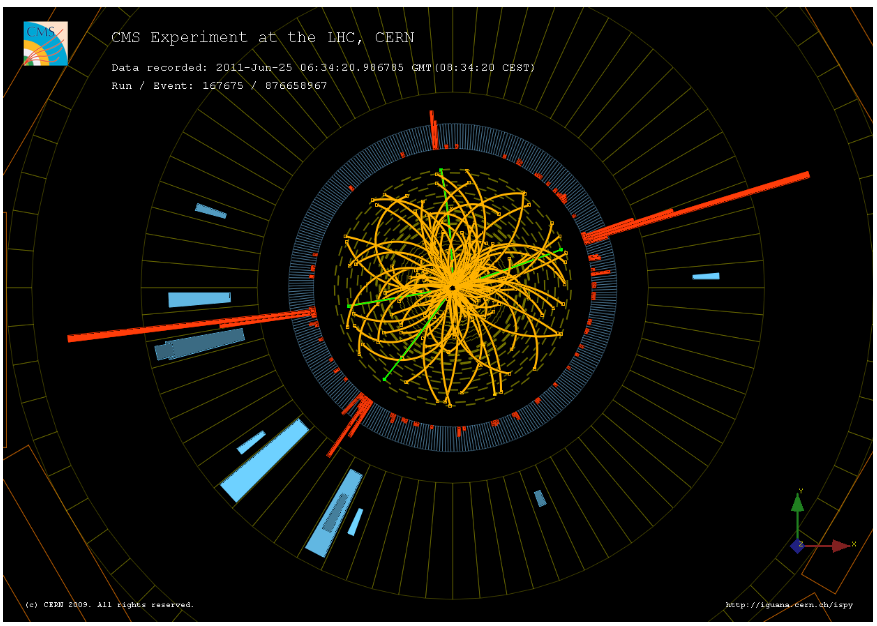
Phenomenology of Particle Physics

ETH Zurich and University of Zurich

HS 2009 - FS 2010



Prof. Dr. Vincenzo Chiochia [University of Zurich]
Prof. Dr. Günther Dissertori [ETH Zurich]
Prof. Dr. Thomas Gehrmann [University of Zurich]
Typeset by Julián Cancino and Julian Schrenk

June 13, 2014

This script is based on the lecture given originally by Prof. Dr. Chiochia (University of Zurich), Prof. Dr. Günther Dissertori (ETH Zurich) and Prof. Dr. Thomas Gehrmann (University of Zurich) in the fall semester 2009 (PPP1) and the spring semester 2010 (PPP2).

Please feel free to send feedback to julian.cancino@gmail.com and kenjschrenk@gmail.com.

We would like to thank the lecturers who used the script in the later semesters, as well as in particular Romain Müller and Murad Tovmasyan for making valuable suggestions and pointing out numerous typos.

Contents

Part I

Fall semester

Chapter 1

Introduction

Literature:

- Halzen/Martin [1]
- Aitchison/Hey [2] (rigorous)
- Seiden [3] (experimental, up to date)
- Nachtmann [4] (difficult to purchase)

Elementary particles are the smallest constituents of matter. Therefore the notion "elementary" changes with scientific progress (cf. Tab. 1.1).

We can define "elementary" as "having no resolvable inner structure". This also means that there can be no excited states. Elementary particles interact in a well-defined way through fundamental interactions. These are

- gravity,
- electromagnetic interaction,
- weak interaction, and
- strong interaction,

where only the last three are relevant, at the elementary particle level, at energies currently available. Range of phenomena:

- structure of matter
- stability of matter

1869	Mendeleev/Meyer	periodic system	atom
1890	J. Thomson	electron	
1910	Bequerel/Curie	radioactivity	atomic nucleus & electron
	Rutherford	scattering	
1932	Chadwick	neutron	proton, neutron, electron
	Anderson	positron	& their antiparticles
1947	Blackett/Powell	pion, muon	"particle zoo"
1956	Cowan/Reines	neutrino	
1967	Glashow/Weinberg/Salam	**electroweak theory**	
1968	SLAC	deep inelastic scattering	quarks & leptons
1972	Fritzsch/Gell-Mann/Leutwyler	**quantum chromodynamics**	
1974	SLAC/BNL	c quark, τ lepton	
1979	DESY	gluon	
1977	Fermilab	b quark	
1983	CERN	W, Z bosons	
1995	Fermilab	t quark, ν_τ	

Table 1.1: Historical outline of the concept of "elementarity"

- instability of matter, radioactivity: decay of elementary particles
- scattering of elementary particles
- production of new particles
- indirect implications
  - early history of the universe
  - fuel cycle in stars
  - astrophysical phenomena: supernovae, very high energy cosmic rays

1.1 Units

The Planck constant

$$\hbar = \frac{h}{2\pi} = 1.0546 \cdot 10^{-34}\,\mathrm{Js} \tag{1.1}$$

has dimension of action and angular momentum. Another important physical constant is the speed of light

$$c = 2.998 \cdot 10^{8}\,\frac{\mathrm{m}}{\mathrm{s}}. \tag{1.2}$$

Because we are dealing with constants, Eq. (1.1) and (1.2) establish a relationship among the units for energy, time, and length. Using so-called natural units, i. e. setting $\hbar = c = 1$, we find

$$\begin{aligned}
[c] &= [\text{length}] \cdot [\text{time}]^{-1} = [L][T]^{-1} \Rightarrow [L] = [T] && (1.3)\\
[\hbar] &= [\text{energy}] \cdot [\text{time}] = [M][L]^2[T]^{-1} \Rightarrow [M] = [L]^{-1} && (1.4)\\
\Rightarrow [M] &= [L]^{-1} = [T]^{-1} \text{ and } [E] = [M]. && (1.5)
\end{aligned}$$

This raises the question of a suitable fundamental unit for energy. One electron volt is the energy acquired by an electron passing a potential difference of $1\,\text{V}$:

$$\begin{aligned}
1\,\text{eV} &= 1.602 \cdot 10^{-19}\,\text{J}\\
\text{keV} &= 10^3\,\text{eV}\\
\text{MeV} &= 10^6\,\text{eV}\\
\text{GeV} &= 10^9\,\text{eV}\\
\text{TeV} &= 10^{12}\,\text{eV}.
\end{aligned}$$

Examples of some orders of magnitude are

$$\begin{aligned}
m_e &= 511\,\text{keV}\\
m_p &= 938\,\text{MeV}\\
m_n &= 939\,\text{MeV}\\
E_e(\text{LEP}) &= 104.5\,\text{GeV}\\
E_p(\text{Tevatron}) &= 980\,\text{GeV}\\
E_p(\text{LHC}) &= 7\,\text{TeV}.
\end{aligned}$$

Converting the units for energy, time, and length into each other yields, in agreement with Eq. (1.5),

$$\hbar = 6.58 \cdot 10^{-25}\,\text{GeV} \cdot \text{s} \stackrel{!}{=} 1 \Rightarrow \boxed{1\,\text{GeV}^{-1} \simeq 6.58 \cdot 10^{-25}\,\text{s}}, \tag{1.6}$$

(recall lifetime $\tau = \frac{1}{\Gamma}$ with Γ the resonance width), and

$$c = 2.998 \cdot 10^8\,\frac{\text{m}}{\text{s}} \stackrel{!}{=} 1 \Rightarrow \boxed{1\,\text{fm} = 10^{-15}\,\text{m} \simeq \frac{1}{200\,\text{MeV}}}. \tag{1.7}$$

Cross sections have dimensions of area:

$$[\sigma] = [L]^2 = [M]^{-2} = \frac{1}{(\text{eV})^2}. \tag{1.8}$$

As unit we choose

$$\frac{1}{(1\,\text{GeV})^2} = 389379\,\text{nb} = 389379 \cdot 10^{-9}\,\text{b}$$

with $1\,\mathrm{b} : 1\,\mathrm{barn} = 10^{-24}\,\mathrm{cm}^2$ the typical scale of nuclear absorption.

The unit of electrical charge can be defined in different ways. The dimensionless fine structure constant α is accordingly expressed differently in terms of e in different systems of units,

$$\begin{aligned}\alpha &= \left.\frac{e^2}{4\pi\varepsilon_0\hbar c}\right|_{\mathrm{SI}} = 7.2972\cdot 10^{-3} \simeq \frac{1}{137}\\ &= \left.\frac{e^2}{\hbar c}\right|_{\mathrm{CGS}}\\ &= \left.\frac{e^2}{4\pi\hbar c}\right|_{\text{Heaviside-Lorentz}},\end{aligned}$$

and determines the strength of the electromagnetic interaction. Therefore, in Heaviside-Lorentz units, the electron charge is fixed to be

$$e = \left.\sqrt{4\pi\alpha}\right|_{\mathrm{HL}}. \tag{1.9}$$

1.2 Elementary interactions

Gravitation. Since

$$Gm_p^2 \approx 10^{-39}$$

and because of the fact that gravity's range is infinite, it is relevant for macroscopic systems (and can be neglected here).

Electromagnetic interaction. Recall that $\alpha \simeq \frac{1}{137}$. The range of the electromagnetic interaction is infinite and typical lifetimes of particles decaying through electromagnetic interactions range from $\tau_{\Sigma^0\to\Lambda^0\gamma} = 10^{-20}\,\mathrm{s}$ to $\tau_{\pi^0\to\gamma\gamma} = 10^{-16}\,\mathrm{s}$. Typical cross sections are of order $\sigma_{ep\to ep} = 1\,\mu\mathrm{b}$. QED's (quantum electrodynamics') predictions have been tested to high theoretical and experimental precision. Consider for example the anomalous magnetic moment of the electron:

$$\begin{aligned}\mu_e^{\mathrm{QED}} &= \frac{e}{2m_e}\frac{g}{2} = \frac{e}{2m_e}\Bigg\{\underbrace{1}_{\text{Dirac}} + \underbrace{\frac{1}{2}\frac{\alpha}{\pi}}_{\text{Schwinger}} \underbrace{-0.388\frac{\alpha^2}{\pi^2}}_{\text{Petermann}} + \underbrace{1.18\frac{\alpha^3}{\pi^3}}_{\text{Laporta/Remiddi}}\Bigg\}\\ &= \frac{e}{2m_e}\{1.0011596521465(270)\}\\ \mu_e^{\mathrm{exp.}} &= \frac{e}{2m_e}\{1.0011596521883(42)\},\end{aligned}$$

where the experimental value was obtained by Van Dyck, Schwinberg and Dehmelt.

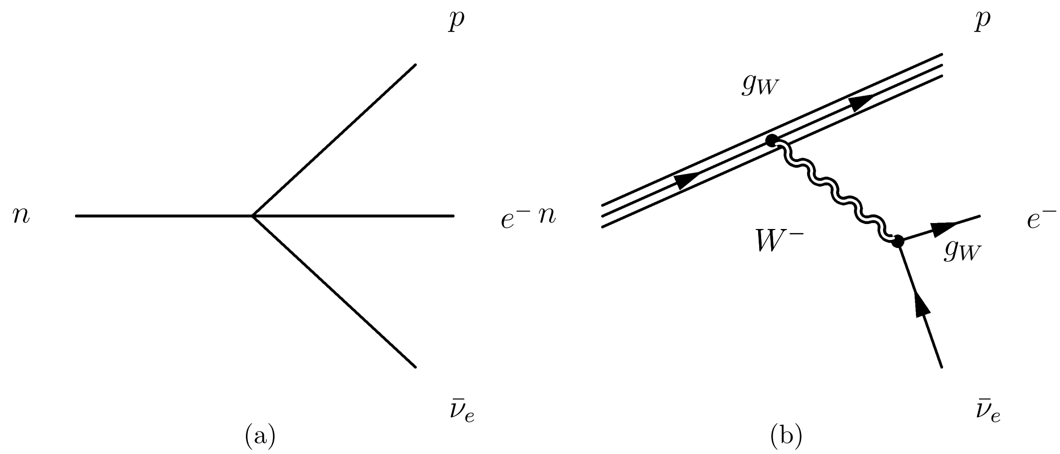


Figure 1.1: *Beta decay of neutron.* Depicted as a point like process, as described by Fermi's constant (a) and via W^- boson exchange (b).

Weak interaction. As an example for weak interactions consider β decay: $n \to pe\bar{\nu}_e$: see Fig. 1.1(a). The range is about 1 fm and for the coupling we have

$$G_F m_p^2 \approx 10^{-5}.$$

The lifetimes go from 10^{-10} s to 10^3 s and cross sections are of order $\sigma \approx 1$ fb. Theoretically, the process is explained by W^- boson exchange, see Fig. 1.1(b), which yields for Fermi's constant $G_F = \frac{g_w^2}{8M_W^2}$.

Strong interaction. At the nuclear level, the Yukawa theory of pion exchange (see figures 1.2(a) and 1.2(b)) is still used. It explains the bonding of protons and neutrons by exchange of massive pions: $m_\pi = 130\,\mathrm{MeV} \Rightarrow$ range $\simeq \frac{1}{m_\pi} = \frac{1}{130\,\mathrm{MeV}} \simeq 1.4\,\mathrm{fm}$. QCD (quantum chromodynamics) states that particles like p, n, and π consist of quarks which interact through gluons. Gluons (in contrast to photons) carry themselves the charges they are coupling to which influences the strong interaction's potential, see fig.1.3. The QCD coupling constant is approximately given by $\alpha_s \simeq 0.12$.

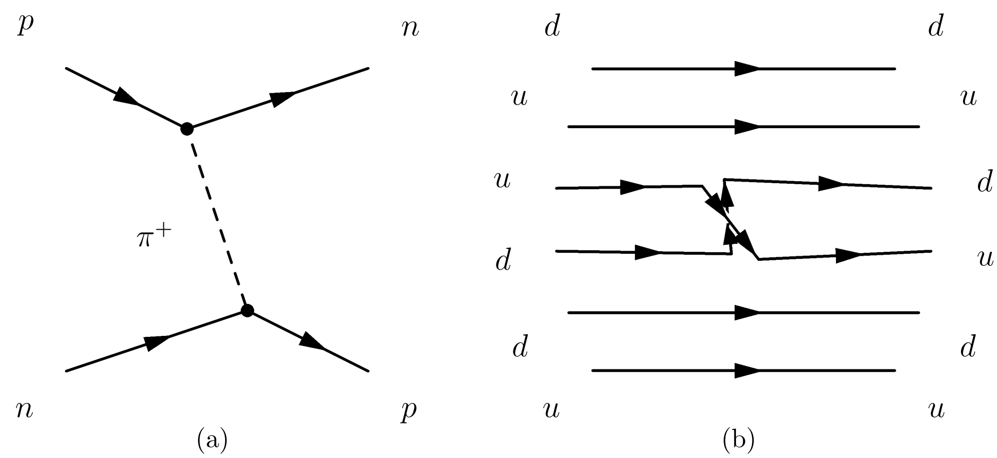


Figure 1.2: *Yukawa theory.* Interaction by pion exchange (a) and exchange of quark and anti-quark (b).

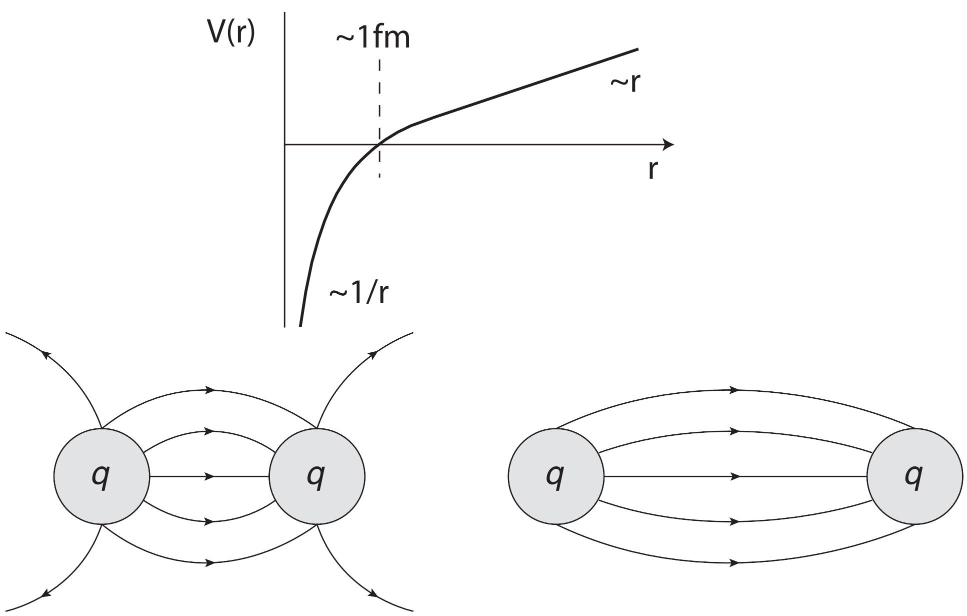


Figure 1.3: *Potential of the strong interaction.*

Chapter 2

Relativistic kinematics

Literature:

- Nachtmann [4]
- Hagedorn [5]
- Byckling/Kajantie [6]

We state some notation concerning special relativity:

$$x^\mu = (x^0 = t, x^1, x^2, x^3) = (t, \vec{x}) \qquad \text{contravariant four-vector} \tag{2.1}$$

$$x_\mu = (t, -\vec{x}) \qquad \text{covariant four-vector} \tag{2.2}$$

$$g^{\mu\nu} = g_{\mu\nu} = \begin{pmatrix} 1 & & & \\ & -1 & & \\ & & -1 & \\ & & & -1 \end{pmatrix} \qquad \text{metric tensor} \tag{2.3}$$

$$\tau^2 = t^2 - \vec{x}^2 = g_{\mu\nu} x^\mu x^\nu = x^\mu x_\mu = x^2 \qquad \text{Lorentz invariant} \tag{2.4}$$

$$d\tau = dt\sqrt{1 - \left(\frac{d\vec{x}}{dt}\right)^2} = \frac{dt}{\gamma} \qquad \text{proper time.} \tag{2.5}$$

Combining Eq. (2.1) and (2.5) we arrive at the four-velocity

$$u^\mu = \frac{dx^\mu}{d\tau} = \frac{dx^\mu}{dt}\frac{dt}{d\tau} = \gamma(1, \vec{v}).$$

Since

$$u^2 = \gamma^2(1 - \vec{v}^2) = 1 > 0,$$

u is a time-like four-vector. The four-momentum is then defined as

$$p^\mu = mu^\mu = m\gamma(1, \vec{v}) = (p^0 = E, \vec{p}).$$

By calculating the corresponding Lorentz invariant,

$$p^2 = m^2 u^2 = m^2 = E^2 - \vec{p}^2,$$

we find the energy-momentum relation

$$E = \sqrt{m^2 + \vec{p}^2}. \tag{2.6}$$

A particle is said to be relativistic if $\vec{p}^2 \not\ll m^2$. Conversely, for a non-relativistic particle, $\vec{p}^2 \ll m^2$, and therefore

$$E = \sqrt{m^2 + \vec{p}^2} = m\left(1 + \frac{1}{2}\frac{\vec{p}^2}{m^2} + \ldots\right) = m + \frac{1}{2}\frac{\vec{p}^2}{m} + \ldots$$

so that we recover the expression for $|\vec{v}| \ll 1$ of Newtonian mechanics.

2.1 Particle decay

The decaying particle's four-momentum is, in the rest frame, given by $p = (M, 0, 0, 0)$, see Fig. 2.1. The decay time (lifetime) is

$$d\tau^2 = dt^2(1 - \vec{v}^2)$$

where dt^2 is the lifetime in the laboratory frame:

$$dt = \gamma d\tau > d\tau. \tag{2.7}$$

The result stated in equation (2.7) has been verified experimentally:

$$\begin{aligned}
&\tau_{\pi^+ \to \mu^+ \nu_\mu} = 2.6 \cdot 10^{-8}\,\mathrm{s} \\
&E_\pi = 20\,\mathrm{GeV},\ \gamma = \frac{E_\pi}{m_\pi} = 143 \Leftrightarrow v = 0.9999 \\
&\Rightarrow \frac{t'_\pi}{t_\pi} = 143.
\end{aligned}$$

Constraints are (i) conservation of energy and momentum, $p = p_1 + p_2$ (4 equations), and (ii) the mass-shell condition, $p_i^2 = m_i^2$:

$$\begin{array}{lll}
p^2 = M^2 & p_1^2 = m_1^2 & p_2^2 = m_2^2 \\
p = (M, \vec{0}) & p_1 = (E_1, \vec{p}_1) & p_2 = (E_2, \vec{p}_2).
\end{array}$$

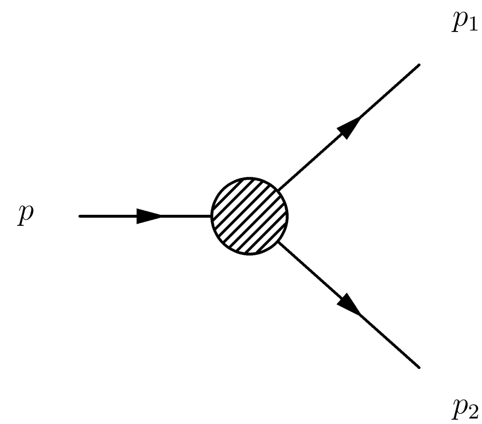


Figure 2.1: *Particle decay.* Dynamics will be discussed later on; at the moment we are dealing with kinematics.

It therefore follows that

$$p \cdot p_i = ME_i \;\Rightarrow E_i = \frac{1}{M} p \cdot p_i = \frac{1}{M}(p_1 \cdot p_i + p_2 \cdot p_i).$$

And, by using $p_1 \cdot p_2 = \frac{1}{2}[(p_1+p_2)^2 - p_1^2 - p_2^2] = \frac{1}{2}[M^2 - m_1^2 - m_2^2]$, we find

$$\begin{aligned} E_1 &= \frac{1}{M}(p_1^2 + p_1 \cdot p_2) = \frac{1}{2M}(M^2 + m_1^2 - m_2^2) \\ E_2 &= \frac{1}{2M}(M^2 - m_1^2 + m_2^2). \end{aligned}$$

By using equation (2.6) and $\vec{p}_1 + \vec{p}_2 = 0$, the absolute value of the three-momenta,

$$\vec{p}_1^{\,2} = E_1^2 - m_1^2 = \frac{1}{4M^2}\Big(M^4 - 2M^2(m_1^2 + m_2^2) + (m_1^2 - m_2^2)^2\Big) = \vec{p}_2^{\,2},$$

is also fixed. This means that only the directions of $\vec{p}_1$ and $\vec{p}_2$ remain unknown, while the energies and the absolute values of the momenta can be calculated directly.

