Particle Physics 1+2
Theory Chapters

Tilman Plehn

Institut für Theoretische Physik

Abstract

This is a set of lecture notes for the theory modules in master-level particle physics 1 and 2, where the second class used to be taught as Standard Model. It assumes a solid theory background from the theory bachelor courses, for instance analytical mechanics, ellectromagnetism, and quantum mechanics. Aspects of quantum field theory are sketched and then skipped, we refer to the specialized lectures for more details. The goal of the lecture is to understand the theory background of modern particle physics, from low to high energies, and be able to do simple calculations.

Contents

1 A simple scattering process

When we compute transition amplitudes for collider like LEP or LHC, we usually combine building blocks defined by Feynman rules in a way which does not make it obvious that we are dealing with a quantum field theory. One of the easiest processes we can look at is

$$e^+e^- \to \gamma^* \to q\bar{q}\,, \tag{1.1}$$

through a photon, all starting from these Feynman rules. Let is see what we start from and how we can compute this process using so-called Feynman rules. For this scattering process we need to describe four external fermions, their coupling to a photon, and the propagation of this boson from the e^+e^- annihilation to the point where is splits into a quark and antiquark pair.

From theoretical mechanics we remember that there are several ways to describe a physical system and calculate its time evolution. Assuming one degree of freedom or a real scalar field ϕ, we can for instance start with the action

$$S = \int d^4x\, \mathscr{L}(\phi, \partial_\mu\phi) \qquad \text{with} \qquad x = \begin{pmatrix} x_0 \\ \vec{x} \end{pmatrix}\,. \tag{1.2}$$

The position x is given by a space-time 4-vector, as we know if from special relativity. Under the integral there is a Lagrange density, which works exactly like the Lagrange function in mechanics, just that the object ϕ now is a quantum field, and that in particle physics we use space-time and the Minkoswki metric (+ - - -). The action has to be invariant under a variation $\delta S = 0$. We can translate this condition into the Euler-Lagrange equations

$$\partial_\mu \left(\frac{\partial \mathscr{L}}{\partial(\partial_\mu\phi)} \right) = \frac{\partial \mathscr{L}}{\partial \phi} \qquad \text{with} \qquad \partial_\mu = \frac{\partial}{\partial x^\mu}\,. \tag{1.3}$$

The second field for our switch from Lagrangian to the Hamiltonian is the (conjugate) momentum, which we can calculate just like in a classical field theory. It is

$$\pi(x) = \frac{\partial \mathscr{L}}{\partial(\partial_0\phi)} = \dot{\phi}\,. \tag{1.4}$$

With these two field we define the third object which we can use to describe the dynamics of a system, the Hamiltonian or energy functional

$$\mathcal{H}(t) = \int d^3x\, \left(\pi\dot{\phi} - \mathscr{L} \right)\,. \tag{1.5}$$

While for example in quantum mechanics this Hamiltonian is the basis of most calculations, in field theory we usually start from the Lagrangian. This also means that at the end of the day we never really use the time-dependence given by the conjugate momentum.

1.1 Boson field

We already know that for our scattering process we need to compute a transition amplitude between two kinds of matter particles, namely incoming electrons and outgoing quarks, interacting via their electric charges. The interaction is classically described by the electromagnetic Lagrangian based on the abelian $U(1)$ field theory,

$$\mathscr{L}_{\text{photon}} = -\frac{1}{4} F_{\mu\nu}F^{\mu\nu} \qquad \text{with} \qquad F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu\,, \tag{1.6}$$

in terms of a photon 4-vector field A_μ. This is exactly what we know from classical electrodynamics written in a covariant way. The Maxwell equations

$$0 = \partial^\mu F_{\mu\nu} = \partial^\mu\partial_\mu A_\nu - \partial^\mu\partial_\nu A_\mu = \Box A_\nu \qquad \text{with} \qquad \Box = \partial_\mu\partial^\mu \tag{1.7}$$

are the equations of motion for this photon field. In the last step we assume the Lorenz gauge condition $\partial_\mu A^\mu = 0$ and find the d'Alembert equation for the vector potential A_μ.

To omit the vector index of the photon field, let us instead use the real scalar ϕ from Eq.(1.2) to illustrate bosonic fields. Including a mass for this real scalar field we can write down its equation of motion which is the same for a spin-zero scalar boson as for the spin-one vector boson of Eq.(1.7)

$$\left(\Box + m^2\right)\, \phi(x) = 0 \; . \tag{1.8}$$

This Klein–Gordon equation corresponds to the d'Alembert equation for the electromagnetic vector potential in Lorentz gauge. This equation of motion for a scalar field with a mass m corresponds to a Lagrangian

$$\mathscr{L}_{\text{scalar}} = \frac{1}{2}(\partial_\mu \phi)(\partial^\mu \phi) - \frac{m^2}{2}\phi^2 \; , \tag{1.9}$$

which we can confirm using the Euler-Lagrange equation Eq.(1.3). If we want to compute the scattering amplitude in momentum space, we need to Fourier-transform the scalar field into momentum space and then quantize it, *i.e.* define commutation properties for the field in position and momentum space.

For our scattering process we need an object that describes the propagation of the (virtual) photon from its production from an e^+e^- pair to its splitting into a $q\bar{q}$ pair. This so-called propagator in position space is defined as a time–ordered product of two field operators sandwiched between vacuum states. We can think of it as describing a photon starting from its birth out of an e^+e^- pair to its death as a $q\bar{q}$ pair,

$$\Delta(x - x') \equiv i \, \langle 0 | T \left(\phi(x)\phi(x')\right) | 0 \rangle \; . \tag{1.10}$$

The time–ordered product of two operators is defined as

$$T\left(A(x)B(x')\right) = \begin{cases} A(x)B(x') & x_0 > x_0' \\ B(x')A(x) & x_0' > x_0 \end{cases} \; . \tag{1.11}$$

We can transform this propagator into Fourier space and find

$$\Delta(x - x') = -\int \frac{d^4k}{(2\pi)^4} \, e^{-ik(x-x')} \, \frac{1}{k^2 - m^2 + i\varepsilon} \; . \tag{1.12}$$

The propagator is the Green function for the Klein–Gordon equation Eq.(1.8), as we can explicitly confirm

$$\begin{aligned} \left(\Box + m^2\right) \Delta(x - x') &= -\int \frac{d^4k}{(2\pi)^4} \, \left(\Box + m^2\right) e^{-ik\cdot(x-x')} \, \frac{1}{k^2 - m^2} \\ &= \int \frac{d^4k}{(2\pi)^4} \, \left((-ik)^2 + m^2\right) e^{-ik\cdot(x-x')} \, \frac{(-1)}{k^2 - m^2} \\ &= \int \frac{d^4k}{(2\pi)^4} \, e^{-ik\cdot(x-x')} \\ &= \delta^4(x - x') \; . \end{aligned} \tag{1.13}$$

All these properties we will later use for the photon field A^μ, a vector field, where each component obeys the Klein–Gordon equation. The propagator and quantization aspects like commutation relations for the field operators will not change. The propagator only gets dressed by factors $g_{\mu\nu}$ where appropriate. For the propagator this generalization is strictly speaking gauge dependent, $g^{\mu\nu}$ corresponds to Feynman gauge.

1.2 Fermion field

Next, we need to describe (external) fermion fields. Matter particles or fermions, like leptons or quarks, have a different equation of motion and a different contribution to the Lagrangian. A field describing a fermionic particle has

to include two spin states of this particles. Moreover, in quantum field theory every fermion $\chi^\dagger$ has an antiparticle with the same mass. The altogether four degrees of freedom naturally combine to one equation with the same mass and the particle and the antiparticle described by one field ψ.

The form of the fermion field is given by the transformation property under the Lorentz transformation. We remind ourselves that a scalar field $\phi(x)$ transforms under a Lorentz transformation via a unitary operator $U(\Lambda)$ as

$$U(\Lambda)^{-1}\phi(x)U(\Lambda) = \phi(\Lambda^{-1}x) \; . \tag{1.14}$$

The fermion field has to live in a different, the so-called spinor representation. It transforms under the Lorentz transformation as

$$U(\Lambda)^{-1}\psi(x)U(\Lambda) = \Lambda_{1/2}\,\psi(\Lambda^{-1}x) \; , \tag{1.15}$$

where $\Lambda_{1/2}$ is this special representation of the Lorentz transformation. We can define it using the four Dirac matrices γ^μ with their anti–commutator Clifford algebra

$$\{\gamma^\mu, \gamma^\nu\} = 2g^{\mu\nu}\,\mathbb{1} \; . \tag{1.16}$$

The unit matrix has the same size as the γ matrices. That we usually write them as 4×4 matrices has nothing to do with the number of — also four — γ matrices. The explicit form of the γ_μ matrices is not relevant because it never appears in actual calculations. All we need is a few trace relations arising from their commutators. A representation of the Lorentz algebra in terms of the Dirac matrices is

$$\Lambda_{1/2} = \exp\left(\frac{\omega_{\mu\nu}}{8}\,[\gamma^\mu, \gamma^\nu]\right) \; . \tag{1.17}$$

This give us the transformation rule for the Dirac matrices

$$\Lambda_{1/2}^{-1}\,\gamma^\mu\,\Lambda_{1/2} = \Lambda^\mu{}_\nu \gamma^\nu \; . \tag{1.18}$$

We now postulate an equation of motion for the fermions, the Dirac equation, which describes fermions in Nature perfectly. One way to motivate this form is by taking some kind of square root of the Klein-Gordon equation. Another way is to write 2-spinors separately, for the two-spin states of the particle or of the antiparticle. Because of the spin, each of these two Dirac equation is then written with the help of the Pauli matrices as generators of the spin group $SU(2)$. For the full Dirac spinor it reads

$$(i\gamma^\mu\partial_\mu - m\mathbb{1})\;\psi(x) \equiv (i\partial\!\!\!/ - m\mathbb{1})\;\psi(x) = 0 \; . \tag{1.19}$$

The unit matrix in the mass term is a four-by-four matrix, just like the Dirac matrices. Note that the size of the γ matrix and the number of Lorentz indices μ is only coincidentally the same. We want to check that this equation is invariant under Lorentz transformations, keeping in mind that $\Lambda_{1/2}$ commutes with everything except for the Dirac matrices

$$\begin{aligned}
(i\gamma^\mu\partial_\mu - m\mathbb{1})\;\psi(x) &\to \left(i\gamma^\mu(\Lambda^{-1})^\nu{}_\mu\partial_\nu - m\mathbb{1}\right)\;\Lambda_{1/2}\psi(\Lambda^{-1}x) \\
&= \Lambda_{1/2}\Lambda_{1/2}^{-1}\left(i\gamma^\mu(\Lambda^{-1})^\nu{}_\mu\partial_\nu - m\mathbb{1}\right)\;\Lambda_{1/2}\psi(\Lambda^{-1}x) \\
&= \Lambda_{1/2}\left(i\Lambda_{1/2}^{-1}\gamma^\mu\Lambda_{1/2}(\Lambda^{-1})^\nu{}_\mu\partial_\nu - m\mathbb{1}\right)\;\psi(\Lambda^{-1}x) \\
&= \Lambda_{1/2}\left(i\Lambda^\mu{}_\rho\gamma^\rho(\Lambda^{-1})^\nu{}_\mu\partial_\nu - m\mathbb{1}\right)\;\psi(\Lambda^{-1}x) \\
&= \Lambda_{1/2}\left(ig^\nu{}_\rho\gamma^\rho\partial_\nu - m\mathbb{1}\right)\;\psi(\Lambda^{-1}x) \\
&= \Lambda_{1/2}\left(i\gamma^\nu\partial_\nu - m\mathbb{1}\right)\;\psi(\Lambda^{-1}x) = 0 \; .
\end{aligned} \tag{1.20}$$

We also see that we can multiply the Dirac equation with $(-i\gamma^\mu\partial_\mu - m\mathbb{1})$ and obtain

$$\begin{aligned}
(-i\gamma^\mu\partial_\mu - m\mathbb{1})\,(i\gamma^\nu\partial_\nu - m\mathbb{1})\;\psi(x) &= \left(\gamma^\mu\gamma^\nu\partial_\mu\partial_\nu + m^2\mathbb{1}\right)\;\psi(x) \\
&= \left(\frac{\gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu}{2}\partial_\mu\partial_\nu + m^2\mathbb{1}\right)\;\psi(x) \\
&= \left(\partial^2 + m^2\mathbb{1}\right)\;\psi(x) \; ,
\end{aligned} \tag{1.21}$$

symbol	meaning	operator
$u_s(k)$	incoming fermion (e^-, q) with momentum k and spin s	a_s
$\bar{v}_s(k)$	incoming anti–fermion $(e^+, \bar{q})$	b_s
$\bar{u}_s(k)$	outgoing fermion (e^-, q)	$a_s^\dagger$
$v_s(k)$	outgoing anti–fermion $(e^+, \bar{q})$	$b_s^\dagger$

Table 1: Assignment of generation and annihilating operators for particle and antiparticle spinors in momentum space, as given in Eq.(1.25).

which means that every fermion field that obeys the Dirac equation also fulfills a Klein–Gordon equation. We will see this similarity when we construct the fermion propagator.

To define a mass term in the Lagrangian we need to form Lorentz scalars or invariants out of the fermion fields ψ. Naively, $(\psi^\dagger\psi)$ would work if the Lorentz transformations in $(\psi^\dagger\Lambda_{1/2}^\dagger\Lambda_{1/2}\psi)$ cancelled. Unfortunately $\Lambda_{1/2}$ is not a unitary transformation. Instead, one can show that the Dirac adjoint

$$\overline{\psi} = \psi^\dagger\gamma^0 \qquad \text{with} \quad \overline{\psi}\psi \to \overline{\psi}\psi \tag{1.22}$$

has the correct transformation property. This allows us to write down the Lagrangian which corresponds to the Dirac equation for a fermion field

$$\mathscr{L}_{\text{fermion}} = \overline{\psi}(i\partial\!\!\!/ - m\mathbb{1})\psi \; . \tag{1.23}$$

Because we will later need the fermion–photon interaction in the Lagrangian, we introduce the convenient form of the covariant derivative

$$\begin{aligned}\mathscr{L}_{\text{fermion-photon}} &= \overline{\psi}\,(iD\!\!\!/ - m\mathbb{1})\,\psi \\ &\equiv \overline{\psi}\,(i(\partial\!\!\!/ + ieA\!\!\!/) - m\mathbb{1})\,\psi = \overline{\psi}\,(i(\partial\!\!\!/ - m\mathbb{1})\,\psi + eqA_\mu\,\overline{\psi}\gamma^\mu\psi\end{aligned} \tag{1.24}$$

The last term describes the coupling of a vector photon field A_μ to a vector-like expression $\overline{\psi}\gamma^\mu\psi$ which we call a vector current of a spinor field.

Just like in the bosonic case we now Fourier-transform the Dirac field operators, which we know have to include four degrees of freedom, particle and anti-particle with two spins each,

$$\begin{aligned}\psi(x) &= \int \frac{d^3k}{(2\pi)^3 2k_0} \sum_{\text{spin } s\,=\,\pm 1/2} \left(e^{ikx}v_s(k)b_s^\dagger(\vec{k}) + e^{-ikx}u_s(k)a_s(\vec{k})\right) \\ \overline{\psi}(x) &= \int \frac{d^3k}{(2\pi)^2 2k_0} \sum_{\text{spin } s\,=\,\pm 1/2} \left(e^{ikx}\bar{u}_s(k)a_s^\dagger(\vec{k}) + e^{-ikx}\bar{v}_s(k)b_s(\vec{k})\right) \; .\end{aligned} \tag{1.25}$$

The 4-dimensional spinors u and v in Fourier space create or annihilate the particle, while $\bar{u}$ and $\bar{v}$ create or annihilate the antiparticle in Fourier space, as listed in Tab. 1. As before, we skip the quantization steps and directly quote the spin sums for the spinors u and v and their Dirac adjoints

$$\begin{aligned}\sum_{\text{spin } s\,=\,\pm 1/2} u_s(k)\bar{u}_s(k) &= k\!\!\!/ + m\mathbb{1} \\ \sum_{\text{spin } s\,=\,\pm 1/2} v_s(k)\bar{v}_s(k) &= k\!\!\!/ - m\mathbb{1} \; .\end{aligned} \tag{1.26}$$

These spin sums do not combine spinors to scalars, but to Dirac matrices. For many applications in LHC physics the masses are negligible, and since they really complicate formulas we evaluate the spin sums without masses whenever we can.

1.3 Scattering

Now we have everything we need to compute a transition amplitude for our scattering process

$$e^-(k_1, s_1) + e^+(k_2, s_2) \to q(k_3, s_3) + \bar{q}(k_4, s_4) \ , \tag{1.27}$$

where k_j and s_j are the four-momenta and spin orientations of the external fermions, and $k_1 + k_2 + k_3 + k_4 = 0$. We neglect the electron and quark masses of the external particles. In the future, or more specifically asymptotically for $t \to +\infty$, the initial state $\lim_{t\to-\infty} |t\rangle \equiv |i\rangle$ will have evolved into the final state $\lim_{t\to\infty} |t\rangle = \mathcal{S}|i\rangle$ via a yet unknown linear operator $\mathcal{S}$. To describe this scattering into a final state $\langle f|$ we need to compute the transition amplitude

$$S \equiv \langle f|\mathcal{S}|i\rangle = \langle q_3\bar{q}_4|\mathcal{S}|e_1^+e_2^-\rangle = \langle 0|a_3^\dagger b_4^\dagger \, \mathcal{S} \, a_1 b_2|0\rangle \ . \tag{1.28}$$

We use a single iindex to indicate the momenta and spins of the external particles. This transition amplitude is not a vacuum expectation value, but the operator $\mathcal{S}$ sandwiched between physically measurable states made from the vacuum using the generation and annihilation field operators $a, a^\dagger, b, b^\dagger$ defined in Eq.(1.25).

The transition matrix element $\mathcal{S}$ can be computed from the time evolution of the initial state $i\partial_t|t\rangle = \mathcal{H}(t)|t\rangle$ in the interaction picture with a time-dependent Hamilton operator,

$$\mathcal{S} = T\left(e^{-i\int dt\, \mathcal{H}(t)}\right) \ , \tag{1.29}$$

again with time ordering T. This form ensures that it generates a unitary transformation. For our computation we will be fine with the interaction Hamiltonian for two incoming and two outgoing fermios, each pair involving a different particle species j with charge q_j,

$$\mathcal{H}_{\text{int}}(t) = -\int d^3x \, \mathscr{L}_{\text{int}}(x) \supset \sum_j -eq_j \int d^3x \, A_\mu \, \overline{\psi}_j \gamma^\mu \psi_j \ , \tag{1.30}$$

in terms of the four-vector x including its first entry $t = x_0$. We skip the corresponding calculation and just give the result for the numbering of the incoming and outgoing particles defined in Eq.(1.27)

$$S = \sum_{\text{spins}} i(2\pi)^4 \delta^4(k_1 + k_2 - k_3 - k_4) \ e^2 q_e q_q \ \bar{u}_3 \gamma_\mu v_4 \frac{1}{(k_1+k_2)^2} \bar{v}_2 \gamma^\mu u_1 \ . \tag{1.31}$$

Stripping off unwanted prefactors we define the transition matrix element for quark–antiquark production in QED as

$$\mathcal{M} = e^2 q_e q_q \ (\bar{u}_3 \gamma_\mu v_4) \ \frac{1}{(k_1+k_2)^2} \ (\bar{v}_2 \gamma^\mu u_1) \ , \tag{1.32}$$

We have to square this matrix element or transition amplitude to compute the transition probability. Part of the squaring is the sum over all spins which uses the spin sums Eq.(1.26) to get rid of the spinors and then some trace rules to get rid of all Dirac matrices. For neither the spinors nor the Dirac matrices we need to know their explicit form

$$\begin{aligned} |\mathcal{M}|^2 =& \sum_{\text{spin, color}} e^4 q_e^2 q_q^2 \ \frac{1}{(k_1+k_2)^4} \ (\bar{v}_4\gamma_\nu u_3)(\bar{u}_1\gamma^\nu v_2) \ (\bar{u}_3\gamma_\mu v_4) \ (\bar{v}_2\gamma^\mu u_1) \\ =& e^4 q_e^2 q_q^2 N_c \ \frac{1}{(k_1+k_2)^4} \ \sum_{\text{spin}} (\bar{v}_4\gamma_\nu u_3)(\bar{u}_1\gamma^\nu v_2) \ (\bar{u}_3\gamma_\mu v_4) \ (\bar{v}_2\gamma^\mu u_1) \ . \end{aligned} \tag{1.33}$$

The color factor N_c is the number of outgoing color singlet states we can form out of a quark and an antiquark with opposite color charges. Because color only appears in the final state we sum over all possible color states or multiply by N_c. In the next step we can observe how the crucial structure of transition amplitudes with external fermions,

namely traces of chains of Dirac matrices, magically form:

$$\begin{aligned}
|\mathcal{M}|^2 =& e^4 q_e^2 q_q^2 N_c \; \frac{1}{(k_1+k_2)^4} \; \sum_{\text{spin}} (\bar v_4)_i (\gamma_\nu)_{ij} (u_3)_j (\bar u_3)_k (\gamma_\mu)_{kl} (v_4)_l \cdots && \text{for one trace} \\
=& e^4 q_e^2 q_q^2 N_c \; \frac{1}{(k_1+k_2)^4} \; \left(\sum_{\text{spin}} (v_4)_l (\bar v_4)_i \right) \left(\sum_{\text{spin}} (u_3)_j (\bar u_3)_k \right) \; (\gamma_\nu)_{ij} (\gamma_\mu)_{kl} \cdots \\
=& e^4 q_e^2 q_q^2 N_c \; \frac{1}{(k_1+k_2)^4} \; (\not k_4)_{li} (\not k_3)_{jk} (\gamma_\nu)_{ij} (\gamma_\mu)_{kl} \cdots && \text{using Eq.(1.26), no masses} \\
=& e^4 q_e^2 q_q^2 N_c \; \frac{1}{(k_1+k_2)^4} \; \text{Tr}\,(\not k_4 \gamma_\nu \not k_3 \gamma_\mu) \; \text{Tr}\,(\not k_1 \gamma^\nu \not k_2 \gamma^\mu) && \text{both traces again.} \qquad (1.34)
\end{aligned}$$

In the final step we need a standard expression for the Dirac trace. Longer traces become very complicated very fast, and we evaluate them using symbolic manipulation on the computer. We find

$$\begin{aligned}
|\mathcal{M}|^2 =& e^4 q_e^2 q_q^2 N_c \; \frac{1}{(k_1+k_2)^4} \; 4\,(k_{3\nu} k_{4\mu} + k_{3\mu} k_{4\nu} - g_{\mu\nu}(k_3 k_4)) \; 4\,(k_1^\nu k_2^\mu + k_1^\mu k_2^\nu - g_{\mu\nu}(k_1 k_2)) \\
=& 16 e^4 q_e^2 q_q^2 N_c \; \frac{1}{(k_1+k_2)^4} \; [2(k_1 k_3)(k_2 k_4) + 2(k_1 k_4)(k_2 k_3) + 0 \times (k_3 k_4)(k_1 k_2)] \qquad \text{with} \quad g_{\mu\nu} g^{\mu\nu} = 4 \\
=& 32 e^4 q_e^2 q_q^2 N_c \; \frac{1}{(k_1+k_2)^4} \; [(k_1 k_3)(k_2 k_4) + (k_1 k_4)(k_2 k_3)] \; , \qquad (1.35)
\end{aligned}$$

To evaluate this matrix element we first introduce Mandelstam variables as squares of sums of 4-vectors,

$$s = (k_1+k_2)^2 \approx 2(k_1 k_2) > 0 \qquad t = (k_1+k_3)^2 \approx 2(k_1 k_3) < 0 \qquad u = (k_1+k_4)^2 \approx 2(k_1 k_4) \; , \qquad (1.36)$$

where in this sign convention all momenta are incoming, $k_1+k_2+k_3+k_4 = 0$, and in the second step we neglect the masses. The second Mandelstam variable can be expressed through the polar or scattering angle

$$t = \frac{s}{2}(-1+\cos\theta) \in [-s, 0] \; . \qquad (1.37)$$

Allowing the external particles have a finite mass we can use this 4-momentum conservation to show

$$s+t+u = k_1^2 + k_2^2 + k_3^2 + k_4^2 \equiv m_1^2 + m_2^2 + m_3^3 + m_4^2 \; . \qquad (1.38)$$

In our case all masses are zero and we find the compact form

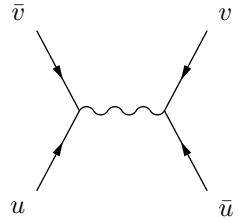
$$\begin{aligned}
|\mathcal{M}|^2 &= 32 e^4 \; q_e^2 q_q^2 N_c \; \frac{1}{s^2} \; \left[\frac{t^2}{4} + \frac{u^2}{4} \right] \\
&= 8 e^4 \; q_e^2 q_q^2 N_c \; \frac{1}{s^2} \; \left[s^2 + 2st + 2t^2 \right] \\
&= 8 e^4 \; q_e^2 q_q^2 N_c \; \left[1 + 2\frac{t}{s} + 2\frac{t^2}{s^2} \right] \; . \qquad (1.39)
\end{aligned}$$

We can briefly check if this number is indeed positive, using the definition of the Mandelstam variable t for massless external particles in terms of the polar angle: the upper phase space boundary $t = 0$ inserted into the brackets in Eq.(1.39) gives $[\cdots] = 1$, just as the lower boundary $t = -s$ with $[\cdots] = 1 - 2 + 2 = 1$. For the central value $t = -s/2$ the minimum value of the brackets is $[\cdots] = 1 - 1 + 0.5 = 0.5$.

The azimuthal angle ϕ plays no role at colliders, unless you want to compute gravitational effects on Higgs production at ATLAS and CMS. Any LHC Monte Carlo will either random-generate a reference angle ϕ for the partonic process or pick one and keep it fixed.

1.4 Feynman rules

Feynman rules are calculational rules which we can extract from the Lagrangian and which allow us to derive Eq.(1.32) directly. We start by drawing Feynman diagrams for all ways we can link the given initial and final states through interaction vertices and internal propagators. For our scattering process there is only one diagram:



It consist of four external fermions labeled according to Tab. 1, one internal photon, and two interaction vertices. From Eq.(1.25) and Tab. 1 we know how to describe external fermions in terms of spinors.

