frac{-i\not{p} + m_B}{p^2 + m_B^2} \quad (\text{C.0.15})$$

- Gluon propagator (we redefine $T_q^{\mu\nu} = \delta^{\mu\nu} - q^\mu q^\nu / q^2$ and $L_q^{\mu\nu} = q^\mu q^\nu / q^2$):

$$(D_0^{-1})^{\mu\nu}(q) = q^2 \left(Z_A T_q^{\mu\nu} + \frac{1}{\xi} L_q^{\mu\nu} \right), \quad D_0^{\mu\nu}(q) = \frac{Z(q^2) T_q^{\mu\nu} + \xi L_q^{\mu\nu}}{q^2} \quad (\text{C.0.16})$$

- Ghost propagator:

$$D_{G,0}^{-1}(q) = -Z_c q^2, \quad D_{G,0}(q) = -\frac{1}{Z_c q^2} \quad (\text{C.0.17})$$

- Quark-gluon vertex:

$$\Gamma_0^\mu = ig t_a Z_\Gamma \gamma^\mu \quad (\text{C.0.18})$$

- Ghost-gluon vertex:

$$\Gamma_{\text{gh},0}^\mu(p) = -ig f_{abc} \tilde{Z}_\Gamma p^\mu \quad (\text{C.0.19})$$

- Three-gluon vertex:

$$\Gamma_{3g,0}^{\mu\nu\rho}(p_1, p_2, p_3) = ig f_{abc} Z_{3g} [(p_1 - p_2)^\rho \delta^{\mu\nu} + (p_2 - p_3)^\mu \delta^{\nu\rho} + (p_3 - p_1)^\nu \delta^{\rho\mu}]. \quad (\text{C.0.20})$$

- Four-gluon vertex:

$$\Gamma_{4g,0}^{\mu\nu\rho\sigma} = -g^2 Z_{4g} \left[f_{abe} f_{cde} (\delta^{\mu\rho} \delta^{\nu\sigma} - \delta^{\nu\rho} \delta^{\mu\sigma}) + f_{ace} f_{bde} (\delta^{\mu\nu} \delta^{\rho\sigma} - \delta^{\nu\rho} \delta^{\mu\sigma}) + f_{ade} f_{cbe} (\delta^{\mu\rho} \delta^{\nu\sigma} - \delta^{\mu\nu} \delta^{\rho\sigma}) \right]. \quad (\text{C.0.21})$$

The Lorentz-invariant dressing functions are identical except that the arguments pick up minus signs. This is often indicated by capital letters such as $Q^2 = q_E^2 = -q_M^2$. In general, if one defines

$$\{s_1, s_2, s_3, \dots\} = \{p_E^2, q_E^2, p_E \cdot q_E, \dots\} = \{-p_M^2, -q_M^2, -p_M \cdot q_M, \dots\}, \quad (\text{C.0.22})$$

then the quantities $F(s_1, s_2, s_3, \dots)$ are the same in Euclidean and Minkowski space. Also for this reason it is convenient to break down Lorentz-*covariant* relations to Lorentz-*invariant* relations, because then the transformations from Minkowski to Euclidean space and vice versa become trivial (assuming that the correct integration paths for loop integrals are chosen such as in Fig. 2.7).

Euclidean formulas. We suppress again the index ‘ E ’ and collect some useful Euclidean formulas. The γ_5 matrix is defined by

$$\gamma^5 = -\gamma^1 \gamma^2 \gamma^3 \gamma^4 = -\frac{1}{24} \varepsilon^{\mu\nu\rho\sigma} \gamma^\mu \gamma^\nu \gamma^\rho \gamma^\sigma \quad (\text{C.0.23})$$

with $\varepsilon^{1234} = 1$. It is convenient to define the fully antisymmetric combinations of Dirac matrices by the commutators

$$[A, B] = AB - BA, \quad (\text{C.0.24})$$

$$[A, B, C] = [A, B]C + [B, C]A + [C, A]B, \quad (\text{C.0.25})$$

$$[A, B, C, D] = [A, B, C]D + [B, C, D]A + [C, D, A]B + [D, A, B]C. \quad (\text{C.0.26})$$

Inserting γ -matrices, this yields the antisymmetric combinations

$$[\gamma^\mu, \gamma^\nu] = \gamma_5 \varepsilon^{\mu\nu\alpha\beta} \gamma^\alpha \gamma^\beta, \quad (\text{C.0.27})$$

$$\frac{1}{6} [\gamma^\mu, \gamma^\nu, \gamma^\rho] = \frac{1}{2} (\gamma^\mu \gamma^\nu \gamma^\rho - \gamma^\rho \gamma^\nu \gamma^\mu) = \frac{1}{4} \{[\gamma^\mu, \gamma^\nu], \gamma^\rho\} = -\gamma_5 \varepsilon^{\mu\nu\rho\sigma} \gamma^\sigma, \quad (\text{C.0.28})$$

$$\frac{1}{24} [\gamma^\mu, \gamma^\nu, \gamma^\alpha, \gamma^\beta] = -\gamma_5 \varepsilon^{\mu\nu\alpha\beta}. \quad (\text{C.0.29})$$