2.2 Two-particle scattering

For a visualisation of the process see Fig. 2.2(a). Once again, the constraints are

$$\begin{aligned} &p_i^2 = m_i^2 \; (i = 1, \ldots, 4) \\ &p_1 + p_2 = p_3 + p_4. \end{aligned}$$

We talk of elastic scattering if $m_1 = m_3$ and $m_2 = m_4$. Consider the Lorentz invariants

$$p_i^2 = m_i^2 \text{ and } \underbrace{p_1 \cdot p_2,\; p_1 \cdot p_3,\; p_1 \cdot p_4,\; p_2 \cdot p_3,\; p_2 \cdot p_4,\; p_3 \cdot p_4}_{\text{6 invariants, 2 linearly independent, 4 linearly dependent}}.$$

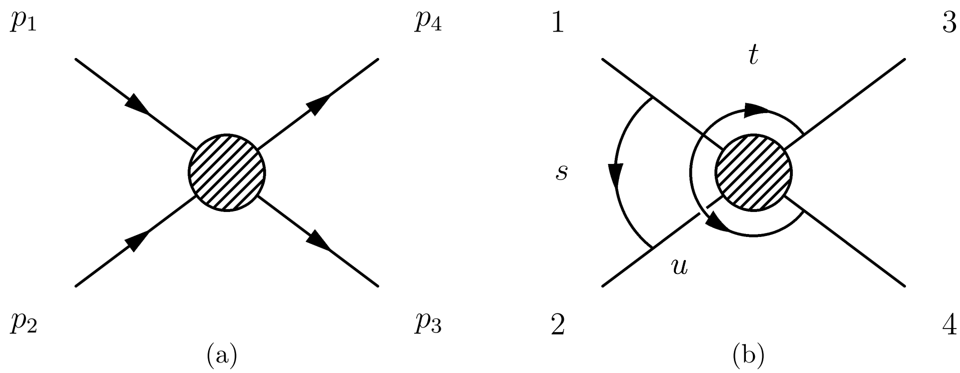


Figure 2.2: *Two-particle scattering.* The kinematical constraints are energy-momentum conservation and the mass shell condition (a). Visualization of Mandelstam variables (b).

Four of them have to be linearly dependent, since there are only two degrees of freedom in the system (center of mass energy and scattering angle).

We now define the Mandelstam variables (see Fig. 2.2(b))

$$\begin{aligned} s &= (p_1 + p_2)^2 \\ t &= (p_1 - p_3)^2 \\ u &= (p_1 - p_4)^2, \end{aligned}$$

where s denotes total center of mass energy squared (positive) and t is the four-momentum transfer squared (negative). Note also that $s + t + u = \sum_{i=1}^{4} m_i^2$.

The center of mass frame is defined by

$$\vec{p}_1 + \vec{p}_2 = 0 = \vec{p}_3 + \vec{p}_4. \tag{2.8}$$

One usually denote variables in this frame with an asterisk: (cm., $p_i = p_i^*$). The laboratory frame is defined by $\vec{p}_2 = 0$ (fixed target) and variables are labelled with an L : (lab., $p_i = p_i^L$). In deep inelastic scattering the Breit system ($p_i = p_i^B$) is used, which is defined by $\vec{p}_1 + \vec{p}_3 = 0$.

In the following we take a closer look at the center of mass frame, see Fig. 2.3. Equation (2.8) leads to

$$\begin{aligned} \vec{p}_1^{\,*} &= -\vec{p}_2^{\,*} = \vec{p} \\ \vec{p}_3^{\,*} &= -\vec{p}_4^{\,*} = \vec{p}\,' \\ p_1 &= \left(E_1^* = \sqrt{\vec{p}^{\,2} + m_1^2}, \vec{p}\right) \\ p_2 &= \left(E_2^* = \sqrt{\vec{p}^{\,2} + m_2^2}, -\vec{p}\right) \\ p_3 &= (E_3^*, \vec{p}\,') \\ p_4 &= (E_4^*, -\vec{p}\,'). \end{aligned}$$

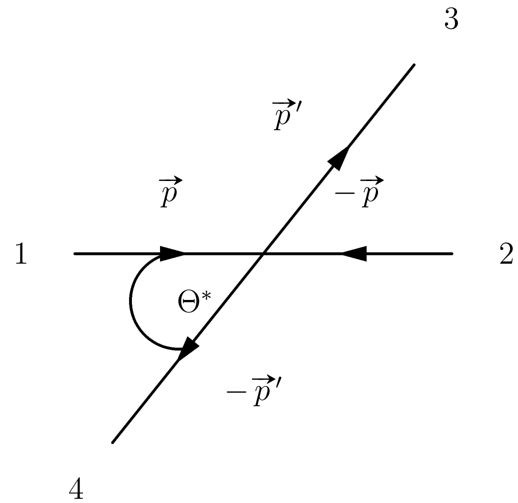


Figure 2.3: *Two-particle scattering in center of mass frame.* For the constraints on the scattering angle Θ^* see section 2.2.1.

The sum

$$p_1 + p_2 = (\underbrace{E_1^* + E_2^*}_{\sqrt{s}}, \vec{0})$$

is no Lorentz invariant, whereas

$$s = (p_1 + p_2)^2 = (E_1^* + E_2^*)^2$$

is one. Now we can express E_i^*, $|\vec{p}|$, and $|\vec{p}'|$ in terms of s (see exercise sheet 1):

$$E_{1,3}^* = \frac{1}{2\sqrt{s}}(s + m_{1,3}^2 - m_{2,4}^2) \tag{2.9}$$

$$\vec{p}^2 = (E_1^*)^2 - m_1^2 = \frac{1}{4s}\lambda(s, m_1^2, m_2^2), \tag{2.10}$$

where we have used the Källén function (triangle function) which is defined by

$$\begin{aligned}\lambda(a,b,c) &= a^2 + b^2 + c^2 - 2ab - 2ac - 2bc \\ &= \left[a - (\sqrt{b} + \sqrt{c})^2\right]\left[a - (\sqrt{b} - \sqrt{c})^2\right] \\ &= a^2 - 2a(b+c) + (b-c)^2.\end{aligned}$$

We can see that the Källén function has the following properties:

- symmetric under $a \leftrightarrow b \leftrightarrow c$ and
- asymptotic behavior: $a \gg b, c : \lambda(a,b,c,) \to a^2$.

This enables us to determine some properties of scattering processes. From $\vec{p}^{\,2}, \vec{p}^{\,\prime 2} > 0$ it follows that

$$s_{\min} = \max\Big\{(m_1+m_2)^2, (m_3+m_4)^2\Big\} \geq 0$$

is the threshold of the process in the s-channel. In the high energy limit ($s \gg m_i^2$) Eq. (2.9) and (2.10) simplify because of the asymptotic behavior of λ and one obtains:

$$E_1^* = E_2^* = E_3^* = E_4^* = |\vec{p}| = |\vec{p}^{\,\prime}| = \frac{\sqrt{s}}{2}.$$

2.2.1 Scattering angle

In the center of mass frame, the scattering angle Θ^* is defined by

$$\vec{p}\cdot\vec{p}^{\,\prime} = |\vec{p}|\cdot|\vec{p}^{\,\prime}|\cos\Theta^*.$$

We also know that

$$\begin{aligned} p_1\cdot p_3 &= E_1^*E_3^* - |\vec{p}_1^{\,*}||\vec{p}_3^{\,*}|\cos\Theta^* \\ t = (p_1-p_3)^2 &= m_1^2 + m_3^2 - 2p_1p_3 = (p_2-p_4)^2 \end{aligned}$$

and can derive $\cos\Theta^* = \text{function}(s,t,m_i^2)$:

$$\cos\Theta^* = \frac{s(t-u) + (m_1^2-m_2^2)(m_3^2-m_4^2)}{\sqrt{\lambda(s,m_1^2,m_2^2)}\sqrt{\lambda(s,m_3^2,m_4^2)}}.$$

This means that $2\to 2$ scattering is described by two independent variables:

$\sqrt{s}$ and Θ^* or $\sqrt{s}$ and t.

2.2.2 Elastic scattering

We now consider the case of elastic scattering. This means that $m_1 = m_3$ and $m_2 = m_4$ (e. g. $ep \to ep$). Therefore Eq. (2.9) and (2.10) simplify:

$$\begin{aligned} &E_1^* = E_3^*,\ E_2^* = E_4^* \\ |\vec{p}|^2 = |\vec{p}^{\,\prime}|^2 &= \frac{1}{4s}\Big(s-(m_1+m_2)^2\Big)\Big(s-(m_1-m_2)^2\Big) \end{aligned}$$

and we find for the scattering angle (in the case of elastic scattering)

$$\begin{aligned} t = (p_1-p_3)^2 &= -(\vec{p}_1 - \vec{p}_3)^2 = -2\vec{p}^{\,2}(1-\cos\Theta^*) \\ \Rightarrow &\boxed{\cos\Theta^* = 1 + \frac{t}{2|\vec{p}|^2}}. \end{aligned}$$

Restriction to the physically valid region yields

$$\left.\begin{array}{c} -1 \leq \cos\Theta^* \leq 1 \\ \vec{p}^{\,2} \geq 0 \end{array}\right\} \Leftrightarrow \left\{\begin{array}{l} -4|\vec{p}|^2 \leq t \leq 0 \\ s \geq (m_1+m_2)^2 \end{array}\right. .$$

2.2.3 Angular distribution

Finally, we find for the angular distribution (bearing in mind that the distribution is rotationally invariant with respect to the "$\vec{p}$-axis" such that $\int d\phi = 2\pi$)

$$\begin{aligned} d\Omega^* &= 2\pi d\cos\Theta^* \\ \frac{d\Omega^*}{dt} &= \frac{4\pi s}{\sqrt{\lambda(s,m_1^2,m_2^2)}\sqrt{\lambda(s,m_3^2,m_4^2)}} = \frac{\pi}{|\vec{p}||\vec{p}\,'|}. \end{aligned} \tag{2.11}$$

2.2.4 Relative velocity

At this point, we introduce the relative velocity, which we will see to be of relevance in defining the particle flux and hence the collider construction,

$$v_{12} = |\vec{v}_1 - \vec{v}_2| = \left|\frac{\vec{p}_1}{E_1} - \frac{\vec{p}_2}{E_2}\right| = \left|\frac{\vec{p}_1^{\,*}}{E_1^*} - \frac{\vec{p}_2^{\,*}}{E_2^*}\right| = \frac{|\vec{p}_1^{\,*}|}{E_1^*E_2^*}\underbrace{(E_1^* + E_2^*)}_{\sqrt{s}}, \tag{2.12}$$

from which we get,

$$\begin{aligned} v_{12}E_1^*E_2^* &= \sqrt{s}|\vec{p}_1^{\,*}| = \sqrt{s}\sqrt{E_1^{*2} - m_1^2} \\ &= \sqrt{s}\sqrt{\frac{1}{4s}(s + m_1^2 - m_2^2)^2 - m_1^2} \\ &= \sqrt{(p_1\cdot p_2)^2 - m_1^2m_2^2}, \end{aligned} \tag{2.13}$$

the so called Møller flux factor. In going from the first line to the second, we used the definition of the Källén function and in going to the third the fact that $s = m_1^2 + m_2^2 + 2p_1\cdot p_2$.

We stress here that $v_{12}E_1^*E_2^*$ is a frame independent quantity. It appears in the definition of the incoming particle flux, an thus in the cross section. It also plays an important role in the normalization issues, since the classical volume element is not Lorentz invariant.

2.2.5 Center of mass and laboratory systems

For the center of mass and the laboratory systems respectively, we have,

$$\begin{aligned} CM : s &= (E_1^* + E_2^*)^2 = (\text{total energy})^2 \\ L : s &= m_1^2 + m_2^2 + 2m_2E_1^L \overset{E_1^L \gg m_1,m_2}{\longrightarrow} 2m_2E_1^L. \end{aligned}$$

As an example for the difference, we look at the two operating modes of the Tevatron at Fermilab (Figure 2.4). The energy of the beam particles is $E_{\text{beam}} = 980\,\text{GeV}$.

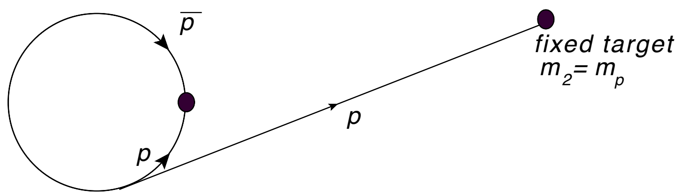


Figure 2.4: *Sketch of the Tevatron accelerator at Fermilab.*

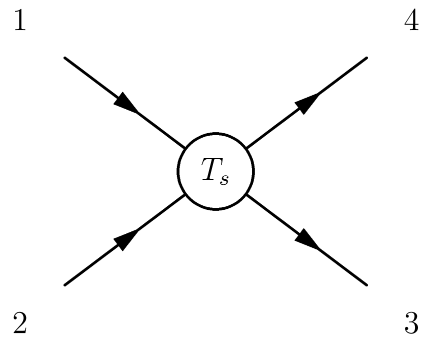


Figure 2.5: *s-channel.*

Used in the $p\bar{p}$-mode, the collision is head on and we are allowed to consider the problem in the center of mass frame and,

$$\sqrt{s_{p\bar{p}}}(\text{Collider}) = 1960\,\text{GeV},$$

which is ideal for discovering new phenomena with the highest possible energy.

If on the other hand, the pN-mode is chosen (N is a nucleus in the target), we need to consider the laboratory frame and we get

$$\sqrt{s_{pN}}(\text{Fixed target}) = 42.7\,\text{GeV} < m_W.$$

Although this mode is less energetic, it is then possible to create a secondary beam. With this method, the existence of ν_τ could be proven.

2.3 Crossing symmetry

The $2 \to 2$ scattering process has some underlying symmetries, which we shall explore now.

Example When we exchange p_3 and p_4, s is not affected but t and u interchange their roles.

We take now a look at the reaction (Figure 2.5), $1+2 \to 3+4$, for which the 4-momentum is conserved :

$$p_1 + p_2 = p_3 + p_4.$$

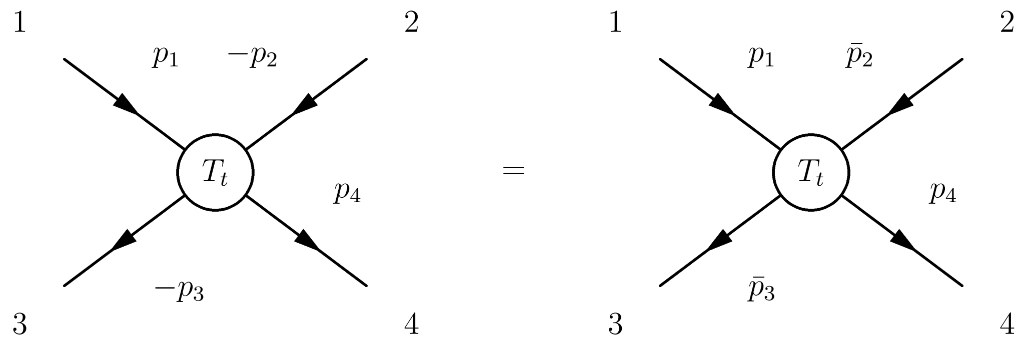


Figure 2.6: *t-channel.*

It is called "s-channel" reaction, because the only positive Mandelstam variable is s. T_s describes the scattering dynamics of the process and will be treated later. It depends on the three Mandelstam variables and is predicted by theoretically (QED, QCD, EW, SUSY,...),

$$T_s(s,t,u) = T(s,t,u)|_{s>0,t\leq 0,u\leq 0}. \tag{2.14}$$

T can then be extended analytically to the whole range $s,t,u \in \mathbb{R}$. Depending on the region, it can then describe different **crossed reactions**.

For instance, suppose we exchange p_2 and p_3, we then get naively (Figure 2.6),

$$p_1 + (-p_3) = (-p_2) + p_4.$$

We now make the interpretation

$$-p_n = p_{\bar{n}},$$

in which $\bar{n}$ stands for the antiparticle of the particle n, leading to the expression (Figure 2.6),

$$p_1 + p_{\bar{3}} = p_{\bar{2}} + p_4.$$

Since 1 and $\bar{3}$ are the incoming particles, we speak of the "t-channel" process. One has

$$T_t(s,t,u) = T(s,t,u)|_{s\leq 0,t>0,u\leq 0}. \tag{2.15}$$

2.3.1 Interpretation of antiparticle-states

As stated above, we interpret particles with 4-momentum $-p$ to be antiparticles with 4-momentum p. The reason for that becomes clear when we look at the 4-current,

$$j^\mu \stackrel{ED}{=} \begin{pmatrix} \rho \\ \vec{j} \end{pmatrix} \stackrel{QM}{=} \underbrace{\underbrace{-e}_{\text{electron charge}} \underbrace{i(\varphi^*\partial^\mu\varphi - \varphi\partial^\mu\varphi^*)}_{\text{probability density}}}_{\text{charge density}}. \tag{2.16}$$

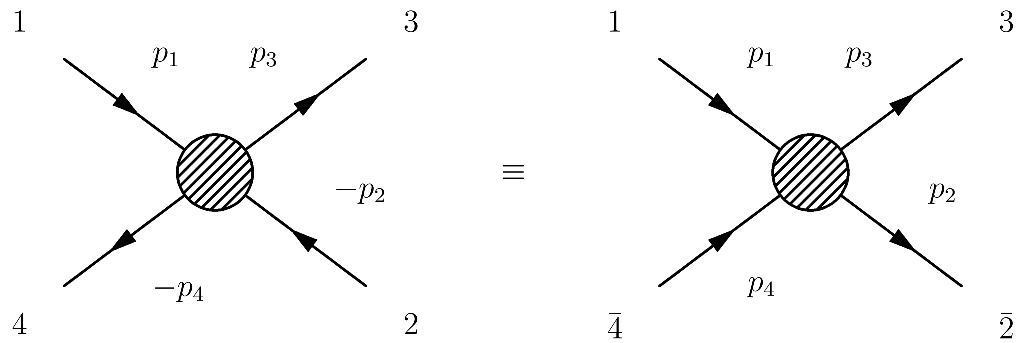


Figure 2.7: *Emission of a positron and absorption of an electron.* The emission of a positron with energy $+E$ is equivalent to the absorption of an electron with energy $-E$.

Inserting the wave function of the free electron,

$$\varphi = N\mathrm{e}^{-ip\cdot x}, \tag{2.17}$$

in the definition of the 4-current Eq. (2.16), one gets

$$\begin{aligned}
e^- \text{ with 4-momentum} + p^\mu &: j^\mu(e^-) = -2e|N|^2 p^\mu = -2e|N|^2 \begin{pmatrix} +E \\ +\vec{p} \end{pmatrix}, \\
e^+ \text{ with 4-momentum} + p^\mu &: j^\mu(e^+) = +2e|N|^2 p^\mu = -2e|N|^2 (-p)^\mu, \\
e^- \text{ with 4-momentum} - p^\mu &: j^\mu(e^-) = -2e|N|^2 (-p)^\mu = -2e|N|^2 \begin{pmatrix} -E \\ -\vec{p} \end{pmatrix},
\end{aligned}$$

and hence the rule,

$$\boxed{j^\mu(e^+) = j^\mu(e^-) \text{ with the subsititution } p^\mu \to -p^\mu}. \tag{2.18}$$

We stress here the fact that the *whole* 4-vector p^μ takes a minus sign, and not only the spatial part $\vec{p}$.

What we effectively used here is the fact that in the phase of Eq. (2.17) we can flip the signs of both p^μ and x^μ without changing the wave function. There is no place here for particle travelling backwards in time!

A particle with 4-momentum $-p^\mu$ is a *representation* for the corresponding *anti*particle with 4-momentum p^μ. Alternatively, one can say that the emission of a positron with energy $+E$ corresponds to the absorption of an electron with energy $-E$. Figure 2.7 restates the last sentence as a Feynman diagram.

The three reactions (s-, t- and u-channels) are described by a single function $T(s,t,u)$ evaluated in the relevant kinematical region ($s \geq 0$ or $t \geq 0$ or $u \geq 0$).

In order to represent the situation, one usually refers to the **Dalitz plot** [1] (Figure 2.8).

[1] or equilateral coordinates

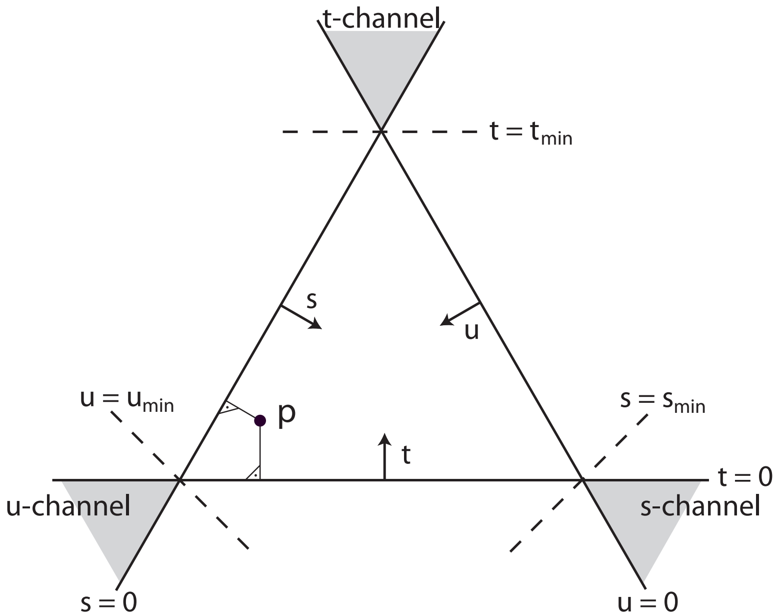


Figure 2.8: *Dalitz plot of s-, t-, and u-channels.*

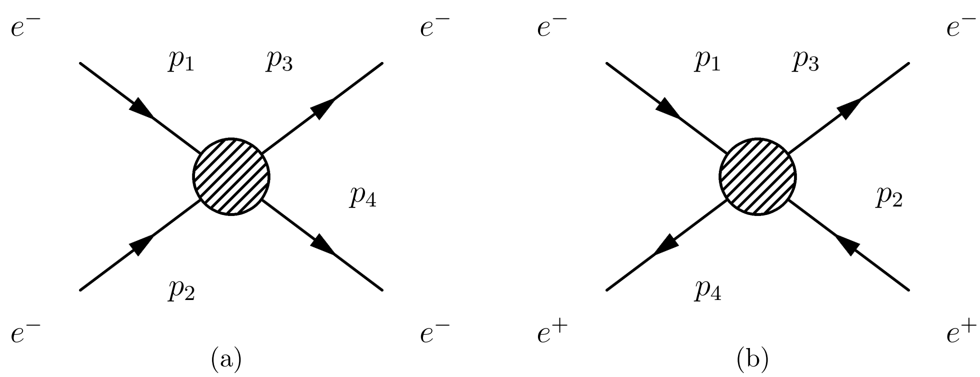


Figure 2.9: *Møller scattering (a) and Bhabha scattering (b).*

Example We take a look at the **Møller scattering**,

$$e^- e^- \to e^- e^-,$$

which is the s-channel of the reaction depicted on Figure 2.9(a). By crossing, we get as u-channel reaction the **Bhabha scattering**,

$$e^+ e^- \to e^+ e^-,$$

which is the reaction depicted on Figure 2.9(b).