Spin sums are the only way to get rid of spinors in the computation. Equation (1.26) shows that as long as we neglect fermion masses the two spinors u and v for particles and antiparticles are identical. To link external particles to each other and to internal propagators we need vertices. If two fermions and a gauge boson interact via a vector current proportional to γ^μ, and adding a conventional factor i, the one vertex rule in QED reads

$$ieq_f\,\gamma^\mu \qquad\qquad (f-\bar{f}-\gamma). \tag{1.40}$$

This factor i we can consistently change for all three-point and four-point vertices in our theory. Finally, there is the intermediate photon which propagates between the γ^μ and the γ^ν vertices. The wave line in the Feynman diagram corresponds to

$$-i\frac{g^{\mu\nu}}{p^2+i\epsilon}\;. \tag{1.41}$$

Again, the factor $-i$ is conventional. For a bosonic propagator it does not matter in which direction the momentum flows. Blindly combining these Feynman rules gives us directly Eq.(1.32), so all we need to do is square the matrix element, insert the spin sums and compute the Dirac trace.

We do not need it for our QED calculation, but for instance process $e^-\gamma \to e^-\gamma$ is described by an intermediate fermion propagator in the s-channel. This propagator is described by the Feynman rule

$$i\frac{\not{p}+m\mathbb{1}}{p^2-m^2} = i\frac{\not{p}+m\mathbb{1}}{\not{p}^2-m^2} = i\frac{\not{p}+m\mathbb{1}}{(\not{p}+m\mathbb{1})(\not{p}-m\mathbb{1})} = i\,(\not{p}-m\mathbb{1})^{-1}\;. \tag{1.42}$$

It lives in the same space as gamma matrices. Because the sign of the 4-momentum matters we asign it in parallel to the fermion arrow. This will work fine until we have to deal with Majorana particles in the neutrino sector or supersymmetry (for those who still remember that).

Whenever we compute such a matrix element starting from a Feynman diagram nothing tells us that the lines in the Feynman diagrams are not actual physical states propagating from the left to the right. Even including loop diagrams will still look completely reasonably from a semi–classical point of view. Feynman rules define an algorithm which hides all field theory input in the calculation of scattering amplitudes and are therefore perfectly suited to compute the differential and total cross sections on the computer.

1.5 Chirality

The vector structure of the QED couplings, for example mediated by a covariant derivative Eq.(1.24) we did not actually motivate. It happens to work on the Lagrangian level and agrees with data, so it is correct. We can write a completely general interaction of two fermions with a boson in terms of basis elements

$$g\,\overline{\psi}M\psi = \sum_{\text{basis } j} g_j\,\overline{\psi}M_j\psi\;. \tag{1.43}$$

For a real (4×4) matrix M the necessary 16 basis elements can be organized such that they are easy to keep track of using Lorentz transformation properties. This eventually leads to the so-called Fierz transformation. The vector γ^μ from the QED interaction gives us four such basis elements, the unit matrix a fifth. Another six we already know as well, they are the generators of the spinor representation $[\gamma^\mu, \gamma^\nu]$. All of them are linearly independent.

Five basis elements are still missing. To define them, we start with another (4×4) matrix which is invariant under proper Lorentz transformations. We can write it in two equivalent forms

$$\gamma_5 = i\gamma^0\gamma^1\gamma^2\gamma^3 \equiv \frac{i}{4!}\,\epsilon_{\mu\nu\rho\sigma}\gamma^\mu\gamma^\nu\gamma^\rho\gamma^\sigma \;, \tag{1.44}$$

using the totally anti–symmetric Levi–Civita tensor $\epsilon_{\mu\nu\rho\sigma}$. This form already shows a major technical complication in dealing with γ_5: in other than four space–time dimensions we do not know how to define the Levi–Civita tensor, which means that for example for regularization purposes we cannot analytically continue our calculation to $n = 4 - 2\epsilon$ dimensions. The main properties of γ_5 are

$$\gamma_5^2 = \mathbb{1} \qquad \text{and} \qquad \{\gamma_\mu, \gamma_5\} = 0 \;. \tag{1.45}$$

It gives us all 16 basis element for the interaction of two spinors:

	degrees of freedom	basis elements M_j
scalar	1	$\mathbb{1}$
vector	4	γ^μ
pseudoscalar	1	$i\gamma_5$
axialvector	4	$\gamma^\mu\gamma_5$
tensor	6	$\frac{i}{2}[\gamma^\mu, \gamma^\nu]$

In the renormalizable Standard Model as a fundamental theory, tensor interactions do not play a role.

An obvious question is: what does it mean to include a factor γ_5 in the interaction, *i.e.* what distinguishes a scalar from a pseudoscalar and a vector from an axialvector? We can give an easy answer by defining three transformations of our field in space and time. The first one is the parity transformation P which mirrors the three spatial coordinates $(t, \vec{x}) \to (t, -\vec{x})$. The second is charge conjugation C which converts particles into their anti–particles. Both of them leave the Dirac equation intact and can be represented by a unitary transformation. The third transformation is time reversal T which converts $(t, \vec{x}) \to (-t, \vec{x})$, also leaves the Dirac equation intact, but only has an anti–unitary representation. Every single one of them is violated in our Standard Model.

Instead of writing out the representation of these transformations in terms of Dirac matrices we characterize them using the basic interactions from Eq.(1.43). Parity symmetry does not allow any interaction including γ_5, which means it forbids pseudoscalars and axialvectors. Time reversal symmetry does not allow any complex couplings g_j. Because any field theory described by a Lagrangian not including some kind of external field is invariant under CPT, and we have never observed CPT violation, a combined CP-invariance is essentially the same as T invariance.

To look at the parity and CP symmetry more systematically, we rotate the $\{\mathbb{1}, \gamma_5\}$ plane and define the two matrices

$$\mathbb{P}_{R,L} = \frac{1}{2}\,(\mathbb{1} \pm \gamma_5) \;. \tag{1.46}$$

It is easy to show that the two are orthogonal projectors,

$$\begin{aligned} \mathbb{P}_L\mathbb{P}_R &= \frac{1}{4}\,(\mathbb{1} - \gamma_5)\,(\mathbb{1} + \gamma_5) = \frac{1}{4}\,\left(\mathbb{1} - \gamma_5^2\right) = 0 \\ \mathbb{P}_{R,L}^2 &= \frac{1}{4}\,\left(\mathbb{1} \pm 2\gamma_5 + \gamma_5^2\right) = \frac{1}{4}\,(2\mathbb{1} \pm 2\gamma_5) = \frac{1}{2}\,(\mathbb{1} \pm \gamma_5) = \mathbb{P}_{R,L} \;. \end{aligned} \tag{1.47}$$

We first look at what happens when we write a kinetic term with left-handed and right-handed projectors or fermion fields,

$$\begin{aligned}
\overline{\psi}\,\partial\!\!\!/\psi &= \overline{\psi}\,\partial\!\!\!/\left(\mathbb{P}_L^2+\mathbb{P}_R^2\right)\psi \\
&= \overline{\psi}\,\left(\mathbb{P}_R\partial\!\!\!/\mathbb{P}_L+\mathbb{P}_L\partial\!\!\!/\mathbb{P}_R\right)\psi \\
&= (\overline{\mathbb{P}_L\psi})\partial\!\!\!/(\mathbb{P}_L\psi)+(\overline{\mathbb{P}_R\psi})\partial\!\!\!/(\mathbb{P}_R\psi) \\
&= \overline{\psi}_L\,\partial\!\!\!/\psi_L+\overline{\psi}_R\,\partial\!\!\!/\psi_R\;. \qquad (1.48)
\end{aligned}$$

The general kinetic term covers left-handed and right-handed fields. Their effect on a mass term is different,

$$\begin{aligned}
\overline{\psi}\,\mathbb{1}\!\!\!/\psi &= \overline{\psi}\,\left(\mathbb{P}_L+\mathbb{P}_R\right)\psi \\
&= \overline{\psi}\,\left(\mathbb{P}_L^2+\mathbb{P}_R^2\right)\psi \\
&= \psi^\dagger\gamma_0\,\left(\mathbb{P}_L^2+\mathbb{P}_R^2\right)\psi && \text{with} \quad \overline{\psi}=\psi^\dagger\gamma^0 \\
&= \psi^\dagger\left(\mathbb{P}_R\gamma^0\mathbb{P}_L+\mathbb{P}_L\gamma^0\mathbb{P}_R\right)\psi && \text{with} \quad \{\gamma_5,\gamma_\mu\}=0 \\
&= (\mathbb{P}_R\psi)^\dagger\gamma^0(\mathbb{P}_L\psi)+(\mathbb{P}_L\psi)^\dagger\gamma^0(\mathbb{P}_R\psi) && \text{with} \quad \gamma_5^\dagger=\gamma_5, \mathbb{P}_{L,R}^\dagger=\mathbb{P}_{L,R} \\
&= (\overline{\mathbb{P}_R\psi})\mathbb{1}\!\!\!/(\mathbb{P}_L\psi)+(\overline{\mathbb{P}_L\psi})\mathbb{1}\!\!\!/(\mathbb{P}_R\psi) \\
&= \overline{\psi}_R\,\mathbb{1}\!\!\!/\psi_L+\overline{\psi}_L\,\mathbb{1}\!\!\!/\psi_R\;. \qquad (1.49)
\end{aligned}$$

To include a fermion mass we need to combine left-handed and right-handed projectors and fermion fields,

$$\begin{aligned}
\overline{\psi}\,\partial\!\!\!/\,\psi &= \overline{\psi}_R\,\partial\!\!\!/\,\psi_R+\overline{\psi}_L\,\partial\!\!\!/\,\psi_L \\
\overline{\psi}\,\mathbb{1}\!\!\!/\,\psi &= \overline{\psi}_R\,\mathbb{1}\!\!\!/\,\psi_L+\overline{\psi}_L\,\mathbb{1}\!\!\!/\,\psi_R\;. \qquad (1.50)
\end{aligned}$$

In other words, we can write for example QED in terms of independent left and right handed fields as long as we neglect all fermion masses. This defines the chiral limit where the Lagrangian is symmetric under $\psi_L \leftrightarrow \psi_R$. Introducing fermion masses breaks this chiral symmetry, or turning the argument around, to introduce fermion masses we need to combine a left-handed and a right-handed fermion fields and give them one common Dirac mass.

Moving on to interactions, we define a combined vector–axialvector coupling as $\gamma_\mu \pm \gamma_\mu\gamma_5 = 2\gamma_\mu\mathbb{P}_{R,L}$. Sandwiching this coupling between fermion fields gives for example

$$\begin{aligned}
\overline{\psi}\,\gamma_\mu\mathbb{P}_L\,\psi &= \overline{\psi}\,\gamma_\mu\mathbb{P}_L^2\,\psi \\
&= \psi^\dagger\mathbb{P}_L\,\gamma_0\,\gamma_\mu\,\mathbb{P}_L\psi && \text{with} \quad \{\gamma_5,\gamma_\mu\}=0 \\
&= \left(\mathbb{P}_L\psi\right)^\dagger\gamma_0\,\gamma_\mu\,\mathbb{P}_L\psi && \text{with} \quad \gamma_5^\dagger=\gamma_5 \\
&= \overline{\psi}_L\,\gamma_\mu\,\psi_L && \text{with} \quad \psi_{L,R}\equiv\mathbb{P}_{L,R}\,\psi\;. \qquad (1.51)
\end{aligned}$$

If we call the eigenstates of $\mathbb{P}_{R,L}$ right-handed and left-handed fermions $\psi_{L,R}$ this chirality allows us to define a vector coupling between only left handed fermions by combining the vector and the axialvector couplings with a relative minus sign. The same is of course true for right handed couplings. We can now describe the $\gamma\bar{f}f$ coupling in QED using the Feynman rule

$$-i\gamma^\mu\left(\ell\mathbb{P}_L+r\mathbb{P}_R\right) \qquad \text{with} \quad \ell=r=qe\;. \qquad (1.52)$$

At this stage it is not obvious at all what chirality means in physics terms. However, we will see that in the Standard Model the left handed fermions play a special role: the massive W bosons only couple to them and not to their right handed counter parts. So chirality is a property of fermions known to one gauge interaction of the Standard Model as part of the corresponding charge.

1.6 Helicity and spin

There exists a property which is identical to chirality for massless fermions and has an easy physical interpretation, helicity. It is defined as the projection of the particle spin onto its three-momentum direction

$$h = \vec{s} \cdot \frac{\vec{p}}{|\vec{p}|} = \left(\vec{s} + \vec{L}\right) \cdot \frac{\vec{p}}{|\vec{p}|} = \vec{J} \cdot \frac{\vec{p}}{|\vec{p}|} \qquad \text{with} \quad \vec{p} \perp \vec{L} \,, \tag{1.53}$$

or equivalently the projection of the combined orbital angular momentum and the spin on the momentum direction. From quantum mechanics we know that there exist discrete eigenvalues for the z component of the angular momentum operator, symmetric around zero. Applied to fermions this gives us two spin states with the eigenvalues of h being $\pm 1/2$. Unfortunately, there is no really nice way to show this identity. What we need to know is that the spin operator is in general given by

$$\vec{s} = \gamma_5 \gamma^0 \, \vec{\gamma} \,. \tag{1.54}$$

We can show this by writing it out in terms of Pauli matrices, but we will skip this here and instead just accept this general form. We then write the solution ψ to the massless Dirac equation after transforming it into momentum space $\psi(\vec{x}) = u(\vec{p}) \exp(-ip \cdot x)$

$$\begin{aligned}
\left(\gamma^0 p_0 - \vec{\gamma}\vec{p}\right) \; u(\vec{p}) &= 0 \\
\gamma_5 \gamma^0 \; \gamma^0 p_0 \; u(\vec{p}) &= \gamma_5 \gamma^0 \; \vec{\gamma}\vec{p} \, u(\vec{p}) \\
\gamma_5 p_0 \; u(\vec{p}) &= \vec{s} \cdot \vec{p} \, u(\vec{p}) \qquad\qquad \text{with } \left(\gamma^0\right)^2 = \mathbb{1} \\
\gamma_5 \; u(\vec{p}) &= \frac{\vec{s} \cdot \vec{p}}{p_0} \; u(\vec{p}) \\
\gamma_5 \; u(\vec{p}) &= \pm \frac{\vec{s} \cdot \vec{p}}{|\vec{p}|} \; u(\vec{p}) = \pm h \; u(\vec{p}) \,.
\end{aligned} \tag{1.55}$$

In other words, the chirality operator γ_5 indeed gives us the helicity of a particle, modulo a sign depending on the sign of the energy. For the helicity it is easy to argue why for massive particles this property is not Lorentz invariant and hence not a well defined property: massless particles propagate with the speed of light, which means we can never boost into their rest frame or pass them. For massive particles we can do that and this way switch the sign of $\vec{p}$ and the sign of h. Luckily, for almost all Standard Model fermions we can neglect their masses at the LHC.

1.7 Cross section measurements

To compute a $2 \to 2$ scattering rate we combine the scattering matrix element from Eq.(1.39) with a two-particle phase space integration for massless particles,

$$s^2 \frac{d\sigma}{dt}\bigg|_{2\to 2} = \frac{\pi}{(4\pi)^2} \; K_{ij} \; |\mathcal{M}|^2 \tag{1.56}$$

with an averaging factor K_{ij} for initial–state spins and colors, as only the sum is included in Eq.(1.39). For incoming electrons as well as incoming quarks this factor K_{ij} includes $1/4$ for the spins. For an incoming $q\bar{q}$ pair we would also average over the color, $1/N_c^2$.

For our QED process we then find the differential cross section in four space–time dimensions, using $\alpha = e^2/(4\pi)$

$$\begin{aligned}
\frac{d\sigma}{dt} &= \frac{1}{s^2} \frac{\pi}{(4\pi)^2} \; \frac{1}{4} \; 8 \; q_e^2 q_q^2 (4\pi\alpha)^2 \; N_c \; \left[1 + 2\frac{t}{s} + 2\frac{t^2}{s^2}\right] \\
&= \frac{1}{s^2} 2\pi\alpha^2 \; N_c \; q_e^2 q_q^2 \; \left[1 + 2\frac{t}{s} + 2\frac{t^2}{s^2}\right] \,.
\end{aligned} \tag{1.57}$$

We integrate this expression over the polar angle or the Mandelstam variable t to compute the total cross section

$$\begin{aligned}
\sigma &= \frac{1}{s^2} 2\pi\alpha^2 \; N_c \; q_e^2 q_q^2 \int_{-s}^{0} dt \; \left[1 + 2\frac{t}{s} + 2\frac{t^2}{s^2}\right] \\
&= \frac{1}{s^2} 2\pi\alpha^2 \; N_c \; q_e^2 q_q^2 \left[t + \frac{t^2}{s} + \frac{2t^3}{3s^2}\right]_{-s}^{0} \\
&= \frac{1}{s^2} 2\pi\alpha^2 \; N_c \; q_e^2 q_q^2 \left[s - \frac{s^2}{s} + \frac{2s^3}{3s^2}\right] \\
&= \frac{1}{s} 2\pi\alpha^2 \; N_c \; q_e^2 q_q^2 \; \frac{2}{3} \qquad\qquad \Rightarrow \qquad\qquad \sigma(e^+e^- \to q\bar{q})\Big|_{\text{QED}} = \frac{4\pi\alpha^2 N_c}{3s} \; q_e^2 q_q^2 \; . \qquad (1.58)
\end{aligned}$$

In the history of QCD, this process played a crucial role, namely the production rate of quarks in e^+e^- scattering. For small enough energies we can neglect the Z exchange contribution. At leading order we can then compute the corresponding production cross sections for muon pairs and for quark pairs in e^+e^- collisions.

$$R \equiv \frac{\sigma(e^+e^- \to \text{hadrons})}{\sigma(e^+e^- \to \ell^+\ell^-)} = \frac{\sum_{\text{quarks}} \frac{4\pi\alpha^2 N_c}{3s} \; q_e^2 q_q^2}{\frac{4\pi\alpha^2}{3s} \; q_e^2 q_\ell^2} = N_c \left(3\frac{1}{9} + 2\frac{4}{9}\right) = \frac{11 N_c}{9} \; , \qquad (1.59)$$

for example for five quark flavors where the top quark is too heavy to be produced at the given e^+e^- collider energy. For those interested in the details we did take one short cut: hadrons are also produced in the hadronic decays of $e^+e^- \to \tau^+\tau^-$ which we strictly speaking need to subtract. This way, R as a function of the collider energy is a beautiful measurement of the weak and color charges of the quarks in QCD.

Finally, if we face the fact that most particle physicists nowadays work on precision hadron colliders, and high-energy e^+e^--colliders are either a thing of the past or a dream for the future, we want to compute our QED process the other way around. This means we move the quarks into the initial state and include a color-averaging factor. The corresponding process is called the Drell–Yan process

$$\sigma(q\bar{q} \to \ell^+\ell^-)\Big|_{\text{QED}} = \frac{4\pi\alpha^2}{3N_c s} \; q_\ell^2 q_q^2 \; . \qquad (1.60)$$

It will be the process which guides us through the discussion of modern collider physics. Obviously, to describe lepton pair production at the LHC, we need to include the massive electroweak gauge bosons, not just the photons. We will get to that later.

2 Extra: helicity amplitudes

In research, transition amplitudes are not computed by squaring matrix elements and computing gamma matrix traces. Instead, all elements of the Feynman rules are evaluated numerically, combined, and then squared. From the previous section we know how to compute the cross section for Z production by writing down all external spinors, external polarization vectors, interaction vertices and propagators and squaring the amplitude analytically. The amplitude itself inherits external indices for example from the polarization vectors, while $|\mathcal{M}|^2$ is a real positive number with a fixed mass dimension depending on the number of external particles.

As an example for a more modern computation of transition amplitudes, we consider again lepton pair production in QED,

$$u\bar{u} \to \gamma^* \to \mu^+\mu^- \; . \tag{2.1}$$

The structure of the amplitude $\mathcal{M}$ with two internal Dirac indices μ and ν involves one vector current on each side $(\bar{u}_f \gamma_\mu u_f)$ where $f = u, \mu$ are to good approximation massless, so we do not have to be careful with the different spinors u and v. The entries in the external spinors are given by the spin of the massless fermions obeying the Dirac equation. For each value of $\mu = 0 \cdots 3$ each current is a complex number, computed from the four component of each spinor and the respective 4×4 gamma matrix γ^μ. The intermediate photon propagator has the form $g_{\mu\nu}/s$, which is a real number for each value of $\mu = \nu$. Summing over μ and ν in both currents forms the matrix element. To square this matrix element we need to sum $\mathcal{M}^* \times \mathcal{M}$ over all possible spin directions of the external fermions.

Instead of squaring this amplitude symbolically we can follow exactly the steps described above and compute an array of numbers for different spin and helicity combinations numerically. Summing over the internal Dirac indices we compute the matrix element; however, to compute the matrix element squared we need to sum over external fermion spin directions or gauge boson polarizations. The helicity basis we have to specify externally. This is why this method is called helicity amplitude approach. To explain the way this method works, we illustrate it for muon pair production based on the implementation in the Madgraph/Helas package.

Madgraph is a tool to compute matrix elements this way. Other event generators have corresponding codes serving the same purposes. In our case, Madgraph5 automatically produces a Fortran (!!) routine which then calls functions to compute spinors, polarization vectors, currents of all kinds, etc. These functions are available as the Helas library. For our toy process Eq.(2.1) the slightly shortened Madgraph5 output reads

```
      REAL*8 FUNCTION MATRIX1(P,NHEL,IC)
C
C     Generated by Madgraph 5
C
C     Returns amplitude squared summed/avg over colors
C     for the point with external lines W(0:6,NEXTERNAL)
C
C     Process: u u~ > mu+ mu- / z WEIGHTED=4 @1
C
      INTEGER    NGRAPHS, NWAVEFUNCS, NCOLOR
      PARAMETER (NGRAPHS=1, NWAVEFUNCS=5, NCOLOR=1)

      REAL*8 P(0:3,NEXTERNAL)
      INTEGER NHEL(NEXTERNAL), IC(NEXTERNAL)

      INCLUDE 'coupl.inc'

      DATA DENOM(1)/1/
      DATA (CF(I,  1),I=  1,  1) /    3/

      CALL IXXXXX(P(0,1),ZERO,NHEL(1),+1*IC(1),W(1,1))
      CALL OXXXXX(P(0,2),ZERO,NHEL(2),-1*IC(2),W(1,2))
      CALL IXXXXX(P(0,3),ZERO,NHEL(3),-1*IC(3),W(1,3))
      CALL OXXXXX(P(0,4),ZERO,NHEL(4),+1*IC(4),W(1,4))
      CALL FFV1_3(W(1,1),W(1,2),GC_2,ZERO, ZERO, W(1,5))
      CALL FFV1_0(W(1,3),W(1,4),W(1,5),GC_3,AMP(1))
      JAMP(1)=+AMP(1)

      DO I = 1, NCOLOR
        DO J = 1, NCOLOR
          ZTEMP = ZTEMP + CF(J,I)*JAMP(J)
```

```
  ENDDO
  MATRIX1 = MATRIX1 + ZTEMP*DCONJG(JAMP(I))/DENOM(I)
ENDDO

END
```

The input to this function are the external four-momenta $p(0:3,1:4)$ and the helicities of all fermions $n_{\rm hel}(1:4)$ in the process. Remember that helicity and chirality are identical only for massless fermions because chirality is defined as the eigenvalue of the projectors $(\mathbb{1} \pm \gamma_5)/2$, while helicity is defined as the projection of the spin onto the momentum direction, *i.e.* as the left or right handedness. We give the exact definition of these two properties in Section 1. The entries of $n_{\rm hel}$ will be either $+1$ or -1. For each point in phase space and each helicity combination the Madgraph subroutine `MATRIX1` computes the matrix element using standard Helas routines.

- `IXXXXX`$(p, m, n_{\rm hel}, n_{\rm sf}, F)$ computes the wave function of a fermion with incoming fermion number, so either an incoming fermion or an outgoing anti–fermion. As input it requires the four-momentum, the mass and the helicity of this fermion. Moreover, $n_{\rm sf} = +1$ marks the incoming fermion u and $n_{\rm sf} = -1$ the outgoing anti–fermion μ^+, because by convention Madgraph defines its particles as u and μ^-.

  The fermion wave function output is a complex array $F(1:6)$. Its first two entries are the left–chiral part of the fermionic spinor, *i.e.* $F(1:2) = (\mathbb{1} - \gamma_5)/2\ u$ or $F(1:2) = (\mathbb{1} - \gamma_5)/2\ v$ for $n_{\rm sf} = \pm 1$. The entries $F(3:4)$ are the right–chiral spinor. These four numbers can directly be computed from the four-momentum if we know the helicity. The four entries correspond to the size of one γ matrix, so we can compute the trace of the chain of gamma matrices. Because for massless particles helicity and chirality are identical, our quarks and leptons will only have finite entries $F(1:2)$ for $n_{\rm hel} = -1$ and $F(3:4)$ for $n_{\rm hel} = +1$.

  The last two entries of F contain the four-momentum in the direction of the fermion flow, namely $F(5) = n_{\rm sf}(p(0) + ip(3))$ and $F(6) = n_{\rm sf}(p(1) + ip(2))$.

- `OXXXXX`$(p, m, n_{\rm hel}, n_{\rm sf}, F)$ does the same for a fermion with outgoing fermion flow, *i.e.* our incoming $\bar{u}$ and our outgoing μ^-. The left–chiral and right–chiral components now read $F(1:2) = \bar{u}(\mathbb{1} - \gamma_5)/2$ and $F(3:4) = \bar{u}(\mathbb{1} + \gamma_5)/2$, and similarly for the spinor $\bar{v}$. The last two entries are $F(5) = n_{\rm sf}(p(0) + ip(3))$ and $F(6) = n_{\rm sf}(p(1) + ip(2))$.

- `FFV1_3`$(F_i, F_o, g, m, \Gamma, J_{io})$ computes the (off–shell) current for the vector boson attached to the two external fermions F_i and F_o. The coupling $g(1:2)$ is a complex array with the interaction of the left–chiral and right–chiral fermion in the upper and lower index. For a general Breit–Wigner propagator we need to know the mass m and the width Γ of the intermediate vector boson. The output array J_{io} again has six components which for the photon with momentum q are

$$\begin{aligned} J_{io}(\mu+1) &= -\frac{i}{q^2}\ F_o^T\ \gamma^\mu \left(g(1)\ \frac{\mathbb{1} - \gamma_5}{2}\ + g(2)\ \frac{\mathbb{1} + \gamma_5}{2} \right)\ F_i \qquad\qquad \mu = 0,1,2,3 \\ J_{io}(5) &= -F_i(5) + F_o(5) \sim -p_i(0) + p_o(0) + i\,(-p_i(3) - p_o(3)) \\ J_{io}(6) &= -F_i(6) + F_o(6) \sim -p_i(1) + p_o(1) + i\,(-p_i(2) + p_o(2))\ . \end{aligned} \tag{2.2}$$

  The first four entries in J_{io} correspond to the index μ or the dimensionality of the Dirac matrices in this vector current. The spinor index is contracted between F_o^T and F_i.