The various contractions of ε -tensors are given by

$$\begin{aligned} \varepsilon^{\mu\nu\rho\lambda} \varepsilon^{\alpha\beta\gamma\lambda} &= \delta^{\mu\alpha} (\delta^{\nu\beta} \delta^{\rho\gamma} - \delta^{\nu\gamma} \delta^{\rho\beta}) + \delta^{\mu\beta} (\delta^{\nu\gamma} \delta^{\rho\alpha} - \delta^{\nu\alpha} \delta^{\rho\gamma}) \\ &\quad + \delta^{\mu\gamma} (\delta^{\rho\beta} \delta^{\nu\alpha} - \delta^{\rho\alpha} \delta^{\nu\beta}), \\ \frac{1}{2} \varepsilon^{\mu\nu\lambda\sigma} \varepsilon^{\alpha\beta\lambda\sigma} &= \delta^{\mu\alpha} \delta^{\nu\beta} - \delta^{\mu\beta} \delta^{\nu\alpha}, \\ \frac{1}{6} \varepsilon^{\mu\lambda\sigma\tau} \varepsilon^{\alpha\lambda\sigma\tau} &= \delta^{\mu\alpha}, \\ \frac{1}{24} \varepsilon^{\lambda\sigma\tau\omega} \varepsilon^{\lambda\sigma\tau\omega} &= 1. \end{aligned} \quad (\text{C.0.30})$$

The ε -tensor satisfies $a^{\{\mu} \varepsilon^{\alpha\beta\gamma\delta\}} = 0$, where a^μ is an arbitrary four-vector and $\{\dots\}$ denotes a symmetrization of indices.

Momentum integrations. Four-momenta are conveniently expressed through hyperspherical coordinates:

$$p^\mu = \sqrt{p^2} \begin{bmatrix} \sqrt{1-z^2} \sqrt{1-y^2} \sin \phi \\ \sqrt{1-z^2} \sqrt{1-y^2} \cos \phi \\ \sqrt{1-z^2} y \\ z \end{bmatrix}. \quad (\text{C.0.31})$$

For a particle in its rest frame, this corresponds to $p^\mu = (\mathbf{0}, im)$. (Actually, in Euclidean space it does not matter *where* we put the mass since each direction is treated equally.) A four-momentum integration reads

$$\int \frac{d^4 p}{(2\pi)^4} = \frac{1}{(2\pi)^4} \frac{1}{2} \int_0^\infty dp^2 p^2 \int_{-1}^1 dz \sqrt{1-z^2} \int_{-1}^1 dy \int_0^{2\pi} d\phi, \quad (\text{C.0.32})$$

where $\frac{1}{2} dp^2 p^2 = dp p^3$ and

$$\int d\Omega_4 = \int_{-1}^1 dz \sqrt{1-z^2} \int_{-1}^1 dy \int_0^{2\pi} d\phi = 2\pi^2 \quad (\text{C.0.33})$$

is the integral over the unit sphere in four dimensions.

Spinors. The positive- and negative-energy onshell spinors for spin-1/2 particles satisfy the Dirac equations

$$\begin{aligned}(i\not{p} + m) u(\mathbf{p}) &= 0 = \bar{u}(\mathbf{p}) (i\not{p} + m), \\(i\not{p} - m) v(\mathbf{p}) &= 0 = \bar{v}(\mathbf{p}) (i\not{p} - m),\end{aligned}\tag{C.0.34}$$

where the conjugate spinor is $\bar{u}(\mathbf{p}) = u(\mathbf{p})^\dagger \gamma^4$. Since the onshell spinors only depend on \mathbf{p} they are the same as in Minkowski space; for example in the standard representation:

$$u_s(\mathbf{p}) = \sqrt{\frac{E_p + m}{2m}} \begin{pmatrix} \xi_s \\ \frac{\mathbf{p} \cdot \boldsymbol{\tau}}{E_p + m} \xi_s \end{pmatrix}\tag{C.0.35}$$

with

$$\xi_+ = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \xi_- = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad E_p = \sqrt{\mathbf{p}^2 + m^2}.$$

We have normalized them to unity,

$$\begin{aligned}\bar{u}_s(\mathbf{p}) u_{s'}(\mathbf{p}) &= -\bar{v}_s(\mathbf{p}) v_{s'}(\mathbf{p}) = \delta_{ss'}, \\ \bar{u}_s(\mathbf{p}) v_{s'}(\mathbf{p}) &= \bar{v}_s(\mathbf{p}) u_{s'}(\mathbf{p}) = 0,\end{aligned}\tag{C.0.36}$$

and their completeness relations define the positive- and negative-energy projectors:

$$\begin{aligned}\sum_s u_s(\mathbf{p}) \bar{u}_s(\mathbf{p}) &= \frac{-i\not{p} + m}{2m} = \Lambda_+(p), \\ \sum_s v_s(\mathbf{p}) \bar{v}_s(\mathbf{p}) &= \frac{-i\not{p} - m}{2m} = -\Lambda_-(p).\end{aligned}\tag{C.0.37}$$

Therefore, $\Lambda_+(p) u(\mathbf{p}) = u(\mathbf{p})$ and $\Lambda_-(p) u(\mathbf{p}) = 0$.