The considerations of this chapter enable us to derive *constraints* on the possible dynamics but are not sufficient to *decide* on the dynamics. To “get” the dynamics we must calculate and compare to experiments decay rates and scattering cross-sections.

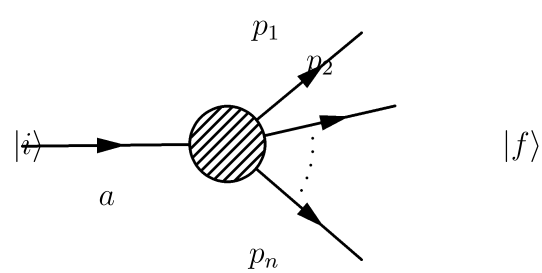
Chapter 3

Lorentz invariant scattering cross section and phase space

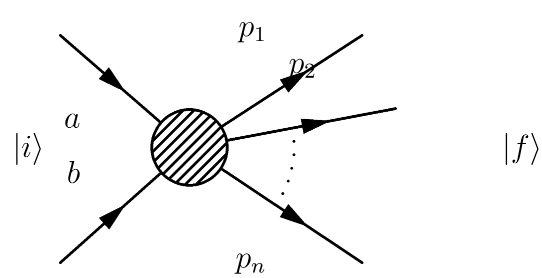
In particle physics, there are basically two observable quantities :

- Decay rates,
- Scattering cross-sections.

Decay:



Scattering:



3.1 $\mathcal{S}$-operator

In both cases $|i\rangle$ denotes the initial state, $|f\rangle$ denotes a multiparticle final state in a Fock space and the box represents the dynamics/interactions and is called the **$\mathcal{S}$-operator**. The last is predicted by the theory describing the interaction.

Example In QM I/II, $\mathcal{S} \propto H'(t) \propto V(t)$ in the first order perturbation theory of the Schrödinger equation.

$\mathcal{S}$ is usually a very complicated object : it contains the information about *all* possible transitions $|i\rangle \to |f\rangle$. Another way to state this is to remark that $\mathcal{S}$ contains all the dynamics of the process.

In experiments one does not get/need/want the full $\mathcal{S}$-operator. Instead, one restricts oneself to specific $|i\rangle$ and $|f\rangle$ e.g. by choosing the beam particles (muon beam,...) for the first and looking only at specific outcomes (3-jets events,...) for the latter.

One represents the $\mathcal{S}$-operator by looking at its matrix elements,

$$\underbrace{\sum_{f'} |f'\rangle \langle f'|}_{\mathbb{1}} \mathcal{S} |i\rangle = \sum_{f'} |f'\rangle \mathcal{S}_{f'i} \tag{3.1}$$

where

$$\mathcal{S}_{f'i} = \langle f'| \mathcal{S} |i\rangle \tag{3.2}$$

To isolate a specific outcome $|f\rangle$, one multiplies Eq. (3.1) by $\langle f|$, and gets,

$$\langle f| \sum_{f'} |f'\rangle \mathcal{S}_{f'i} = \sum_{f'} \underbrace{\langle f\,|f'\rangle}_{=\delta_{ff'}} \mathcal{S}_{f'i} = \mathcal{S}_{fi}. \tag{3.3}$$

Hence, the probability for the process $|i\rangle \to |f\rangle$ is,

$$P(|i\rangle \to |f\rangle) = |\mathcal{S}_{fi}|^2 \tag{3.4}$$

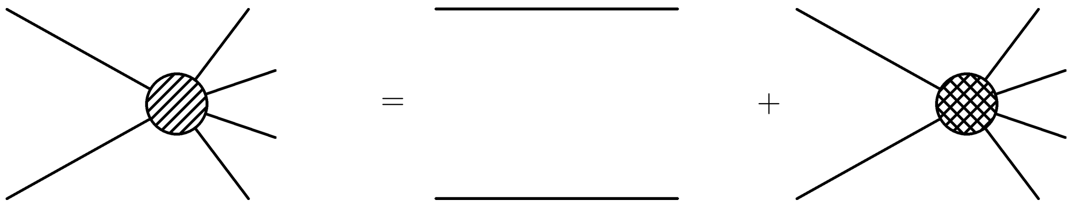
In general we can write,

$$\mathcal{S}_{fi} = \underbrace{\delta_{fi}}_{\text{no int.}} + \underbrace{i(2\pi)^4 \delta^{(4)}(p_f - p_i)}_{\text{4-momentum cons.}} \cdot \underbrace{\mathcal{T}_{fi}}_{\text{scat. amplitude}}, \tag{3.5}$$

or using a shorthand notation

$$\mathcal{S} = \mathbb{1} + i\mathcal{T},$$

in Feynman diagrams:



In the discussion of particle physics, a frequently used quantity is the transition probability per unit time,

$$w_{fi} = \frac{|\mathcal{S}_{fi}|^2}{T}. \tag{3.6}$$

3.2 Fermi's golden rule

From Eqs. (3.4) and (3.5), we see that we must address the issue of defining the value of a squared Dirac δ-function. To do this we use the rather pragmatic approach due to Fermi:

$$\begin{aligned}\left[2\pi\delta(p_f^0 - p_i^0)\right]^2 &= \int dt\, \mathrm{e}^{i(p_f^0 - p_i^0)t} \cdot 2\pi\delta(p_f^0 - p_i^0) \\ &= T \cdot 2\pi\delta(p_f^0 - p_i^0) \end{aligned} \tag{3.7}$$

$$\begin{aligned}\left[(2\pi)^3\delta^{(3)}(\vec{p}_f - \vec{p}_i)\right]^2 &= \iiint d^3x\, \mathrm{e}^{i(\vec{p}_f - \vec{p}_i)\cdot\vec{x}} \cdot (2\pi)^3\delta^{(3)}(\vec{p}_f - \vec{p}_i) \\ &= V \cdot (2\pi)^3\delta^{(3)}(\vec{p}_f - \vec{p}_i) \end{aligned} \tag{3.8}$$

$$\Rightarrow w_{fi} = \frac{|\mathcal{S}_{fi}|^2}{T} = V \cdot (2\pi)^4\delta^{(4)}(p_f - p_i) \cdot |\mathcal{T}_{fi}|^2 \tag{3.9}$$

To talk about the transition rate, we look at a Fock-space with a fixed number of particles.

Experimentally, the angle and energy-momentum is only accessible up to a given accuracy. We therefore use differential cross-sections in angle $d\Omega$ and energy-momentum dp near Ω, p respectively.

Motivating example In a cubic box of volume $V = L^3$ with infinitely high potential wells, the authorized momentum-values are discretely distributed.

$$p = \frac{2\pi}{L}n \Rightarrow dn = \frac{L}{2\pi}dp \Rightarrow d^3n = \left(\frac{L}{2\pi}\right)^3 d^3p,$$

and hence,

$$dw_{fi} = V \cdot (2\pi)^4\delta^{(4)}(p_f - p_i) \cdot |\mathcal{T}_{fi}|^2 \cdot \prod_{f=1}^{n_f} \frac{V}{(2\pi)^3} d^3p_f, \tag{3.10}$$

where n_f stands for the number of particles in the final state.

In order to get rid of normalization factors, we define a new matrix element $\mathcal{M}_{fi}$ by,

$$\mathcal{T}_{fi} \stackrel{!}{=} \left(\prod_{i=1}^{n_i} \frac{1}{\sqrt{2E_i V}}\right)\left(\prod_{f=1}^{n_f} \frac{1}{\sqrt{2E_f V}}\right)\mathcal{M}_{fi}. \tag{3.11}$$

At first sight, the apparation of the energies of both the initial and final states might be surprising. It is however needed in order to compensate the noninvariance of the volume, so that EV is a Lorentz invariant quantity. From now on we will always normalize our states to $2E$ (instead of 1 as is usually the case in nonrelativistic quantum mechanics).

We now substitute the definition (3.11) in Eq. (3.10) to get the fundamentally important expression,

$$\boxed{dw_{fi} = \frac{V^{1-n_i}}{(2\pi)^{3n_f-4}}\delta^{(4)}(p_f - p_i)\cdot|\mathcal{M}_{fi}|^2 \cdot \prod_{i=1}^{n_i}\frac{1}{2E_i}\prod_{f=1}^{n_f}\frac{d^3p_f}{2E_f}}. \tag{3.12}$$

We can then specify this result for the two cases of interest, as we do in the following subsections.

3.2.1 Total decay rate

In the case where $n_i = 1$, we view w_{fi} as a **decay rate** for the reaction,

$$a \to 1 + 2 + \cdots + n_f.$$

We have

$$\Gamma_{a\to\{n_f\}} = w_{\{f\}a} \qquad \text{(decay width)}, \tag{3.13}$$

$$\tau_{a\to\{n_f\}} = \frac{1}{\Gamma_{a\to\{n_f\}}} \qquad \text{(lifetime)}, \tag{3.14}$$

where $\{n_f\}$ stands for the n_f-particle final state $1 + 2 + \cdots + n_f$.

The next step is the definition of the **total decay width**,

$$\boxed{\Gamma_a = \sum_{\{n_f\}}\Gamma_{a\to\{n_f\}} = \frac{1}{2E_a}\frac{1}{(2\pi)^{3n_f-4}}\cdot\int\frac{d^3p_1}{2E_1}\cdots\frac{d^3p_{n_f}}{2E_{n_f}}\delta^{(4)}(p_f - p_i)|\mathcal{M}_{fi}|^2}, \tag{3.15}$$

and the lifetime

$$\tau_a = \frac{1}{\Gamma_a} \tag{3.16}$$

We remark that since E_a is not a Lorentz invariant quantity, Γ_a also depends on the reference frame. The quantity stated under the name "lifetime" in particle physics listings is always the lifetime as measured in the rest frame of the particle and is hence always the shortest one.

Example Without relativistic time dilation, one would expect the μ leptons generated by cosmic rays in the high atmosphere and traveling almost at the speed of light to be able to travel $c\tau_\mu \approx 600\,\mathrm{m}$ before decaying, making their detection on the earth surface almost impossible. When one takes time dilation into account, the distance becomes $c\tau_\mu \approx 10\,\mathrm{km}$, which is in accordance with the observed μ leptons number reaching the earth. This was actually for long the only available test of special relativity.

3.2.2 Scattering cross section

We now analyze the case of $n_i = 2$, i.e. the case of two particles interacting via the reaction,

$$a + b \to 1 + 2 + \cdots + n_f,$$

thus getting the **scattering cross section** $\sigma(a + b \to 1 + 2 + \cdots n_f)$ defined by,

$$\sigma = \frac{\#\text{ of transitions } a + b \to 1 + 2 + \cdots n_f \text{ per unit time}}{\#\text{ of incoming particles per unit surface and time}} = \frac{w_{fi}}{\text{incoming flux}}. \tag{3.17}$$

The denominator can also be stated as,

$$\text{incoming flux} = (\text{number density}) \cdot (\text{relative velocity}) = \frac{v_{ab}}{V}.$$

Using Eqs. (2.12) and (3.17) we then find,

$$\boxed{\sigma_{i\to\{n_f\}} = \frac{1}{4F}\frac{1}{(2\pi)^{3n_f-4}}\int\left(\prod_{f=1}^{n_f}\frac{d^3p_f}{2E_f}\right)\delta^{(4)}\left(\sum_{f=1}^{n_f}p_f - p_a - p_b\right)|\mathcal{M}_{fi}|^2}, \tag{3.18}$$

in which we see once more the Lorentz invariant Møller flux factor,

$$\begin{aligned} F &= E_a E_b v_{ab} = \sqrt{(p_a \cdot p_b)^2 - m_a^2 m_b^2} \\ &= \frac{1}{2}\sqrt{(s - (m_a + m_b)^2)(s - (m_a - m_b)^2)} \xrightarrow{s \gg m_a^2, m_b^2} \frac{s}{2}. \end{aligned} \tag{3.19}$$

From the form of (3.18), we see that the total cross section is manifestly a Lorentz invariant quantity, since it only depends on Lorentz invariants.

3.2.3 Invariant phase space for n_f-particles

We have already seen that the scattering angle is related to the Mandelstam t-variable (Section 2.2.2).

In order to make the same statement for multiparticle final states, we define the n_f-particles phase space,

$$R_{n_f} = \int dR_{n_f} = \int \frac{d^3p_1}{2E_1} \cdots \frac{d^3p_{n_f}}{2E_{n_f}} \delta^{(4)} \left(\sum_{f=1}^{n_f} p_f - \sum_{i=1}^{n_i} p_i \right). \tag{3.20}$$

We now prove that R_n is a Lorentz invariant quantity.

$$\frac{d^3p_i}{2E_i} = \int_0^\infty dE_i \delta(p_i^2 - m_i^2) d^3p_i \tag{3.21}$$

$$= \int_{-\infty}^\infty \underbrace{d^4p_i}_{\text{L.I.}} \delta(\underbrace{p_i^2 - m_i^2}_{\text{L.I.}}) \underbrace{\theta(E_i)}_{E_i>0 \text{ is L.I.}} . \tag{3.22}$$

3.2.4 Differential cross section

In order to get the differential cross section, we define,

$$t_{jk} := (p_j - p_k)^2 = f(\angle(\vec{p}_j, \vec{p}_k)), \tag{3.23}$$

and write

$$\frac{d\sigma}{dt_{jk}} = \frac{1}{4F} \frac{1}{(2\pi)^{3n_f-4}} \int dR_{n_f} |\mathcal{M}_{fi}|^2 \delta(t_{jk} - (p_j - p_k)^2). \tag{3.24}$$

Starting from this expression, one can deduce differential distributions in all other kinematical variables (energies, angles) by expressing those through the t_{jk}'s.

3.3 $2 \to 2$ scattering cross section

Next we turn our attention towards the very important special case of $2 \to 2$ scattering, $n_i = n_f = 2$:

$$a + b \to 1 + 2.$$

3.3.1 Phase space

First, we take a look at the phase space R_2, we see that there are 6 integration variables and 4 constraints, i.e. we are left with only 2 free parameters. The goal of the next steps will be to get rid of the δ-functions.

$$\begin{aligned}
R_2 &= \int \frac{d^3\vec{p}_1}{2\tilde{E}_1}\frac{d^3\vec{p}_2}{2\tilde{E}_2}\delta^{(4)}(p_1+p_2-p_a-p_b) \\
&\overset{(3.26)}{=} \int d^4p_1\delta(p_1^2-m_1^2)d^4p_2\delta(p_2^2-m_2^2)\theta(E_1)\theta(E_2)\delta^{(4)}(p_1+p_2-p_a-p_b) \\
&\overset{(3.27)}{=} \int d^4p_1\delta(p_1^2-m_1^2)\delta((p_a+p_b-p_1)^2-m_2^2)\theta(E_1)\theta(E_a+E_b-E_1) \\
&= \int_0^{E_a+E_b} dE_1 \int_0^{\infty} |\vec{p}_1|^2 d|\vec{p}_1| d\Omega\delta(E_1^2-\vec{p}_1^{\,2}-m_1^2)\delta((p_a+p_b-p_1)^2-m_2^2) \\
&\overset{(3.28)}{=} \int_0^{E_a+E_b} dE_1 d\Omega \underbrace{\frac{\sqrt{E_1^2-m_1^2}}{2}}_{\int_0^\infty |\vec{p}_1|^2\delta(E_1^2-\vec{p}_1^{\,2}-m_1^2)d|\vec{p}_1|} \delta\left(s-2(p_a+p_b)\cdot p_1+m_1^2-m_2^2\right)
\end{aligned} \tag{3.25}$$

where we have used,

$$\frac{1}{2\tilde{E}_1} = \int dE_1\delta(E_1^2-\vec{p}_1^{\,2}-m_1^2)\theta(E_1), \quad \tilde{E}_1 = \sqrt{m_1^2+\vec{p}_1^{\,2}} \tag{3.26}$$

$$1 = \int d^4p_2\delta^{(4)}(p_1+p_2-p_a-p_b), \tag{3.27}$$

$$\delta(E_1^2-\vec{p}_1^{\,2}-m_1^2) = \frac{1}{2|\vec{p}_1|}\left(\delta\left(|\vec{p}_1|-\sqrt{E_1^2-m_1^2}\right)+\underbrace{\delta\left(|\vec{p}_1|+\sqrt{E_1^2-m_1^2}\right)}_{=0,\text{ since }|\vec{p}_1|\geq 0}\right). \tag{3.28}$$

We did not make any assumption about the reference frame up to this point. We now specify our calculation for the center of mass frame,

$$\vec{p}_a+\vec{p}_b = 0 \Rightarrow E_a+E_b = \sqrt{s},$$

bringing Eq. (3.25) into,

$$\begin{aligned}
R_2 &= \int_0^{\sqrt{s}} dE_1^* d\Omega^* \frac{|\vec{p}_1^{\,*}|}{2}\delta(s-2\sqrt{s}E_1^*+m_1^2-m_2^2) \\
&= \int d\Omega^* \frac{|\vec{p}_1^{\,*}|}{4\sqrt{s}} \\
&\Rightarrow dR_2 = \frac{1}{8s}\sqrt{\lambda(s,m_1^2,m_2^2)}d\Omega^*.
\end{aligned} \tag{3.29}$$

For the last steps we used Eq. (2.10) and the fact that,

$$\delta(s-2\sqrt{s}E_1^*+m_1^2-m_2^2) = \frac{1}{2\sqrt{s}}\delta\left(E_1^*-\frac{1}{2\sqrt{s}}(s+m_1^2-m_2^2)\right).$$

A last step of the calculation can be made if the integrand has *no* angular dependency: since we are in the center of mass frame, we then have manifestly a 4π-symmetry and the scattering angle can take any value, the only restriction being that the two scattered particles are flying back-to-back in the center of mass frame. Therefore R_2 is then simply the integrand multiplied with the volume of the unit sphere, i.e.

$$R_2 = \int dR_2 = \frac{\pi}{2s}\sqrt{\lambda(s, m_1^2, m_2^2)}. \tag{3.30}$$

This simplification always applies for a $1 \to 2$ decay, but usually not for a $2 \to 2$ scattering reaction, where the incoming beam direction breaks the 4π-symmetry.

3.3.2 Differential cross section

Using Eq. (2.11) and (3.24) for $n_f = 2$, we get,

$$\frac{d\sigma}{d\Omega^*} = \frac{d\sigma}{dt}\frac{dt}{d\Omega^*} = \frac{|\vec{p}_1^{\,*}|}{64\pi^2 F\sqrt{s}}|\mathcal{M}_{fi}|^2, \tag{3.31}$$

resulting in the **differential cross section**,

$$\boxed{\frac{d\sigma}{d\Omega^*} = \frac{1}{64\pi^2 s}\frac{|\vec{p}_1^{\,*}|}{|\vec{p}_a^{\,*}|}|\mathcal{M}_{fi}|^2}, \tag{3.32}$$

since from Eq. (2.12) $F = \sqrt{s}|\vec{p}_a^{\,*}|$.

For the special case of elastic scattering $|\vec{p}_1^{\,*}| = |\vec{p}_a^{\,*}|$, we get,

$$\boxed{\frac{d\sigma^{el.}}{d\Omega^*} = \frac{1}{64\pi^2 s}|\mathcal{M}_{fi}|^2}. \tag{3.33}$$

Finally, we write here the invariant differential cross section of $a + b \to 1 + 2$ in terms of invariants for future references,

$$\boxed{\frac{d\sigma}{dt} = \frac{1}{16\pi\lambda(s, m_a^2, m_b^2)}|\mathcal{M}_{fi}|^2} \xrightarrow{s \gg m_1^2, m_2^2} \frac{1}{16\pi s^2}|\mathcal{M}_{fi}|^2. \tag{3.34}$$

3.4 Unitarity of the $\mathcal{S}$-operator

We can compute the transition probability from the matrix elements for the transition $|i\rangle \to |f\rangle$,

$$|\mathcal{S}_{fi}|^2 = |\langle f|\,\mathcal{S}\,|i\rangle|^2, \tag{3.35}$$

$$\sum_f |\mathcal{S}_{fi}|^2 = 1, \tag{3.36}$$

where $\sum_f$ stands for

$$\sum_{\text{spins, particle types, quantum numbers}} \int \prod_f \left(\frac{V}{(2\pi)^3} d^3 p_f \right).$$

Developing and using the completeness relation,

$$\sum_f |f\rangle \langle f| = \mathbb{1},$$

we obtain

$$\begin{aligned} 1 = \sum_f \langle i| \mathcal{S}^\dagger |f\rangle \langle f| \mathcal{S} |i\rangle = \langle i| \mathcal{S}^\dagger \mathcal{S} |i\rangle \qquad\qquad & \forall\, |i\rangle \\ \Rightarrow \boxed{\mathcal{S}^\dagger \mathcal{S} = \mathbb{1}}, \qquad\qquad & \end{aligned} \tag{3.37}$$

in other words $\mathcal{S}$ is a unitary operator.

This important fact has profound implications. We state here two of them.

First, for two orthogonal states $|i\rangle$ and $|j\rangle$, we have,

$$\langle j| \mathcal{S}^\dagger \mathcal{S} |i\rangle = \langle j|i\rangle = \delta_{ij}.$$

The other implication concerns the expression introduced in Eq. (3.5),

$$\mathcal{S}_{fi} = \delta_{fi} + i(2\pi)^4 \delta^{(4)}(p_f - p_i) \cdot \mathcal{T}_{fi}. \tag{3.38}$$

For a free theory, $\mathcal{T}_{fi} = 0$ and hence $\mathcal{S}_{fi} = \delta_{fi}$. On the other hand, for an interacting theory $\mathrm{Im}\, \mathcal{S}_{fi} \neq 0$.