  As two more arguments J_{io} includes the four-momentum flowing through the gauge boson propagator. They allow us to reconstruct q^μ from the last two entries

$$q^\mu = (\mathrm{Re}J_{io}(5), \mathrm{Re}J_{io}(6), \mathrm{Im}J_{io}(6), \mathrm{Im}J_{io}(5))\ . \tag{2.3}$$

- `FFV1_0`(F_i, F_o, J, g, V) computes the amplitude of a fermion–fermion–vector coupling using the two external fermionic spinors F_i and F_o and an incoming vector current J which in our case comes from `FFV1_3`. Again, the coupling $g(1:2)$ is a complex array, so we numerically compute

$$F_o^T\ \not{J} \left(g(1)\ \frac{\mathbb{1} - \gamma_5}{2}\ + g(2)\ \frac{\mathbb{1} + \gamma_5}{2} \right)\ F_i\ . \tag{2.4}$$

All spinor and Dirac indices of the three input arguments are contracted in the final result. Momentum conservation is not enforced by `FFV1_0`, so we have to take care of it by hand.

Given the list above it is easy to follow how Madgraph computes the amplitude for $u\bar{u} \to \gamma^* \to \mu^+\mu^-$. First, it calls wave functions for all external particles and puts them into the array $W(1:6, 1:4)$. The vectors $W(*,1)$ and $W(*,3)$ correspond to $F_i(u)$ and $F_i(\mu^+)$, while $W(*,2)$ and $W(*,4)$ mean $F_o(\bar{u}$ and $F_o(\mu^-)$.

The first vertex we evaluate is the incoming quark–photon vertex. Given the wave functions $F_i = W(*,1)$ and $F_o = W(*,2)$ `FFV1_3` computes the vector current for the massless photon in the s-channel. Not much changes if we instead choose a massive Z boson, except for the arguments m and Γ in the `FFV1_3` call. Its output is the photon current $J_{io} \equiv W(*,5)$.

The last step combines this current with the two outgoing muons coupling to the photon. Since this number gives the final amplitude, it should return a complex number, not an array. Madgraph calls `FFV1_0` with $F_i = W(*,3)$ and $F_o = W(*,4)$, combined with the photon current $J = W(*,5)$. The result `AMP` is copied into `JAMP` without an additional sign which could have come from the relative ordering of external fermions in different Feynman diagrams contributing to the same process.

The only remaining sum left to compute before we square `JAMP` is the color structure, which in our simple case means one color structure with a color factor $N_c = 3$.

As an added bonus Madgraph produces a file with all Feynman diagrams in which the numbering of the external particles corresponds to the second argument of W and the numbering of the Feynman diagrams corresponds to the argument of `AMP`. This helps us identify intermediate results W, each of which is only computed once, even if is appears several times in the different Feynman diagrams.

As mentioned above, to calculate the transition amplitude Madgraph requires all masses and couplings. They are transferred through common blocks in the file coupl.inc and computed elsewhere. In general, Madgraph uses unitary gauge for all vector bosons, because in the helicity amplitude approach it is easy to accommodate complicated tensors, in exchange for a large number of Feynman diagrams.

The function `MATRIX1` described above is not yet the full story. When we square $\mathcal{M}$ symbolically we need to sum over the spins of the outgoing states to transform a spinor product of the kind $u\bar{u}$ into the residue or numerator of a fermion propagator. To obtain the full transition amplitude numerically we correspondingly sum over all helicity combinations of the external fermions, in our case $2^4 = 16$ combinations.

```
      SUBROUTINE SMATRIX1(P,ANS)
C
C     Generated by Madgraph 5
C
C     Returns amplitude squared summed/avg over colors
C     and helicities for the point in phase space P(0:3,NEXTERNAL)
C
C     Process: u u~ > mu+ mu- / z
C
      INTEGER     NCOMB, NGRAPHS, NDIAGS, THEL
      PARAMETER (NCOMB=16, NGRAPHS=1, NDIAGS=1, THEL=2*NCOMB)

      REAL*8 P(0:3,NEXTERNAL)

      INTEGER I,J,IDEN
      INTEGER NHEL(NEXTERNAL,NCOMB),NTRY(2),ISHEL(2),JHEL(2)
      INTEGER JC(NEXTERNAL),NGOOD(2), IGOOD(NCOMB,2)
      REAL*8 T,MATRIX1
      LOGICAL GOODHEL(NCOMB,2)

      DATA NGOOD /0,0/
      DATA ISHEL/0,0/
      DATA GOODHEL/THEL*.FALSE./

      DATA (NHEL(I,   1),I=1,4) /-1,-1,-1,-1/
      DATA (NHEL(I,   2),I=1,4) /-1,-1,-1, 1/
      DATA (NHEL(I,   3),I=1,4) /-1,-1, 1,-1/
      DATA (NHEL(I,   4),I=1,4) /-1,-1, 1, 1/
      DATA (NHEL(I,   5),I=1,4) /-1, 1,-1,-1/
```

```
DATA (NHEL(I,   6),I=1,4) /-1, 1,-1, 1/
DATA (NHEL(I,   7),I=1,4) /-1, 1, 1,-1/
DATA (NHEL(I,   8),I=1,4) /-1, 1, 1, 1/
DATA (NHEL(I,   9),I=1,4) / 1,-1,-1,-1/
DATA (NHEL(I,  10),I=1,4) / 1,-1,-1, 1/
DATA (NHEL(I,  11),I=1,4) / 1,-1, 1,-1/
DATA (NHEL(I,  12),I=1,4) / 1,-1, 1, 1/
DATA (NHEL(I,  13),I=1,4) / 1, 1,-1,-1/
DATA (NHEL(I,  14),I=1,4) / 1, 1,-1, 1/
DATA (NHEL(I,  15),I=1,4) / 1, 1, 1,-1/
DATA (NHEL(I,  16),I=1,4) / 1, 1, 1, 1/
DATA IDEN/36/

DO I=1,NEXTERNAL
  JC(I) = +1
ENDDO

DO I=1,NCOMB
  IF (GOODHEL(I,IMIRROR) .OR. NTRY(IMIRROR).LE.MAXTRIES) THEN
    T = MATRIX1(P ,NHEL(1,I),JC(1))
    ANS = ANS+T
  ENDIF
ENDDO

ANS = ANS/DBLE(IDEN)
END
```

The important part of this subroutine is the list of possible helicity combinations stored in the array $n_{\text{hel}}(1:4, 1:16)$. Adding all different helicity combinations means a loop over the second argument and a call of `MATRIX1` with the respective helicity combination. Because of the naive helicity combinations many are not allowed the array `GOODHEL` keeps track of valid combinations. After an initialization to all 'false' this array is only switched to 'true' if `MATRIX1` returns a finite value, otherwise Madgraph does not waste time to compute the matrix element. At the very end, a complete spin–color averaging factor is included as `IDEN` and in our case given by $2 \times 2 \times N_c^2 = 36$.

Altogether, Madgraph provides us with the subroutine `SMATRIX1` and the function `MATRIX1` which together compute $\overline{|\mathcal{M}|^2}$ for each phase space point given as an external momentum configuration. This helicity method might not seem particularly appealing for a simple $(2 \to 2)$ process, but it makes it possible to compute processes with many particles in the final state and typically up to 10000 Feynman diagrams which we could never square symbolically, no matter how many graduate students' live times we throw in.

3 Electroweak theory

The fundamental structure describing the interactions between the elementary particles of the Standard Model is gauge interaction. We have already looked at QED, the interaction of leptons and quarks with a photon, leading to interactions between fermion currents. We need to have a closer look at this gauge symmetry structure, before we can describe for instance the weak interaction.

3.1 QED gauge invariance

Even though we already know how to use Feynman rule to compute QED scattering amplitudes, and we even have an idea how these Feynman rules are related to the underlying quantum field theory, let us have another look at the QED Lagrangian. We already know that the Lagrangian in Eq.(1.6) describes the quantum version of the photons from electrodynamics,

$$\mathscr{L}_{\text{photon}} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} \qquad \text{with} \qquad F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \; , \tag{3.1}$$

The field strength is already symmetric under the local gauge transformation with a space-time-dependent parameter $\alpha(x)$,

$$\begin{aligned} A_\mu \to A_\mu - \frac{1}{e}\partial_\mu \alpha \qquad \Rightarrow \qquad F_{\mu\nu} &\to \partial_\mu \left(A_\nu - \frac{1}{e}\partial_\nu \alpha \right) - \partial_\nu \left(A_\mu - \frac{1}{e}\partial_\mu \alpha \right) \\ &= F_{\mu\nu} - \frac{1}{e} \left(\partial_\mu \partial_\nu - \partial_\nu \partial_\mu \right) \alpha = F_{\mu\nu} \; . \end{aligned} \tag{3.2}$$

We know this shift symmetry from electrodynamics, but it is not clear what α really means. This becomes clear when we introduce a generic fermion spinor ψ as in Eq.(1.24),

$$\begin{aligned} \mathscr{L}_{\text{fermion-photon}} &= \overline{\psi} \left(i \not{D} - m \mathbb{1} \right) \psi \\ &\equiv \overline{\psi} \left(i(\not{\partial} + ieq \not{A}) - m \mathbb{1} \right) \psi \; , \end{aligned} \tag{3.3}$$

where q is the fermion charge in units of the electron charge. We can rotate the fermion field by the angle $\alpha(x)$,

$$\begin{aligned} \psi &\to e^{iq\alpha} \psi \\ \overline{\psi} &\to e^{-iq\alpha} \overline{\psi} \; , \end{aligned} \tag{3.4}$$

such that the fermion mass term is symmetric in itself, which means QED has no problem with massive fermions. The kinetic term with the covariant derivative transforms into

$$\begin{aligned} i\overline{\psi} \not{D} \psi \equiv i\overline{\psi} \gamma^\mu (\partial_\mu + ieqA_\mu) \psi &\to i\overline{\psi} e^{-iq\alpha} \gamma^\mu \left(\partial_\mu + ieq \left(A_\mu - \frac{1}{e}\partial_\mu \alpha \right) \right) e^{iq\alpha} \psi \\ &= i\overline{\psi} \gamma^\mu e^{-iq\alpha} \left(\partial_\mu e^{iq\alpha} + e^{iq\alpha} ieq \left(A_\mu - \frac{1}{e}\partial_\mu \alpha \right) \right) \psi \\ &= i\overline{\psi} \gamma^\mu e^{-iq\alpha} \left(e^{iq\alpha} iq \partial_\mu \alpha + e^{iq\alpha} \partial_\mu + e^{iq\alpha} ieq \left(A_\mu - \frac{1}{e}\partial_\mu \alpha \right) \right) \psi \\ &= i\overline{\psi} \gamma^\mu \left(iq\partial_\mu \alpha + \partial_\mu + ieqA_\mu - iq\partial_\mu \alpha \right) \psi \\ &= i\overline{\psi} \gamma^\mu \left(\partial_\mu + ieqA_\mu \right) \psi \\ &\equiv i\overline{\psi} \not{D} \psi \; . \end{aligned} \tag{3.5}$$

The combined QED Lagrangian is invariant under a local rotation of the fermion field(s), if we also shift the photon field the way we know it from electrodynamics. Already there, this shift was referred to as a gauge transformation. In

combination, we say that the QED interaction between photons and fermions is defined by a local $U(1)$ gauge transformation of both fields.

To understand why the covariant derivative is useful, we look at the first and last lines of Eq.(3.5). We immediately see that the gauge transformation of the covariant derivative is

$$\not{D} \to e^{iq\alpha} \not{D} e^{-iq\alpha} \; . \tag{3.6}$$

We can also replace the definition of the field strength in terms of the gauge field by a definition in terms of the covariant derivative, for example acting on a test function $f(x)$ with $[A, f] = 0$,

$$\begin{aligned} F_{\mu\nu} \, f &= \frac{1}{ieq} \, [D_\mu, D_\nu] \, f \\ &= \frac{1}{ieq} \, (\partial_\mu + ieqA_\mu)(\partial_\nu + ieqA_\nu) \, f - \frac{1}{ieq} \, (\partial_\nu + ieqA_\nu)(\partial_\mu + ieqA_\mu) \, f \\ &= (\partial_\mu A_\nu) f + A_\nu (\partial_\mu f) + A_\mu (\partial_\nu f) - (\partial_\nu A_\mu) f - A_\mu (\partial_\nu f) - A_\nu (\partial_\mu f) \\ &= ((\partial_\mu A_\nu) f - (\partial_\nu A_\mu)) \, f \; . \end{aligned} \tag{3.7}$$

In this form the partial derivative acts only on the gauge field, so unlike the first line the definition in the last line is not an operator equation. In this derivation we assume that the gauge field commutes, which we call abelian,

$$[A_\mu, A_\nu] = 0 \; . \tag{3.8}$$

3.2 Massive gauge bosons

One of the shortcomings of QED is that it only defines long-range interactions. From the work of Hideki Yukawa in 1935 we know that the mass of the exchange particles changes the form of the interaction potential in Fourier space,

$$\begin{aligned} V(r) &= -\frac{e^2}{r} && \text{massless particle exchange} \\ V(r) &= -g^2 \, \frac{e^{-mr}}{r} && \text{massive particle exchange with } m. \end{aligned} \tag{3.9}$$

Yukawa did not actually talk about the weak nuclear force at the quark level. His model was based on fundamental protons and neutrons, and his exchange particles were pions. But his argument applies perfectly to the electroweak Fermi interaction between quarks. This leads to the challenge of formally including a photon mass in the gauge-invariant QED Lagrangian.

The first step towards defining a massive version of QED is to include a photon mass in the kinematic Lagrangian of Eq.(3.1). We immediately see that just adding a photon mass term

$$\frac{1}{2} m^2 A^2 \to \frac{1}{2} m^2 \left(A_\mu - \frac{1}{e} \partial_\mu \alpha \right)^2 \tag{3.10}$$

is not allowed by the gauge symmetry. The key idea is to add an innocent looking real scalar field without a mass and without a coupling to the photon, but with a scalar–photon mixing term and a well-chosen gauge transformation. The result is called the Boulware–Gilbert model or Stückelberg mass generation,

$$\begin{aligned} \mathscr{L} &= -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} + \frac{1}{2} e^2 f^2 A_\mu^2 + \frac{1}{2} \, (\partial_\mu \phi)^2 - ef A_\mu \partial^\mu \phi \\ &= -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} + \frac{1}{2} e^2 f^2 \left(A_\mu - \frac{1}{ef} \partial_\mu \phi \right)^2 \; , \end{aligned} \tag{3.11}$$

where f is a common mass scale for the photon mass and the mixing. It ensures that all terms in the Lagrangian have mass dimension four — remembering that bosonic fields like A_μ and ϕ have mass dimension one. If we define the

massive photon field as

$$\tilde{A}_\mu = A_\mu - \frac{1}{ef}\partial_\mu\phi \tag{3.12}$$

the field strength does not change,

$$\begin{aligned} F_{\mu\nu}\Big|_{\tilde{A}} &= \partial_\mu\tilde{A}_\nu - \partial_\nu\tilde{A}_\mu = \partial_\mu\left(A_\nu - \frac{1}{ef}\partial_\nu\phi\right) - \partial_\nu\left(A_\mu - \frac{1}{ef}\partial_\mu\phi\right) \\ &= \partial_\mu A_\nu - \partial_\nu A_\mu = F_{\mu\nu}\Big|_A \,, \end{aligned} \tag{3.13}$$

and we can write the Lagrangian of Eq.(3.11) as

$$\begin{aligned} \mathscr{L} &= -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \frac{1}{2}e^2f^2\tilde{A}_\mu^2 \\ &= -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \frac{1}{2}m_A^2\tilde{A}_\mu^2 \qquad \text{with} \qquad m_A = ef \,. \end{aligned} \tag{3.14}$$

For a gauge invariant theory, a suitable gauge transformation of the scalar field has to cancel the contribution of the explicit mass term in Eq.(3.11). The simple choice

$$\phi \longrightarrow \phi - f\alpha \,. \tag{3.15}$$

indeed gives us

$$A_\mu - \frac{1}{ef}\partial_\mu\phi \to A_\mu - \frac{1}{e}\partial\alpha - \frac{1}{ef}\partial_\mu\phi + \frac{1}{ef}\partial_\mu(f\alpha) = A_\mu - \frac{1}{ef}\partial_\mu\phi \tag{3.16}$$

This Lagrangian describes a massive photon field $\tilde{A}_\mu$, which has absorbed the real scalar ϕ as its additional longitudinal component. This is because a massless gauge boson A_μ has only two on–shell degrees of freedom, the left-handed and right-handed polarizations, while the massive $\tilde{A}_\mu$ has an additional longitudinal polarization. To describe it, the massive photon $\tilde{A}$ has 'eaten' the real scalar field ϕ.

What kind of properties does this field ϕ need to have, so that we can use it to provide a photon mass? From the gauge transformation we immediately see that any additional purely scalar term in the Lagrangian, like a scalar potential $V(\phi)$, needs to be symmetric under the shift $\phi \to \phi - f\alpha$, so it does not spoil gauge invariance. This means that we cannot write down polynomial terms ϕ^n, like a mass, a self coupling, or an interaction term ϕAA. Only derivative interactions proportional to $\partial\phi$ attached to gauge-invariant currents are allowed. For them, we the shift by α turns into a total derivative in the Lagrangian.

This example illustrates a few vital properties of Nambu–Goldstone bosons (NGB). Such massless physical states appear in many areas of physics and are described by Goldstone's theorem. It applies to global continuous symmetries of the Lagrangian which are violated by a non–symmetric vacuum state, a mechanism called spontaneous symmetry breaking. Based on Lorentz invariance and states with a positively definite norm we can then prove: *If a global symmetry group is spontaneously broken into a group of lower rank, its broken generators correspond to physical Goldstone modes. These scalar fields transform non–linearly under the larger and linearly under the smaller group. This way they are massless and cannot form a potential, because the non–linear transformation only allows derivative terms in the Lagrangian.*

For our massive QED case we are breaking the $U(1)$ gauge symmetry, which naively introduces a massless scalar degree of freedom ϕ. Following Eq.(3.12) it provides the missing longitudinal polarization for the massive photon. This combines two problems into one solution — we can break the gauge symmetry without creating unobserved massless particles, and our massive photon gets an additional degree of freedom. We will use the same trick for the Higgs later.

3.3 Fermion doublets

The structural element of the Fermi theory and also of the Standard Model is the $SU(2)$ doublet structure of paired fermions, like protons and neutrons. If a common gauge transformation should link the two double components, an obvious choice is to replace the local $U(1)$ gauge invariance by a local $SU(2)$ gauge invariance. We remind ourselves that an $SU(2)$ transformation is very similar to an $O(3)$ rotation, and the generators of the $SU(2)$ transformation are the Pauli matrices. This means that a representation of $SU(2)$ transformations is given by

$$U = e^{i\alpha_a \tau_a/2} \qquad \text{with} \qquad \tau_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \tau_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \qquad \tau_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{3.17}$$

They satisfy the relation

$$\tau_a \tau_b = \delta_{ab} + i\epsilon_{abc}\tau_c \qquad \Leftrightarrow \qquad [\tau_a, \tau_b] = 2i\epsilon_{abc}\tau_c \; . \tag{3.18}$$

This means that we can write a commutation of two objects with $SU(2)$ indices as an object with one $SU(2)$ index. For later use we also need a sum rule for the Pauli matrices $\tau_{1,2,3}$,

$$\sum_{a,b} \tau_a \tau_b = \sum_{a,b} (\delta_{ab} + i\epsilon_{abc}\tau_c) = \sum \delta_{ab} + i \sum_{a \neq b} \epsilon_{abc}\tau_c = \sum \delta_{ab} + i \sum_{a<b} (\epsilon_{abc} + \epsilon_{bac})\, \tau_c = \sum \delta_{ab} \; . \tag{3.19}$$

The basis of three Pauli matrices we can write in terms of $\tau_{1,2,3}$ or in terms of $\tau_{+,-,3}$ with

$$\tau_+ = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \qquad \tau_- = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \tag{3.20}$$

From the QED Lagrangian we know how to describe fermion masses and interactions with gauge bosons in a dimension-4 Lagrangian. Before we move on, let us briefly look at the dimensionality of the Lagrangian, so we can use it to structure the weak Lagrangian.

The dimensionality is crucial for the renormalizability or fundamental nature of a Lagrangian. The definition of the transition in terms of the interaction Hamiltonian in Eq.(1.29),

$$\mathcal{S} \sim \exp\left[-i \int dt \mathcal{H}(t)\right] = \exp\left[-i \int d^4x \, \left(\pi\dot{\phi} - \mathscr{L}\right)\right] \; , \tag{3.21}$$

suggests that the Hamiltonian should have the unit energy or momentum, and the Lagrangian should have mass dimension four. The elements of the Lagrangian have the mass dimensions.

$$\begin{aligned} &\text{scalar field} && [\phi] = M \\ &\text{explicit mass} && [m] = M \\ &\text{gauge field} && [A_\mu] = M \\ &\text{space derivative} && [\partial_\mu] = M \\ &\text{field strength} && [F_{\mu\nu}] = M^2 \\ &\text{fermion spinor} && [\psi] = M^{3/2} \; . \end{aligned} \tag{3.22}$$

If we want our theory to be renormalizable and valid to arbitrarily large scales, the Lagrangian cannot include inverse masses, because in that case a momentum or energy in the numerator would eventually lead to an exploding ratio in the ultraviolet. This means all terms in a fundamental Lagragian should have mass dimension two to four, complemented by explicit masses. This is the way we will organize the weak Lagrangian.

As a starting point, the interaction of fermions, our case quarks, with gauge bosons is most easily written in terms of covariant derivatives. Just like the kinetic term for the gauge bosons, they have mass dimension four,

$$\mathscr{L}_{D4} = \overline{Q}_L i \not{D} Q_L + \overline{Q}_R i \not{D} Q_R - \frac{1}{4} F_{a,\mu\nu} F_a^{\mu\nu} ... \tag{3.23}$$

From Eq.(3.7) we know that the covariant derivatives can be used to describe the field strengths. Howewer, going from the abelian $U(1)$ transformation to the non-abelian $SU(2)$ transformation the condition $[A_\mu, A_\nu] = 0$ is not true anymore. The good news is that the definition of the field strength in terms of the covariant derivative still holds for the non-ablelian case,

$$\begin{aligned} F_{a,\mu\nu} &\equiv \frac{1}{ieq}[D_\mu, D_\nu]_a \\ &= \partial_\mu A_{a,\nu} - \partial_\nu A_{a,\mu} + ieq[A_\mu, A_\nu]_a \ , \end{aligned} \tag{3.24}$$

now with the $SU(2)$-index a. The gauge invariance of the field strength follows conveniently from Eq.(3.7).

3.4 Weak gauge bosons

The main theme of the electroweak theory is that it combines our known QED with the $U(1)$ gauge transformation of the physical photon and the fermion singlet with the $SU(2)$ gauge transformations of the gauge bosons and fermion doublets. In this combination there appears a mixing which we can describe in two different ways:

1. The neutral gauge bosons can mix, so the observed mass eigenstates are the photon A_μ and the massive Z_μ, but they are related to the interaction eigenstate W_μ^3 and a massless B_μ. We can think of the B_μ as the photon of a proto-QED before we combine QED with the weak symmetry to the actual QED. This description leads to the weak mixing angle.

2. The $U(1)$ rotation of the fermion fields and the neutral $SU(2)$ rotation via τ_3 can be combined to individual $U(1)$ rotations of the left-handed and right-handed fermion spinors. This desciption leads to the Gell-Mann–Nishijima formula.

Starting with the mixing from interaction eigenstates to mass eigenstates, we assume that the two ingredients to the neutral electroweak interactions are a massless proto-photon B_μ and the neutral electroweak gauge boson field W_μ^3. Both are neutral particles with the same quantum numbers, so they can mix to the mass eigenstates A_μ and Z_μ,

$$\begin{pmatrix} A_\mu \\ Z_\mu \end{pmatrix} = \begin{pmatrix} c_w & s_w \\ -s_w & c_w \end{pmatrix} \begin{pmatrix} B_\mu \\ W_\mu^3 \end{pmatrix} \qquad \text{with} \qquad s_w \equiv \sin\theta_w \quad c_w \equiv \cos\theta_w \ . \tag{3.25}$$

The photon describes the $U(1)$ charge transformation and couples to electric charge. The field B_μ couples to the so-called $U(1)$ hypercharge y. The angle θ_w is the weak mixing angle or Weinberg angle. Both B_μ and A_μ are massless, because the determinant of the mass matrix keeps its zero eigenvalue after rotations.

Unlike QED, the weak interaction knows about the chirality of the fermion fields, so we have to distinguish $\not{D}_L$ and $\not{D}_R$. In the interaction basis the covariant derivatives include the massless B_μ and the three massive $W_{a,\mu}$,

$$\begin{aligned} D_{L\mu} =& \partial_\mu + ig'\frac{y}{2}B_\mu + ig\sum_{a=1,2,3} W_{a,\mu}\frac{\tau_a}{2} \\ D_{R\mu} =& D_{L\mu}\Big|_{\tau_{1,2,3}=0} \ . \end{aligned} \tag{3.26}$$

This definition implies that our $SU(2)$ gauge transformation only act on the left-handed doublets, so we refer to it as $SU(2)_L$. The right-handed fields can be written as ntuples, but they do not have a doublet structure under $SU(2)$. The effect of this is structure is that the massive charged W-boson only couples doublet like (u_L, d_L).

In the mass basis for the neutral states the covariant derivative from Eq.(3.26) has to read

$$D_{L\mu} = \partial_\mu + ieqA_\mu + ig_Z\left(-qs_w^2\mathbb{1} + \frac{\tau_3}{2}\right)Z_\mu + ig\left(\frac{\tau_1}{2}W_\mu^1 + \frac{\tau_2}{2}W_\mu^2\right) \tag{3.27}$$

We omit the relations between the couplings e and g_Z to g' and g and the weak mixing angle. The relation of y to q and the τ_3 eigenvalues will be discussed later.