An obvious comparison of the real and imaginary parts of $\mathcal{S}_{fi}$, tells us that,

$$\begin{aligned} \mathrm{Re}\, \mathcal{T}_{fi} &\rightsquigarrow \mathrm{Im}\, \mathcal{S}_{fi} \qquad\qquad && \text{(virtual contribution)}, \\ \mathrm{Im}\, \mathcal{T}_{fi} &\rightsquigarrow \mathrm{Re}\, \mathcal{S}_{fi} \qquad\qquad && \text{(absorbtive contribution)}. \end{aligned}$$

Taking a closer look at the absorbtive contribution, we get,

$$2i \mathrm{Im}\, \mathcal{T}_{fi} = \mathcal{T}_{fi} - \mathcal{T}^*_{fi} = i(2\pi)^4 \delta^{(4)}(p_f - p_i) \sum_n \mathcal{T}_{fn} \mathcal{T}^*_{in},$$

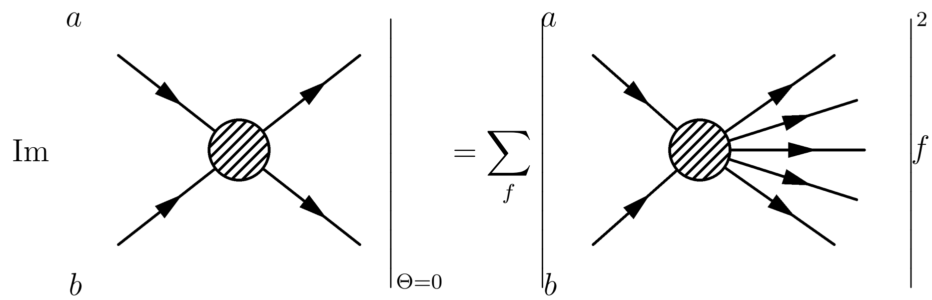
where n denotes an intermediate state.

The special case of elastic forward scattering ($|f\rangle = |i\rangle\,, \Theta^* = 0$) yields the surprising **optical theorem**,

$$\boxed{\mathrm{Im}\, \mathcal{M}_{ii} = \sqrt{\lambda(s, m_a^2, m_b^2)}\sigma_{tot}}, \tag{3.39}$$

relating a very specific element of $\mathcal{S}_{fi}$ with the total cross section for the transition $|i\rangle \to |f\rangle$, which is a measure for the probability for this transition to occur at all.

We can rewrite it symbolically with Feynman diagrams:



The computation of the matrix elements $\mathcal{M}_{fi}$ will be treated from Chapter 5 on.

Chapter 4

Accelerators and collider experiments

This chapter gives an introduction to particle accelerators and detectors as well as to data analysis tools relevant in this context. This involves the definition and application of concepts based on the kinematics developed in Chap. 2. Basic principles of particle accelerators are discussed as well as fixed target and beam collider experiments. The concepts of center of mass energy, luminosity, cross section, and event rates are introduced, followed by the basic building blocks of particle physics experiments. In order to be able to analyze the data gathered with collider experiments, we will introduce the concepts of rapidity, transverse and missing momentum (applications of momentum conservation) and invariant mass.

Modern techniques in experimental particle physics can be classified according to their use of accelerators. Non-accelerator-based experiments (e. g. the setup in Fig. 4.1) include measurements based on cosmic rays, solar and atmospheric neutrinos, and searches for dark matter. The latter, together with dark energy, could account for 95% of the universe. In the case of cosmic rays we can study high energy particles without having to accelerate them. Advances in neutrino physics have been achieved using large targets of (heavy) water surrounded by photomultipliers (e. g. Super-Kamiokande: neutrino oscillations). Accelerator-based experiments, on the other hand, include fixed target experiments and particle colliders, which are the topic of this chapter. As an example for particle colliders, the Large Hadron Collider (LHC) is shown in Fig. 4.2 with its four collision sites.

4.1 Particle accelerators: motivations

Particle accelerators are a fundamental tool for research in physics. Their importance and fields of use can be understood when one considers their main parameter, the beam energy. If we intend to use accelerators as large "microscopes", the spatial resolution increases with beam energy. According to the de Broglie equation, the relation between

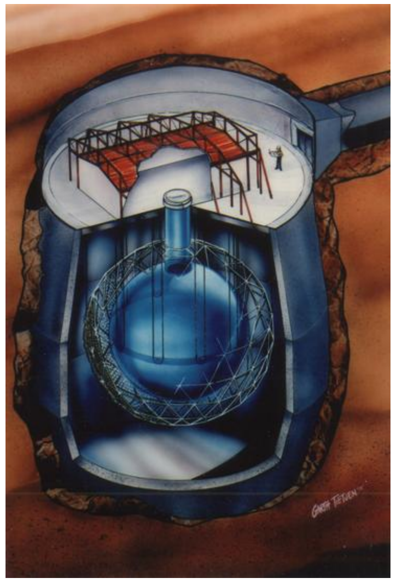


Figure 4.1: *Example of a non-accelerator-based experiment.* Heavy water targets can be protected from radiation background by installing them in deep-underground facilities. The target is surrounded by photomultipliers.

Figure 4.2: *The Large Hadron Collider at CERN with its four experiments CMS, ATLAS, LHCb, and ALICE.*

momentum $|\vec{p}|$ and wavelength λ of a wave packet is given by

$$\lambda = \frac{h}{|\vec{p}|}. \tag{4.1}$$

Therefore, larger momenta correspond to shorter wavelengths and access to smaller structures. In addition, it is possible to use accelerators to produce new particles. As we have seen in Chap. 2, this requires the more energy the heavier the particles are. Because beams are circulated for several hours accelerators are based on beams of stable particles and antiparticles, such as e^+, e^- or $p, \bar{p}$ or e, p (Deutsches Elektronen-Synchrotron, DESY). There are two possibilities as to what to collide a beam of accelerated particles with:

1. collision with another beam;
2. collision with a fixed target.

In both cases one can study the resulting interactions with particle detectors. By using a fixed target, one can furthermore produce a beam of secondary particles that may be stable, unstable, charged or neutral, solving the impossibility of accelerating unstable or neutral particles directly.

In the search for new sub-structures, Eq. (4.1) is the fundamental relation. It tells us that the resolution increases as we go to higher energies. For instance the resolution of 1 GeV/c and 10^3 GeV/c are:

$$\begin{aligned}
|\vec{p}| &= 1\,\frac{\text{GeV}}{\text{c}} \to \lambda = 1.24 \cdot 10^{-15}\,\text{m} \simeq \text{size of a proton} \\
|\vec{p}| &= 10^3\,\frac{\text{GeV}}{\text{c}} \to \lambda = 1.24 \cdot 10^{-18}\,\text{m} \simeq \text{size of proton substructures, e.g. quarks.}
\end{aligned}$$

Consider now the second scenario mentioned above, namely the search for new particles with high mass. For a collision of a particle with mass m_1 and momentum $\vec{p}_1$ with another particle $m_2, \vec{p}_2$ the energy in the laboratory frame is given by[1]

$$\begin{aligned}
E_L &= \sqrt{\vec{p}_1^{\,2} c^2 + m_1^2 c^4} + \sqrt{\vec{p}_2^{\,2} c^2 + m_2^2 c^4} \\
|\vec{p}_L| &= |\vec{p}_1 + \vec{p}_2| \\
E_L^2 - \vec{p}_L^{\,2} c^2 &= E^{*2} - \underbrace{\vec{p}^{\,*2}}_{=0} c^2 \\
\Rightarrow E^* &= \sqrt{E_L^2 - \vec{p}_L^{\,2} c^2}.
\end{aligned}$$

The production energy threshold for particles produced at rest is therefore:

$$E^* = \sum_i m_i c^2, \text{ while } E_{\text{kin}} = 0$$

[1]Recall that we asterisk quantities given in the center of mass frame. See Sect. 2.2 for labeling conventions.

where m_i is the mass of the i-th particle of the final state. We can conclude that, since the center of mass energy E^* grows with the energy in the laboratory frame E_L, we can produce higher masses if we have higher energies at our disposal. This allows to produce particles not contained in ordinary matter.

Example: As an example, consider inelastic proton collisions. Imagine we want to produce three protons and one antiproton by colliding a proton beam against a proton target (e. g. a hydrogen target). The corresponding reaction is

$$pp \to \bar{p}ppp$$

where conservation of the baryon number requires the presence of one antiproton in the final state. What is the minimum momentum of the proton beam for the reaction to take place? Since particles and antiparticles have the same mass and the target is at rest in the laboratory frame, we find

$$\begin{aligned} m_1 = m_2 = m &= 0.9383\,\frac{\mathrm{GeV}}{\mathrm{c}^2} \\ |\vec{p}_L| &= |\vec{p}_1|,\ |\vec{p}_2| = 0 \\ \text{at threshold: } E^* &= 4mc^2 = 3.7532\,\mathrm{GeV} \\ \Rightarrow |\vec{p}_1| &= 6.5\,\frac{\mathrm{GeV}}{\mathrm{c}}. \end{aligned}$$

4.1.1 Center of mass energy

As we have seen, the center of mass energy E^* is the energy available in collision experiments. We therefore want to compare fixed target and colliding beam experiments concerning their available energy. In the case of beam-target collision, E^* is determined by (with m the mass of both the beam and target particles)

$$\begin{aligned} E_L &= \sqrt{\vec{p}_L^{\,2}c^2 + m^2c^4} + mc^2 \\ E^{*2} &= M^2c^4 = E_L^2 - \vec{p}_L^{\,2}c^2 = 2m^2c^4 + 2mc^2\sqrt{\vec{p}_L^{\,2}c^2 + m^2c^4}. \end{aligned}$$

Setting $|\vec{p}_L| = p_{\mathrm{inc}}$ and neglecting the mass of the target we get:

$$E^* = \sqrt{2mc^2p_{\mathrm{inc}}c} = 1.37\,\sqrt{\mathrm{GeV}}\sqrt{p_{\mathrm{inc}}c} = 1.37\,\sqrt{\mathrm{GeV}}\sqrt{E_{\mathrm{inc}}}.$$

This means that, in the case of a fixed target experiment, the center of mass energy grows only with square root of E_{inc} (see Fig. 4.3).

However, in beam-beam collisions, we find $E^* = E_{\mathrm{CM}} = 2E_{\mathrm{inc}}$. Therefore, it is much more efficient to use two beams in opposite directions, as the following examples demonstrate

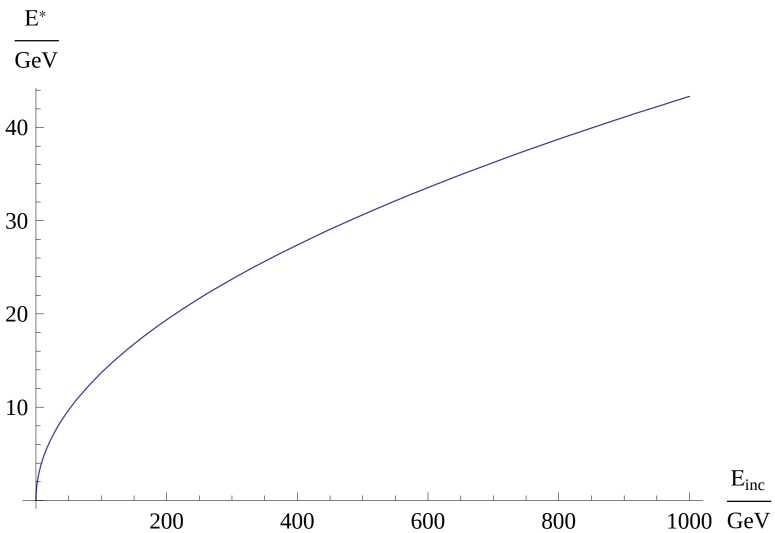


Figure 4.3: *Center of mass energy of the colliding beam for a fixed target experiment.* The energy increases with the square root of the beam energy.

(target is for instance hydrogen):

$$\overrightarrow{22\,\mathrm{GeV}} + \overleftarrow{22\,\mathrm{GeV}} \text{ has the same } E_{\mathrm{CM}} \text{ as } \overrightarrow{1\,\mathrm{TeV}} + m_{\mathrm{target}};$$
$$\overrightarrow{1\,\mathrm{TeV}} + \overleftarrow{1\,\mathrm{TeV}} \text{ has the same } E_{\mathrm{CM}} \text{ as } \overrightarrow{10^3\,\mathrm{TeV}} + m_{\mathrm{target}}.$$

The concept of colliding beams naturally leads to large circular accelerators. But for them to work properly some technical problems have to be solved. For instance, the particle density in a beam is much lower than in a solid or liquid target (see also the concept of luminosity in Sect. 4.3.2). Therefore, one tries to cross the beams many times and maximize the beam intensities (number of particle bunches per beam). As mentioned before, this approach only works with stable particles or antiparticles. Furthermore, in order to avoid beam-gas interactions (unintended fixed target collisions), a high vacuum is needed in the beam-pipe (about 10^{-9} Pa). Two beam lines are needed in the particle-particle case, whereas in the particle-antiparticle case one beam line is sufficient, with the two beams circulating in opposite directions. Finally, electronics represent another crucial part of the setup. At a rate of about $40 \cdot 10^6$ collisions per second a fast electronic system is necessary to decide what collisions to select.

4.2 Acceleration methods

Bearing in mind that an electric field $\vec{E}$ produces an accelerating force $\vec{F}$ on a charge q,

$$\vec{F} = q\vec{E},$$

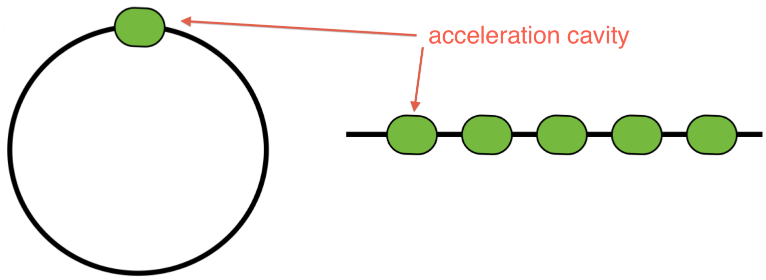


Figure 4.4: *Sketch of a circular (left) and linear (right) accelerator.* A circular machine needs to have one acceleration cavity, while a linear machine needs several cavities in series in order to reach high energies.

one could use an electrostatic field to accelerate charged particles. Since the maximal available potential difference (cf. Van de Graaff accelerator) is about 10 MV, one can accelerate particles up to 10 MeV. However, the fact that the electrostatic field is conservative ($\oint \vec{E} \cdot d\vec{l} = 0$) implies that the energy transfer only depends on the potential difference and not on the path. Therefore, circulating the beam in an electrostatic field does not lead to an increasing acceleration. The problem is solved by using several times a small but variable potential difference. This can be done using circular or linear machines. In a circular accelerator, one can use several times the same acceleration cavity (see Fig. 4.4, left), whereas in a linear accelerator several cavities in series are needed to reach high energies (see Fig. 4.4, right). In the case of a circular accelerator, the particles will receive a certain amount of energy at every turn, provided they are in phase with the accelerating potential. Because of the inertia principle, one further needs a magnetic field providing the centripetal force to keep particles on a circular path. An outline of historical developments in particle accelerators is given in Tab. 4.1. In the following sections, we will take a more detailed look at two types of accelerators: cyclotrons and synchrotrons.

4.2.1 Cyclotron

The sketch of a cyclotron is shown in Fig. 4.5. Particles are injected in the center and accelerated with a variable potential while a magnetic field $\vec{B}$ keeps them on spiral trajectories. Finally, particles are extracted and used in experiments. Cyclotrons are rather compact, as one can also see in Fig. 4.6. The maximal energy is of order 20 MeV for cyclotrons and up to 600 MeV for synchro-cyclotrons. For a particle moving in the cyclotron the centripetal and Lorentz forces are balanced:

$$m\frac{v^2}{\rho} = qvB \tag{4.2}$$

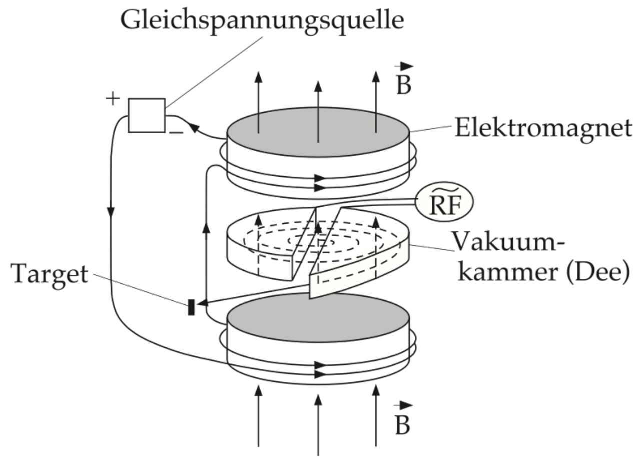


Figure 4.5: *Sketch of a cyclotron accelerator.* Source: [8, p. 108].

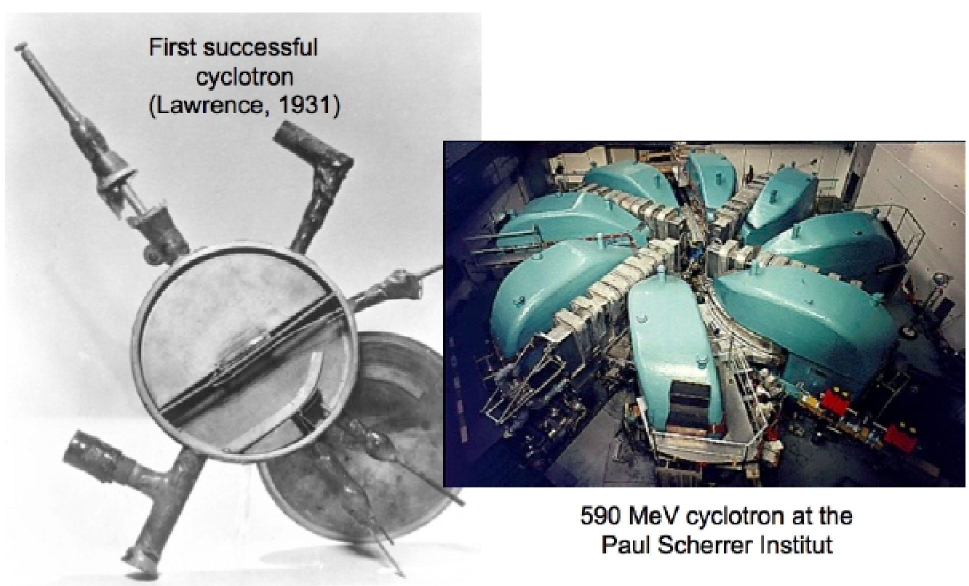


Figure 4.6: *A first prototype of a cyclotron (by Lawrence) and the* 590 MeV *isochronous cyclotron at PSI.*

Year	Accelerator	Beam energy
1921	"Kaskadengenerator" (Greinacher)	
1924–1928	Concept and first prototype of linear accelerator (Ising / Wideröe)	
1932	First nuclear reaction induced by cascade particle accelerator, $p^7\mathrm{Li} \to 2\alpha$ (Cockroft / Walton)	400 keV protons
1930	First Van de Graaff accelerator	1.5 MV
1930–1932	First cyclotron (concept: Lawrence)	1.5 MeV
	Upgraded cyclotrons (Synchrocyclotron)	300 − 700 MeV
1953	First synchrotron at Brookhaven lab—Cosmotron (concept: Oliphant / Veksler / McMillan)	3 GeV
1958	Proton Syncrotron (CERN)	28 GeV
1983	Tevatron (Fermilab)	1000 GeV
1990	HERA (DESY): first and only electron-proton collider	
2008	Large Hadron Collider (CERN)	up to 7000 GeV

Table 4.1: Evolution timeline in particle accelerators (q. v. [7, pp. 9]).

where v is the velocity of the particle, m the mass, q the charge, and ρ the trajectory radius. This yields for the cyclotron frequency ω

$$v = \omega\rho \tag{4.3}$$

$$\Rightarrow \omega = \frac{qB}{m}. \tag{4.4}$$

The alternating high voltage used to accelerate the particles (see Fig. 4.5) matches the cyclotron frequency, such that the particles are accelerated when passing the capacitor between the two half disks, also called as "D's". We can also conclude that the radius of the particle trajectory grows linearly with its momentum. For relativistic particles, Eq. (4.4) has to be modified:

$$\omega' = \frac{qB}{\gamma m}$$

where $\gamma = 1/\sqrt{1 - v^2/c^2}$. This modification has, for example, the following effect on the revolution frequency:

$$\frac{v}{c} = 50\% \Rightarrow \gamma = 1.155 \Rightarrow \omega' = 0.86\omega$$
$$\frac{v}{c} = 99\% \Rightarrow \gamma = 7.1 \Rightarrow \omega' = 0.14\omega.$$

Isochronous cyclotrons compensate for the variation in frequency by increasing the magnetic field (rather than changing the frequency) with the radius.

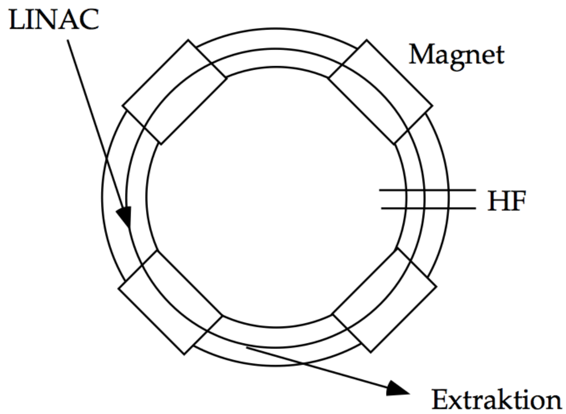


Figure 4.7: *Sketch of a synchrotron accelerator.* High frequency cavities are used to accelerate the particles. Dipole magnets keep them on circular trajectories. Linear accelerators are used for pre-acceleration and injection. Source: [8, p. 110].

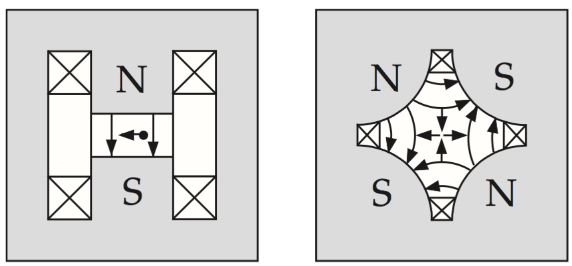


Figure 4.8: *Magnets used in synchrotrons.* Dipole magnets (left) keep the beam on a circular path, while quadrupole magnets (right) focus particles in the vertical or horizontal plane. Source: [8, p. 111].

4.2.2 Synchrotron

In the case of the synchrotron, the trajectory radius is kept constant. This is achieved by dipole magnets (see Fig. 4.8), while high frequency cavities are used to accelerate the particles (see Fig. 4.7). The problem of reducing the cross section to increase the particle density is solved by using quadrupole magnets (see Fig. 4.8). Their focussing and defocussing properties can be combined in a way as to lead to an overall focussing of the beam. Starting from Eq. (4.2), we have for the radius ρ

$$\rho = \frac{p}{qB}.$$

This yields, setting $q = ze$ (with e the unit charge),

$$cp[\mathrm{eV}] = czB\rho = 3 \cdot 10^8 \, \frac{\mathrm{m}}{\mathrm{s}} zB[\mathrm{T}]\rho[\mathrm{m}] \tag{4.5}$$

$$\Rightarrow p\left[\frac{\mathrm{GeV}}{\mathrm{c}}\right] = 0.3zB[\mathrm{T}]\rho[\mathrm{m}] \tag{4.6}$$

for the momentum.

Example: As an example, we consider the LHC at CERN: With a circumference of 27 km, yielding a radius of 4.3 km, an average magnetic field of 5.4 T is needed to keep protons with momentum 7 TeV/c on circular trajectories. Magnetic fields of this magnitude require very large currents and therefore superconductors which only work at low temperatures (about 2° K). The superconducting cables are therefore cooled with liquid helium.

Particle beams are injected into the vacuum pipe at relatively low energy with the magnetic field at its minimal value. Because the particles traverse acceleration cavities at every turn, the momentum grows accordingly. Since the beam has to be kept on the same radius, the magnetic filed also has to grow. On the other hand, rising velocity means changing revolution frequency and the frequency of the potential differences must be kept in phase with the particles. When maximum momentum is reached the accelerating cavities are switched off and the beam can be extracted to be used in experimental areas (see Fig. 4.9) or to be injected in larger synchrotrons (see Fig. 4.10). If the beam remains in the synchrotron ring it can be steered to cross other beams in collision points.

Another possible application of synchrotrons is to use the synchrotron radiation emitted by circulating beams. For this purpose one uses electrons, since they produce more synchrotron radiation than hadrons because of their smaller mass. The highly energetic photons emitted are used for measurements in solid state physics and protein research. An example is the Swiss Light Source at the Paul Scherrer Institute (Villigen, Switzerland), where electrons are pre-accelerated by a 100 MeV linear accelerator, injected into a synchrotron of 288 m circumference, kept on track by 36 dipole magnets with 1.4 T field, focussed by 177 quadrupole magnets, for a total beam energy of 2.8 GeV.

4.3 Particle physics experiments

In the following sections we introduce or recapitulate some basic concepts in particle physics experiments.

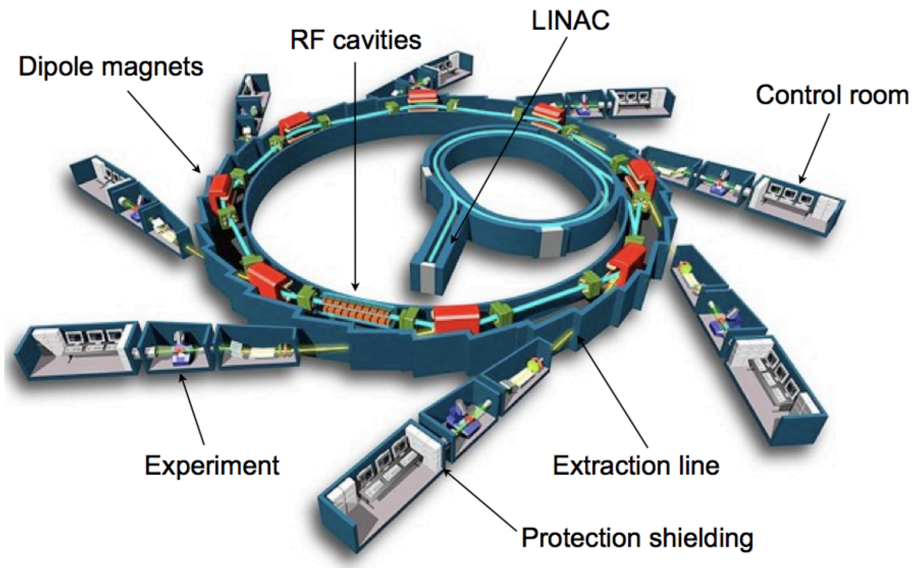


Figure 4.9: *Schematic view of a synchrotron.* Beams can be extracted and used in several experimental areas.

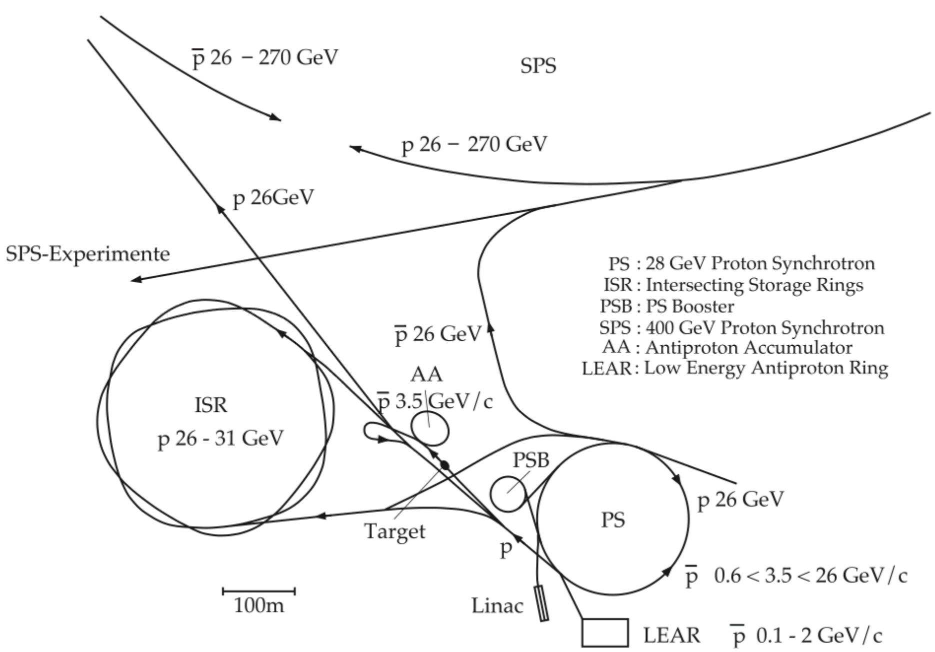


Figure 4.10: *Accelerator system at CERN.* Beams accelerated in linear machines and small synchrotrons are injected into larger synchrotron rings. Source: [8, p. 113].

4.3.1 Cross section

For a detailed introduction to the concept of cross section see Sect. 3.2.2. We recall that cross sections have dimension of area (cm^2). The common unit is barn, defined as

$$1\,\mathrm{b} = 10^{-24}\,\mathrm{cm}^2.$$

Until now, we have used the total cross section σ. This is a sum of contributions by many final states:

$$\sigma_{\mathrm{tot}} = \sum_i \sigma_i.$$

Example: Results of total cross section measurements for pp and $p\bar{p}$ collisions are shown in Fig. 4.11.

4.3.2 Luminosity

While cross sections characterize the scattering process, the luminosity characterizes an accelerators performance. With cross section σ and number of events per second R, the luminosity L is given by

$$R = L\sigma. \tag{4.7}$$

Because the dimension of the cross section is a surface, the units of luminosity are $\mathrm{cm}^{-2}\mathrm{s}^{-1}$.

The meaning of luminosity can be illustrated considering e. g. an e^+e^- accelerator with N particles per beam, revolving f times per second. We assume a Gaussian shaped beam with dimensions s_x and s_y, which yields a transverse size of $4\pi s_x s_y$. In one turn, one electron crosses $N/(4\pi s_x s_y)$ positrons. Because there are N particles revolving in each beam f times per second the number of collisions per second is

$$L = \frac{fN^2}{4\pi s_x s_y}. \tag{4.8}$$

From Eq. 4.7, the number of events per second is

$$R = \frac{\sigma f N^2}{4\pi s_x s_y}. \tag{4.9}$$

From Eq. (4.8) we notice that the luminosity can be increased by reducing the cross section of the beam, by increasing the number of particles in the beam or by increasing the revolution frequency.

In general, the luminosity of an accelerator gradually increases over time, while accelerator physicists learn how to operate the machine and to squeeze the beam size at the

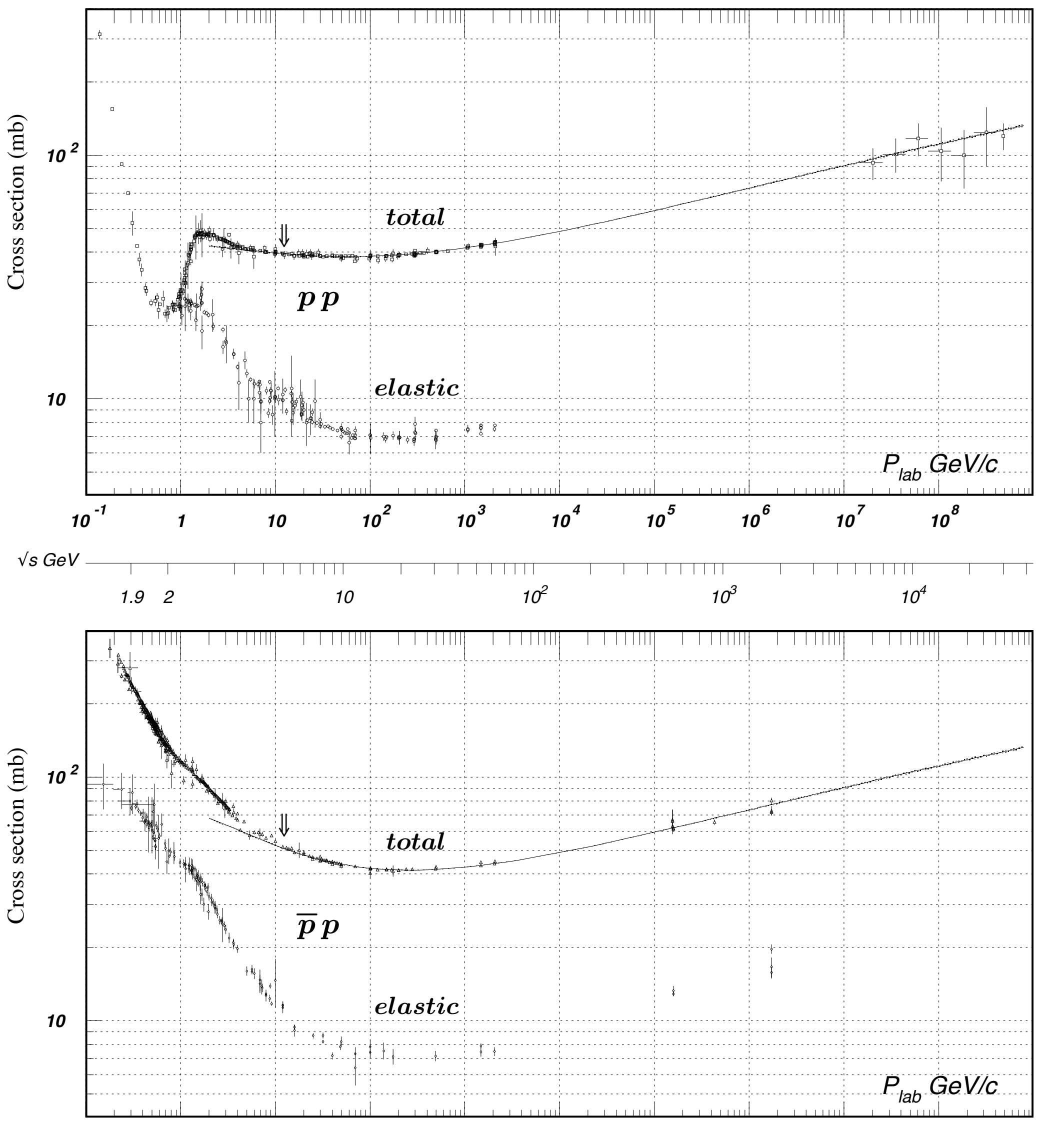


Figure 4.11: *Total and elastic cross sections for pp and $p\bar{p}$ collisions as a function of laboratory beam momentum and total center of mass energy.* Source: [9].

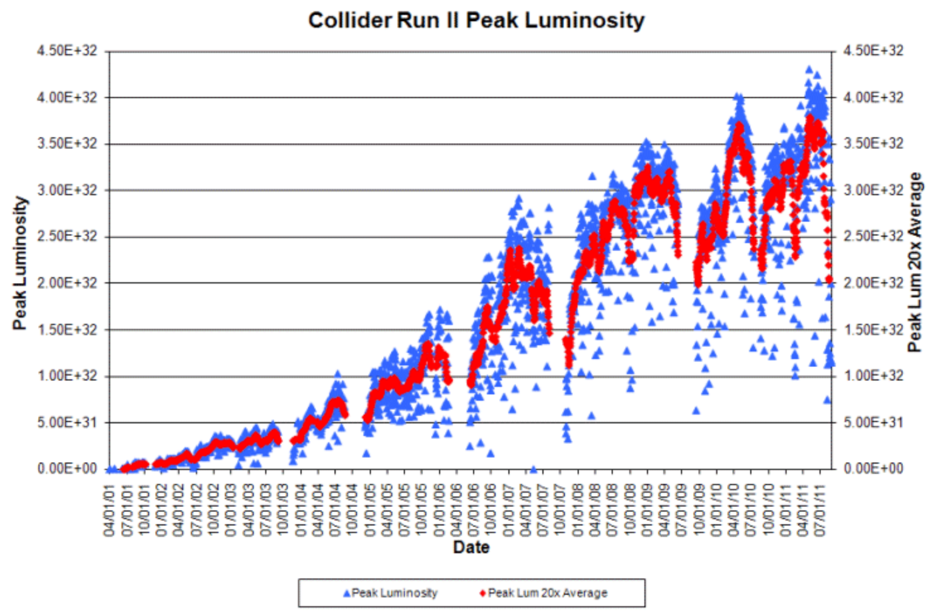


Figure 4.12: *Instantaneous luminosity at Tevatron as a function of time (2001 – 2009).* Note that the target luminosity for LHC is $10^{34}\,\mathrm{cm^{-2}s^{-1}}$.

intersection point. For example, the evolution of instantaneous luminosity over time at Tevatron is shown in Fig. 4.12.

The integral of the delivered luminosity over time is called integrated luminosity and is a measure of the collected data size. The integrated luminosity delivered by Tevatron until early 2009 is shown in Fig. 4.13.

Example: Consider an accelerator ring with the following properties:

- Ring length $= 100\,\mathrm{m}$;
- Revolution frequency $= 3 \cdot 10^6\,\mathrm{Hz} = 3\,\mathrm{MHz}$;
- $N = 10^{10}$ particles;
- $s_x = 0.1\,\mathrm{cm}$, $s_z = 0.01\,\mathrm{cm}$.

Using Eq. (4.9), we can calculate $L = 10^{29}\,\mathrm{cm^{-2}s^{-1}}$. If we are interested in a rare process, for example $e^+e^- \to p\bar{p}$ (the cross section is $\sigma = 1\,\mathrm{nb} = 10^{-33}\,\mathrm{cm}^2$) and have $E_{\mathrm{CM}} \sim 2 - 3\,\mathrm{GeV}$ we only expect $R = 10^{-4}$ events per second or about 0.35 events per hour.

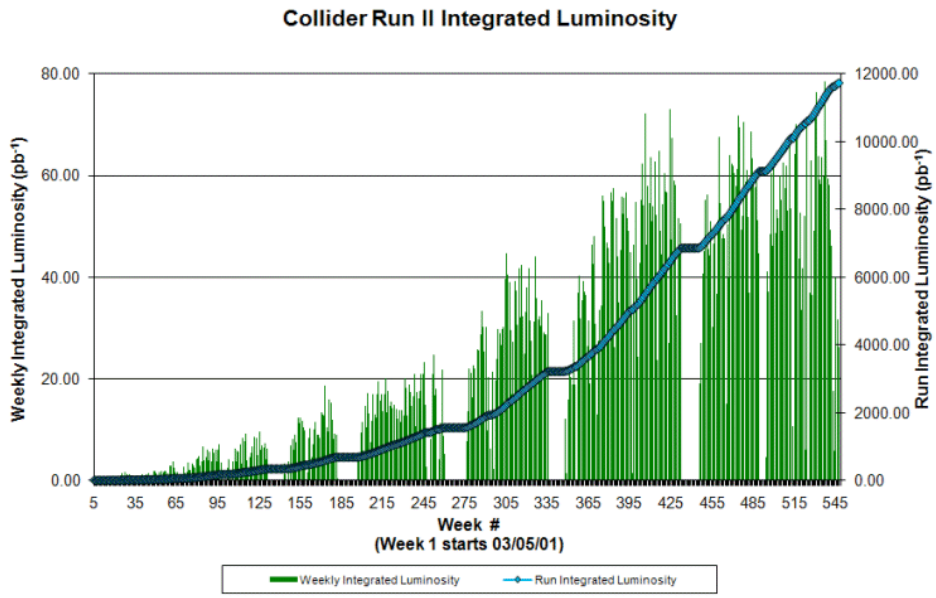


Figure 4.13: *Integrated luminosity at Tevatron as function of time.*

4.3.3 Particle detectors

To gather data from experiments carried out at accelerators, we need particle detectors. They are disposed around the interaction region and detect (directly or indirectly) the reaction products. Typically, the following measurements are performed on final state particles:

- Spatial coordinates and timing of final state;
- Momentum;
- Energy;
- Type of particle (particle ID).

Because of kinematical constraints, for fixed target experiments the production of final states is mainly in the forward direction. Therefore, the detector has to cover only a small solid angle (see Fig. 4.14). In colliding beam experiments, on the other hand, cylindrically symmetric detectors with hermeticity down to small angles are preferred (see Fig. 4.15). A collider physics experiment has in general tracking detectors in a solenoidal field surrounded by calorimeters and particle ID detectors (e. g. muon ID). To allow the momentum measurements, a solenoidal magnetic field is applied parallel to the colliding beams. The particles trajectories in the magnetic field are measured in the inner layers by silicon pixel and silicon strip tracking devices. They are surrounded by calorimeters measuring the particles' energy. The general structure of such a detector, shown in Fig. 4.15,

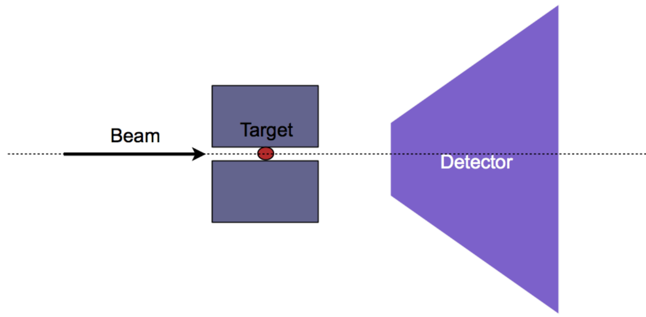


Figure 4.14: *Schematic view of an experimental setup for a fixed target experiment.*

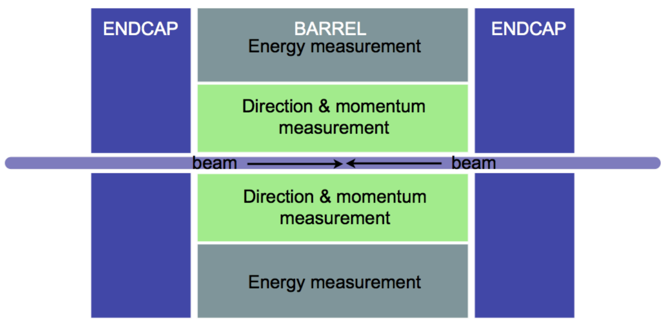


Figure 4.15: *Schematic view of a detector for colliding beam experiments.*

is also visible in the Compact Muon Solenoid (CMS) experiment at LHC. A sketch of the CMS experiment is given in Fig. 4.16.

In high energy experiments the momentum measurement is based on the deflection of charged particles in a magnetic field. Consider a simple case involving a dipole magnet (Fig. 4.17(a)). One can measure the track direction before and after the bending influence of the magnetic field to obtain the angle θ. The momentum is derived from Eq. (4.6):

$$\begin{aligned} p &= 0.3BR \\ \text{length} = l &= 2R\sin\left(\frac{\theta}{2}\right) \sim R\theta \\ \Rightarrow \theta &= \frac{\text{length}}{R} = \frac{0.3Bl}{p} \\ \Rightarrow p &= \frac{0.3Bl}{\theta}. \end{aligned}$$

In collider experiments the B field is parallel to the beams, which means that curvature only happens in the transverse plane (Fig. 4.17(b)). The momentum resolution is given by

$$\frac{\sigma(p_{\mathrm{T}})}{p_{\mathrm{T}}} = \frac{\sigma_{r\phi}p_{\mathrm{T}}}{0.3Bl_R^2}\left[\frac{720}{n+4}\right]^{-\frac{1}{2}}$$

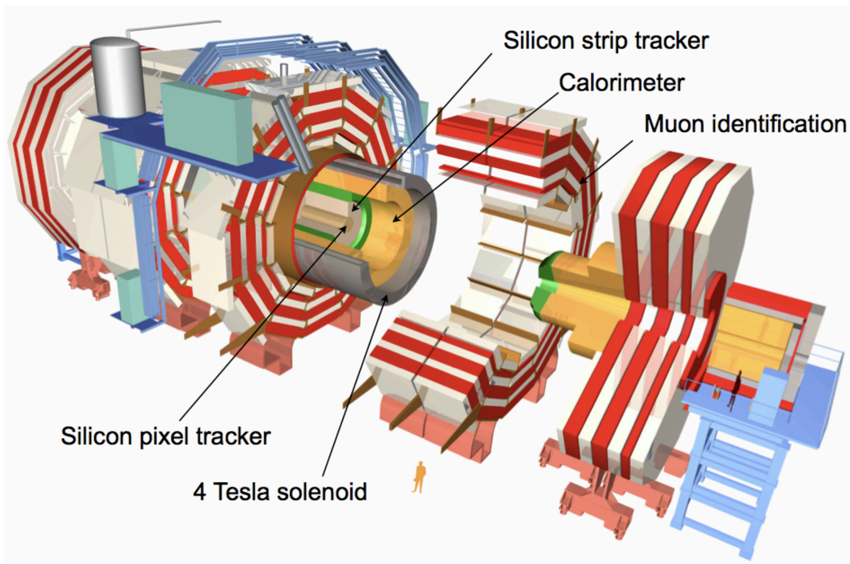
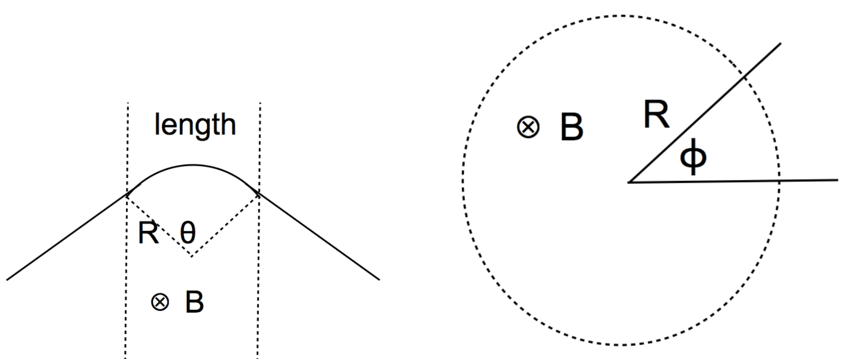
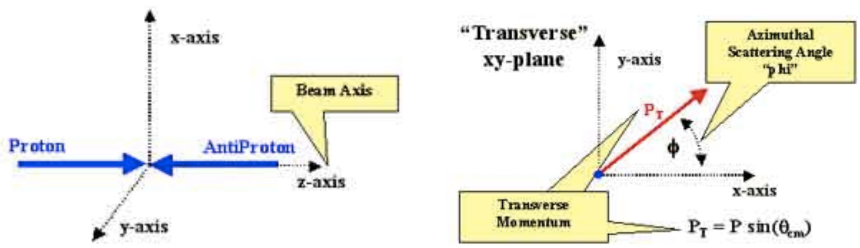


Figure 4.16: *The CMS experiment at the LHC.*



(a) (b)

Figure 4.17: *Momentum measurement in collider experiments using a magnetic field.* The magnetic field is parallel to the beams (orthogonal to the page).