The mass basis of the charged weak bosons is most conveniently written in terms of $\tau_\pm$. To switch bases we only have to make sure we keep the standard normalization of all fields,

$$\begin{aligned}
\sqrt{2}\left(\tau_+ W_\mu^+ + \tau_- W_\mu^-\right) &= \sqrt{2}\begin{pmatrix} 0 & W_\mu^+ \\ 0 & 0 \end{pmatrix} + \sqrt{2}\begin{pmatrix} 0 & 0 \\ W_\mu^- & 0 \end{pmatrix} \\
&\stackrel{!}{=} \tau_1 W_\mu^1 + \tau_2 W_\mu^2 = \begin{pmatrix} 0 & W_\mu^1 \\ W_\mu^1 & 0 \end{pmatrix} + \begin{pmatrix} 0 & -iW_\mu^2 \\ iW_\mu^2 & 0 \end{pmatrix} \\
\Leftrightarrow \qquad & W_\mu^+ = \frac{1}{\sqrt{2}}\left(W_\mu^1 - iW_\mu^2\right) \qquad\qquad W_\mu^- = \frac{1}{\sqrt{2}}\left(W_\mu^1 + iW_\mu^2\right) \\
\Rightarrow \qquad & D_{L\mu} = \partial_\mu + ieqA_\mu + ig_Z\left(-qs_w^2 + \frac{\tau_3}{2}\right)Z_\mu + i\frac{g}{\sqrt{2}}\left(\tau_+ W_\mu^+ + \tau_- W_\mu^-\right)
\end{aligned} \tag{3.28}$$

To check this mass basis, we look at the masses of the gauge bosons, which appear as dimension-2 mass terms in the electroweak Lagrangian. Using the above relation, we write them in terms of the charged W-fields,

$$\mathscr{L}_{D2} = \frac{m_W^2}{2}\left(W^{1,\mu}W_\mu^1 + W^{2,\mu}W_\mu^2\right) + \frac{m_Z^2}{2}Z^\mu Z_\mu = m_W^2 W^{+,\mu}W_\mu^- + \frac{m_Z^2}{2}Z^\mu Z_\mu \; . \tag{3.29}$$

Now both mass terms are proportional to the field combination $\phi^*\phi = |\phi|^2$, as we know it from scalars. The relative factor two in front of the W mass appears because the Z field is neutral and the W field is charged, again the same as for neutral and charged scalars.

Finally, we look at the Lagrangian terms describing the fermion masses with mass dimension three. From Eq.(1.50) we know that it requires a combination of the left-handed doublet Q_L and the right-handed singlet fields Q_R,

$$\mathscr{L}_{D3} = -\overline{Q}_L m_Q Q_R + ... \tag{3.30}$$

This form will require a doublet structure of the Higgs–Goldstone fields, which we will discuss next term. For now we ignore this complication. Moreover, these mass terms can be matrices in generation space, which implies that we might have to rotate the fermion fields from an interaction basis into the mass basis, where these mass matrices are diagonal. Flavor physics dealing with such 3×3 mass matrices is its own field of physics. We will also omit this complication for now.

The problem with the mass term in Eq.(3.30) is that they are not invariant under the $SU(2)_L$ gauge transformation $U(x)$, which only transforms the left-handed fermion fields

$$Q_L \stackrel{U}{\to} UQ_L \qquad\qquad Q_R \stackrel{U}{\to} Q_R \; . \tag{3.31}$$

Obviously, there is no way we can make left–right mixing fermion mass terms in Eq.(3.30) invariant under a left-handed $SU(2)_L$ gauge transformation,

$$\overline{Q}_L m_Q Q_R \stackrel{U}{\to} \overline{Q}_L U^{-1} m_Q Q_R \neq \overline{Q}_L m_Q Q_R \; . \tag{3.32}$$

To see what we need to add to make fermion masses consistent with the electroweak gauge symmetry, we need to combine the local $U(1)$ transformations with the neutral component of the $SU(2)_L$ transformation

$$V = e^{i\beta\tau_3/2} \; . \tag{3.33}$$

From our QED calculation we know that the gauge transformation of the fermion fields is related to the form of the covariant derivative. Equation (3.28) then tell us how to combine V with the $U(1)$ charge transformation to a $U(1)$

hypercharge transformation. This means we need to evaluate

$$\begin{aligned}
\exp(i\beta q)\, V^\dagger &= \exp(i\beta q)\, \exp\left(-\frac{i}{2}\beta\tau_3\right) && \text{with} \quad V = U(x)\Big|_{\tau_3} = \exp\left(\frac{i}{2}\beta\tau_3\right) \\
&= \exp\left(i\beta\frac{y\mathbb{1}+\tau_3}{2}\right)\, \exp\left(-\frac{i}{2}\beta\tau_3\right) && \text{with} \quad q\mathbb{1} \equiv \frac{y\mathbb{1}+\tau_3}{2} \qquad y_Q = \frac{1}{3} \quad y_L = -1 \\
&= \exp\left(i\frac{\beta}{2}y\mathbb{1}\right)\ . && (3.34)
\end{aligned}$$

The relation between the charge q, the hypercharge y, and the isospin τ_3 is the same as in Eq.(3.27), called the Gell-Mann–Nishijima formula. The indices Q and L denote quark and lepton doublets. In this notation we do not distinguish $U(1)$ rotations by a real angle and $SU(2)$ rotations proportional to a unit matrix. This is reflected in the Gell-Mann–Nishijima formula, where τ_3 has to be replaced by its eigenvalue ± 1 for up–type and down–type fermions to relate charge and hypercharge.

In analogy to Eq.(3.31) left-handed and right-handed fermion spinors transform under the combination of V and the two $U(1)$, now denoted as V, as

$$\begin{aligned}
Q_L &\xrightarrow{V} \exp\left(i\beta q_Q\right) V^\dagger Q_L = \exp\left(i\frac{\beta}{2}y_Q\mathbb{1}\right) Q_L \\
Q_R &\xrightarrow{V} \exp\left(i\beta q_Q\right) Q_R\ . && (3.35)
\end{aligned}$$

The right-handed fermions only see the electric charge.

3.5 Sigma model

One way of solving this problem with weak gauge invariance of fermion mass terms is to introduce an additional field $\Sigma(x)$. This field will a similar role as the real scalar field we used for the photon mass generation. Its physical properties will become clear piece by piece from the way it appears in the Lagrangian and from the required gauge invariance.

First, we introduce Σ into the fermion mass term. This will tell us what it takes to make this mass term gauge invariant under the weak transformations defined in Eqs.(3.31) and (3.35)

$$\begin{aligned}
&\overline{Q}_L \Sigma m_Q Q_R \xrightarrow{U} \overline{Q}_L U^{-1}\Sigma^{(U)} m_Q Q_R \stackrel{!}{=} \overline{Q}_L \Sigma m_Q Q_R \\
&\quad \Leftrightarrow \qquad \Sigma \to \Sigma^{(U)} = U\Sigma\ . && (3.36)
\end{aligned}$$

and

$$\begin{aligned}
\overline{Q}_L \Sigma m_Q Q_R &\xrightarrow{V} \overline{Q}_L \exp\left(-i\frac{\beta}{2}y\mathbb{1}\right) \Sigma^{(V)} m_Q \exp\left(i\beta q\right) Q_R \\
&= \overline{Q}_L \Sigma^{(V)} \exp\left(-i\frac{\beta}{2}y\mathbb{1}\right)\, \exp\left(i\beta q\right) m_Q Q_R \qquad\qquad \exp\left(i\frac{\beta}{2}y\mathbb{1}\right) \text{ always commuting} \\
&= \overline{Q}_L \Sigma^{(V)} V m_Q Q_R \\
&\stackrel{!}{=} \overline{Q}_L \Sigma\, m_Q Q_R \qquad \Leftrightarrow \qquad \Sigma \to \Sigma^{(V)} = \Sigma V^\dagger\ . && (3.37)
\end{aligned}$$

Combining both gives us the needed transformation property

$$\Sigma \to U\Sigma V^\dagger\ . \tag{3.38}$$

We see that Σ is a 2×2 matrix with mass dimension zero. The fermion mass Lagrangian is gauge invariant without specifying anything about the relation of Σ with propagating or physical fields

$$\mathscr{L}_{D3} = -\overline{Q}_L \Sigma m_Q Q_R - \overline{L}_L \Sigma m_L L_R + \text{h.c.} + ... \tag{3.39}$$

In a second step, we use Σ to introduce gauge boson masses. We recall the covariant derivative from Eq.(3.26),

$$D_{L\mu} = \partial_\mu + ig'\frac{y}{2}B_\mu + igW_{a,\mu}\frac{\tau_a}{2} \; . \tag{3.40}$$

We first choose the form of the covariant derivative acting on Σ,

$$\begin{aligned} D_\mu\Sigma &= \partial_\mu\Sigma + ig'\Sigma B_\mu \frac{y}{2}\Big|_{q=0} + igW_{a,\mu}\frac{\tau_a}{2}\Sigma \\ &= \partial_\mu\Sigma - ig'\Sigma B_\mu\frac{\tau_3}{2} + igW_{a,\mu}\frac{\tau_a}{2}\Sigma \; , \end{aligned} \tag{3.41}$$

With the abbreviations

$$V_\mu \equiv \Sigma(D_\mu\Sigma)^\dagger \qquad \text{and} \qquad T \equiv \Sigma\tau_3\Sigma^\dagger \; , \tag{3.42}$$

we will show in Sec. 3.6 that we can write the gauge boson mass Lagrangian as

$$\mathscr{L}_{D2} = -\frac{v^2}{4}\;\mathrm{Tr}[V_\mu V^\mu] - \Delta\rho\frac{v^2}{8}\;\mathrm{Tr}[TV_\mu]\;\mathrm{Tr}[TV^\mu] \; . \tag{3.43}$$

The trace acts on the 2×2 $SU(2)$ matrices. The parameter $\Delta\rho$ is conventional and will be the focus of Section 3.8.

Before we compute the weak boson masses, we see which gauge invariant terms of mass dimension four we can write down using Σ. Our first attempt for a building block

$$\Sigma^\dagger\Sigma \overset{U,V}{\to} (U\Sigma V^\dagger)^\dagger(U\Sigma V^\dagger) = V\Sigma^\dagger U^\dagger U\Sigma V^\dagger = V\Sigma^\dagger\Sigma V^\dagger \neq \Sigma^\dagger\Sigma \tag{3.44}$$

is forbidden by invariance under Eq.(3.38). However, a circular trace $\mathrm{Tr}(\Sigma^\dagger\Sigma) \to \mathrm{Tr}(V\Sigma^\dagger\Sigma V^\dagger) = \mathrm{Tr}(\Sigma^\dagger\Sigma)$ allows for the additional potential terms, meaning terms with no derivatives

$$\mathscr{L}_\Sigma = -\frac{\mu^2 v^2}{4}\,\mathrm{Tr}(\Sigma^\dagger\Sigma) - \frac{\lambda v^4}{16}\left(\mathrm{Tr}(\Sigma^\dagger\Sigma)\right)^2 + \cdots \; , \tag{3.45}$$

with properly chosen prefactors μ, v, λ. This finalizes our construction of the weak Lagrangian organized by mass dimension,

$$\mathscr{L} = \mathscr{L}_{D2} + \mathscr{L}_{D3} + \mathscr{L}_{D4} + \mathscr{L}_\Sigma \; . \tag{3.46}$$

3.6 Weak boson masses

To check that Eq.(3.43) gives the correct masses in the Standard Model we assume that Σ acquires a vacuum expectation value. The simplest way to achieve this and obtain the correct fermion masses is to just write

$$\Sigma(x) = \mathbb{1} \; . \tag{3.47}$$

This choice is called unitary gauge. It looks like a dirty trick to first introduce $\Sigma(x) = \mathbb{1}$ and then use this field for a gauge invariant implementation of gauge boson masses. Clearly, a constant does not exhibit the correct transformation property under the U and V symmetries, but we can always work in a specific gauge and only later check the physical predictions for gauge invariance.

We now check $\mathscr{L}_{D2}$ as written in Eq.(3.43) for the correct gauge boson masses. Using the covariant derivative from Eq.(3.41) acting on a now constant field we can compute V_μ in unitary gauge

$$\begin{aligned} V_\mu = \Sigma(D_\mu\Sigma)^\dagger &= \mathbb{1}(D_\mu\Sigma)^\dagger \\ &= -igW_\mu^+\frac{\tau_+}{\sqrt{2}} - igW_\mu^-\frac{\tau_-}{\sqrt{2}} - igW_\mu^3\frac{\tau_3}{2} + ig'B_\mu\frac{\tau_3}{2} \\ &= -i\frac{g}{\sqrt{2}}\left(W_\mu^+\tau_+ + W_\mu^-\tau_-\right) - ig_Z Z_\mu\frac{\tau_3}{2} \, , \end{aligned} \tag{3.48}$$

with $Z_\mu = c_w W^3_\mu - s_w B_\mu$ and the two coupling constants

$$g_Z = \frac{g}{c_w} \qquad \text{and} \qquad g' = \frac{g s_w}{c_w} \; . \tag{3.49}$$

This gives us the first of the two terms in $\mathscr{L}_{D2}$ using $\tau_\pm^2 = 0$ and $\text{Tr}(\tau_3 \tau_\pm) = 0$,

$$\begin{aligned} \text{Tr}[V_\mu V^\mu] &= -2\,\frac{g^2}{2} W^+_\mu W^{-\mu}\,\text{Tr}(\tau_+\tau_-) - \frac{g_Z^2}{4} Z_\mu Z^\mu\,\text{Tr}(\tau_3^2) \\ &= -g^2 W^+_\mu W^{-\mu} - \frac{g_Z^2}{2} Z_\mu Z^\mu \; , \end{aligned} \tag{3.50}$$

In the second step we use $\text{Tr}(\tau_\pm \tau_\mp) = 1$, and $\text{Tr}(\tau_3^2) = \text{Tr}\,\mathbb{1} = 2$. The mass term proportional to $\Delta\rho$ also simplifies in unitary gauge

$$\begin{aligned} & T = \Sigma \tau_3 \Sigma^\dagger = \tau_3 \\ \Rightarrow \qquad & \text{Tr}(T V_\mu) = \text{Tr}\left(-i g_Z Z_\mu \frac{\tau_3^2}{2} \right) = -i g_Z Z_\mu \\ \Rightarrow \quad & \text{Tr}(T V_\mu)\;\text{Tr}(T V^\mu) = -g_Z^2 Z_\mu Z^\mu \; . \end{aligned} \tag{3.51}$$

Combining both terms with the prefactor in Eq.(3.43) yields the complete gauge boson mass term

$$\begin{aligned} \mathscr{L}_{D2} &= -\frac{v^2}{4}\left(-g^2 W^+_\mu W^{-\mu} - \frac{g_Z^2}{2} Z_\mu Z^\mu \right) - \Delta\rho \frac{v^2}{8}\left(-g_Z^2 Z_\mu Z^\mu \right) \\ &= \frac{v^2 g^2}{4} W^+_\mu W^{-\mu} + \frac{v^2 g_Z^2}{8}\,(1+\Delta\rho)\, Z_\mu Z^\mu \; . \end{aligned} \tag{3.52}$$

Identifying the masses with the form given in Eq.(3.29) and assuming universality of neutral and charged current interactions ($\Delta\rho = 0$) we find

$$\begin{aligned} m_W &= \frac{gv}{2} \\ m_Z &= \sqrt{1+\Delta\rho}\,\frac{g_Z v}{2} \overset{\Delta\rho=0}{=} \frac{g_Z v}{2} = \frac{gv}{2c_w} \; . \end{aligned} \tag{3.53}$$

A possible additional and unwanted Z-mass contribution $\Delta\rho$ will come back in Sec. 3.8. From the known gauge boson masses ($m_W \sim 80$ GeV) and weak coupling ($g \sim 0.7$) we find $v \sim 246$ GeV.

3.7 Weak boson propagators

Finally, let us at least mention different gauge choices and the appearance of Goldstone modes. If we break the full electroweak gauge symmetry $SU(2)_L \times U(1)_Y \to U(1)_Q$ we expect three Goldstone bosons which become part of the weak gauge bosons and promote those from massless gauge bosons (with two degrees of freedom each) to massive gauge bosons (with three degrees of freedom each). This is the point of view of the unitary gauge, in which we never see Goldstone modes.

In the general renormalizable R_ξ gauge we can actually see the Goldstone modes in the gauge boson propagators

$$\begin{aligned} \Delta^{\mu\nu}_{VV}(p) &= \frac{-i}{p^2 - m_V^2 + i\epsilon}\left[g^{\mu\nu} + (\xi-1)\frac{p^\mu p^\nu}{p^2 - \xi m_V^2} \right] \\ &= \begin{cases} \dfrac{-i}{p^2 - m_V^2 + i\epsilon}\left[g^{\mu\nu} - \dfrac{p^\mu p^\nu}{m_V^2} \right] & \text{unitary gauge } \xi \to \infty \\ \dfrac{-i}{p^2 - m_V^2 + i\epsilon}\, g^{\mu\nu} & \text{Feynman gauge } \xi = 1 \\ \dfrac{-i}{p^2 - m_V^2 + i\epsilon}\left[g^{\mu\nu} - \dfrac{p^\mu p^\nu}{p^2} \right] & \text{Landau gauge } \xi = 0 \; . \end{cases} \end{aligned} \tag{3.54}$$

If these gauge choices are physically equivalent, something has to compensate for the fact that in Feynman gauge the whole Goldstone term vanishes and the polarization sum looks like a massless gauge boson, while in unitary gauge we can see the effect of these modes. This is done by the Goldstone propagator

$$\Delta_{VV}(p^2) = \frac{-i}{p^2 - \xi m_V^2 + i\epsilon} \,, \tag{3.55}$$

The Goldstone mass $\sqrt{\xi} m_V$ depends on the gauge: in unitary gauge the infinitely heavy Goldstones do not propagate ($\Delta_{VV}(p^2) \to 0$), while in Feynman gauge and in Landau gauge we have to include them as particles. From this form we can guess that they will indeed cancel the second term of the gauge boson propagators.

These different gauges have different Feynman rules and Green's functions, even a different particle content. For a given problem one or the other might be the most efficient to use in computations or proofs. For example, the proof of renormalizability was first formulated in unitary gauge. Loop calculations might be most efficient in Feynman gauge, because of the simplified propagator structure, while many QCD processes benefit from an explicit projection on the physical external gluons. Tree level helicity amplitudes are usually computed in unitary gauge, etc...

3.8 Custodial symmetry

Analyzing the appearance of $\Delta\rho$ in Eq.(3.43) and Eq.(3.53) we will see that not only higher energies, but also higher precision leads to a breakdown of the effective sigma model. The general gauge-symmetric Lagrangian for the gauge boson masses in Eq.(3.43) involves both terms, where $\mathrm{Tr}[V_\mu V^\mu]$ gives m_W and m_Z proportional to $g \equiv g_W$ and g_Z, while $(\mathrm{Tr}[TV_\mu])^2$ only contributes to m_Z.

The the two gauge boson masses can be expressed in terms of the weak mixing angle θ_w, assumping that that G_F or g universally govern charged-current and neutral-current interactions. At tree level this experimentally very well tested relation corresponds to $\Delta\rho = 0$ or

$$\frac{m_W^2}{m_Z^2} = \frac{g^2}{g_Z^2} = c_w^2 \,. \tag{3.56}$$

We can introduce a free parameter ρ, which breaks this relation

$$\begin{aligned} g_Z^2 &\to g_Z^2 \, \rho \\ m_Z &\to m_Z \, \sqrt{\rho} = m_Z \, \sqrt{1 + \Delta\rho} \,, \end{aligned} \tag{3.57}$$

It corresponds the theoretically derived $\Delta\rho$. In experimental reality, we need a reason to ensure $\Delta\rho = 0$, and the $SU(2)_L \times U(1)_Y$ gauge symmetry unfortunately does not do the job.

In the Standard Model $\rho = 1$ is actually violated at the one-loop level. This means we are looking for an approximate symmetry of the Standard Model. What we can hope for is that this symmetry is at least a good symmetry in the $SU(2)_L$ gauge sector and slightly broken elsewhere. One possibility is to replace $SU(2)_L \times U(1)_Y$ symmetry with a larger $SU(2)_L \times SU(2)_R$ symmetry, which could even be global,

$$\begin{aligned} &\Sigma \to U \Sigma V^\dagger \qquad\qquad U \in SU(2)_L \qquad\qquad V \in SU(2)_R \\ &\mathrm{Tr}(\Sigma^\dagger \Sigma) \to \mathrm{Tr}\left(V \Sigma^\dagger U^\dagger U \Sigma V^\dagger\right) = \mathrm{Tr}(\Sigma^\dagger \Sigma) \,. \end{aligned} \tag{3.58}$$

In this setup, the three components of W^μ form a triplet under $SU(2)_L$ and a singlet under $SU(2)_R$, so $\rho = 1$.

In the gauge boson and fermion mass terms computed in unitary gauge the Σ field becomes identical to its vacuum expectation value $\mathbb{1}$. The two $SU(2)$ transformations act on the vacuum expectation value as

$$\langle \Sigma \rangle \to \langle U \Sigma V^\dagger \rangle = \langle U \mathbb{1} V^\dagger \rangle = U V^\dagger \stackrel{!}{=} \mathbb{1} \,. \tag{3.59}$$

The symmetry requirement can only be satisfied if $U = V$, which means that the vacuum expectation value for Σ breaks $SU(2)_L \times SU(2)_R$ to the diagonal or custodial subgroup $SU(2)_{L+R}$.

Even beyond tree level the global $SU(2)_L \times SU(2)_R$ symmetry structure can protect the relation $\rho = 1$. If fermions reside in $SU(2)_L$ and $SU(2)_R$ doublets we cannot generate any difference between up–type and down–type fermions, which implies for instance $m_b = m_t$. The measured masses $m_t \gg m_b$ leads to $\rho \neq 1$, because self energy loops in the W propagator mix a the bottom and top quark, while the Z propagator includes pure bottom and top loops,

$$\begin{aligned}
\Delta\rho &\supset \frac{3G_F}{8\sqrt{2}\pi^2}\left(m_t^2 + m_b^2 - 2\frac{m_t^2 m_b^2}{m_t^2 - m_b^2}\log\frac{m_t^2}{m_b^2}\right) \\
&= \frac{3G_F}{8\sqrt{2}\pi^2}\left(2m_b^2 + m_b^2\delta - 2m_b^2\frac{1+\delta}{\delta}\log(1+\delta)\right) \qquad \text{defining} \quad m_t^2 = m_b^2(1+\delta) \\
&= \frac{3G_F}{8\sqrt{2}\pi^2}\left(2m_b^2 + m_b^2\delta - 2m_b^2\left(\frac{1}{\delta}+1\right)\left(\delta - \frac{\delta^2}{2} + \frac{\delta^3}{3} + \mathcal{O}(\delta^4)\right)\right) \\
&= \frac{3G_F}{8\sqrt{2}\pi^2} m_b^2\left(2 + \delta - 2 - 2\delta + \delta + \delta^2 - \frac{2}{3}\delta^2 + \mathcal{O}(\delta^3)\right) \\
&= \frac{3G_F}{8\sqrt{2}\pi^2} m_b^2\left(\frac{1}{3}\delta^2 + \mathcal{O}(\delta^3)\right) \\
&= \frac{G_F m_W^2}{8\sqrt{2}\pi^2}\left(\frac{\left(m_t^2 - m_b^2\right)^2}{m_W^2 m_b^2} + \cdots\right) .
\end{aligned} \tag{3.60}$$

In the Taylor series above the assumption of δ being small is of course not realistic, but the result is nevertheless instructive: the shift vanishes very rapidly towards the symmetric limit $m_t \sim m_b$. For the realistic Standard Model mass ratios it becomes

$$\Delta\rho \supset \frac{3G_F}{8\sqrt{2}\pi^2} m_t^2\left(1 - 2\frac{m_b^2}{m_t^2}\log\frac{m_t^2}{m_b^2}\right) = \frac{3G_F m_W^2}{8\sqrt{2}\pi^2}\ \frac{m_t^2}{m_W^2}\ \left(1 + \mathcal{O}\left(\frac{m_b^2}{m_t^2}\right)\right) . \tag{3.61}$$

A second contribution to the ρ parameter will arise from Higgs loops,

$$\Delta\rho \supset -\frac{11 G_F m_Z^2 s_w^2}{24\sqrt{2}\pi^2}\ \log\frac{m_H^2}{m_Z^2}. \tag{3.62}$$

We want to mention that is another parameterization of the same effect, the T parameter. It is part of an effective theory parameterization of deviations from the tree level relations between gauge boson masses, mixing angles, and neutral and charged current couplings,

$$\{S, T, U\} \tag{3.63}$$

Two of these so-called Peskin–Takeuchi parameters can be understood fairly easily: the S-parameter corresponds to a shift of the Z mass. The T parameter compares contributions to the W and Z masses. The third parameter U is less important for most models. Again, we quote the contributions from the heavy fermion doublet,

$$\begin{aligned}
\Delta S &= \frac{N_c}{6\pi}\ \left(1 - 2Y\log\frac{m_t^2}{m_b^2}\right) \\
\Delta T &= \frac{N_c}{4\pi s_w^2 c_w^2 m_Z^2}\ \left(m_t^2 + m_b^2 - \frac{2m_t^2 m_b^2}{m_t^2 - m_b^2}\log\frac{m_t^2}{m_b^2}\right) ,
\end{aligned} \tag{3.64}$$

with $Y = 1/6$ for quarks and $Y = -1/2$ for leptons. While the parameter S has nothing to do with our custodial symmetry, ρ and $T \sim \Delta\rho/\alpha$ are closely linked. Their main difference is the reference point, where $\rho = 1$ refers to its tree level value and $T = 0$ is often chosen for some kind of light Higgs mass and including the Standard Model top-bottom corrections.

Typical experimental constraints form an ellipse in the S vs T plane along the diagonal. They are usually quoted as ΔT with respect to a reference Higgs mass. Compared to a 125 GeV Standard Model Higgs boson the measured

values range around $T \sim 0.1$ and $S \sim 0.05$. Additional contributions $\Delta T \sim 0.1$ tend to be within the experimental errors, much larger contributions are in contradiction with experiment.

There are two reasons to discuss these loop contributions breaking the custodial symmetry in the Standard Model. First, $\Delta\rho$ is experimentally very strongly constrained by electroweak precision measurements, which means that alternative models for electroweak symmetry breaking usually include the same kind of approximate custodial symmetry by construction. Second, in the Standard Model we can measure the symmetry violations from the heavy quarks and from the Higgs sector shown in Eqs.(3.60) and (3.62) in electroweak precision measurements. Even though the Higgs contributions depend on the Higgs mass only logarithmically, we can then derive an upper bound on the Higgs mass of the order of $\mathcal{O}(200\ \text{GeV})$. Since the Higgs discovery , studying electroweak precision data given the measured Higgs mass is one of the most sensitive consistency tests of the Standard Model.

4 Quantum Chromodynamics

After introducing QED as a $U(1)$-gauge interaction and the weak force as an $SU(2)$-gauge interaction, the step to the strong interaction or QCD with based on $SU(3)$ gauge invariance is structurally simple. In particular, we assume, in agreement with all existing measurements, that the gluons as the QCD gauge bosons as massless and that their interaction is not affected by the weak quantum numbers or the fermion doublet structure. This means we can build QCD as a non-abelian version of QED. We will see that QCD, in spite of the structural similarity to QED, has especially interesting properties and equally interesting experimental consequences.

4.1 QCD Lagrangian

We start with the $SU(3)$-version of the same argument we made for $SU(2)$ from Eq.(3.17) on. The $SU(3)$ transformations are given by

$$U = e^{i\alpha_a T_a} \qquad \text{with} \qquad T_{1,2,3} = \frac{1}{2}\begin{pmatrix} \tau_{1,2,3} & 0 \\ 0 & 0 \end{pmatrix} \cdots \tag{4.1}$$

with $a = 1 \ldots N_c^2 - 1 = 1 \ldots 8$ and the number of colors in the fundamental representation, $N_c = 3$. Please note the conventional factor $1/2$ relative to the Pauli matrices. The (3×3) matrices T_a are the traceless, hermitian, and unitary Gell-Mann matrices. We give the first three of them in terms of the Pauli matrices, but there is no point in writing them down because we only need their algebraic properties to compute the color factors of scattering amplitudes,

$$[T_a, T_b] = i f_{abc} T_c \qquad \text{and} \qquad \text{Tr}(T_a T_b) = T_R \delta_{ab} \equiv \frac{1}{2}\delta_{ab} \; . \tag{4.2}$$

Here, f_{abc} are the antisymmetric structure constants of $SU(3)$ with

$$f_{acd} f_{bcd} = N_c \delta_{ab} \; . \tag{4.3}$$

The one formula we need to compute color factors for quark processes will be

$$(T_a)_{ij} (T_a)_{k\ell} = \frac{1}{2}\left(\delta_{i\ell}\delta_{jk} - \frac{1}{N_c}\delta_{ij}\delta_{k\ell}\right) \; . \tag{4.4}$$

The QCD Lagrangian can be constructed as a non-abelian massless version of the QED Lagrangian. For the dimension-4 Lagrangian we follow Eq.(3.23), limit ourselves to the quarks as fermions charged under $SU(3)$, and write

$$\mathscr{L}_{\text{QCD}} = \bar{q}_i i\gamma^\mu (D_\mu)_{ij} q_j - \frac{1}{4} F_{a,\mu\nu} F_a^{\mu\nu} \; , \tag{4.5}$$

with the gluon field strength tensor

$$F_{a,\mu\nu} = \partial_\mu A_{a,\nu} - \partial_\nu A_{a,\mu} - i g_s [A_\mu, A_\nu]_a \; , \tag{4.6}$$

and the $SU(3)$-covariant derivative with the appropriate color indices

$$(D_\mu)_{ij} = \partial_\mu \mathbb{1}_{ij} - i g_s A_{\mu,a} (T_a)_{ij} \tag{4.7}$$

From these formulas we see immediately that the quark propagators are the same as for QED, and unlike for the weak interaction we do not have to keep track of the chirality. The quark-quark-gluon interaction is very similar to the quark-quark-photon vertex in Eq.(1.40),

$$-i g_s \, \gamma^\mu \, (T_a)_{ij} \qquad\qquad (q_j - \bar{q}_i - g_{a,\mu}). \tag{4.8}$$

The only additional element is the $SU(3)$-matrix, and we already see that all color matrices in a scattering process will form traces along the quark line(s). The color factors factorize from the Dirac algebra and can be computed independently.

Let us start with the electroweak processes

$$e^+e^- \to q\bar{q} \qquad \text{and} \qquad q\bar{q} \to e^+e^- \tag{4.9}$$

Here no gluons appear at leading order, which means we do not have any matrices T^a in our calculations. Nevertheless, a color factor appears because we sum over the (identical) colors of the two external quarks. We can write this color sum for the squared matrix element formally as

$$\delta_{ij}\delta_{ji} = \delta_{ii} = N_c \tag{4.10}$$

The different between the process $e^+e^- \to q\bar{q}$ and $q\bar{q} \to e^+e^-$ is that for the former we sum over the color states in the final state and for the latter we average over the color states in the initial state, giving us another factor $1/N_c$. Let us now radiate a gluon from any of these processes, for instance

$$e^+e^- \to q\bar{q}g \qquad \text{and} \qquad q\bar{q} \to e^+e^-g \tag{4.11}$$

First, we have to deal with another gluon in the Dirac traces, just like when radiating a photon. In addition, there is the color contribution in Eq.(4.8),

$$(T_a)_{ij}(T_a)_{ji} = \text{Tr}(T_aT_a) = \frac{1}{2}\delta_{aa} = \frac{N_c^2-1}{2} \; . \tag{4.12}$$

It can be expressed in terms of the fundamental Casimir as

$$\text{Tr}(T_aT_a) = N_cC_F \qquad \text{with} \qquad C_F = \frac{N_c^2-1}{2N_c} = \frac{4}{3} \; , \tag{4.13}$$

Another example is the successive radiation of two gluons from a hard quark,

$$q \to qg_ag_b \; . \tag{4.14}$$

Depending on the order of the two gluons, we first find the planar color factor,

$$\begin{aligned}
\text{Tr}(T^aT^aT^bT^b) &= (T^aT^a)_{il}(T^bT^b)_{li} \\
&= \frac{1}{4}\left(\delta_{il}\delta_{jj} - \frac{\delta_{ij}\delta_{jl}}{N_c}\right)\left(\delta_{il}\delta_{jj} - \frac{\delta_{ij}\delta_{jl}}{N_c}\right) \\
&= \frac{1}{4}\left(\delta_{il}N_c - \frac{\delta_{il}}{N_c}\right)\left(\delta_{il}N_c - \frac{\delta_{il}}{N_c}\right) \\
&= N_c\,\left(\frac{N_c^2-1}{2N_c}\right)^2 = N_cC_F^2 = \frac{16}{3} = \mathcal{O}(N_c^3) \; .
\end{aligned} \tag{4.15}$$

When we cross the gluon lines between the diagram and its complex conjugate we get the same way

$$\text{Tr}(T_aT_bT_aT_b) = -\frac{C_F}{2} = -\frac{2}{3} = \mathcal{O}(N_c) \; . \tag{4.16}$$

That contribution is suppressed by a factor eight, which means that two gluons have a significant preference for ordered emission.

We can also compute the color factor for the purely gluonic theory, *i.e.* radiating gluons off two hard gluons in the final state. For instance, planar double gluon emission with the exchanged gluon indices b and f gives us the largest color factor

$$f^{abd}f^{abe}f^{dfg}f^{efg} = N_c\delta^{de}\;N_c\delta^{de} = \mathcal{O}(N_c^3) \; , \tag{4.17}$$

now independent of the ordering.

4.2 Ghosts

The main complication of QCD compared to QED comes from the kinetic term of the gluons, where we keep in mind that for non-abelian gauge groups the term $F_{a,\mu\nu}F_a^{\mu\nu}$ gives rise to the gauge boson propagator and to the self-interactions. We immediately see this when we insert the formula in Eq.(4.6) into the Lagrangian of Eq.(4.5) and encounter up to four powers of the gluon field, and no derivative.

Let us go back to QED, where in Sec. 3.2 we emphasize that photons have only two transverse degree of freedom and then still write the gluon field in terms of a 4-vector A_μ, effectively consisting of four degrees of freedom. How does QED ensure that in the actual calculation the longitudinal and the scalar degrees of freedom do not contribute?

The way to tackle this question is through the gauges introduced in Eq.(3.54), simplified in the massless limit to

$$\Delta^{\mu\nu}(p) = \frac{-i}{p^2+i\epsilon}\left[g^{\mu\nu} + (\xi-1)\frac{p^\mu p^\nu}{p^2}\right] = \begin{cases} \dfrac{-i}{p^2+i\epsilon}\, g^{\mu\nu} & \text{Feynman gauge } \xi = 1 \\ \dfrac{-i}{p^2+i\epsilon}\left[g^{\mu\nu} - \dfrac{p^\mu p^\nu}{p^2}\right] & \text{Landau/Lorenz gauge } \xi = 0\ . \end{cases} \tag{4.18}$$

Here we need to be careful with the word 'gauge', because the different forms of the propagator have nothing to do with the gauge symmetry of the Lagrangian. The unitary gauge makes no sense for massless particles. For the weak bosons the unitary gauge was the only way to decouple and get right of the Goldstone modes.

To learn how to remove the unwanted degrees of freedom we write the photon Lagrangian from Eq.(3.1) such that it gives us the photon propagator from Eq.(4.18) in the R_ξ gauge. We only quote the result as

$$\mathscr{L}_{\text{photon, gf}} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} - \frac{1}{2\xi}(\partial_\mu A^\mu)^2\ . \tag{4.19}$$

In analogy to Eq.(1.7) this Lagrangian gives us the equation of motion

$$\partial^\mu\partial_\mu A_\nu - \left(1+\frac{1}{\xi}\right)\partial^\mu\partial_\nu A_\mu = 0\ . \tag{4.20}$$

It is equivalent to the version without the gauge fixing term, in that it requires the d'Alembert equation for the photon and the Lorenz gauge condition.

From the massive photon example we know how to turn Eq.(4.19) gauge-invariant. From the weak gauge bosons we also know how to switch from one gluon propagator to another — again we need to introduce another field. For the massless gauge bosons we will refer to these new fields as ghosts. Let us start with the QED Lagrangian in Eq.(4.19), including the gauge fixing term corresponding to the general photon propagator. Because we know from Eq.(3.2) that $F_{\mu\nu}$ is $U(1)$-gauge invariant, we also know that the Lagrangian in R_ξ gauge is not locally $U(1)$-symmetric. From Eq.(3.2) we remember the gauge transformation of the photon as

$$\begin{aligned} A^\mu &\to A^\mu - \frac{1}{e}\partial^\mu\alpha \\ -\frac{1}{2\xi}(\partial_\mu A^\mu)^2 &\to -\frac{1}{2\xi}\left(\partial_\mu A^\mu - \frac{1}{e}\partial^2\alpha\right)^2 \approx -\frac{1}{2\xi}(\partial_\mu A^\mu)^2 + \frac{1}{e\xi}(\partial_\mu A^\mu)(\partial^2\alpha)\ . \end{aligned} \tag{4.21}$$

In the last step we ignore higher powers of α, because we are working with infinitesimal gauge transformations. To turn this Lagrangian gauge invariant we add an auxiliary field $\bar{c}$ such that

$$\mathscr{L}_{\text{photon-ghost}} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} - \frac{1}{2\xi}(\partial_\mu A^\mu)^2 + \bar{c}\,\partial^2\alpha\ , \tag{4.22}$$

We see that the combination of the gauge fixing term and this new term is gauge-invariant if

$$\bar{c} \to \bar{c} - \frac{1}{e\xi}(\partial_\mu A^\mu)\ . \tag{4.23}$$

To give a meaning to this auxiliary term we move one of the derivatives through an integration by parts. In this form the additional term in the Lagrangian makes sense if we upgrade α to another field, $\alpha \to c$, such that the auxiliary $\bar{c}$ and the upgraded c form a complex scalar,

$$\begin{aligned}\mathscr{L}_{\text{photon-ghost}} &= -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} - \frac{1}{2\xi}(\partial_\mu A^\mu)^2 + (\partial_\mu \bar{c})\,(\partial^\mu \alpha) \\ &\to -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} - \frac{1}{2\xi}(\partial_\mu A^\mu)^2 + (\partial_\mu \bar{c})\,(\partial^\mu c)\;. \end{aligned} \tag{4.24}$$

In Eq.(3.11) we have seen something similar, namely a scalar field added to the QED Lagrangian to make sure the massive degrees of freedom are correctly described. Because the unitary gauge is not defined in Eq.(4.18), there is no gauge choice for which we can decouple and neglect the ghosts, they have to be computed in the Lorenz and in the Feynman gauges. What saves us in QED is that they do not appear elsewhere in the Lagrangian, so they are propagating but non-interacting fields. For our QED calculations this means that they are irrelevant and we can ignore them.

The situation changes because of the non-abelian structure of QCD, which replaces the standard derivative with a covariant derivative in the gauge transformation and also in the kinetic term for the ghosts. The gauge-fixed Lagrangian with the compensating ghost fields then reads

$$\mathscr{L}_{\text{gluon-ghost}} = -\frac{1}{4}F_{a,\mu\nu}F_a^{\mu\nu} - \frac{1}{2\xi}(\partial_\mu A_a^\mu)^2 - (\partial^\mu \bar{c}_a)(\partial_\mu c_a) + g_s f_{abc} A_a^\mu (\partial_\mu \bar{c}_b) c_c\;. \tag{4.25}$$

The gluon propagator in the R_ξ gauge is the same as the photon propagator in Eq.(4.18), just with a factor δ_{ab} in the numerator. We are skipping the crucial triple and quartic gluon self-interactions, because they are lengthy. However, we can write down the ghost Feynman rules, including the ghost propagator following from Eq.(4.22)

$$-\frac{\delta_{ab}}{p^2 + i\epsilon} \tag{4.26}$$

and the ghost-ghost-gluon interaction

$$-igf_{abc}p_\mu \qquad\qquad (c - \bar{c} - g)\;, \tag{4.27}$$

where p_μ is the 4-momentum of the incoming ghost field.

Because we derived the ghosts through gauge invariance, we can at least for external gluons follow a slightly different direction to ensure the correct degrees of freedom contribute to the matrix element. For a matrix element with two external gluons $\mathcal{M}^{\mu\nu}$ the explicit condition is

$$p_\mu \mathcal{M}^{\mu\nu} = 0 = p_\nu \mathcal{M}^{\mu\nu}\;, \tag{4.28}$$

and we can enforce it by projecting $\mathcal{M}^{\mu\mu}$ onto the allowed tensor structures. Following the same line of thought for propagators leads us back to Eq.(4.18), where we find for the so-called transverse tensor

$$T^{\mu\nu} = g^{\mu\nu} - \frac{p^\nu p^\mu}{p^2} \qquad \Rightarrow \qquad p_\mu T^{\mu\nu} = 0 = p_\nu T^{\mu\nu}\;. \tag{4.29}$$

It is a projector on the transverse degrees of freedom because

$$\begin{aligned} T^{\mu\nu}T^\rho_\nu &= \left(g^{\mu\nu} - \frac{p^\nu p^\mu}{p^2}\right)\left(g^\rho_\nu - \frac{p_\nu p^\rho}{p^2}\right) \\ &= g^{\mu\rho} - \frac{p^\mu p^\rho}{p^2} - \frac{p^\rho p^\mu}{p^2} + \frac{p^2 p^\mu p^\rho}{p^4} = T^{\mu\rho}\;. \end{aligned} \tag{4.30}$$

This means that the gauge propagator in Lorenz gauge is guaranteed to be transverse in the covariant sense. The problem with this argument is that propagators define individual Feynman diagrams, and gauge invariance only holds for all Feynman diagrams combined. This means that for internal gluons, including loop integrals, even the Lorenz gauge does not ensure the correct gluon polarization and we always have to include the ghosts explicitly.

4.3 Ultraviolet divergences

Renormalization as the proper treatment of ultraviolet divergences is one of the most important things to understand about quantum field theory. It is driven by the appearance of ultraviolet divergences, which we first regularize, *i.e.* describe in a well-defined manner, and then renormalize away through counter terms. This renormalization leads to the appearance of the renormalization scale.

Scales automatically arise from infrared or ultraviolet divergences. We can see this by writing down a simple scalar loop integral, with to two virtual scalar propagators with masses $m_{1,2}$ and an external momentum p flowing through a diagram,

$$B(p^2; m_1, m_2) \equiv \int \frac{d^4q}{16\pi^2} \, \frac{1}{q^2 - m_1^2} \frac{1}{(q+p)^2 - m_2^2} \, . \tag{4.31}$$

Such two-point functions appear for example in the gluon self energy with virtual gluons, with massless ghost scalars, with a Dirac trace in the numerator for quarks, and with massive scalars for supersymmetric scalar quarks. In those cases the two masses are identical $m_1 = m_2$. The integration measure $1/(16\pi^2)$ is dictated by the Feynman rule for the integration over loop momenta. Counting powers of q, we see that the integral behaves like

$$B(p^2; m_1, m_2) \sim \frac{1}{16\pi^2} \int \frac{d^4q}{q^4} \tag{4.32}$$

in the ultraviolet, so it is logarithmically divergent.

One regularization scheme is a cutoff into the momentum integral Λ, for example through the so-called Pauli—Villars regularization. Because the ultraviolet behavior of the integrand or integral cannot depend on any parameter living at a small energy scales, the parameterization of the ultraviolet divergence in Eq.(4.31) cannot involve the mass m or the external momentum p^2. The scalar two-point function has mass dimension zero, so its divergence has to be proportional to $\log(\Lambda/\mu_R)$ with a dimensionless prefactor and some scale μ_R^2 which is an artifact of the regularization of such a Feynman diagram. Because it is an artifact, this scale μ_R has to eventually vanish from our theory prediction.

A more elegant regularization scheme is dimensional regularization. It is designed not to break gauge invariance and naively seems to not introduce a mass scale μ_R. When we shift the momentum integration from 4 to $4 - 2\epsilon$ dimensions and use analytic continuation in the number of space–time dimensions to renormalize the theory, a renormalization scale μ_R appears when we ensure the two-point function and with it observables like cross sections keep their correct mass dimension

$$\begin{aligned} \int \frac{d^4q}{16\pi^2} \, \frac{1}{q^2 - m_1^2} \frac{1}{(q+p)^2 - m_2^2} &\to \mu_R^{2\epsilon} \int \frac{d^{4-2\epsilon}q}{16\pi^2} \, \frac{1}{q^2 - m_1^2} \frac{1}{(q+p)^2 - m_2^2} \\ &= \frac{i\mu_R^{2\epsilon}}{(4\pi)^2} \left[\frac{C_{-1}}{\epsilon} + C_0 + C_1 \, \epsilon + \mathcal{O}(\epsilon^2) \right] \, . \end{aligned} \tag{4.33}$$

The constants C_i in the series in $1/\epsilon$ depend on the loop integral. To regularize the ultraviolet divergence we go into the limit $\epsilon > 0$ and find mathematically well defined poles $1/\epsilon$. Defining scalar integrals with the integration measure $1/(i\pi^2)$ will make for example C_{-1} come out as of the order $\mathcal{O}(1)$. This is the reason we usually find factors $1/(4\pi)^2 = \pi^2/(2\pi)^4$ in front of the loop integrals.

4.4 Counter terms

The ultraviolet poles in $1/\epsilon$ will cancel with universal counter terms once we renormalize the theory. We include counter terms by shifting parameters in the Lagrangian and the leading order matrix element. They cancel the poles for example from virtual one-loop diagrams,

$$\begin{aligned} |\mathcal{M}_{\text{LO}}(g) + \mathcal{M}_{\text{virt}}|^2 &= |\mathcal{M}_{\text{LO}}(g)|^2 + 2 \, \text{Re} \, \mathcal{M}_{\text{LO}}(g) \mathcal{M}_{\text{virt}} + \cdots \\ &\to |\mathcal{M}_{\text{LO}}(g + \delta g)|^2 + 2 \, \text{Re} \, \mathcal{M}_{\text{LO}}(g) \mathcal{M}_{\text{virt}} + \cdots \\ \text{with} \qquad g \to g^{\text{bare}} &= g + \delta g \qquad \text{and} \quad \delta g \propto \alpha_s/\epsilon \, . \end{aligned} \tag{4.34}$$

The dots indicate higher orders in α_s.

The counter terms do not come with a factor $\mu_R^{2\epsilon}$, so this factor will not be matched between the actual ultraviolet divergence and the counter term. We can keep track of the renormalization scale best by expanding the prefactor of the regularized but not yet renormalized integral in Eq.(4.33) in a Taylor series in ϵ, no question asked about convergence radii

$$\begin{aligned}
\mu_R^{2\epsilon} \left[\frac{C_{-1}}{\epsilon} + C_0 + \mathcal{O}(\epsilon)\right] &= e^{2\epsilon \log \mu_R} \left[\frac{C_{-1}}{\epsilon} + C_0 + \mathcal{O}(\epsilon)\right] \\
&= \left[1 + 2\epsilon \log \mu_R + \mathcal{O}(\epsilon^2)\right] \left[\frac{C_{-1}}{\epsilon} + C_0 + \mathcal{O}(\epsilon)\right] \\
&= \frac{C_{-1}}{\epsilon} + C_0 + C_{-1} \log \mu_R^2 + \mathcal{O}(\epsilon) \\
&\to \frac{C_{-1}}{\epsilon} + C_0 + C_{-1} \log \frac{\mu_R^2}{M^2} + \mathcal{O}(\epsilon) \; . \qquad (4.35)
\end{aligned}$$

In the last step we correct by hand for the fact that $\log \mu_R^2$ with a mass dimension inside the logarithm cannot appear in our calculations. From somewhere else in our calculation the logarithm will be matched with a $\log M^2$ where M^2 is the typical mass or energy scale in our process. This little argument shows that also in dimensional regularization we introduce a mass scale μ_R which appears as $\log(\mu_R^2/M^2)$ in the renormalized expression for our observables.

In Eq.(4.35) there appear two finite contributions to a given observable, the expected C_0 and the renormalization–induced C_{-1}. Because the factors C_{-1} are linked to the counter terms in the theory we can often guess them without actually computing the complete loop integral, which is very useful in cases where they numerically dominate.

Counter terms are not uniquely defined. They remove a given divergence to return finite observables, but we are free to add any finite contribution we want. This opens many ways to define a counter term for example based on physical processes where counter terms do not only cancel the pole but also finite contributions at a given order in perturbation theory. An example for such a physical renormalization scheme is the on–shell scheme for masses, where we define a counter term such that external on–shell particles do not receive any corrections to their masses. For the top mass this means

$$\begin{aligned}
m_t^{\text{bare}} &= m_t + \delta m_t \\
&= m_t + m_t \frac{\alpha_s C_F}{4\pi} \left(3\left(-\frac{1}{\epsilon} + \gamma_E - \log(4\pi) - \log\frac{\mu_R^2}{M^2}\right) - 4 + 3\log\frac{m_t^2}{M^2}\right) \\
&\equiv m_t + m_t \frac{\alpha_s C_F}{4\pi} \left(-\frac{3}{\tilde{\epsilon}} - 4 + 3\log\frac{m_t^2}{M^2}\right) \qquad \Leftrightarrow \qquad \frac{1}{\tilde{\epsilon}\left(\frac{\mu_R}{M}\right)} \equiv \frac{1}{\epsilon} - \gamma_E + \log\frac{4\pi\mu_R^2}{M^2} \; , \qquad (4.36)
\end{aligned}$$

with the color factor $C_F = (N^2-1)/(2N)$ and the Euler constant $\gamma_E \approx 0.577$ coming from the evaluation of the Gamma function $\Gamma(\epsilon) = 1/\epsilon + \gamma_E + \mathcal{O}(\epsilon)$. The convenient scale dependent pole $1/\tilde{\epsilon}$ includes the universal additional terms like the Euler gamma function and the scaling logarithm. This logarithm is the big problem in this universality argument, since we need to introduce the arbitrary energy scale M to separate the universal logarithm of the renormalization scale and the parameter-dependent logarithm of the physical process.

Another example for a process dependent renormalization scheme is the mixing of γ and Z propagators. There we choose the counter term of the weak mixing angle such that an on–shell Z boson cannot oscillate into a photon, and vice versa.

One common feature of all mass counter terms is $\delta m \propto m$, which means that our renormalization is multiplicative,

$$m^{\text{bare}} = Z_m\, m = (1 + \delta Z_m)\, m = \left(1 + \frac{\delta m}{m}\right) m = m + \delta m \qquad \text{with} \qquad \delta Z_m = \frac{\delta m}{m} \; , \qquad (4.37)$$

This form implies that particles with zero mass will not obtain a finite mass through renormalization. If we remember that chiral symmetry protects a Lagrangian from acquiring fermion masses this means that on–shell renormalization

does not break this symmetry. A massless theory cannot become massive by mass renormalization. Regularization and renormalization schemes which do not break symmetries of the Lagrangian are ideal.

Another way of introducing counter terms is by defining a renormalization point. This can be the energy scale at which the counter terms cancels all higher order contributions, divergent as well as finite. The best known example is the electric charge which we renormalize in the Thomson limit of zero momentum transfer through the photon propagator

$$e \to e^{\text{bare}} = e + \delta e \; . \tag{4.38}$$

Finally, there is a way to define a completely general counter term: if dimensional regularization does not break any of the symmetries of our Lagrangian, we can simply subtract the ultraviolet pole. The only question is: do we subtract $1/\epsilon$ in the MS scheme or $1/\bar{\epsilon}$ in the $\overline{\text{MS}}$ scheme. In the $\overline{\text{MS}}$ scheme the counter term becomes scale dependent.

Carefully counting, there are three scales we need to consider:

1. the physical scale in the process, like the top mass m_t in the matrix element for the top decay;

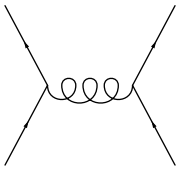
2. the renormalization scale μ_R, a reference scale as part of the definition of the counter term;

3. The scale M separating the counter term from the process dependent result, which we can choose however we want. The role of M will become clear when we go through the example of the running strong coupling α_s.

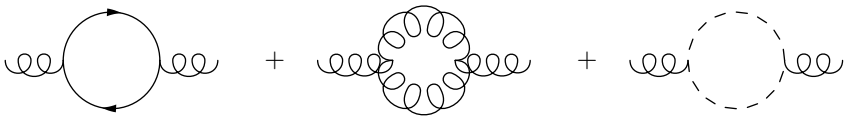
Of course, we would prefer to choose all three scales the same, but in a complex physical process this will not be possible. For example, any massive $(2 \to 3)$ production process naturally involves several external physical scales.