(a) (b)

Figure 4.18: *Axes labelling conventions (a) and definition of transverse momentum (b).* Source: [10].

where $\sigma_{r\phi}$ is the error on each measurement point, l_R the radial length of the track, and n the number of equidistant points.

4.4 Kinematics and data analysis methods

In this section we describe the data analysis tools used in collider particle physics experiments discussed in Sect. 4.3. We introduce variables in the laboratory frame and methods based on momentum conservation and invariant mass. Momentum conservation leads to the concepts of transverse momentum and missing mass. As examples, we discuss two- and three-jet events as well as the W boson discovery.

4.4.1 Pseudorapidity and transverse momentum

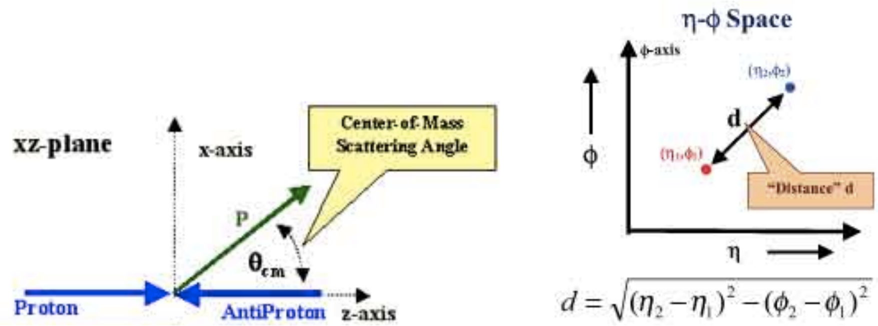
Consider the collision of two beams in the laboratory frame. The axes labelling conventions are given in Fig. 4.18(a) ($p\bar{p}$ scattering). The momentum of each particle produced in a collision can be decomposed in a component parallel to the beams (longitudinal, along the z direction) and one perpendicular to the beams (transverse, in the xy plane) as shown in Fig. 4.18(b). The transverse component of the momentum is given by ($\Theta^* \equiv \theta_{\mathrm{CM}}$)

$$p_{\mathrm{T}} = p \sin(\theta_{\mathrm{CM}})$$

and spans an angle ϕ with the x axis. To measure the longitudinal angle of the emerging particle jet one usually uses a variable called pseudorapidity η. It is defined by

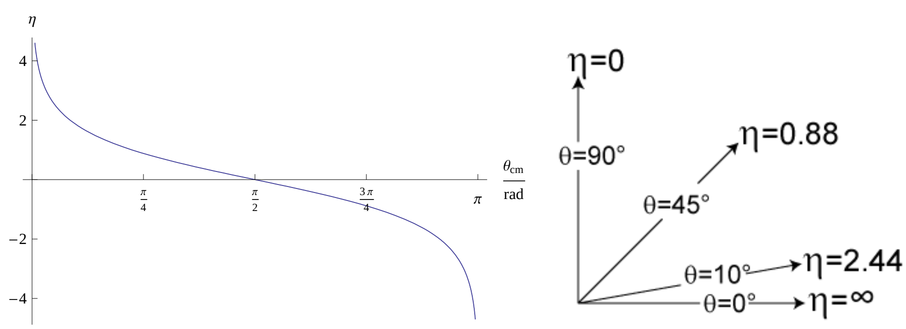
$$\eta = -\ln\left[\tan\left(\frac{\theta_{\mathrm{CM}}}{2}\right)\right]$$

and is Lorentz invariant under longitudinal boosts (see Fig. 4.19(a)). Momenta in the transverse plane are also invariant under longitudinal relativistic transformations. Therefore, the distance between single particles or jets of particles is usually measured in the $\eta\phi$ plane, as shown in Fig. 4.19(b).



(a) (b)

Figure 4.19: *Definition of the longitudinal scattering angle θ_{CM} (a) and definition of particle distance in the η-ϕ plane (b).* Source: [10].



(a) (b)

Figure 4.20: *Pseudorapidity as a function of θ_{CM} (a) and pseudorapidity for various values of θ_{CM} (b).* Source (b): [11].

Particles produced at $\theta_{\mathrm{CM}} = 90°$ have zero pseudorapidity. As visualized by Fig. 4.20(a) and 4.20(b), high $|\eta|$ values are equivalent to very shallow scattering angles. Typical coverage of central detectors extends to $|\eta| \sim 3$. Coverage of high rapidities ($\theta_{\mathrm{CM}} < 5°$) can be achieved with detectors placed at large z positions.

4.4.2 Momentum conservation in particle jets

Experiments in hadron colliders usually deal with particles at high transverse momentum. This is because the incoming particles collide head-on and have no transverse momentum before scattering and therefore, the final state particles must have zero *total* transverse momentum. Processes involving large momentum transfer produce particles in the center of the detector (small pseudorapidity). An example of such a process is given in Fig. 4.21. The experimental signature of a two jet event is shown in Fig. 4.22. The calorimeter measures the deposited energy in cells of the η-ϕ plane. Both charged and neutral particles are detected. The histogram shows the energy measured in each cell. Note that the main signals are symmetric in azimuth and at about zero pseudorapidity. The momentum of

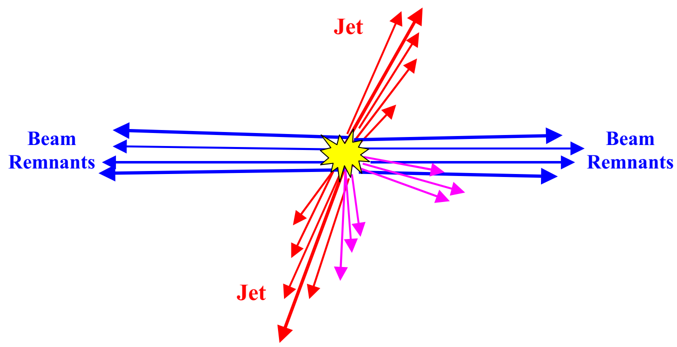


Figure 4.21: *Two jet event production at hadron colliders.* Source: [12].

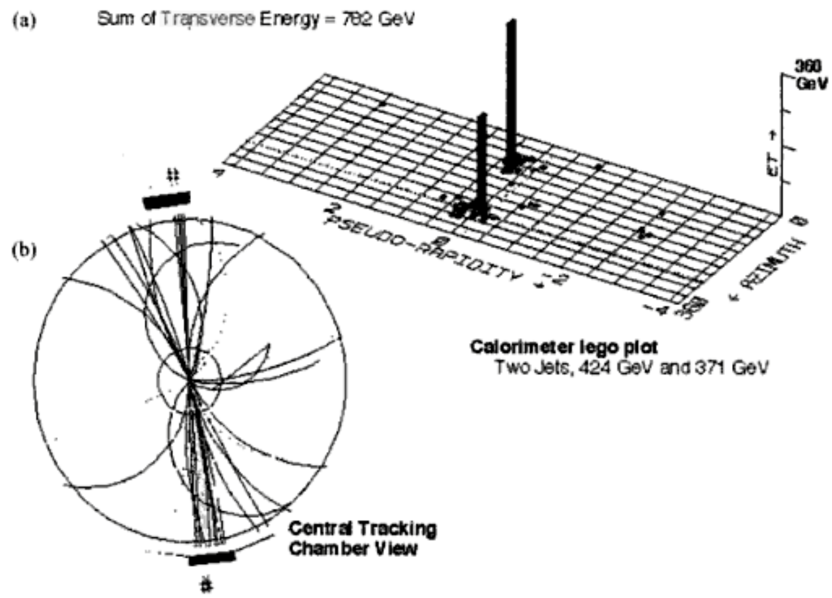


Figure 4.22: *Two jet event, reconstructed in the tracking chamber (b) and calorimeter signals (a) of the DØ experiment.*

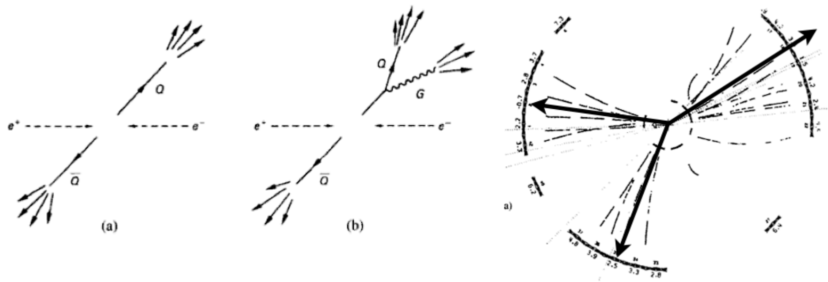


Figure 4.23: *Two- and three-jet events in e^+e^- collisions.* The rightmost sketch shows the tracks reconstructed in the central tracking detector.

each *charged* particle in a jet is measured by the central tracking chamber. Low momentum components yield smaller bending radii and the total transverse momentum has to be zero.

Electron-positron pairs can annihilate producing quark pairs (see Fig. 4.23(a)). This was studied for example at the Large Electron-Positron Collider (LEP). In some cases, a gluon can be radiated from one of the outgoing quarks (see Fig. 4.23(b)). In the latter case one observes three particle jets in the final state: two quark jets and one gluon jet. If no particle escapes the detector the three jets must have total transverse energy equal to zero. In the next section we discuss the case of particles escaping the experiment undetected. This topic is discussed more thoroughly in Chap. 8.

4.4.3 Missing mass method

A collision is characterized by an initial total energy and momentum ($E_{\text{in}}, \vec{p}_{\text{in}}$). In the final state we have n particles with total energy and momentum given by:

$$E = \sum_i^n E_i, \tag{4.10}$$

$$\vec{p} = \sum_i^n \vec{p}_i. \tag{4.11}$$

Sometimes an experiment may measure $E < E_{\text{in}}$ and $\vec{p} \neq \vec{p}_{in}$. In this case one or more particles have not been detected. Typically this happens with neutral particles, most often neutrinos, but also with neutrons, π^0, or K^0_{L}. The latter have a long lifetime and may decay outside the sensitive volume. To quantify this process, we introduce the concept of missing mass:

$$\text{missing mass} \times c^2 = \sqrt{(E_{\text{in}} - E)^2 - (\vec{p}_{\text{in}} - \vec{p})^2 c^2}. \tag{4.12}$$

The missing mass is measured for every collision and its spectrum is plotted. If the spectrum has a well-defined peak one particle has escaped our detector.

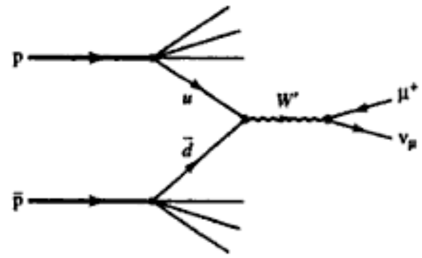


Figure 4.24: *Production and decay of a W^+ boson in a $p\bar{p}$ collision.*

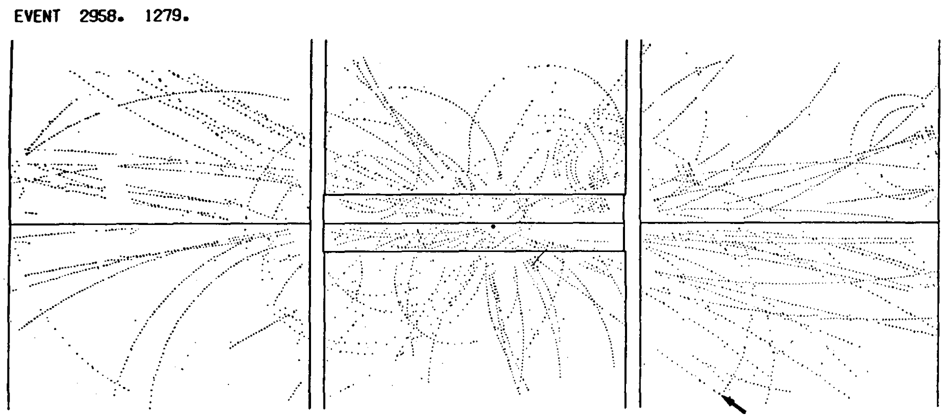


Figure 4.25: *Event with a W boson decay candidate via $W^+ \to e^+ + \nu_e$.* The event was recorded by the UA1 experiment (CERN). Source: [13, p. 112].

Example: Consider the decay of W bosons. They can be produced in proton-anti-proton collisions mainly via the process shown in Fig. 4.24; a u quark collides with an anti-d quark producing a W^+ boson. The W^+ then decays into a neutrino-lepton pair. The muon is detected and its momentum can be measured. The neutrino escapes the detector undetected. The total sum of the transverse momenta is therefore not zero! In other words, the experimental signature of the neutrino in the experiment is the *missing transverse momentum*. One of the first events [13, p. 112] attributed to production and decay of a W^+ boson is shown in Fig. 4.25. The arrow shows the lepton (e^+) and the missing momentum is compatible with the e^+ transverse momentum.

4.4.4 Invariant mass method

The invariant mass is a characteristic of the total energy and momentum of an object or a system of objects that is the same in all frames of reference. When the system as a whole is at rest, the invariant mass is equal to the total energy of the system divided by c^2. If

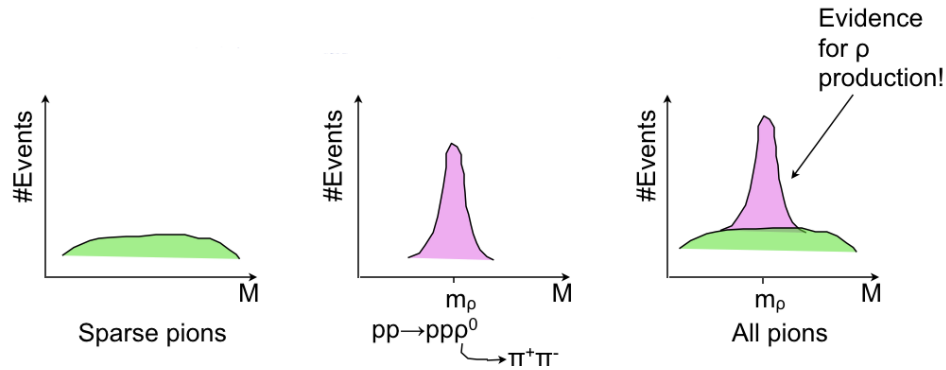


Figure 4.26: *Event distribution for invariant mass of the pion pair in the process $pp \to pp\pi^+\pi^-$.* The sparse pions' (left) distribution is broad and can be predicted using simulation techniques. The invariant mass of the pion pairs stemming from ρ^0 decay (center) is peaked around m_ρ. All pions contribute to the recorded events (right).

the system is one particle, the invariant mass may also be called the rest mass:

$$m^2c^4 = E^2 - \vec{p}^{\,2}c^2.$$

For a system of N particles we have

$$W^2c^4 = \left(\sum_i^N E_i\right)^2 - \left(\sum_i^N \vec{p}_i c\right)^2 \tag{4.13}$$

where W is the invariant mass of the decaying particle. For a particle of Mass M decaying into two particles, $M \to 1+2$, Eq. 4.13 becomes:

$$M^2c^4 = (E_1+E_2)^2 - (\vec{p}_1+\vec{p}_2)^2c^2 = m_1^2c^4 + m_2^2c^4 + 2(E_1E_2 - \vec{p}_1\cdot\vec{p}_2c^2) = (p_1+p_2)^2.$$

Example: Particles like ρ, ω, ϕ have average lifetime of $10^{-22} - 10^{-23}$ s. How do we know of their existence if they live so shortly? Consider, for example, the reaction $pp \to pp\pi^+\pi^-$. We identify all four particles in the final state and measure their momentum. Let's focus on the pion pair, the total energy and momentum of the pair are:

$$E = E_+ + E_-$$
$$\vec{p} = \vec{p}_+ + \vec{p}_-.$$

The corresponding invariant mass is

$$Mc^2 = \sqrt{E^2 - \vec{p}^{\,2}c^2}.$$

The event distribution for the variable M will look like the plot in Fig. 4.26. The peak in the event rate at m_ρ is evidence for ρ production.

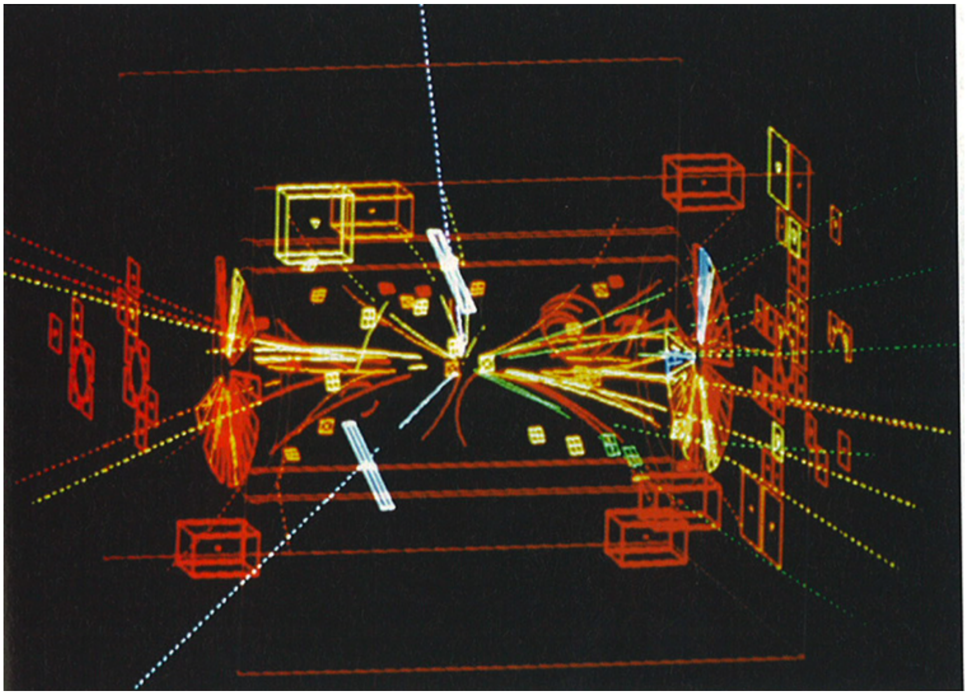


Figure 4.27: *Z^0 boson discovery at the UA1 experiment (CERN).* The Z^0 boson decays into a e^+e^- pair, shown as white dashed lines.

Example: Another example illustrating this point is the Z discovery in 1984. Fig. 4.27 shows an event where the Z boson, after production by proton-proton collision decays into an e^+e^- pair (white dashed lines). The invariant mass of the pair is about 92 GeV.

Example: Consider now the π^0 reconstruction. Neutral pions decay in photon pairs in about 99% of the cases. By measuring the angle and energy of the emitted photons (see Fig. 4.28) one can reconstruct the mass of the decaying pion (see Fig. 4.29).

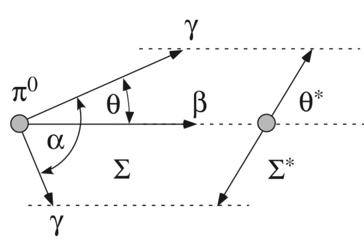


Figure 4.28: *π^0 decay in two photons.* Σ denotes the laboratory frame (left) and Σ^* denotes the pion rest frame (right). Source [8, p. 95].

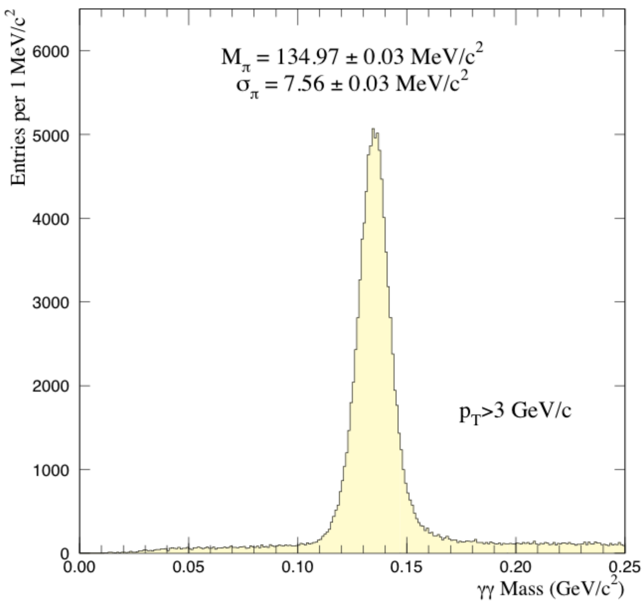


Figure 4.29: *Invariant mass spectrum for photon pairs.* The π^0 appears as a peak at the pion mass.

Example: In case of three body decays, $R \to 1+2+3$, one can define three invariant masses:

$$m_{12}^2 c^4 \equiv (p_1 + p_2)^2$$
$$m_{13}^2 c^4 \equiv (p_1 + p_3)^2$$
$$m_{23}^2 c^4 \equiv (p_2 + p_3)^2.$$

This yields

$$\begin{aligned} m_{12}^2 + m_{13}^2 + m_{23}^2 &= m_1^2 + m_2^2 + m_3^2 + (p_1 + p_2 + p_3)^2 \frac{1}{c^4} \\ &= m_1^2 + m_2^2 + m_3^2 + M^2. \end{aligned}$$

This means that there are only two independent invariant masses.