4.5 Running coupling

To get an idea what these different scales mean we compute the running strong coupling $\alpha_s(\mu_R^2)$. A simple process where we can study it is bottom pair production, where at some energy range we will be dominated by valence quarks: $q\bar{q} \to b\bar{b}$. The only Feynman diagram is an s-channel off–shell gluon with momentum $p^2 \equiv s$,



At next–to–leading order this gluon propagator will be corrected by self energy loops, where the gluon splits into two quarks or gluons and re-combines before it produces the two final–state bottoms. Let us for now assume that all quarks are massless. The Feynman diagrams for the gluon self energy include a quark look, a gluon loop, and the ghost loop which removes the unphysical degrees of freedom of the gluon inside the loop:



The gluon self energy correction or so-called vacuum polarization will be a scalar. All fermion lines close in the Feynman diagram and the Dirac trace is computed inside the loop. In color space the self energy will (hopefully) be diagonal, just like the gluon propagator itself, so we can ignore the color indices for now. In Lorenz gauge the gluon propagator is proportional to the transverse tensor defined in Eq.(4.29). The same should be true for the gluon self

energy, which we therefore write as $\Pi^{\mu\nu} \equiv \Pi\, T^{\mu\nu}$. Including the gluon, quark, and ghost loops the regularized gluon self energy with a momentum flow p^2 through the propagator reads

$$\begin{aligned} -\frac{1}{p^2}\,\Pi\left(\frac{\mu_R^2}{p^2}\right) &= \frac{\alpha_s}{4\pi}\,\left(-\frac{1}{\tilde{\epsilon}}+\log\frac{p^2}{M^2}\right)\,\left(\frac{13}{6}N_c-\frac{2}{3}n_f\right)+\mathcal{O}(\log m_t^2) \\ &\equiv \alpha_s\,\left(-\frac{1}{\tilde{\epsilon}}+\log\frac{p^2}{M^2}\right)\,b_0+\mathcal{O}(\log m_t^2) \\ &\quad \text{with} \qquad b_0=\frac{1}{4\pi}\left(\frac{13}{6}N_c-\frac{2}{3}n_f\right) \\ &\quad \text{but really} \qquad b_0=\frac{1}{4\pi}\left(\frac{11}{3}N_c-\frac{2}{3}n_f\right) \overset{\text{SM}}{>} 0\,. \end{aligned} \tag{4.39}$$

The number of fermions coupling to the gluons is n_f. The factor b_0 reflects the one-loop corrections. Strictly speaking, it gives the first term in a perturbative series in the strong coupling $\alpha_s = g_s^2/(4\pi)$.

In the last step of Eq.(4.39) we have snuck in additional contributions by replacing the factor 13/6 by a factor 11/3. This is related to the fact that there are actually three types of divergent virtual gluon diagrams in the physical process $q\bar{q} \to b\bar{b}$: the external quark self energies with renormalization factors $Z_f^{1/2}$, the internal gluon self energy Z_A, and the vertex corrections Z_{Aff}. The physical parameters we can renormalize in this process are the strong coupling and the quark masses. Wave function renormalization constants are not physical. The entire divergence which eventually needs to be absorbed in Z_g is given by the combination

$$Z_{Aff} = Z_g Z_A^{1/2} Z_f \qquad \Leftrightarrow \qquad \frac{Z_{Aff}}{Z_A^{1/2} Z_f} \equiv Z_g\,. \tag{4.40}$$

This changes the factor from 13/6 to 11/3 in the running of the strong coupling.

We can check this definition of Z_g by comparing all vertices in which the strong coupling g_s appears, namely the gluon coupling to quarks and ghosts, as well as the triple and quartic gluon vertex. All of them need to have the same divergence structure

$$\frac{Z_{Aff}}{Z_A^{1/2} Z_f} \overset{!}{=} \frac{Z_{Acc}}{Z_A^{1/2} Z_c} \overset{!}{=} \frac{Z_{3A}}{Z_A^{3/2}} \overset{!}{=} \sqrt{\frac{Z_{4A}}{Z_A^2}}\,. \tag{4.41}$$

If we had done the same calculation in QED and looked for a running electric charge, we would have found that the vacuum polarization diagrams for the photon do account for the entire counter term of the electric charge. The other two renormalization constants Z_{Aff} and Z_f cancel because of gauge invariance.

In contrast to QED, the strong coupling diverges in the Thomson limit because QCD is confined towards large distances and weakly coupled at small distances. Lacking a well enough motivated reference point we are lead to renormalize $\alpha_s = g_s^2/(4\pi)$ in the $\overline{\text{MS}}$ scheme. From Eq.(4.39) we know that the ultraviolet pole which needs to be cancelled by the counter term is proportional to the function b_0

$$\begin{aligned} \alpha_s^{\text{bare}} = \left(1+\frac{\delta\alpha_s}{\alpha_s}\right)\alpha_s &\overset{\overline{\text{MS}}}{=} \left(1-\left.\frac{\Pi}{p^2}\right|_{\text{pole}}\right)\alpha_s(M^2) \\ &\overset{\text{Eq.(4.39)}}{=} \left(1-\frac{\alpha_s}{\tilde{\epsilon}\left(\frac{\mu_R}{M}\right)}\,b_0\right)\alpha_s(M^2)\,. \end{aligned} \tag{4.42}$$

Here we explicitly include the scale dependence of the counter term. Because the bare coupling does not depend on any scales, this means that the renormalized α_s depends on an unphysical scale M. We can also evaluate it at the momentum flowing through the gluon propagator p^2 and write according to Eq.(4.39)

$$\alpha_s^{\text{bare}} = \alpha_s(p^2)\left(1-\frac{\alpha_s(p^2)b_0}{\tilde{\epsilon}}+\alpha_s(p^2)b_0\log\frac{p^2}{M^2}\right)\,. \tag{4.43}$$

On the right-hand side α_s is evaluated at the physical scale p^2. The logarithm shifts the argument of $\tilde{\epsilon}$ from M^2 to p^2. This way the formula defines a running coupling $\alpha_s(p^2)$ and accounts for shifts between the physical scale p^2 and the general scale M^2 coming out of the $\overline{\text{MS}}$ scheme. Identifying the right-hand sides of Eqs.(4.42) and (4.43) we find

$$\begin{aligned}\alpha_s(M^2) &= \alpha_s(p^2) + \alpha_s^2(p^2) b_0 \log \frac{p^2}{M^2} \\ &= \alpha_s(p^2) \left(1 + \alpha_s(p^2) b_0 \log \frac{p^2}{M^2}\right) \; . \end{aligned} \tag{4.44}$$

To the given loop order the argument of the strong coupling squared on the right side can be neglected — its effect is of higher order. We nevertheless keep the argument as a higher order effect to later distinguish different approaches to the running coupling.

4.6 Resummation

We can do better than fixed order in perturbation theory: instead of simply including the gluon self energy bubble at a given order in perturbation theory we can include chains of one-loop diagrams with Π appearing many times in the gluon propagator. It means we replace the off–shell gluon propagator by

$$\begin{aligned}\frac{T^{\mu\nu}}{p^2} \to & \frac{T^{\mu\nu}}{p^2} + \left(\frac{T}{p^2} \cdot (-T\,\Pi) \cdot \frac{T}{p^2}\right)^{\mu\nu} \\ & + \left(\frac{T}{p^2} \cdot (-T\,\Pi) \cdot \frac{T}{p^2} \cdot (-T\,\Pi) \cdot \frac{T}{p^2}\right)^{\mu\nu} + \cdots \\ = & \frac{T^{\mu\nu}}{p^2} \sum_{j=0}^{\infty} \left(-\frac{\Pi}{p^2}\right)^j = \frac{T^{\mu\nu}}{p^2} \; \frac{1}{1 + \Pi/p^2} \; , \end{aligned} \tag{4.45}$$

schematically written without the factors i. To avoid indices we abbreviate $T^{\mu\nu} T^\rho_\nu = T \cdot T$ and

$$(T \cdot T \cdot T)^{\mu\nu} = T^{\mu\rho} T^\sigma_\rho T^\nu_\sigma = T^{\mu\sigma} T^\nu_\sigma = T^{\mu\nu} \; . \tag{4.46}$$

This (re-)summation moves the finite shift in α_s from Eqs.(4.39) and (4.43) into the denominator, while we assume that the pole will be properly taken care of at any given order,

$$\alpha_s^{\text{bare}} = \alpha_s(M^2) - \frac{\alpha_s^2 b_0}{\tilde{\epsilon}} \equiv \frac{\alpha_s(p^2)}{1 - \alpha_s(p^2)\; b_0 \; \log \dfrac{p^2}{M^2}} - \frac{\alpha_s^2 b_0}{\tilde{\epsilon}} \; . \tag{4.47}$$

As before, we can relate the values of α_s at two reference points, *i.e.* we consider it a renormalization group equation (RGE) which evolves physical parameters from one scale to another in analogy to the fixed order version in Eq.(4.44)

$$\frac{1}{\alpha_s(M^2)} = \frac{1}{\alpha_s(p^2)} \left(1 - \alpha_s(p^2)\; b_0 \; \log \frac{p^2}{M^2}\right) = \frac{1}{\alpha_s(p^2)} - b_0 \; \log \frac{p^2}{M^2} \; . \tag{4.48}$$

The factor α_s in the parentheses can be evaluated at either of the two scales, the difference is a higher order effect. If we keep it at p^2 we see how Eq.(4.48) is different from Eq.(4.43) and how resumming the vacuum expectation bubbles differs from the un-resummed result in higher order contributions. When we differentiate $\alpha_s(p^2)$ with respect to p^2 we find with $d/dx(1/\alpha_s) = -1/\alpha_s^2 \; d\alpha_s/dx$

$$\begin{aligned} p^2 \frac{d\alpha_s(p^2)}{dp^2} = \frac{d\alpha_s(p^2)}{d \log p^2} &= -\alpha_s^2 \frac{d}{d \log p^2} \frac{1}{\alpha_s(p^2)} \\ &= -\alpha_s^2 \frac{d}{d \log p^2} \; \left[\frac{1}{\alpha_s(M^2)} + b_0 \log \frac{p^2}{M^2}\right] \\ &= -\alpha_s^2 \, b_0 + \mathcal{O}(\alpha_s^2) \\ &\to -\alpha_s^2 \sum_{n=0} b_n \alpha_s^n \equiv \beta \; . \end{aligned} \tag{4.49}$$

This is the running of the strong coupling constant including all higher order terms b_n. From Eq.(4.39) we know that $b_0 > 0$, which means that towards larger scales the strong coupling has a negative slope. The ultraviolet limit of the strong coupling is zero and QCD is asymptotically free, while in the infrared it hits a Landau pole.

If we are interested in the running of the strong coupling as a function of the physical measurement scale, $\alpha_s(p^2)$, we can solve Eq.(4.48) for

$$\frac{1}{\alpha_s(p^2)} = \frac{1}{\alpha_s(M^2)} \left(1 + \alpha_s(M^2) b_0 \log \frac{p^2}{M^2}\right) . \tag{4.50}$$

For $b_0 > 0$ this multiplicative factor can become zero at $p^2 \ll M^2$. This implies $\alpha_s(p^2) \to \infty$, the so-called Landau pole defining a small energy scale at which the strong coupling diverges and our theory switches degrees of freedom from quarks and gluon to mesons and baryons. We turn the problem into a feature by turning the Landau pole into a reference scale Λ_{QCD} at a given loop order through

$$\begin{aligned} 1 + \alpha_s(M^2)\, b_0\, \log \frac{\Lambda^2_{\text{QCD}}}{M^2} \stackrel{!}{=} 0 \quad &\Leftrightarrow \quad \log \frac{\Lambda^2_{\text{QCD}}}{M^2} = -\frac{1}{\alpha_s(M^2) b_0} \\ &\Leftrightarrow \quad \log \frac{p^2}{M^2} = \log \frac{p^2}{\Lambda^2_{\text{QCD}}} - \frac{1}{\alpha_s(M^2) b_0} , \end{aligned} \tag{4.51}$$

and then include it in the running

$$\begin{aligned} \frac{1}{\alpha_s(p^2)} &\stackrel{\text{Eq.(4.48)}}{=} \frac{1}{\alpha_s(M^2)} + b_0\, \log \frac{p^2}{M^2} \\ &= \frac{1}{\alpha_s(M^2)} + b_0 \log \frac{p^2}{\Lambda^2_{\text{QCD}}} - \frac{1}{\alpha_s(M^2)} = b_0 \log \frac{p^2}{\Lambda^2_{\text{QCD}}} \qquad \Leftrightarrow \qquad \alpha_s(p^2) = \frac{1}{b_0 \log \dfrac{p^2}{\Lambda^2_{\text{QCD}}}} . \end{aligned} \tag{4.52}$$

An interesting aspect of Λ_{QCD} is that we introduce a scale into our theory without ever setting it. All we do is renormalize a coupling which becomes strong at large energies and search for the mass scale of this strong interaction. This trick is called dimensional transmutation.

In terms of language, there is a little bit of confusion between field theorists and phenomenologists: we have introduced the renormalization scale μ_R as the renormalization point, for example of the strong coupling constant. In the $\overline{\text{MS}}$ scheme, the subtraction of $1/\tilde{\epsilon}$ shifts the scale dependence of the strong coupling to M^2 and moves the logarithm $\log M^2/\Lambda^2_{\text{QCD}}$ into the definition of the renormalized parameter. This is what we call the renormalization scale in the phenomenological sense, *i.e.* the argument we evaluate α_s at.

4.7 Resumming scaling logarithms

Up to now we have introduced the running strong coupling in a fairly abstract manner and did not link the resummation of diagrams and the running of α_s in Eqs.(4.44) and (4.49) to physics. In what way does the resummation of the one-loop diagrams for the s-channel gluon improve our prediction of the LHC observables?

As an illustration we look at a simple observable which depends on just one physical energy scale p^2. The first observable coming to mind is again the Drell–Yan cross section $\sigma(q\bar{q} \to \mu^+\mu^-)$, but since we are not really sure what to do with the parton densities we resort to simpler e^+e^- collisions. A simple observable which includes α_s at least in the one-loop corrections is

$$R = \frac{\sigma(e^+e^- \to \text{hadrons})}{\sigma(e^+e^- \to \mu^+\mu^-)} = N_c \sum_{\text{quarks}} Q_q^2 = \frac{11 N_c}{9} . \tag{4.53}$$

The numerical value at leading order assumes five quarks. Including higher order corrections we can express the result in a power series in α_s. In the $\overline{\text{MS}}$ scheme we introduce an unphysical scale dependence on M in the individual r_n

$$R\left(\frac{p^2}{M^2},\alpha_s\right) = \sum_{n=0} r_n\left(\frac{p^2}{M^2}\right)\,\alpha_s^n(M^2) \qquad \text{with} \qquad r_0 = \frac{11N_c}{9}\,. \tag{4.54}$$

The r_n we can assume to be dimensionless. This implies that the calculated r_n only depend on ratios of two scales, the external p^2 and the artificial M^2.

At the same time R is an observable, so including all orders in perturbation theory it cannot depend on any artificial scale M. Writing this dependence as a total derivative and setting it to zero we find an equation which would be called a Callan–Symanzik equation if instead of the running coupling we had included a running mass

$$\begin{aligned}
0 &\stackrel{!}{=} M^2\frac{d}{dM^2}R\left(\frac{p^2}{M^2},\alpha_s(M^2)\right) \\
&= \left[M^2\frac{\partial}{\partial M^2}+\beta\,\frac{\partial}{\partial\alpha_s}\right]\sum_{n=0} r_n\left(\frac{p^2}{M^2}\right)\,\alpha_s^n && \text{with} \quad \beta = M^2\frac{\partial\alpha_s}{\partial M^2} \\
&= \sum_{n=1} M^2\frac{\partial r_n}{\partial M^2}\,\alpha_s^n + \sum_{n=1}\beta\, r_n\, n\alpha_s^{n-1} && \text{with} \quad r_0 = \frac{11N_c}{9} = \text{const} \\
&= M^2\sum_{n=1}\frac{\partial r_n}{\partial M^2}\,\alpha_s^n - \sum_{n=1}\sum_{m=0} n r_n\,\alpha_s^{n+m+1}\,b_m && \text{with} \quad \beta = -\alpha_s^2\sum_{m=0} b_m\alpha_s^m \\
&= M^2\frac{\partial r_1}{\partial M^2}\,\alpha_s + \left(M^2\frac{\partial r_2}{\partial M^2} - r_1 b_0\right)\alpha_s^2 + \left(M^2\frac{\partial r_3}{\partial M^2} - 2r_2 b_0 - r_1 b_1\right)\alpha_s^3 + \mathcal{O}(\alpha_s^4)\,.
\end{aligned} \tag{4.55}$$

This perturbative series in α_s has to vanish in each order of perturbation theory,

$$\begin{aligned}
\frac{\partial r_1}{\partial\log M^2} &= 0 \\
\frac{\partial r_2}{\partial\log M^2} &= r_1 b_0 \\
\frac{\partial r_3}{\partial\log M^2} &= r_1 b_1 + 2r_2(M^2) b_0 \\
&\vdots
\end{aligned} \tag{4.56}$$

The mix of r_n derivatives and the perturbative terms in the β function can be seen for α_s^3: first, we have the appropriate NNNLO corrections r_3; next, we have one loop in the gluon propagator b_0 and two loops for example in the vertex r_2; and finally, we need the two-loop diagram for the gluon propagator b_1 and a one-loop vertex correction r_1.

The M^2-dependence vanishes for r_0 and r_1. Keeping in mind the integration constants c_n we find the solutions

$$\begin{aligned}
r_0 &= c_0 = \frac{11N_c}{9} \\
r_1 &= c_1 \\
r_2 &= c_2 + r_1 b_0\log\frac{M^2}{p^2} = c_2 + c_1 b_0\log\frac{M^2}{p^2} \\
r_3 &= \int d\log\frac{M'^2}{p^2}\left(c_1 b_1 + 2\left(c_2 + c_1 b_0\log\frac{M'^2}{p^2}\right)b_0\right) = c_3 + (c_1 b_1 + 2c_2 b_0)\log\frac{M^2}{p^2} + c_1 b_0^2\log^2\frac{M^2}{p^2} \\
&\vdots
\end{aligned} \tag{4.57}$$

This chain of r_n values suggests to interpret the fixed-order perturbative series in Eq.(4.54) as implicitly including terms $\log^{n-1} M^2/p^2$ in each r_n. They become problematic if the logarithm becomes large enough to spoil the

convergence in terms of $\alpha_s \sim 0.1$, *i.e.* measuring R at scales p^2 far away from the scale choice for the strong coupling constant. M^2.

Interestinglty, we can use Eq.(4.57) to express R in terms of the c_n,

$$R = \sum_n r_n \left(\frac{p^2}{M^2}\right) \alpha_s^n(M^2) = c_0 + c_1 \left(1 + \alpha_s(M^2) b_0 \log \frac{M^2}{p^2} + \alpha_s^2(M^2) b_0^2 \log^2 \frac{M^2}{p^2} + \cdots\right) \alpha_s(M^2) + c_2 \left(1 + 2\alpha_s(M^2) b_0 \log \frac{M^2}{p^2} + \cdots\right) \alpha_s^2(M^2) + \cdots \tag{4.58}$$

We can resum this geometric series to

$$R = c_0 + c_1 \frac{\alpha_s(M^2)}{1 - \alpha_s(M^2) b_0 \log \frac{M^2}{p^2}} + c_2 \left(\frac{\alpha_s(M^2)}{1 - \alpha_s(M^2) b_0 \log \frac{M^2}{p^2}}\right)^2 + \cdots \equiv \sum c_n \, \alpha_s^n(p^2) \, . \tag{4.59}$$

In the last step we use Eq.(4.48) with flipped arguments p^2 and M^2, derived from the resummation of the vacuum polarization bubbles. In contrast to the r_n, the c_n are by definition independent of p^2/M^2 and therefore more suitable as a perturbative series in the presence of potentially large logarithms. This re-organization of the perturbation series for R can be interpreted as resumming all logarithms of the kind $\log M^2/p^2$ in a new organization of the perturbative series. Some higher–order factors c_n are known, for example inserting $N_c = 3$ and five quark flavors just as we assume in Eq.(4.53)

$$R = \frac{11}{3} \left(1 + \frac{\alpha_s(p^2)}{\pi} + 1.4 \left(\frac{\alpha_s(p^2)}{\pi}\right)^2 - 12 \left(\frac{\alpha_s(p^2)}{\pi}\right)^3 + \mathcal{O}\left(\frac{\alpha_s(p^2)}{\pi}\right)^4\right) \, . \tag{4.60}$$

This alternating series with increasing perturbative prefactors indicates the asymptotic instead of convergent behavior of perturbative QCD. At the bottom mass scale the relevant coupling factor is only $\alpha_s(m_b^2)/\pi \sim 1/14$, so a further increase of the c_n would become dangerous. However, a detailed look into the calculation shows that the dominant contributions to c_n arise from the analytic continuation of logarithms, which are large finite terms for example from $\text{Re}(\log^2(-E^2)) = \log^2 E^2 + \pi^2$. In the literature such π^2 terms arising from the analytic continuation of loop integrals are often phrased in terms of $\zeta_2 = \pi^2/6$.

Before moving on we collect the logic of the argument given in this section: when we regularize an ultraviolet divergence we automatically introduce a reference scale μ_R. Naively, this could be an ultraviolet cutoff scale, but even the seemingly scale invariant dimensional regularization in the conformal limit of our field theory cannot avoid the introduction of a scale. There are several ways of dealing with such a scale: first, we can renormalize our parameter at a reference point. Secondly, we can define a running parameter and this way absorb the scale logarithm into the $\overline{\text{MS}}$ counter term. For the asymptotically free strong coupling Λ_{QCD} leaves us with a compact form of the running coupling $\alpha_s(M^2)$.

Strictly speaking, at each order in perturbation theory the scale dependence should vanish together with the ultraviolet poles, as long as there is only one scale affecting a given observable. However, defining the running strong coupling we sum one-loop vacuum polarization graphs. Even when we compute an observable at a given loop order, we implicitly include higher order contributions. They lead to a dependence of our perturbative result on the artificial scale M^2, which phenomenologists refer to as renormalization scale dependence.

Using the R ratio we see what our definition of the running coupling means in terms of resumming logarithms: reorganizing our perturbative series to get rid of the ultraviolet divergence $\alpha_s(p^2)$ resums the scale logarithms $\log p^2/M^2$ to all orders in perturbation theory.

5 Partons and DGLAP equation

During our introduction to QCD and the running coupling we focused on the ultraviolet or high-energy behavior of the theory and the running coupling relating subtracted divergences to resummed logarithms. In this section we will follow a similar approach to infrared divergences, where we will find that collinear divergences lead to the DGLAP equations for parton densities, which resum collinear logarithms.

5.1 Incoming partons

To predict for instance the Drell–Yan process at a hadron collider, we need to introduce parton distribution functions, describing the probability of finding a collinear parton with momentum fraction x in a proton. A pdf is not an observable, only a distribution in the mathematical sense: it has to produce reasonable results when we integrate it together with a test function. Different parton densities have very different behavior — for the valence quarks (uud) they peak (quite a bit) below $x < 1/3$, while the gluon pdf is small at $x \sim 1$ and grows very rapidly towards small x. For some typical part of the relevant parameter space ($x = 10^{-3} \cdots 10^{-1}$) the gluon density roughly scales like $f_g(x) \propto x^{-2}$. Towards smaller x values it becomes even steeper.

While we cannot compute parton distribution functions $f_i(x)$ as a function of the momentum fraction x there are a few predictions we can make based on symmetries and properties of the hadrons, leading to sum rules:

1. The parton distributions in an antiproton are linked to those inside a proton through the CP-symmetry, which is exact for QCD. Therefore,
$$f_q^{\bar{p}}(x) = f_{\bar{q}}(x) \qquad f_{\bar{q}}^{\bar{p}}(x) = f_q(x) \qquad f_g^{\bar{p}}(x) = f_g(x) \; . \tag{5.1}$$

2. If the proton consists of three valence quarks uud, plus quantum fluctuations from the vacuum which can either involve gluons or quark–antiquark pairs, the contribution from the sea quarks has to be symmetric in quarks and antiquarks. The expectation values for the signed numbers of up and down quarks inside a proton have to fulfill
$$\langle N_u \rangle = \int_0^1 dx \; (f_u(x) - f_{\bar{u}}(x)) = 2 \qquad \langle N_d \rangle = \int_0^1 dx \; (f_d(x) - f_{\bar{d}}(x)) = 1 \; . \tag{5.2}$$

3. The total momentum of the proton has to consist of sum of all parton momenta. We can write this as the expectation value
$$\left\langle \sum x_i \right\rangle = \int_0^1 dx \; x \; \left(\sum_q f_q(x) + \sum_{\bar{q}} f_{\bar{q}}(x) + f_g(x) \right) = 1 \; . \tag{5.3}$$
What makes this prediction interesting is that we can compute the same sum only taking into account the measured quark and antiquark parton densities. We find
$$\int_0^1 dx \; x \; \left(\sum_q f_q(x) + \sum_{\bar{q}} f_{\bar{q}}(x) \right) \approx \frac{1}{2} \; . \tag{5.4}$$
Half of the proton momentum is then carried by gluons.

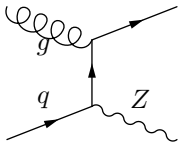
With this pdf we can compute a hadronic cross section from its partonic counterpart,
$$\boxed{\sigma_{\text{tot}} = \int_0^1 dx_1 \int_0^1 dx_2 \; \sum_{ij} f_i(x_1) \, f_j(x_2) \; \hat{\sigma}_{ij}(x_1 x_2 S)} \; , \tag{5.5}$$
where i, j are the incoming partons with the momentum factions $x_{i,j}$. The partonic energy of the scattering process is $s = x_1 x_2 S$ with the LHC proton energy of $\sqrt{S} = 13.6$ TeV. The partonic cross section $\hat{\sigma}$ includes all the necessary θ and δ functions for energy–momentum conservation. When we express a general n–particle cross section $\hat{\sigma}$ including the phase space integration, the x_i integrations and the phase space integrations can of course be interchanged, but Jacobians will make life hard.

5.2 Infrared divergences

Let us look at the radiation of additional partons in the Drell–Yan process. We can start for example by computing the cross section for the partonic process

$$q\bar{q} \to Zg \; . \tag{5.6}$$

This partonic process involves renormalization of ultraviolet divergences as well as loop diagrams which we have to include before we can say anything reasonable, *i.e.* ultraviolet and infrared finite. To make life easier we study collinear infrared divergences for the crossed process



It should behave like any other $(2 \to 2)$ jet radiation process, except that it has a different incoming state than the leading order Drell–Yan process and hence does not involve virtual corrections. This means we do not have to deal with ultraviolet divergences and renormalization, and can concentrate on parton or jet radiation from the initial state.