As an example, let's study the reaction:

$$K^- p \to \Lambda \pi^+ \pi^- \ (\Lambda \to \pi^- p).$$

We can measure two invariant masses:

$$m_{12} \equiv m(\Lambda\pi^-) \text{ and } m_{13} \equiv m(\Lambda\pi^+).$$

The so-called "Dalitz plot" given in Fig. 4.30 shows the relation between m_{13}^2 and m_{12}^2. The $\Sigma^\pm$ resonances appear as two bands in the Dalitz plot around 1.4 GeV.

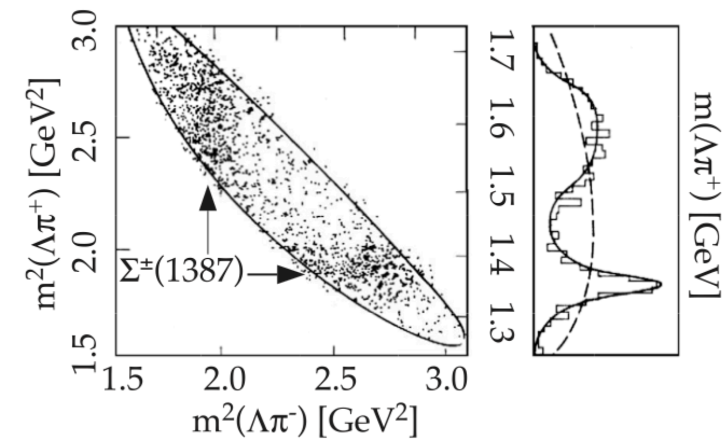


Figure 4.30: *Dalitz plot for* $K^-p \to \pi^+\pi^-\Lambda$ $(\Lambda \to \pi^-p)$. Source: [8, p. 200].

Chapter 5

Elements of quantum electrodynamics

5.1 Quantum mechanical equations of motion

In quantum mechanics I & II, the correspondence principle played a central role. It is in a sense the recipe to quantize a system whose Hamiltonian is known. It consists in the following two substitution rules :

$$\vec{p} \longmapsto -i\nabla \qquad \text{(momentum)}, \tag{5.1}$$
$$E \longmapsto i\partial_t \qquad \text{(energy)}. \tag{5.2}$$

For nonrelativistic quantum mechanics we get the celebrated **Schrödinger equation**,

$$i\partial_t\psi = H\psi, \qquad \text{with } H = -\frac{1}{2m}\triangle + V(\vec{x}), \tag{5.3}$$

whose free solution ($V(\vec{x}) \equiv 0$) is,

$$\psi(\vec{x}, t) = C\mathrm{e}^{-i(Et-\vec{p}\cdot\vec{x})} \qquad \text{with } E = \frac{\vec{p}^{\,2}}{2m}.$$

The relativistic version of the energy-momentum relationship is however,

$$E^2 = \vec{p}^{\,2} + m^2, \tag{5.4}$$

from which we get, using again the correspondence principle Eq. (5.1),

$$-\partial_t^2\psi = (-\triangle + m^2)\psi. \tag{5.5}$$

At this point we define some important symbols which will follow us troughout the rest of this lecture,

$$\partial_\mu := (\partial_t, \nabla), \tag{5.6}$$
$$\partial^\mu := g^{\mu\nu}\partial_\nu = (\partial_t, -\nabla), \tag{5.7}$$
$$\Box := \partial_\mu\partial^\mu = \partial_t^2 - \triangle. \tag{5.8}$$

With this notation we can then reformulate Eq. (5.5) to get the **Klein-Gordon equation**,

$$\boxed{\left(\Box + m^2\right)\psi = 0}, \tag{5.9}$$

with solutions,

$$\psi(\vec{x}, t) = C\mathrm{e}^{-i(Et - \vec{p}\cdot\vec{x})} \qquad \text{with } E = \pm\sqrt{\vec{p}^{\,2} + m^2}.$$

We see that in this case it is possible to have negative energy eigenvalues, a fact not arising with the nonrelativistic case.

As in the case of the Schrödinger equation (5.3) we can formulate a continuity equation. To do so we multiply the Klein-Gordon equation (5.9) by the left with ψ^* and its conjugate, $(\Box + m^2)\,\psi^* = 0$, by ψ and then subtract both equations to get,

$$\begin{aligned} 0 &= \psi^*\partial^\mu\partial_\mu\psi - \psi\partial^\mu\partial_\mu\psi^* \\ &= \partial^\mu(\psi^*\partial_\mu\psi - \psi\partial_\mu\psi^*) \\ &\Rightarrow \partial_t(\psi^*\partial_t\psi - \psi\partial_t\psi^*) + \nabla\cdot(\psi^*\nabla\psi - \psi\nabla\psi^*) = 0. \end{aligned} \tag{5.10}$$

We would like to interpret $\psi^*\partial_t\psi - \psi\partial_t\psi^*$ in Eq. (5.10) as a probability density, or more exactly,

$$\rho = i(\psi^*\partial_t\psi - \psi\partial_t\psi^*),$$

which is *not* a positive definite quantity (as we can convince ourselves by computing ρ for the plane wave solution), and hence cannot be interpreted as a probability density as it was the case in QM.

When computing the continuity equation for the Schrödinger equation, where such a problem does not arise, we see that the problem lies essentially in the presence of a *second* order time derivative in the Klein-Gordon equation.

We now make a big step, by imposing that our equation of motion only contains a *first* order time derivative. Since we want a Lorentz invariant equation of motion, we conclude that only a *linear* dependence on ∇ is allowed. Following Dirac's intuition, we make the ansatz,

$$(i\gamma^\mu\partial_\mu - m)\psi = (i\gamma^0\partial_t + i\vec{\gamma}\cdot\nabla - m)\psi = 0. \tag{5.11}$$

Turning back to the correspondence principle we remark that,

$$(\gamma^0 E - \vec{\gamma} \cdot \vec{p} - m)\psi = 0 \qquad (5.12)$$
$$\Rightarrow (\gamma^0 E - \vec{\gamma} \cdot \vec{p} - m)^2 \psi = 0,$$

which must stay compatible with the mass-shell relation, $E^2 = \vec{p}^2 + m^2$.

This implies that the γ^μ's cannot be numbers since it would then be impossible to satisfy,

$$(\vec{\gamma} \cdot \vec{p})^2 = \left(\sum_{i=1}^{3} \gamma^i p_i \right)^2 \overset{!}{\propto} \vec{p}^2,$$

so we let them be $n \times n$ matrices, for an n which is still to be determined.

We now derive relations that the γ^μ's must fullfill, so that the mass-shell relation remains true. From Eq. (5.12), and again with the correspondence principle, we must have,

$$\underbrace{i\partial_t}_{E} = (\gamma^0)^{-1} \vec{\gamma} \cdot \underbrace{(-i\nabla)}_{\vec{p}} + (\gamma^0)^{-1} m$$

$$\Rightarrow \underbrace{-\partial_t^2}_{E^2} = - \sum_{i,j=1}^{3} \frac{1}{2} \left((\gamma^0)^{-1}\gamma^i(\gamma^0)^{-1}\gamma^j + (\gamma^0)^{-1}\gamma^j(\gamma^0)^{-1}\gamma^i \right) \partial_i\partial_j \qquad (5.13)$$

$$- i \cdot m \sum_{i=1}^{3} \left((\gamma^0)^{-1}\gamma^i(\gamma^0)^{-1} + (\gamma^0)^{-1}(\gamma^0)^{-1}\gamma^i \right) \partial_i \qquad (5.14)$$

$$+ m^2 (\gamma^0)^{-1}(\gamma^0)^{-1} \qquad (5.15)$$

$$\overset{!}{=} (\partial_i\partial_i + m^2) \qquad (5.16)$$

$$= (\underbrace{-\triangle}_{\vec{p}^2} + m^2).$$

Comparing Eqs. (5.16) and (5.15) we conclude that,

$$(\gamma^0)^{-1}(\gamma^0)^{-1} \overset{!}{=} \mathbb{1} \Rightarrow (\gamma^0)^{-1} = \gamma^0. \qquad (5.17)$$

Defining $\{a, b\} := ab + ba$ and comparing Eq. (5.16) and Eq. (5.14) we get

$$\gamma^0\gamma^i(\gamma^0)^{-1} \overset{!}{=} 0 \Rightarrow \{\gamma^i, \gamma^0\} = 0. \qquad (5.18)$$

Finally, comparing Eq. (5.16) and Eq. (5.13), we have,

$$-\frac{1}{2} \left((\gamma^0)^{-1}\gamma^i(\gamma^0)^{-1}\gamma^j + (\gamma^0)^{-1}\gamma^j(\gamma^0)^{-1}\gamma^i \right) \overset{!}{=} \delta_{ij} \Rightarrow \{\gamma^i, \gamma^j\} = -2\delta_{ij}. \qquad (5.19)$$

We can summarize Eqs. (5.17), (5.18) and (5.19) in the **Clifford algebra of the γ-matrices**,

$$\boxed{\{\gamma^\mu, \gamma^\nu\} = 2g^{\mu\nu}\mathbb{1}}, \tag{5.20}$$

where $g^{00} = 1$, $g^{ii} = -1$ and all the other elements vanish.

Important facts : The eigenvalues of γ^0 can only be ± 1 and those of γ^i $\pm i$ and the γ-matrices have vanishing trace :

$$\begin{aligned}
\operatorname{Tr}\gamma^i &= \operatorname{Tr}(\gamma^0\gamma^0\gamma^i) = -\operatorname{Tr}(\gamma^0\gamma^i\gamma^0) = -\operatorname{Tr}\gamma^i \Rightarrow \operatorname{Tr}\gamma^i = 0,\\
\operatorname{Tr}\gamma^0 &= \operatorname{Tr}(\gamma^0\gamma^i(\gamma^i)^{-1}) = -\operatorname{Tr}(\gamma^i\gamma^0(\gamma^i)^{-1}) = -\operatorname{Tr}(\gamma^0) \Rightarrow \operatorname{Tr}\gamma^0 = 0.
\end{aligned}$$

The eigenvalue property of γ^0 implies with the last equation that the dimension n of the γ-matrices must be even.

For $n = 2$ there are no matrices satisfying Eq. (5.20), as can be checked by direct computation.

For $n = 4$ there are many possibilities. The most common choice in textbooks is the Dirac-Pauli representation :

$$\gamma^0 = \mathbb{1} \otimes \sigma^3 = \begin{pmatrix} \mathbb{1} & 0 \\ 0 & -\mathbb{1} \end{pmatrix}, \qquad \gamma^i = \sigma^i \otimes (i\sigma^2) = \begin{pmatrix} 0 & \sigma^i \\ -\sigma^i & 0 \end{pmatrix}, \tag{5.21}$$

with the Pauli matrices,

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

and the Kronecker product of 2×2-matrices,

$$A \otimes B = \begin{pmatrix} b_{11}A & b_{12}A \\ b_{21}A & b_{22}A \end{pmatrix}.$$

Looking at the **Dirac equation**

$$\boxed{(i\gamma^\mu\partial_\mu - m)\psi = 0} \tag{5.22}$$

we see that ψ is no longer a function but a vector, called **(4-)spinor**,

$$\psi = \begin{pmatrix} \psi_1 \\ \psi_2 \\ \psi_3 \\ \psi_4 \end{pmatrix}.$$

For 4-spinors, there are two types of adjoints, namely,

- the hermitian adjoint $\psi^\dagger = (\psi_1^*, \psi_2^*, \psi_3^*, \psi_4^*)$, and
- the Dirac adjoint $\bar{\psi} := \psi^\dagger\gamma^0 = (\psi_1^*, \psi_2^*, -\psi_3^*, -\psi_4^*)$.

Note that $\bar{\psi}$ satisfies a dirac equation of its own,

$$i\partial_\mu\bar{\psi}\gamma^\mu + m\bar{\psi} = 0. \tag{5.23}$$

We now focus our attention on the continuity equation for the Dirac field. From Eqs. (5.22) and (5.23),

$$i\psi^\dagger(\partial_t\psi) = \left(-i\psi^\dagger\gamma^0\gamma^i\partial_i + \psi^\dagger\gamma^0 m\right)\psi,$$

and its hermitian conjugate,

$$-i(\partial_t\psi^\dagger)\psi = \left(i(\partial_i\psi^\dagger)\gamma^0\gamma^i + \psi^\dagger\gamma^0 m\right)\psi,$$

we get the difference,

$$\begin{aligned}\partial_t(\psi^\dagger\psi) &= -\left[(\partial_i\psi^\dagger)\gamma^0\gamma^i\psi + \psi^\dagger\gamma^0\gamma^i i(\partial_i\psi)\right],\\ \partial_t(\bar{\psi}\gamma^0\psi) &= -\partial_i(\bar{\psi}\gamma^i\psi).\end{aligned} \tag{5.24}$$

We identify the components as,

$$\rho = \bar{\psi}\gamma^0\psi, \qquad \vec{j} = \bar{\psi}\vec{\gamma}\psi,$$

or interpreting them as components of a 4-vector as in classical electrodynamics,

$$j^\mu = \bar{\psi}\gamma^\mu\psi, \tag{5.25}$$

we see that Eq. (5.24) can be reexpressed in the manifestly covariant form,

$$\boxed{\partial_\mu j^\mu = 0}. \tag{5.26}$$

5.2 Solutions of the Dirac equation

Before we look at the solutions of the free Dirac equation, we introduce the slash notation for contraction with the γ-matrices : $\not{a} := \gamma^\mu a_\mu$. The Dirac equation then reads $(i\not{\partial} - m)\psi = 0$.

5.2.1 Free particle at rest

In the rest frame of a particle, the Dirac equation reduces to,

$$i\gamma^0\partial_t\psi = m\psi,$$

for which we find four linearly indepedent solutions, namely,

$$\psi_1 = \mathrm{e}^{-imt}\begin{pmatrix}1\\0\\0\\0\end{pmatrix}, \quad \psi_2 = \mathrm{e}^{-imt}\begin{pmatrix}0\\1\\0\\0\end{pmatrix}, \quad E = m \qquad \text{(particles)}$$
$$\psi_3 = \mathrm{e}^{+imt}\begin{pmatrix}0\\0\\1\\0\end{pmatrix}, \quad \psi_4 = \mathrm{e}^{+imt}\begin{pmatrix}0\\0\\0\\1\end{pmatrix}, \quad E = -m \qquad \text{(antiparticles).}$$

5.2.2 Free particle

In order to preserve the Lorentz invariance of a solution, it must only depend on Lorentz scalars – quantities which are invariant under Lorentz transformations – like $p \cdot x = p_\mu x^\mu$. We make the ansatz,

$$\begin{aligned}\psi_{1,2} &= \mathrm{e}^{-ip\cdot x}u_\pm(p), \qquad p^0 > 0\\ \psi_{3,4} &= \mathrm{e}^{+ip\cdot x}v_\mp(-p), \qquad p^0 < 0.\end{aligned}$$

Pluging those ansatz in the Dirac equation, we get,

$$(\not{p} - m)u_\pm(p) = \bar{u}_\pm(\not{p} - m) = 0, \tag{5.27}$$
$$(\not{p} + m)v_\pm(p) = \bar{v}_\mp(\not{p} + m) = 0, \tag{5.28}$$

where we replaced $-p \mapsto p$ in the second equation, having thus $p^0 > 0$ in both cases now.

5.2.3 Explicit form of u and v

As checked in the exercices, the explicit form for the u and v functions are,

$$u_\pm(p) = \sqrt{p^0 + m}\begin{pmatrix}\chi_\pm\\ \frac{\vec{\sigma}\cdot\vec{p}}{p^0+m}\chi_\pm\end{pmatrix}, \tag{5.29}$$
$$v_\pm(p) = \sqrt{p^0 + m}\begin{pmatrix}\frac{\vec{\sigma}\cdot\vec{p}}{p^0+m}\chi_\mp\\ \chi_\mp\end{pmatrix}, \tag{5.30}$$

where $\chi_+ = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ corresponds to a "spin up" state and $\chi_- = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ to a "spin down" state.

We note on the way that the application,

$$\vec{p} \longmapsto \vec{\sigma} \cdot \vec{p} = \sigma^i p_i = \begin{pmatrix} p_3 & p_1 - ip_2 \\ p_1 + ip_2 & -p_3 \end{pmatrix},$$

defines an isomorphism between the vector spaces of 3-vectors and hermitian 2×2-matrices.

5.2.4 Operators on spinor spaces

Hamiltonian The Hamiltonian is defined by $i\partial_t \psi = H\psi$. Isolating the time derivative in the Dirac equation, Eq. (5.22), we read out,

$$H = -i\gamma^0\gamma^i\partial_i + \gamma^0 m = \begin{pmatrix} m\mathbb{1} & \vec{\sigma} \cdot \vec{p} \\ \vec{\sigma} \cdot \vec{p} & -m\mathbb{1} \end{pmatrix}. \tag{5.31}$$

Helicity The helicity is the compenent of the spin in the direction of motion $\hat{\vec{p}} := \frac{\vec{p}}{|\vec{p}|}$, and is defined by,

$$h = \frac{1}{2}\vec{\sigma} \cdot \hat{\vec{p}} \otimes \mathbb{1} = \frac{1}{2}\begin{pmatrix} \vec{\sigma} \cdot \hat{\vec{p}} & 0 \\ 0 & \vec{\sigma} \cdot \hat{\vec{p}} \end{pmatrix}. \tag{5.32}$$

By direct computation, one can check that $[H, h] = 0$, and thus there exist a set of eigenfunctions diagonalizing H and h simultaneously. The eigenvalues of h are then constants of the motion and hence good quantum numbers to label the corresponding states.

This quantum number λ can take two values,

$$\lambda = \begin{cases} +\frac{1}{2} & \text{positive helicity} \leftrightsquigarrow \vec{s} \upharpoonleft\upharpoonright \vec{p}, \\ -\frac{1}{2} & \text{negative helicity} \leftrightsquigarrow \vec{s} \upharpoonleft\downharpoonright \vec{p}. \end{cases} \tag{5.33}$$

We stress here that helicity/handedness is *not* a Lorentz invariant quantity for massive particles.

Consider $\vec{p}$ in the z-direction, then,

$$\frac{1}{2}\vec{\sigma} \cdot \hat{\vec{p}}\chi_\pm = \frac{1}{2}\sigma^3\chi_\pm = \pm\frac{1}{2}\chi_\pm.$$

From the last argumentative steps, we are not surprised with the statement that the Dirac equation describes spin-$\frac{1}{2}$ particles.

Chirality Consider the Dirac equation for the case of massless particles. This is a good approximation for $E \gg m$, which is often the case in accelerator experiments. Setting $m = 0$ simplifies Eq. (5.22) leading to

$$i\gamma^\mu \partial_\mu \psi = 0.$$

Eq. (5.29) and (5.30) change accordingly. We consider for now the particle solutions $u_\pm$:

$$u_\pm(p) = \sqrt{|\vec{p}|}\begin{pmatrix} \chi_\pm \\ \frac{\vec{\sigma}\cdot\vec{p}}{|\vec{p}|}\chi_\pm \end{pmatrix} = \sqrt{|\vec{p}|}\begin{pmatrix} \chi_\pm \\ \pm\chi_\pm \end{pmatrix}. \tag{5.34}$$

It is convenient to define the so-called chirality matrix γ_5:

$$\gamma_5 = i\gamma^0\gamma^1\gamma^2\gamma^3$$

which in the Dirac-Pauli representation reads

$$\gamma_5 = \begin{pmatrix} 0 & \mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix}.$$

Using that the γ-matrices fulfill $\{\gamma^\mu, \gamma^\nu\} = 2g^{\mu\nu}$ (see Eq. (5.20)), one can show that

$$\{\gamma_5, \gamma^\mu\} = 0 \text{ and} \tag{5.35}$$
$$\gamma_5^2 = \mathbb{1}. \tag{5.36}$$

These properties of γ_5 imply that if ψ is a solution of the Dirac equation then so is $\gamma_5\psi$. Furthermore, since $\gamma_5^2 = \mathbb{1}$ the eigenvalues of the chirality matrix are ± 1:

$$\gamma_5\psi_\pm = \pm\psi_\pm$$

which defines the chirality basis $\psi_\pm$.

Let us apply the γ_5 matrix to the spinor part of particle solutions of the free Dirac equation given in Eq. (5.34):

$$\gamma_5 u_\pm(p) = \begin{pmatrix} 0 & \mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix}\sqrt{|\vec{p}|}\begin{pmatrix} \chi_\pm \\ \pm\chi_\pm \end{pmatrix} = \sqrt{|\vec{p}|}\begin{pmatrix} \pm\chi_\pm \\ \chi_\pm \end{pmatrix} \tag{5.37}$$
$$= \pm\sqrt{|\vec{p}|}\begin{pmatrix} \chi_\pm \\ \pm\chi_\pm \end{pmatrix} = \pm u_\pm(p). \tag{5.38}$$

A similar calculation shows that for the antiparticle solutions

$$\gamma_5 v_\pm(p) = \mp v_\pm(p). \tag{5.39}$$

Therefore, the helicity eigenstates for $m = 0$ are equivalent to the chirality eigenstates. Results (5.38) and (5.39) lead to the notion of handedness (which is borrowed from chemistry):

- u_+ describes a right handed particle: $\overrightarrow{\text{spin}} \upharpoonleft\upharpoonright \vec{p}_{e^-}$ and
- v_+ describes a left handed antiparticle: $\overleftarrow{\text{spin}} \downharpoonleft\upharpoonright \vec{p}_{e^+}$

where the converse holds for u_- and v_-.