The amplitude for this $(2 \to 2)$ process is — modulo charges and averaging factors, but including all Mandelstam variables

$$|\mathcal{M}|^2 \sim -\frac{t}{s} - \frac{s^2 - 2m_Z^2(s+t-m_Z^2)}{st} \; . \tag{5.7}$$

The Mandelstam variable t for one massless final–state particle can be expressed in terms of the rescaled emission angle

$$t = -s(1-\tau)y \qquad \text{with} \qquad y = \frac{1-\cos\theta}{2} \in [0,1] \qquad \text{and} \qquad \tau = \frac{m_Z^2}{s} < 1 \; . \tag{5.8}$$

Similarly, we obtain $u = -s(1-\tau)(1-y)$, so as a first check we can confirm that $t + u = -s(1-\tau) = -s + m_Z^2$. The collinear limit when the gluon splits in the beam direction is given by

$$\begin{aligned} y \to 0 \qquad &\Leftrightarrow \qquad t \to 0 \qquad \Leftrightarrow \qquad u = -s + m_Z^2 < 0 \\ |\mathcal{M}|^2 &\to \frac{s^2 - 2sm_Z^2 + 2m_Z^4}{s(s-m_Z^2)} \; \frac{1}{y} + \mathcal{O}(y^0) \; . \end{aligned} \tag{5.9}$$

This expression is divergent for collinear gluon radiation or gluon splitting, *i.e.* for small angles y. We can translate this $1/y$ divergence for example into the transverse momentum of the gluon or Z

$$sp_T^2 = tu = s^2(1-\tau)^2 \; y(1-y) = (s-m_Z^2)^2 y + \mathcal{O}(y^2) \tag{5.10}$$

In terms of p_T, the collinear limit our matrix element squared in Eq.(5.9) becomes

$$|\mathcal{M}|^2 \sim \frac{s^2 - 2sm_Z^2 + 2m_Z^4}{s^2} \; \frac{s-m_Z^2}{p_T^2} + \mathcal{O}(p_T^0) \; . \tag{5.11}$$

The matrix element for the tree level process $qg \to Zq$ has a leading divergence proportional to $1/p_T^2$. To compute the total cross section for this process we need to integrate the matrix element over the two-particle phase space. Approximating the matrix element as C'/y or C/p_T^2 this gives us

$$\int_{y^{\min}}^{y^{\max}} dy \frac{C'}{y} = \int_{p_T^{\min}}^{p_T^{\max}} dp_T^2 \frac{C}{p_T^2} = \; 2 \int_{p_T^{\min}}^{p_T^{\max}} dp_T \; p_T \; \frac{C}{p_T^2} \simeq 2C \int_{p_T^{\min}}^{p_T^{\max}} dp_T \frac{1}{p_T} = 2C \; \log \frac{p_T^{\max}}{p_T^{\min}} \tag{5.12}$$

The form C/p_T^2 for the matrix element is of course only valid in the collinear limit; in the non–collinear phase space C is not a constant.

For this divergence we can follow the same strategy as for the ultraviolet divergence. First, we regularize it for example using dimensional regularization. Then, we find a well–defined way to get rid of it. Dimensional regularization means writing the two-particle phase space in $n = 4 - 2\epsilon$ dimensions. Just for reference, the complete formula for the y-distribution reads

$$s\,\frac{d\sigma}{dy} = \frac{\pi(4\pi)^{-2+\epsilon}}{\Gamma(1-\epsilon)}\,\left(\frac{\mu_F^2}{m_Z^2}\right)^\epsilon\,\frac{\tau^\epsilon(1-\tau)^{1-2\epsilon}}{y^\epsilon(1-y)^\epsilon}\,|\mathcal{M}|^2 \sim \left(\frac{\mu_F^2}{m_Z^2}\right)^\epsilon\,\frac{|\mathcal{M}|^2}{y^\epsilon(1-y)^\epsilon}\,. \tag{5.13}$$

In the second step we only keep the factors we are interested in. The additional factor $1/y^\epsilon$ regularizes the integral at $y \to 0$, as long as $\epsilon < 0$ by slightly increasing the suppression of the integrand in the infrared regime. This means that for infrared divergences we choose $n = 4 + 2|\epsilon|$ space–time dimensions. After integrating the leading collinear divergence $1/y^{1+\epsilon}$ we are left with a pole $1/(-\epsilon)$.

What is important to notice is again the appearance of a scale $\mu_F^{2\epsilon}$ with the n-dimensional integral. Now it arises from an infrared regularization and is referred to as factorization scale. The actual removal of the infrared pole — corresponding to the renormalization in the ultraviolet case — is called mass factorization and works exactly the same way as renormalizing a parameter: in a well–defined scheme we subtract the pole from the fixed-order matrix element squared.

5.3 Parton splitting

Infrared divergences occur for massless particles in the initial or final state, so we need to go through all ways incoming or outgoing gluons and quark can split into each other. The factorized phase space is common to all different channels. The first and at the LHC most important case is the splitting of one gluon into two,

$$g(p_a) \to g(p_b) + g(p_c) \qquad \text{with} \qquad p_a^2 \gg p_b^2, p_c^2\,. \tag{5.14}$$

The two daughter gluons are close to mass shell while the mother has to have a finite positive invariant mass. We assign the direction of the momenta as $p_a = -p_b - p_c$ and describe the kinematics of this approximately collinear process in terms of the energy fractions z and $1-z$ defined as

$$\begin{aligned} z = \frac{|E_b|}{|E_a|} = 1 - \frac{|E_c|}{|E_a|} \qquad & p_a^2 = (-p_b-p_c)^2 = 2(p_b p_c) = 2z(1-z)(1-\cos\theta)E_a^2 = z(1-z)E_a^2\theta^2 + \mathcal{O}(\theta^4) \\ \Leftrightarrow \qquad & \theta \equiv \theta_b + \theta_c \simeq \frac{1}{|E_a|}\,\sqrt{\frac{p_a^2}{z(1-z)}}\,, \end{aligned} \tag{5.15}$$

in the collinear limit and in terms of the opening angle θ between $\vec{p}_b$ and $\vec{p}_c$. Using this phase space parameterization we divide an $(n+1)$-particle process into an n-particle process and a splitting process of quarks and gluons. The phase space factorization is easy to define when we look at the splitting of an outgoing gluon as part of an n-particle hard production process. This turns this process into $(n+1)$-particle production, with the two gluons in the final state, $g(p_b)$ and $g(p_c)$. The complete phase space integration has the form

$$\begin{aligned} d\Phi_{n+1} &= \cdots \frac{d^3\vec{p}_b}{2(2\pi)^3|E_b|}\,\frac{d^3\vec{p}_c}{2(2\pi)^3|E_c|} = \cdots \frac{d^3\vec{p}_a}{2(2\pi)^3|E_a|}\,\frac{d^3\vec{p}_c}{2(2\pi)^3|E_c|}\,\frac{|E_a|}{|E_b|} \\ &\equiv d\Phi_n\,\frac{dp_{c,3}dp_T p_T d\phi}{2(2\pi)^3|E_c|}\,\frac{1}{z} \\ &= d\Phi_n\,\frac{dp_{c,3}dp_T^2 d\phi}{4(2\pi)^3|E_c|}\,\frac{1}{z} \end{aligned} \tag{5.16}$$

We can separate the $(n+1)$-particle space into an n-particle phase space and a $(1 \to 2)$ splitting phase space without any approximation.

Our next task is to translate $p_{c,3}$ and p_T^2 into z and $p_a^2 \neq 0$. This can be done if we assume approximately collinear collinear splittings, where we find that

$$\frac{dp_{c,3}}{|E_c|} = \frac{dz}{1-z}\,(1+\mathcal{O}(\theta)) \qquad \text{and} \qquad dp_T^2 = z(1-z)dp_a^2\,. \tag{5.17}$$

This gives us the final result for the separated collinear phase space, assuming azimuthal symmetry in an addition step,

$$d\Phi_{n+1} = d\Phi_n\,\frac{dz\,dp_a^2\,d\phi}{4(2\pi)^3}\,(1+\mathcal{O}(\theta)) = d\Phi_n\,\frac{dz\,dp_a^2}{4(2\pi)^2}\,(1+\mathcal{O}(\theta))\,. \tag{5.18}$$

Adding the transition matrix elements we can write a full factorization in the collinear approximation as

$$\begin{aligned}
d\sigma_{n+1} &= \overline{|\mathcal{M}_{n+1}|^2}\; d\Phi_{n+1} \\
&= \overline{|\mathcal{M}_{n+1}|^2}\; d\Phi_n \frac{dp_a^2\,dz}{4(2\pi)^2}\;\;(1+\mathcal{O}(\theta)) \\
&\simeq \frac{2g_s^2}{p_a^2}\;\hat{P}(z)\;\overline{|\mathcal{M}_n|^2}\; d\Phi_n \frac{dp_a^2\,dz}{16\pi^2} \qquad \text{assuming} \quad \boxed{\overline{|\mathcal{M}_{n+1}|^2} \simeq \frac{2g_s^2}{p_a^2}\;\hat{P}(z)\;\overline{|\mathcal{M}_n|^2}} \\
&= d\sigma_n\;\frac{dp_a^2}{p_a^2}\;dz\;\frac{\alpha_s}{2\pi}\hat{P}(z)\,.
\end{aligned} \tag{5.19}$$

For splitting incoming partons we replace $p_a^2 \to t$, the usual Mandelstam variable. We can show this assumed factorization by constructing the appropriate splitting kernels $\hat{P}(z)$ for all quark and gluon configurations:

- First comes gluon splitting into two gluons. To compute its transition amplitude we need to write down all gluon momenta and polarizations in a specific frame. We skip the derivation and just quote the result

$$\begin{aligned}
\overline{|\mathcal{M}_{n+1}|^2} &= \frac{2g_s^2}{p_a^2}\;\frac{N_c}{2}\;2\left[\frac{z}{1-z} + z(1-z) + \frac{1-z}{z}\right]\;\overline{|\mathcal{M}_n|^2} \\
&\equiv \frac{2g_s^2}{p_a^2}\;\hat{P}_{g\leftarrow g}(z)\;\overline{|\mathcal{M}_n|^2} \\
&\Leftrightarrow \qquad \hat{P}_{g\leftarrow g}(z) = C_A\left[\frac{z}{1-z} + \frac{1-z}{z} + z(1-z)\right]\,,
\end{aligned} \tag{5.20}$$

  using $C_A = N_c$. The splitting kernel is symmetric when we exchange the two gluons z and $(1-z)$. It diverges if either gluon becomes soft. The notation $\hat{P}_{i\leftarrow j} \sim \hat{P}_{ij}$ is inspired by a matrix notation which we can use to multiply the splitting matrix from the right with the incoming parton vector to get the final parton vector.

- A second kernel describes the splitting of a gluon into two quarks. Again, we omit the calculation and quote

$$\hat{P}_{q\leftarrow g}(z) = T_R\left[z^2 + (1-z)^2\right]\,. \tag{5.21}$$

  It is symmetric under $z \leftrightarrow (1-z)$ because QCD does not distinguish between the outgoing quark and antiquark.

- The third splitting is gluon radiation off a quark line,

$$\hat{P}_{q\leftarrow q}(z) = C_F\frac{1+z^2}{1-z}\,. \tag{5.22}$$

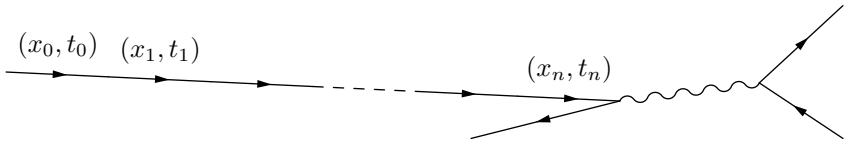
- Just switching $z \leftrightarrow (1-z)$ we can read off the kernel for a quark splitting into the final–state gluon

$$\hat{P}_{g\leftarrow q}(z) = C_F\frac{1+(1-z)^2}{z}\,. \tag{5.23}$$

Similar to ultraviolet divergences these splitting kernels are universal. Crucially, they do not depend on the hard n-particle matrix element as part of the full $(n+1)$-particle process.

5.4 DGLAP equation

To describe successive splittings we start with a quark inside the proton with an energy fraction x_0, as it enters the hadronic phase space integral. As this quark is confined inside the proton, it can only have small transverse momentum, which means its four-momentum squared t_0 is negative and its absolute value is small, $|t_0| = p^2 \approx 0$. Let us simplify our argument by assuming that there exists only successive gluon radiation off an incoming quark,



In that case each collinear gluon radiation will decrease the quark energy and increase its virtuality through recoil,

$$x_{j+1} < x_j \qquad \text{and} \qquad |t_{j+1}| = -t_{j+1} > -t_j = |t_j| \; . \tag{5.24}$$

We know what the successive splitting means in terms of splitting probabilities and can describe how the parton density $f(x, -t)$ evolves in the (x, t) plane as depicted in Figure 1. We start at (x_0, t_0) and interpret each branching as a step down in $x_j \to x_{j+1}$ and a step up in $|t_{j+1}|$. The splitting path in the $(x - t)$ plane is marked by discrete points. The probability of a splitting to occur is given by Eq.(5.19),

$$\frac{\alpha_s}{2\pi} \, \hat{P}(z) \, \frac{dt}{t} \, dz \equiv \frac{\alpha_s}{2\pi} \, \hat{P}_{q \leftarrow q}(z) \, \frac{dt}{t} \, dz \; . \tag{5.25}$$

The evolved parton density at (x_n, t_n) enters the hard scattering, including its energy–momentum conservation.

The link between partonic and hadronic cross section is described by the probability or parton density $f(x, -t)$ over an infinitesimal square,

$$[x_j, x_j + \delta x] \qquad \text{and} \qquad [|t_j|, |t_j| + \delta t] \; . \tag{5.26}$$

In our (x, t) plane we can compute the flows into this square and out of this square, which together define the net change in f in the sense of a differential equation,

$$\delta f_{\text{in}} - \delta f_{\text{out}} = \delta f(x, -t) \; . \tag{5.27}$$

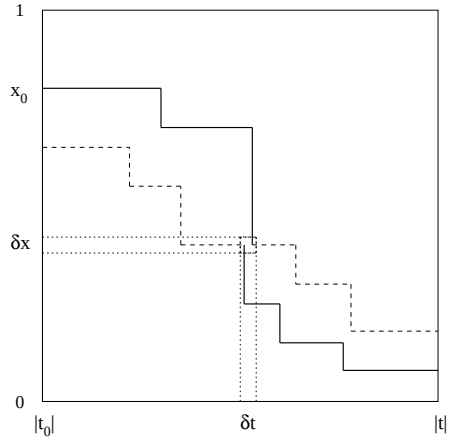


Figure 1: Path of an incoming parton in the $(x - t)$ plane. Because we define t as a negative number its axis is labelled $|t|$.

We compute the incoming and outgoing flows from the history of the (x,t) evolution. At this stage our picture becomes a little subtle; the way we define the path between two splittings in Figure 1 it can enter and leave the square either vertically or horizontally. Because we want to arrive at a differential equation in t we choose the vertical drop, such that the area the incoming and outgoing flows see is given by δt. If we define a splitting as a vertical drop in x at the target value t_{j+1}, an incoming path hitting the square can come from any x-value above the square. Using this convention and following the fat solid lines in Figure 1 the vertical flow into the square (x,t) is

$$\begin{aligned}
\delta f_{\text{in}}(-t) &= \delta t \left(\frac{\alpha_s \hat{P}}{2\pi t} \otimes f\right)(x,-t) \\
&= \frac{\delta t}{t} \int_x^1 \frac{dz}{z} \frac{\alpha_s}{2\pi} \hat{P}(z) f\left(\frac{x}{z},-t\right) \\
&\equiv \frac{\delta t}{t} \int_0^1 \frac{dz}{z} \frac{\alpha_s}{2\pi} \hat{P}(z) f\left(\frac{x}{z},-t\right) \qquad \text{assuming } f(x',-t)=0 \text{ for } x'>1 \,. \qquad (5.28)
\end{aligned}$$

We use the definition of a convolution

$$(f\otimes g)(x) = \int_0^1 dx_1 dx_2 f(x_1) g(x_2)\, \delta(x-x_1 x_2) = \int_0^1 \frac{dx_1}{x_1} f(x_1) g\left(\frac{x}{x_1}\right) = \int_0^1 \frac{dx_2}{x_2} f\left(\frac{x}{x_2}\right) g(x_2) \,. \qquad (5.29)$$

The outgoing flow also leaves the infinitesimal square vertically, again following the fat solid line in Figure 1

$$\delta f_{\text{out}}(-t) = \delta t \int_0^1 dy \frac{\alpha_s \hat{P}(y)}{2\pi t} f(x,-t) = \frac{\delta t}{t} f(x,-t) \int_0^1 dy \frac{\alpha_s}{2\pi} \hat{P}(y) \,. \qquad (5.30)$$

The y-integration is not a convolution, because we know the starting condition and integrate over all final configurations. Combining Eq.(5.28) and Eq.(5.30) we can compute the change in the quark pdf as

$$\begin{aligned}
\delta f(x,-t) = \delta f_{\text{in}} - \delta f_{\text{out}} &= \frac{\delta t}{t} \left[\int_0^1 \frac{dz}{z} \frac{\alpha_s}{2\pi} \hat{P}(z)\, f\left(\frac{x}{z},-t\right) - \int_0^1 dy \frac{\alpha_s}{2\pi} \hat{P}(y)\, f(x,-t)\right] \\
&= \frac{\delta t}{t} \int_0^1 \frac{dz}{z} \frac{\alpha_s}{2\pi} \left[\hat{P}(z) - \delta(1-z) \int_0^1 dy \hat{P}(y)\right] f\left(\frac{x}{z},-t\right) \\
&\equiv \frac{\delta t}{t} \int_x^1 \frac{dz}{z} \frac{\alpha_s}{2\pi} \hat{P}(z)_+ f\left(\frac{x}{z},-t\right) \\
\Leftrightarrow \qquad \frac{\delta f(x,-t)}{\delta(-t)} &= \frac{1}{(-t)} \int_x^1 \frac{dz}{z} \frac{\alpha_s}{2\pi} \hat{P}(z)_+ f\left(\frac{x}{z},-t\right) \qquad (5.31)
\end{aligned}$$

Strictly speaking, we require α_s to only depend on t and introduce the plus subtraction

$$F(z)_+ \equiv F(z) - \delta(1-z) \int_0^1 dy\, F(y) \qquad \text{or} \qquad \int_0^1 dz \frac{f(z)}{(1-z)_+} = \int_0^1 dz \left(\frac{f(z)}{1-z} - \frac{f(1)}{1-z}\right) \,. \qquad (5.32)$$

For the second definition we choose $F(z) = 1/(1-z)$, multiply it with an arbitrary test function $f(z)$ and integrate over z. The plus–subtracted integral is by definition finite in the soft limit $z \to 1$, where some splitting kernels diverge. It is related to dimensional regularization, defined as

$$\int_0^1 dz \frac{1}{(1-z)^{1-\epsilon}} = \int_0^1 dz \frac{1}{z^{1-\epsilon}} = \left.\frac{z^\epsilon}{\epsilon}\right|_0^1 = \frac{1}{\epsilon} \qquad \text{with } \epsilon > 0 \,, \qquad (5.33)$$

corresponding to $4+2\epsilon$ dimensions. We can relate the dimensionally regularized integral to the plus subtraction as

$$\begin{aligned}\int_0^1 dz\ \frac{f(z)}{(1-z)^{1-\epsilon}} &= \int_0^1 dz\ \frac{f(z)-f(1)}{(1-z)^{1-\epsilon}} + f(1)\int_0^1 dz\ \frac{1}{(1-z)^{1-\epsilon}} \\ &= \int_0^1 dz\ \frac{f(z)-f(1)}{1-z}\left(1+\mathcal{O}(\epsilon)\right) + \frac{f(1)}{\epsilon} \\ &= \int_0^1 dz\ \frac{f(z)}{(1-z)_+}\left(1+\mathcal{O}(\epsilon)\right) + \frac{f(1)}{\epsilon} \qquad \text{by definition} \\ \Leftrightarrow \qquad \int_0^1 dz\ \frac{f(z)}{(1-z)^{1-\epsilon}} - \frac{f(1)}{\epsilon} &= \int_0^1 dz\ \frac{f(z)}{(1-z)_+}\left(1+\mathcal{O}(\epsilon)\right)\ . \end{aligned} \tag{5.34}$$

The dimensionally regularized integral minus the pole, *i.e.* the finite part of the dimensionally regularized integral, is the same as the plus–subtracted integral modulo terms of the order ϵ. This means that the plus-subtraction plays a similar role as $\overline{\text{MS}}$ renormalization for UV-divergences.

To regularize our splitting kernel $\hat{P}_{q\leftarrow q}$ in Eq.(5.22) we can, actually, define two subtraction schemes,

$$\begin{aligned}\left(\frac{1+z^2}{1-z}\right)_+ - (1+z^2)\ \left(\frac{1}{1-z}\right)_+ &= \frac{1+z^2}{1-z} - \delta(1-z)\int_0^1 dy\ \frac{1+y^2}{1-y} - \frac{1+z^2}{1-z} + \delta(1-z)\int_0^1 dy\ \frac{1+z^2}{1-y} \\ &= -\delta(1-z)\int_0^1 dy\ \left(\frac{1+y^2}{1-y} - \frac{2}{1-y}\right) \\ &= \delta(1-z)\int_0^1 dy\ \frac{y^2-1}{y-1} = \delta(1-z)\int_0^1 dy\ (y+1) = \frac{3}{2}\delta(1-z)\ . \end{aligned} \tag{5.35}$$

This means we can write the quark splitting kernel in two equivalent ways, where $\delta(1-z)$ implies that the difference is a contribution to a soft–radiation phase space integral. We simplify the notation as $\hat{P}_+ \equiv P$ and write

$$\begin{aligned} P_{q\leftarrow q}(z) &\equiv C_F\left(\frac{1+z^2}{1-z}\right)_+ = C_F\left[\frac{1+z^2}{(1-z)_+} + \frac{3}{2}\delta(1-z)\right] \\ \Rightarrow \qquad \boxed{\frac{df_q(x,-t)}{d\log(-t)} = -t\ \frac{df_q(x,-t)}{d(-t)} &= \int_x^1 \frac{dz}{z}\ \frac{\alpha_s}{2\pi}\ P_{q\leftarrow q}(z)\ f_q\left(\frac{x}{z},-t\right)}\ . \end{aligned} \tag{5.36}$$

This is the Dokshitzer–Gribov–Lipatov–Altarelli–Parisi or DGLAP equation for quarks radiating gluons. We have to generalize this DGLAP equation to quarks and gluons. For the quark density on the left hand side we write

$$\frac{df_q(x,-t)}{d\log(-t)} = -t\ \frac{df_q(x,-t)}{d(-t)} = \sum_{j=q,g}\int_x^1 \frac{dz}{z}\ \frac{\alpha_s}{2\pi}\ P_{q\leftarrow j}(z)\ f_j\left(\frac{x}{z},-t\right)\ . \tag{5.37}$$

The splitting from gluon to quark can be included as

$$\delta f_q(x,-t) = \frac{\delta t}{t}\ \left[\int_0^1 \frac{dz}{z}\ \frac{\alpha_s}{2\pi}\ \hat{P}_{q\leftarrow q}(z)\ f_q\left(\frac{x}{z},-t\right) + \int_0^1 \frac{dz}{z}\ \frac{\alpha_s}{2\pi}\ \hat{P}_{q\leftarrow g}(z)\ f_g\left(\frac{x}{z},-t\right)\right]\ . \tag{5.38}$$

The second term is a convolution proportional to the gluon pdf. Quarks can be produced in gluon splitting but cannot vanish into it. Therefore Eq.(5.38) $P_{q\leftarrow g}$ does not include a plus–regulator

$$P_{q\leftarrow g}(z) \equiv \hat{P}_{q\leftarrow g}(z) = T_R\left[z^2+(1-z)^2\right]\ . \tag{5.39}$$

This kernel is indeed missing a soft-radiation divergence for $z\to 1$.

The second parton density we have to study is the gluon density. The incoming contribution to the infinitesimal square is given by the sum of four splitting scenarios each leading to a gluon with virtuality $-t_{j+1}$

$$\begin{aligned}\delta f_{\text{in}}(-t) &= \frac{\delta t}{t}\int_0^1 \frac{dz}{z}\ \frac{\alpha_s}{2\pi}\ \left[\hat{P}_{g\leftarrow g}(z)\left(f_g\left(\frac{x}{z},-t\right) + f_g\left(\frac{x}{1-z},-t\right)\right) + \hat{P}_{g\leftarrow q}(z)\left(f_q\left(\frac{x}{z},-t\right) + f_{\bar{q}}\left(\frac{x}{z},-t\right)\right)\right] \\ &= \frac{\delta t}{t}\int_0^1 \frac{dz}{z}\ \frac{\alpha_s}{2\pi}\ \left[2\hat{P}_{g\leftarrow g}(z) f_g\left(\frac{x}{z},-t\right) + \hat{P}_{g\leftarrow q}(z)\left(f_q\left(\frac{x}{z},-t\right) + f_{\bar{q}}\left(\frac{x}{z},-t\right)\right)\right]\ , \end{aligned} \tag{5.40}$$

using $P_{g\leftarrow\bar{q}} = P_{g\leftarrow q}$ in the first line and $P_{g\leftarrow g}(1-z) = P_{g\leftarrow g}(z)$ in the second. To leave the volume element in (x,t)-space a gluon can either split into two gluons or radiate one of n_f light-quark flavors. Combining the incoming and outgoing flows we find

$$\begin{aligned}\delta f_g(x,-t) =& \frac{\delta t}{t}\int_0^1 \frac{dz}{z}\,\frac{\alpha_s}{2\pi}\,\left[2\hat{P}_{g\leftarrow g}(z) f_g\left(\frac{x}{z},-t\right) + \hat{P}_{g\leftarrow q}(z)\left(f_q\left(\frac{x}{z},-t\right) + f_{\bar{q}}\left(\frac{x}{z},-t\right)\right)\right] \\ & -\frac{\delta t}{t}\int_0^1 dy\,\frac{\alpha_s}{2\pi}\,\left[\hat{P}_{g\leftarrow g}(y) + n_f \hat{P}_{q\leftarrow g}(y)\right] f_g(x,-t) \end{aligned} \tag{5.41}$$

Unlike in the quark case these terms do not immediately correspond to regularizing the diagonal splitting kernel using the plus prescription.

First, the contribution to δf_{in} proportional to f_q or $f_{\bar{q}}$ which is not matched by the outgoing flow. From the quark case we already know how to deal with it. The corresponding splitting kernel does not need any regularization, so we define

$$P_{g\leftarrow q}(z) \equiv \hat{P}_{g\leftarrow q}(z) = C_F \frac{1+(1-z)^2}{z}\,. \tag{5.42}$$

We see that the structure of the DGLAP equation implies that the two off-diagonal splitting kernels do not include any plus prescription $\hat{P}_{i\leftarrow j} = P_{i\leftarrow j}$. This is expected because the kernels are finite in the soft limit, $z\to 1$.

Next, we can compute the y-integral describing the gluon splitting into a quark pair directly,

$$\begin{aligned} -\int_0^1 dy\,\frac{\alpha_s}{2\pi}\,n_f\,\hat{P}_{q\leftarrow g}(y) &= -\frac{\alpha_s}{2\pi}\,n_f\,T_R\int_0^1 dy\,\left[1-2y+2y^2\right] && \text{using Eq.(5.39)} \\ &= -\frac{\alpha_s}{2\pi}\,n_f\,T_R\left[y-y^2+\frac{2y^3}{3}\right]_0^1 \\ &= -\frac{2}{3}\,\frac{\alpha_s}{2\pi}\,n_f\,T_R\,. \end{aligned} \tag{5.43}$$

Finally, the two terms proportional to the pure gluon splitting $P_{g\leftarrow g}$ in Eq.(5.41) require some work. The y-integral from the outgoing flow has to consist of a finite term and a term we can use to define the plus prescription for $\hat{P}_{g\leftarrow g}$. The integral gives

$$\begin{aligned} -\int_0^1 dy\,\frac{\alpha_s}{2\pi}\,\hat{P}_{g\leftarrow g}(y) &= -\frac{\alpha_s}{2\pi}\,C_A\int_0^1 dy\,\left[\frac{y}{1-y}+\frac{1-y}{y}+y(1-y)\right] && \text{using Eq.(5.20)} \\ &= -\frac{\alpha_s}{2\pi}\,C_A\int_0^1 dy\,\left[\frac{2y}{1-y}+y(1-y)\right] \\ &= -\frac{\alpha_s}{2\pi}\,C_A\int_0^1 dy\,\left[\frac{2(y-1)}{1-y}+y(1-y)\right] - \frac{\alpha_s}{2\pi}\,C_A\int_0^1 dy\,\frac{2}{1-y} \\ &= -\frac{\alpha_s}{2\pi}\,C_A\int_0^1 dy\,\left[-2+y-y^2\right] - \frac{\alpha_s}{2\pi}\,2C_A\int_0^1 dz\,\frac{1}{1-z} \\ &= -\frac{\alpha_s}{2\pi}\,C_A\left[-2+\frac{1}{2}-\frac{1}{3}\right] - \frac{\alpha_s}{2\pi}\,2C_A\int_0^1 dz\,\frac{1}{1-z} \\ &= \frac{\alpha_s}{2\pi}\,\frac{11}{6}\,C_A - \frac{\alpha_s}{2\pi}\,2C_A\int_0^1 dz\,\frac{1}{1-z}\,. \end{aligned} \tag{5.44}$$

The second term in this result is what we need to replace the first term in the splitting kernel of Eq.(5.20) proportional to $1/(1-z)$ by $1/(1-z)_+$. We can see this using $f(z)=z$ and correspondingly $f(1)=1$ in Eq.(5.32). The two finite terms in Eq.(5.43) and Eq.(5.44) are included in the definition of $\hat{P}_{g\leftarrow g}$ ad hoc. Because the regularized splitting kernel appears in a convolution, the two finite terms require an explicit factor $\delta(1-z)$. Collecting all of them we arrive at

$$P_{g\leftarrow g}(z) = 2C_A\,\left(\frac{z}{(1-z)_+}+\frac{1-z}{z}+z(1-z)\right) + \frac{11}{6}\,C_A\,\delta(1-z) - \frac{2}{3}\,n_f\,T_R\,\delta(1-z)\,. \tag{5.45}$$

In its standard form and in terms of the factorization scale $\mu_F^2 \equiv -t$ the complete DGLAP equation reads

$$\boxed{\frac{df_i(x,\mu_F)}{d\log\mu_F^2} = \sum_j \int_x^1 \frac{dz}{z}\,\frac{\alpha_s}{2\pi}\,P_{i\leftarrow j}(z)\,f_j\left(\frac{x}{z},\mu_F\right) = \frac{\alpha_s}{2\pi}\sum_j \left(P_{i\leftarrow j}\otimes f_j\right)(x,\mu_F)}\,. \tag{5.46}$$

Let us briefly recapitulate: for the full quark and gluon particle content of QCD we have derived the DGLAP equation which describes a factorization scale dependence of the quark and gluon parton densities. The universality of the splitting kernels is obvious from the way we derive them — no information on the n-particle process ever enters the derivation.

The DGLAP equation is formulated in terms of four splitting kernels of gluons and quarks which are linked to the splitting probabilities, but which for the DGLAP equation have to be regularized. With the help of a plus–subtraction all kernels $P_{i\leftarrow j}(z)$ become finite, including in the soft limit $z \to 1$. However, splitting kernels are only regularized when needed, so the finite off-diagonal quark–gluon and gluon–quark splittings are unchanged. This means the plus prescription really acts as an infrared renormalization, moving universal infrared divergences into the definition of the parton densities. The original collinear divergence has vanished as well.

5.5 Solving the DGLAP equation

While it is hard to solve the DGLAP equation in Eq.(5.46) for vectors of pdfs, we can simplify our life by solving it for eigenvalues in parton space. One such eigenvalue is the non–singlet parton density,

$$f_q^{\text{NS}} = (f_q - f_{\bar{q}})\,. \tag{5.47}$$

Since gluons cannot distinguish between quarks and antiquarks, the gluon contribution to their evolution cancels, at least in the massless limit, at arbitrary loop order. The corresponding DGLAP equation is

$$\frac{df_q^{\text{NS}}(x,\mu_F)}{d\log\mu_F^2} = \int_x^1 \frac{dz}{z}\,\frac{\alpha_s}{2\pi}\,P_{q\leftarrow q}(z)\,f_q^{\text{NS}}\left(\frac{x}{z},\mu_F\right)\,. \tag{5.48}$$

To solve it we need a transformation which simplifies a convolution, leading us to the Mellin transform. Starting from a function $f(x)$ of a real variable x we define the Mellin transform into moment space m as

$$\mathcal{M}[f](m) \equiv \int_0^1 dx\,x^{m-1}f(x) \qquad\qquad f(x) = \frac{1}{2\pi i}\int_{c-i\infty}^{c+i\infty} dm\,\frac{\mathcal{M}[f](m)}{x^m}\,, \tag{5.49}$$

where for the back transformation we choose an arbitrary appropriate constant $c > 0$, such that the integration contour for the inverse transformation lies to the right of all singularities of the analytic continuation of $\mathcal{M}[f](m)$. The key property is that the Mellin transform of a convolution is the product of the two Mellin transforms, which gives us the transformed DGLAP equation

$$\begin{aligned}\frac{d\mathcal{M}[f_q^{\text{NS}}](m,\mu_F)}{d\log\mu_F^2} &= \frac{\alpha_s}{2\pi}\,\mathcal{M}\left[\int_0^1 \frac{dz}{z}P_{q\leftarrow q}\left(\frac{x}{z}\right)f_q^{\text{NS}}(z)\right](m)\\ &= \frac{\alpha_s}{2\pi}\,\mathcal{M}[P_{q\leftarrow q}\otimes f_q^{\text{NS}}](m)\\ &= \frac{\alpha_s}{2\pi}\,\mathcal{M}[P_{q\leftarrow q}](m)\,\mathcal{M}[f_q^{\text{NS}}](m,\mu_F)\,.\end{aligned} \tag{5.50}$$

In Mellin space it has the simple solution

$$\begin{aligned}\mathcal{M}[f_q^{\text{NS}}](m,\mu_F) &= \mathcal{M}[f_q^{\text{NS}}](m,\mu_{F,0}) \ \exp\left(\frac{\alpha_s}{2\pi}\ \mathcal{M}[P_{q\leftarrow q}](m)\log\frac{\mu_F^2}{\mu_{F,0}^2}\right) \\ &= \mathcal{M}[f_q^{\text{NS}}](m,\mu_{F,0}) \ \left(\frac{\mu_F^2}{\mu_{F,0}^2}\right)^{\frac{\alpha_s}{2\pi}\mathcal{M}[P_{q\leftarrow q}](m)} \\ &\equiv \mathcal{M}[f_q^{\text{NS}}](m,\mu_{F,0}) \ \left(\frac{\mu_F^2}{\mu_{F,0}^2}\right)^{\frac{\alpha_s}{2\pi}\gamma(m)} , \end{aligned} \tag{5.51}$$

defining $\gamma(m) = \mathcal{M}[P](m)$.

This solution still includes μ_F and α_s as two free parameters. To simplify this form we can include $\alpha_s(\mu_R^2)$ in the running of the DGLAP equation and identify the renormalization and factorization scales

$$\mu_F \equiv \mu_R \equiv \mu \ . \tag{5.52}$$

Physically, this identification is motivated by one-scale problems where we have no freedom to choose either of the two scales. In the DGLAP equation it allows us to replace $\log\mu^2$ by α_s using Eq.(4.52)

$$\frac{d}{d\log\mu^2} = \frac{d\log\alpha_s}{d\log\mu^2}\ \frac{d}{d\log\alpha_s} = \frac{1}{\alpha_s}\ \frac{d\alpha_s}{d\log\mu^2}\ \frac{d}{d\log\alpha_s} = -\alpha_s b_0\ \frac{d}{d\log\alpha_s}\ . \tag{5.53}$$

The pre-factor α_s cancels the factor α_s on the right hand side of the DGLAP equation Eq.(5.50)

$$\begin{aligned}\frac{d\mathcal{M}[f_q^{\text{NS}}](m,\mu)}{d\log\alpha_s} &= -\frac{1}{2\pi b_0}\gamma(m)\ \mathcal{M}[f_q^{\text{NS}}](m,\mu) \\ \mathcal{M}[f_q^{\text{NS}}](m,\mu) &= \mathcal{M}[f_q^{\text{NS}}](m,\mu_0)\ \exp\left(-\frac{1}{2\pi b_0}\ \gamma(m)\log\frac{\alpha_s(\mu^2)}{\alpha_s(\mu_0^2)}\right) \\ &= \mathcal{M}[f_q^{\text{NS}}](m,\mu_{F,0})\ \left(\frac{\alpha_s(\mu_0^2)}{\alpha_s(\mu^2)}\right)^{\frac{\gamma(m)}{2\pi b_0}} \ . \end{aligned} \tag{5.54}$$

Among other things, in this derivation we neglect that some splitting functions have singularities and therefore the Mellin transform is not obviously well defined. Our convolution is not really a convolution either, because we cut it off at Q_0^2 etc; but the final structure in Eq.(5.54) really holds.

We can find the same solution in pure Yang–Mills theory, *i.e.* in QCD without quarks. Looking at the different color factors in QCD this limit can also be derived as the leading terms in N_c. In that case there also exists only one splitting kernel defining an anomalous dimension γ, and we find the same solution as in Eq.(5.54).

Remembering how we arrive at the DGLAP equation we notice an analogy to the case of ultraviolet divergences and the running coupling. We start from universal infrared divergences. We describe them in terms of splitting functions which we regularize using the plus prescription. The DGLAP equation plays the role of a renormalization group equation for example for the running coupling. It links parton densities evaluated at different scales μ_F. In analogy to the scaling logarithms considered in Section 4.7 we should test if we can point to a type of logarithm the DGLAP equation resums by reorganizing our perturbative series of parton splitting.

5.6 Resumming collinear logarithms

In our discussion of the DGLAP equation and its solution we encounter the splitting probability in the exponent. To make sense of such a structure we remind ourselves that ratios of α_s values to some power can appear as a result of a resummed series. Such a series would need to include powers of $(\mathcal{M}[\hat{P}])^n$ summed over n which corresponds to a sum over splittings with a varying number of partons in the final state. This suggests that parton densities cannot be

formulated in terms of a fixed final state because they include effects from any number of collinear partons summed over the number of such partons, like

$$\boxed{pp \to \mu^+\mu^- + X} \qquad \text{where } X \text{ includes any number of collinear jets.} \tag{5.55}$$

The same argument leads us towards the logarithms the running parton densities re-sum. To identify them we build a physical model based on collinear splitting, but without using the DGLAP equation. We then solve it to see the resulting structure of the solutions and compare it to the structure of the DGLAP solutions in Eq.(5.54).

We start from the basic equation defining the physical picture of parton splitting in Eq.(5.19). To treat initial state splittings, we need a definition of the virtuality t. We introduce a positive transverse momentum variable

$$-t = \frac{\vec{p}_T^{\,2}}{1-z} > 0 \qquad \Rightarrow \qquad \frac{dt}{t} = \frac{d\vec{p}_T^{\,2}}{\vec{p}_T^{\,2}} \; , \tag{5.56}$$

where $\vec{p}_T^{\,2}$ is the transverse three-momentum of of the parton pair after splitting. Equation(5.19) with a running strong coupling then becomes

$$\sigma_{n+1}(x,\mu_F) \approx \int_{x_0}^1 \frac{dx_n}{x_n} \; P_{g\leftarrow g}\left(\frac{x}{x_n}\right) \sigma_n(x_n,\mu_0) \int_{\mu_0}^{\mu_F} \frac{d\vec{p}_{T,n}^{\,2}}{\vec{p}_{T,n}^{\,2}} \frac{\alpha_s(\mu_R^2)}{2\pi} \; . \tag{5.57}$$

Because the splitting kernel is infrared divergent we cut off the convolution integral at x_0. Similarly, the transverse momentum integral is bounded by an infrared cutoff μ_0 and the physical external scale μ_F. They give the range in which an additional collinear radiation is included in σ_{n+1}. For splitting the two integrals in Eq.(5.57) it is crucial that μ_0 is the only scale the matrix element σ_n depends on. The other integration variable, the transverse momentum, does not feature in σ_n because collinear factorization is defined in the limit $\vec{p}_T^{\,2} \to 0$.

If μ_F is the global upper boundary of the transverse momentum integration for collinear splitting, we can apply the recursion formula in Eq.(5.57) iteratively

$$\begin{aligned}\sigma_{n+1}(x,\mu_F) \approx & \int_{x_0}^1 \frac{dx_n}{x_n} \; P_{g\leftarrow g}\left(\frac{x}{x_n}\right) \; \cdots \; \int_{x_0}^1 \frac{dx_1}{x_1} \; P_{g\leftarrow g}\left(\frac{x_2}{x_1}\right) \; \sigma_1(x_1,\mu_0) \\ & \times \int_{\mu_0}^{\mu_F} \frac{d\vec{p}_{T,n}^{\,2}}{\vec{p}_{T,n}^{\,2}} \frac{\alpha_s(\mu_R^2)}{2\pi} \; \cdots \; \int_{\mu_0}^{\mu_F} \frac{d\vec{p}_{T,1}^{\,2}}{\vec{p}_{T,1}^{\,2}} \frac{\alpha_s(\mu_R^2)}{2\pi} \; .\end{aligned} \tag{5.58}$$

Next, we will look for assumptions which allow us to solve Eq.(5.57) and compare the result to the solution of the DGLAP equation. To develop this physics picture of the DGLAP equation we make two assumptions:

1. We identify the scale of the strong coupling α_s with the transverse momentum scale of the splitting,
$$\mu_R^2 = \vec{p}_T^{\,2} \; . \tag{5.59}$$

2. Finally, we assume strongly ordered splittings in the transverse momentum. If the ordering of the splitting is fixed externally by the chain of momentum fractions x_j, this means
$$\mu_0^2 < \vec{p}_{T,1}^{\,2} < \vec{p}_{T,2}^{\,2} < \cdots < \mu_F^2 \tag{5.60}$$

Under these two assumptions the transverse momentum integrals in Eq.(5.58) become

$$\begin{aligned}& \int_{\mu_0}^{\mu_F} \frac{d\vec{p}_{T,n}^{\,2}}{\vec{p}_{T,n}^{\,2}} \frac{\alpha_s(\vec{p}_{T,n}^{\,2})}{2\pi} \; \cdots \int_{\mu_0}^{p_{T,3}} \frac{d\vec{p}_{T,2}^{\,2}}{\vec{p}_{T,2}^{\,2}} \frac{\alpha_s(\vec{p}_{T,2}^{\,2})}{2\pi} \int_{\mu_0}^{p_{T,2}} \frac{d\vec{p}_{T,1}^{\,2}}{\vec{p}_{T,1}^{\,2}} \frac{\alpha_s(\vec{p}_{T,1}^{\,2})}{2\pi} \\ = & \int_{\mu_0}^{\mu_F} \frac{d\vec{p}_{T,n}^{\,2}}{\vec{p}_{T,n}^{\,2}} \frac{1}{2\pi b_0 \log \dfrac{\vec{p}_{T,n}^{\,2}}{\Lambda^2_{\rm QCD}}} \; \cdots \int_{\mu_0}^{p_{T,3}} \frac{d\vec{p}_{T,2}^{\,2}}{\vec{p}_{T,2}^{\,2}} \frac{1}{2\pi b_0 \log \dfrac{\vec{p}_{T,2}^{\,2}}{\Lambda^2_{\rm QCD}}} \int_{\mu_0}^{p_{T,2}} \frac{d\vec{p}_{T,1}^{\,2}}{\vec{p}_{T,1}^{\,2}} \frac{1}{2\pi b_0 \log \dfrac{\vec{p}_{T,1}^{\,2}}{\Lambda^2_{\rm QCD}}} \\ = & \frac{1}{(2\pi b_0)^n} \int_{\mu_0}^{\mu_F} \frac{d\vec{p}_{T,n}^{\,2}}{\vec{p}_{T,n}^{\,2}} \frac{1}{\log \dfrac{\vec{p}_{T,n}^{\,2}}{\Lambda^2_{\rm QCD}}} \; \cdots \int_{\mu_0}^{p_{T,3}} \frac{d\vec{p}_{T,2}^{\,2}}{\vec{p}_{T,2}^{\,2}} \frac{1}{\log \dfrac{\vec{p}_{T,2}^{\,2}}{\Lambda^2_{\rm QCD}}} \int_{\mu_0}^{p_{T,2}} \frac{d\vec{p}_{T,1}^{\,2}}{\vec{p}_{T,1}^{\,2}} \frac{1}{\log \dfrac{\vec{p}_{T,1}^{\,2}}{\Lambda^2_{\rm QCD}}} \; .\end{aligned} \tag{5.61}$$

We can solve the individual integrals by switching variables, for example in the last integral

$$\int_{\mu_0}^{p_{T,2}} \frac{d\vec{p}_{T,1}^2}{\vec{p}_{T,1}^2} \frac{1}{\log \frac{\vec{p}_{T,1}^2}{\Lambda_{\rm QCD}^2}} = \int_{\log\log \mu_0^2/\Lambda^2}^{\log\log p_{T,2}^2/\Lambda^2} d\log\log \frac{\vec{p}_{T,1}^2}{\Lambda_{\rm QCD}^2} \qquad \text{with} \qquad \frac{d(ax)}{(ax)\log x} = d\log\log x$$

$$= \log \frac{\log \vec{p}_{T,2}^2/\Lambda_{\rm QCD}^2}{\log \mu_0^2/\Lambda_{\rm QCD}^2} \; . \tag{5.62}$$

This gives us for the chain of transverse momentum integrals, shifted to get rid of the lower boundaries,

$$\begin{aligned}
&\int^{p_{T,n}\equiv\mu_F} d\log \frac{\log \vec{p}_{T,n}^2/\Lambda_{\rm QCD}^2}{\log \mu_0^2/\Lambda_{\rm QCD}^2} \; \cdots \int^{p_{T,2}\equiv p_{T,3}} d\log \frac{\log \vec{p}_{T,2}^2/\Lambda_{\rm QCD}^2}{\log \mu_0^2/\Lambda_{\rm QCD}^2} \int^{p_{T,1}\equiv p_{T,2}} d\log \frac{\log \vec{p}_{T,1}^2/\Lambda_{\rm QCD}^2}{\log \mu_0^2/\Lambda_{\rm QCD}^2} \\
&= \int^{p_{T,n}\equiv\mu_F} d\log \frac{\log \vec{p}_{T,n}^2/\Lambda_{\rm QCD}^2}{\log \mu_0^2/\Lambda_{\rm QCD}^2} \; \cdots \int^{p_{T,2}\equiv p_{T,3}} d\log \frac{\log \vec{p}_{T,2}^2/\Lambda_{\rm QCD}^2}{\log \mu_0^2/\Lambda_{\rm QCD}^2} \left(\log \frac{\log \vec{p}_{T,2}^2/\Lambda_{\rm QCD}^2}{\log \mu_0^2/\Lambda_{\rm QCD}^2} \right) \\
&= \int^{p_{T,n}\equiv\mu_F} d\log \frac{\log \vec{p}_{T,n}^2/\Lambda_{\rm QCD}^2}{\log \mu_0^2/\Lambda_{\rm QCD}^2} \; \cdots \frac{1}{2} \left(\log \frac{\log \vec{p}_{T,3}^2/\Lambda_{\rm QCD}^2}{\log \mu_0^2/\Lambda_{\rm QCD}^2} \right)^2 \\
&= \int^{p_{T,n}\equiv\mu_F} d\log \frac{\log \vec{p}_{T,n}^2/\Lambda_{\rm QCD}^2}{\log \mu_0^2/\Lambda_{\rm QCD}^2} \left(\frac{1}{2} \cdots \frac{1}{n-1} \right) \left(\log \frac{\log \vec{p}_{T,n}^2/\Lambda_{\rm QCD}^2}{\log \mu_0^2/\Lambda_{\rm QCD}^2} \right)^{n-1} \\
&= \frac{1}{n!} \left(\log \frac{\log \mu_F^2/\Lambda_{\rm QCD}^2}{\log \mu_0^2/\Lambda_{\rm QCD}^2} \right)^n = \frac{1}{n!} \left(\log \frac{\alpha_s(\mu_0^2)}{\alpha_s(\mu_F^2)} \right)^n \; .
\end{aligned} \tag{5.63}$$

This is the final result for the chain of transverse momentum integrals in Eq.(5.58). After integrating over the transverse momenta, the strong coupling is evaluated at $\mu_R \equiv \mu_F$.

Next, we look at the convolution integrals in Eq.(5.57),

$$\sigma_{n+1}(x,\mu) \approx \frac{1}{n!} \left(\frac{1}{2\pi b_0} \log \frac{\alpha_s(\mu_0^2)}{\alpha_s(\mu^2)} \right)^n \int_{x_0}^1 \frac{dx_n}{x_n} \, P_{g\leftarrow g}\left(\frac{x}{x_n}\right) \cdots \int_{x_0}^1 \frac{dx_1}{x_1} \, P_{g\leftarrow g}\left(\frac{x_2}{x_1}\right) \, \sigma_1(x_1,\mu_0) \; . \tag{5.64}$$

As before, we Mellin-transform the equation into moment space

$$\begin{aligned}
\mathcal{M}[\sigma_{n+1}](m,\mu) \sim & \frac{1}{n!} \left(\frac{1}{2\pi b_0} \log \frac{\alpha_s(\mu_0^2)}{\alpha_s(\mu^2)} \right)^n \mathcal{M}\left[\int_{x_0}^1 \frac{dx_n}{x_n} \, P_{g\leftarrow g}\left(\frac{x}{x_n}\right) \cdots \int_{x_0}^1 \frac{dx_1}{x_1} \, P_{g\leftarrow g}\left(\frac{x_2}{x_1}\right) \, \sigma_1(x_1,\mu_0) \right](m) \\
= & \frac{1}{n!} \left(\frac{1}{2\pi b_0} \log \frac{\alpha_s(\mu_0^2)}{\alpha_s(\mu^2)} \right)^n \gamma(m)^n \, \mathcal{M}[\sigma_1](m,\mu_0) \qquad\qquad \text{using } \gamma(m) \equiv \mathcal{M}[P](m) \\
= & \frac{1}{n!} \left(\frac{1}{2\pi b_0} \log \frac{\alpha_s(\mu_0^2)}{\alpha_s(\mu^2)} \, \gamma(m) \right)^n \mathcal{M}[\sigma_1](m,\mu_0) \; .
\end{aligned} \tag{5.65}$$

Finally, we sum the production cross sections for up to n collinear jets,

$$\begin{aligned}
\sum_{n=0}^\infty \mathcal{M}[\sigma_{n+1}](m,\mu) = & \mathcal{M}[\sigma_1](m,\mu_0) \; \sum_n \frac{1}{n!} \left(\frac{1}{2\pi b_0} \log \frac{\alpha_s(\mu_0^2)}{\alpha_s(\mu^2)} \, \gamma(m) \right)^n \\
= & \mathcal{M}[\sigma_1](m,\mu_0) \; \exp\left(\frac{\gamma(m)}{2\pi b_0} \; \log \frac{\alpha_s(\mu_0^2)}{\alpha_s(\mu^2)} \right) \\
= & \mathcal{M}[\sigma_1](m,\mu_0) \; \left(\frac{\alpha_s(\mu_0^2)}{\alpha_s(\mu^2)} \right)^{\frac{\gamma(m)}{2\pi b_0}} \; .
\end{aligned} \tag{5.66}$$

This is the same structure as the DGLAP equation's solution in Eq.(5.54). This means that we can understand the physics of the DGLAP equation using our model calculation of a successive gluon emission, including the generically

	renormalization scale μ_R	factorization scale μ_F
source	ultraviolet divergence	collinear (infrared) divergence
poles cancelled	counter terms (renormalization)	parton densities (mass factorization)
summation	resum self energy bubbles	resum parton splittings
parameter	running coupling $\alpha_s(\mu_R^2)$	running parton density $f_j(x, \mu_F)$
evolution	RGE for α_s	DGLAP equation
large scales	decrease of σ_{tot}	increase of σ_{tot} for gluons/sea quarks
theory background	renormalizability proven for gauge theories	factorization proven all orders for DIS proven order-by-order DY...

Table 2: Comparison of renormalization and factorization scales appearing in LHC cross sections.

variable number of collinear jets in the form of $pp \to \mu^+\mu^- + X$, as shown in Eq.(5.55). On the left hand side of Eq.(5.66) we have the sum over any number of additional collinear partons; on the right hand side we see fixed order Drell–Yan production without any additional partons, but with an exponentiated correction factor. Comparing this to the running parton densities we can draw the analogy that any process computed with a scale dependent parton density where the scale dependence is governed by the DGLAP equation includes any number of collinear partons.

We can also identify the logarithms which are resummed by scale dependent parton densities. Going back to Eq.(5.12) reminds us that we start from the divergent collinear logarithms $\log p_T^{\text{max}}/p_T^{\text{min}}$ arising from the collinear phase space integration. In our model for successive splitting we replace the upper boundary by μ_F. The collinear logarithm of successive initial–state parton splitting diverges for $\mu_0 \to 0$, but it gets absorbed into the parton densities and determines the structure of the DGLAP equation and its solutions. The upper boundary μ_F tells us to what extent we assume incoming quarks and gluons to be a coupled system of splitting partons and what the maximum momentum scale of these splittings is. Transverse momenta $p_T > \mu_F$ generated by hard parton splitting are not covered by the DGLAP equation and hence not a feature of the incoming partons anymore. They belong to the hard process and have to be consistently simulated. While this scale can be chosen freely we have to make sure that it does not become too large, because at some point the collinear approximation $C \simeq$ constant in Eq.(5.12) ceases to hold.