Exploiting the eigenvalue equations (5.38) and (5.39), one can define the projectors

$$P_{\substack{R\\L}} = \frac{1}{2}(\mathbb{1} \pm \gamma_5). \tag{5.40}$$

They project to $u_\pm, v_\pm$ for arbitrary spinors. For example we have

$$P_L u_\pm = \frac{1}{2}(\mathbb{1} - \gamma_5)u_\pm = \frac{1}{2}(\mathbb{1} \mp \mathbb{1})u_\pm = \begin{cases} 0 \\ \mathbb{1}u_- \end{cases}.$$

To show that Eq. (5.40) indeed defines projectors, we check (using Eq. 5.36) for idempotence,

$$P^2_{\substack{R\\L}} = \frac{1}{4}(\mathbb{1} \pm \gamma_5)(\mathbb{1} \pm \gamma_5) = \frac{1}{4}(\mathbb{1} \pm 2\gamma_5 + \gamma_5^2) = \frac{1}{2}(\mathbb{1} \pm \gamma_5) = P_{\substack{R\\L}},$$

orthogonality,

$$P_R P_L = \frac{1}{4}(\mathbb{1} + \gamma_5)(\mathbb{1} - \gamma_5) = \frac{1}{4}(\mathbb{1} - \gamma_5^2) = 0,$$

and completeness,

$$P_R + P_L = \mathbb{1}.$$

Note that the projectors P_L and P_R are often used to indicate the chirality basis:

$$\begin{aligned} u_{L,R} &= P_{L,R}u \\ v_{L,R} &= P_{L,R}v. \end{aligned}$$

What has been derived so far rests on the assumption that the mass be zero. In this case, chirality is equivalent to helicity which is also Lorentz invariant. If, on the other hand $m \neq 0$, chirality and helicity are not equivalent: In this case chirality, while Lorentz invariant, is not a constant of the motion,

$$[\gamma_5, H_{\text{Dirac}}] \neq 0,$$

and therefore not a good quantum number. Helicity though is a constant of the motion, but, since spin is unaffected by boosts, it is not Lorentz invariant for non-vanishing mass: For every possible momentum $\vec{p}$ in one frame of reference there is another frame in

Figure 5.1: *Helicity for the case of non-vanishing mass.*

		Chirality $\gamma_5 = i\gamma^0\gamma^1\gamma^2\gamma^3$	**Helicity** $h(\hat{\vec{p}}) = \frac{1}{2}\vec{\sigma}\cdot\hat{\vec{p}}\otimes\mathbb{1}$
$m = 0$	Constant of motion	✓	✓
	Lorentz invariant	✓	✓
$m \neq 0$	Constant of motion	✗	✓
	Lorentz invariant	✓	✗

Table 5.1: Chirality and helicity.

which the particle moves in direction $-\vec{p}/|\vec{p}|$ (see Fig. 5.1). A comparison of chirality and helicity is given in Tab. 5.1.

Although chirality is not a constant of the motion for $m \neq 0$, it is still a useful concept (and becomes important when one considers weak interactions). A solution of the Dirac equation ψ can be decomposed:

$$\psi = \psi_L + \psi_R$$

where ψ_L and ψ_R are not solutions of the Dirac equation. The W vector boson of the weak interaction only couples to ψ_L.

As for the normalization of the orthogonal spinors (5.29) and (5.30), the most convenient choice is:

$$\begin{aligned} \bar{u}_s(p)u_{s'}(p) &= 2m\delta_{ss'} \\ \bar{v}_s(p)v_{s'}(p) &= -2m\delta_{ss'} \end{aligned}$$

where $s, s' = \pm$.

Using $\bar{\psi} = \psi^\dagger\gamma^0$, one can show that the following completeness relations (or polarization sum rules) hold:

$$\sum_{s=\pm} u_s(p)\bar{u}_s(p) = \not{p} + m \tag{5.41}$$

$$\sum_{s=\pm} v_s(p)\bar{v}_s(p) = \not{p} - m. \tag{5.42}$$

Comparing these polarization sums with, for instance, the Dirac equation for u, Eq. (5.27), one sees that $\not{p} + m$ projects on the subspace of particle solutions.

5.3 Field operator of the Dirac field

The spinors

$$\begin{aligned} &u_s(p)e^{-ip\cdot x}, \text{ eigenvalues } E_p = +\sqrt{|\vec{p}|^2+m^2}, \text{ and} \\ &v_s(-p)e^{ip\cdot x}, \text{ eigenvalues } E_p = -\sqrt{|\vec{p}|^2+m^2}, \end{aligned}$$

are eigenfunctions of the Dirac Hamiltonian and therefore solutions of the Dirac equation. From these solutions we can deduce the field operator of the Dirac field (which fulfills the Dirac equation) [1]:

$$\psi(x) = \int \frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2p^0}}\sum_{s=\pm}\left\{a_s(\vec{p})u_s(p)e^{-ip\cdot x} + b_s^\dagger(\vec{p})v_s(p)e^{ip\cdot x}\right\} \tag{5.43}$$

$$\bar{\psi}(x) = \int \frac{d^3p}{(2\pi)^3}\frac{1}{\sqrt{2p^0}}\sum_{s=\pm}\left\{a_s^\dagger(\vec{p})\bar{u}_s(p)e^{ip\cdot x} + b_s(\vec{p})\bar{v}_s(p)e^{-ip\cdot x}\right\} \tag{5.44}$$

where

$a_s^\dagger(\vec{p})$: creation operator of particle with momentum $\vec{p}$
$b_s^\dagger(\vec{p})$: creation operator of antiparticle with momentum $\vec{p}$
$a_s(\vec{p})$: annihilation operator of particle with momentum $\vec{p}$
$b_s(\vec{p})$: annihilation operator of antiparticle with momentum $\vec{p}$.

In advanced quantum mechanics we have seen that field operators create or annihilate position eigenstates. The field operator in Eq. (5.44) does the same thing while furthermore consistently combining the equivalent possibilities for particle creation and antiparticle annihilation: $a_s^\dagger(\vec{p})$ creates individual particle momentum eigenstates from which a weighted superposition is formed, the integral over $b_s(\vec{p})$ on the other hand, annihilates a weighted superposition of antiparticles. Since the creation of a particle at position x is equivalent to the annihilation of its antiparticle at position x, both terms have to appear in the field operator $\bar{\psi}(x)$. Because we have to consider particles and antiparticles, here the energy spectrum is more complicated than in the pure particle case. The creation terms come with a positive-sign plane wave factor $e^{ip\cdot x}$ while the annihilation terms contribute $e^{-ip\cdot x}$. The equivalence of particle creation and antiparticle annihilation is to be understood in the sense that they lead to the same change in a given field configuration.

The Dirac field is a spin-1/2 field. Therefore, the Pauli exclusion principle must hold, imposing anti-commutation relations on the field operators:

$$\begin{aligned} &\{\psi(\vec{x},t),\psi(\vec{x}',t)\} = \{\bar{\psi}(\vec{x},t),\bar{\psi}(\vec{x}',t)\} = 0 \\ &\{\psi(\vec{x},t),\bar{\psi}(\vec{x}',t)\} = \gamma^0\delta^3(\vec{x}-\vec{x}'). \end{aligned}$$

[1] The normalization is chosen to avoid an explicit factor $2p^0$ in the anticommutators of the fields and of the creation and annihilation operators.

Because of Eq. (5.43) and (5.44), this implies for the creation and annihilation operators

$$\begin{aligned}
\{a_r^\dagger(\vec{p}), a_s^\dagger(\vec{p}\,')\} &= \{a_r(\vec{p}), a_s(\vec{p}\,')\} = 0\\
\{b_r^\dagger(\vec{p}), b_s^\dagger(\vec{p}\,')\} &= \{b_r(\vec{p}), b_s(\vec{p}\,')\} = 0\\
\{a_r(\vec{p}), a_s^\dagger(\vec{p}\,')\} &= \delta_{rs}(2\pi)^3\delta^3(\vec{p}-\vec{p}\,')\\
\{b_r(\vec{p}), b_s^\dagger(\vec{p}\,')\} &= \delta_{rs}(2\pi)^3\delta^3(\vec{p}-\vec{p}\,').
\end{aligned}$$

As an example for the relation of field operator and ladder operator anti-commutation relations, we calculate $\{\psi(\vec{x},t), \bar{\psi}(\vec{x}\,',t)\}$, assuming anti-commutation relations for the creation and annihilation operators:

$$\begin{aligned}
&\{\psi(\vec{x},t), \bar{\psi}(\vec{y},t)\}\\
&= \int \frac{d^3p d^3\vec{q}}{(2\pi)^6}\frac{1}{\sqrt{2p^0 2q^0}}\sum_{r,s}\Big[e^{ip\cdot x}e^{-iqy}v_r(p)\bar{v}_s(q)\{b_r^\dagger(\vec{p}), b_s(\vec{q})\}\\
&\quad + e^{-ip\cdot x}e^{iqy}u_r(p)\bar{u}_s(q)\{a_r(\vec{p}), a_s^\dagger(\vec{q})\}\Big]\\
&= \int \frac{d^3p}{(2\pi)^3}\frac{1}{2p^0}\Big[e^{-i\vec{p}(\vec{x}-\vec{y})}\sum_s v_s(p)\bar{v}_s(p) + e^{i\vec{p}(\vec{x}-\vec{y})}\sum_s u_s(p)\bar{u}_s(p)\Big]
\end{aligned}$$

which, using the completeness relations, Eq. (5.41) and (5.42),

$$\begin{aligned}
&= \int \frac{d^3p}{(2\pi)^3}\frac{1}{2p^0}\Big[e^{-i\vec{p}(\vec{x}-\vec{y})}(\underbrace{p^0\gamma^0}_{\text{even}} - \underbrace{\vec{p}\cdot\vec{\gamma}}_{\text{odd}} - m) + e^{i\vec{p}(\vec{x}-\vec{y})}(\underbrace{p^0\gamma^0}_{\text{even}} - \underbrace{\vec{p}\cdot\vec{\gamma}}_{\text{odd}} + m)\Big]\\
&= \gamma^0\int\frac{d^3p}{(2\pi^3)}e^{i\vec{p}(\vec{x}-\vec{y})} = \gamma^0\delta^3(\vec{x}-\vec{y}).
\end{aligned}$$

However, in the laboratory one prepares in general (to a first approximation) momentum eigenstates, rather than position eigenstates. Therefore, we give the expression[2] for the momentum operator:

$$P^\mu = \int\frac{d^3k}{(2\pi)^3}k^\mu\sum_s\Big(a_s^\dagger(\vec{k})a_s(\vec{k}) + b_s^\dagger(\vec{k})b_s(\vec{k})\Big)$$

which is just the momentum weighted with the number operator $N = a^\dagger a + b^\dagger b$. Using the anti-commutation relations for the ladder operators, one can show that the momentum operator fulfills the following useful commutation relations:

$$\begin{aligned}
[P^\mu, a_s^\dagger(\vec{p})] &= p^\mu a_s^\dagger(\vec{p})\\
[P^\mu, b_s^\dagger(\vec{p})] &= p^\mu b_s^\dagger(\vec{p})\\
[P^\mu, a_s(\vec{p})] &= -p^\mu a_s(\vec{p})\\
[P^\mu, b_s(\vec{p})] &= -p^\mu b_s(\vec{p}).
\end{aligned}$$

[2]This expression is obtained from Noether's theorem using the technique of normal ordering. These topics are discussed in text books on quantum field theory, e. g. by Peskin/Schroeder [14].

Vacuum state The vacuum state is denoted by $|0\rangle$ and has the property [3],

$$P^\mu |0\rangle = 0, \tag{5.45}$$

i.e. the vacuum has no momentum.

Using the commutation relations stated above and the property (5.45), we conclude that,

$$P^\mu a_s^\dagger(\vec{p})\, |0\rangle = p^\mu a_s^\dagger(\vec{p})\, |0\rangle\,, \tag{5.46}$$

in other words, the state $a_s^\dagger(\vec{p})\,|0\rangle$ is an eigenstate of P^μ with momentum p^μ.

With this fact in mind, we define the following states,

$$\begin{aligned} \left|e^-(p,s)\right\rangle &= \sqrt{2E_{\vec{p}}}\, a_s^\dagger(\vec{p})\,|0\rangle\,, &(5.47)\\ \left|e^+(p,s)\right\rangle &= \sqrt{2E_{\vec{p}}}\, b_s^\dagger(\vec{p})\,|0\rangle\,, &(5.48)\end{aligned}$$

of a particle respectively antiparticle with momentum eigenstate p and spin s.

The factor $\sqrt{2E_{\vec{p}}}$ is there in order to ensure a Lorentz invariant normalization,

$$\begin{aligned} \langle e^-(q,r)|e^-(p,s)\rangle &= 2\sqrt{E_{\vec{q}}E_{\vec{p}}}\,\langle 0|\, a_r(\vec{q})a_s^\dagger(\vec{p})\,|0\rangle \\ &= 2\sqrt{E_{\vec{q}}E_{\vec{p}}}\,\langle 0|\left\{a_r(\vec{q}), a_s^\dagger(\vec{p})\right\} - a_s^\dagger(\vec{p})\,\underbrace{a_r(\vec{q})\,|0\rangle}_{=0} \\ &= \delta_{rs} 2E_{\vec{p}}(2\pi)^3\delta^{(3)}(\vec{q}-\vec{p}). \end{aligned}$$

The definition of states (5.47) and (5.48) corresponds to a continuum normalization in infinite volume. From the above equation, it can be seen that the dimensionality of the one-particle norm $\langle e^-(p,s)|e^-(p,s)\rangle$ is,

$$\frac{(energy)}{(momentum)^3} = (energy)\cdot(volume),$$

meaning that we have a constant particle density of $2E$ particles per unit volume. To obtain single particle states in a given volume V, one must therefore multiply $|e^-(p,s)\rangle$ with a normalization factor $1/\sqrt{2EV}$:

$$\begin{aligned} \left|e^-(p,s)\right\rangle_{\text{single-particle}} &= \frac{1}{\sqrt{2EV}}\left|e^-(p,s)\right\rangle &(5.49)\\ \left|e^+(p,s)\right\rangle_{\text{single-particle}} &= \frac{1}{\sqrt{2EV}}\left|e^+(p,s)\right\rangle &(5.50)\end{aligned}$$

[3] After applying the nontrivial concept of normal ordering, here only motivated by the number interpretation in the operator P^μ.

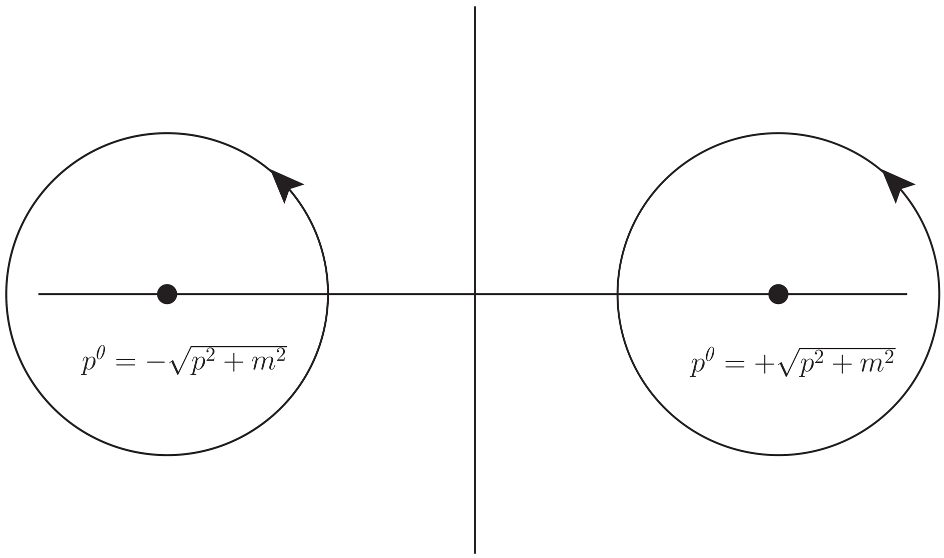


Figure 5.2: *Integration paths for Dirac propagator.*

5.4 Dirac propagator

In order to solve general Dirac equations, we want to apply a formalism similar to the one used in classical electrodynamics, namely Green's functions.

We introduce the scalar propagator,

$$\begin{aligned}\Delta^{\pm}(x) &= \pm\frac{1}{i}\int\frac{d^3p}{(2\pi)^3 2p^0}\mathrm{e}^{\mp ip\cdot x}\\ &= \pm\frac{1}{i}\int\frac{d^4p}{(2\pi)^3}\delta(p^2-m^2)\mathrm{e}^{\mp ip\cdot x},\end{aligned} \tag{5.51}$$

which satisfies the Klein-Gordon equation,

$$(\Box + m^2)\Delta^{\pm}(x) = 0.$$

Representation as a contour integral

$$\Delta^{\pm}(x) = -\int\limits_{C^{\pm}}\frac{d^4p}{(2\pi)^4}\frac{\mathrm{e}^{-ip\cdot x}}{p^2-m^2}, \tag{5.52}$$

where the paths $C^{\pm}$ are depicted in Fig. 5.2.

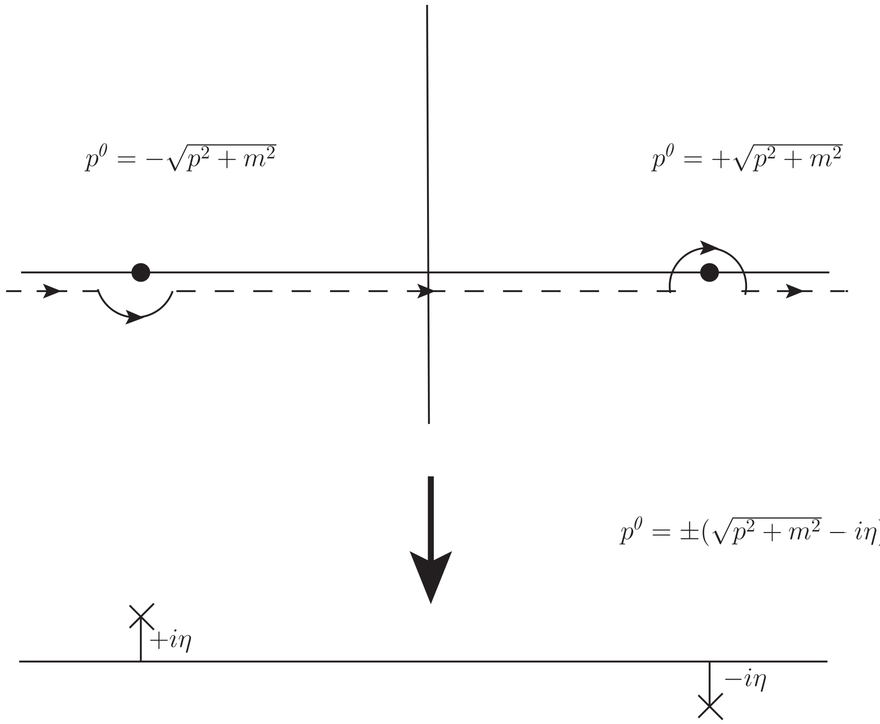


Figure 5.3: *Deformed integration paths and $+i\varepsilon$ convention.*

5.4.1 Feynman propagator

To get a "true" Green's function for the operator $\Box + m^2$, we need to introduce a discontinuity, and define the Feynman propagator

$$\Delta_F(x) = \theta(t)\Delta^+(x) - \theta(-t)\Delta^-(x), \tag{5.53}$$

where we deform the paths of Fig. 5.3 according to the sign of $t = x^0$ to get convergent integrals over the real line (details can be found in a complex analysis book, see e.g. Freitag & Busam [15]) :

- $x^0 > 0, \operatorname{Im} p^0 < 0 \Rightarrow \mathrm{e}^{-ip^0x^0} \overset{R\to\infty}{\longrightarrow} 0 : C^+$,
- $x^0 < 0, \operatorname{Im} p^0 > 0 \Rightarrow \mathrm{e}^{-ip^0x^0} \overset{R\to\infty}{\longrightarrow} 0 : C^-$.

$+i\varepsilon$ convention Instead of deforming the integration path, one can also shift the two poles and integrate over the whole real p^0-axis, without having to worry about the poles,

$$p^0 = \pm\sqrt{\vec{p}^2 + m^2} \longrightarrow \pm(\sqrt{\vec{p}^2 + m^2} - i\eta),$$

yielding

$$\Delta_F(x) = \lim_{\varepsilon\to 0+} \int \frac{d^4p}{(2\pi)^4} \frac{\mathrm{e}^{-ip\cdot x}}{p^2 - m^2 + i\varepsilon}, \tag{5.54}$$

the Green's function of the Klein-Gordon equation,

$$(\Box + m^2)\Delta_F(x) = \int \frac{d^4p}{(2\pi)^4} \mathrm{e}^{-ip\cdot x} \frac{-p^2 + m^2}{p^2 - m^2} = -\delta^{(4)}(x). \tag{5.55}$$

Propagator A propagator is the transition amplitude of a particle between creation at x^μ and annihilation at x'^μ (or vice-versa). It is a fundamental tool of quantum field theory.

After getting the Feynman propagator for the Klein-Gordon field (spin 0), we want to focus on the propagator for fermions (spin 1/2).

We compute the anticommutation relations for the field in this case getting,

$$\begin{aligned}\{\psi(x), \bar{\psi}(x')\} =& \int \frac{d^3p d^3p'}{(2\pi)^6 \sqrt{2p^0}\sqrt{2p'^0}} \sum_{r,s} \Big[\mathrm{e}^{i(p\cdot x - p'\cdot x')} v_r(p)\bar{v}_s(p')\{b_r^\dagger(p), b_s(p')\} \\ & \mathrm{e}^{-i(p\cdot x - p'\cdot x')} u_r(p)\bar{u}_s(p')\{a_r(p), a_s^\dagger(p')\}\Big] \\ =& \int \frac{d^3p}{(2\pi)^3 2p^0} \left[\mathrm{e}^{ip\cdot(x-x')}(\not{p} - m) + \mathrm{e}^{-ip\cdot(x-x')}(\not{p} + m)\right] \\ =& (i\not{\partial} + m) \int \frac{d^3p}{(2\pi)^3 2p^0} \left(\mathrm{e}^{-ip\cdot(x-x')} - \mathrm{e}^{ip\cdot(x-x')}\right),\end{aligned} \tag{5.56}$$

where we made use of the completeness relations (5.41) and (5.42) in going from the first to the second line.

We now define the **Feynman fermion propagator**,

$$iS(x - x') \equiv (i\not{\partial} + m)(\Delta^+(x - x') + \Delta^-(x - x')). \tag{5.57}$$

Splitting ψ and $\bar{\psi}$ in their creation $\psi^-, \bar{\psi}^-$ and annihilation $\psi^+, \bar{\psi}^+$ parts (looking only at the operators $a_s^\dagger, b_s^\dagger$ and a_s, b_s respectively), we get the comutation relations,

$$\{\psi^+(x), \bar{\psi}^-(x')\} = (i\not{\partial} + m)\Delta^+(x - x') = iS^+(x - x'), \tag{5.58}$$

$$\{\psi^-(x), \bar{\psi}^+(x')\} = (i\not{\partial} + m)\Delta^-(x - x') = iS^-(x - x'). \tag{5.59}$$

$S^\pm(x - x')$ can as well be represented as contour integrals,

$$S^\pm(x) = \int_{C^\pm} \frac{d^4p}{(2\pi)^4} \mathrm{e}^{-ip\cdot x} \frac{\not{p} + m}{p^2 - m^2} = \int_{C^\pm} \frac{d^4p}{(2\pi)^4} \mathrm{e}^{-ip\cdot x} \frac{1}{\not{p} - m}, \tag{5.60}$$

which is well defined because $(\not{p}+m)(\not{p}-m) = (p^2 - m^2)\mathbb{1}$.

We take a look at the time ordered product of fermion operators,

$$T(\psi(x)\bar{\psi}(x')) = \begin{cases} \psi(x)\bar{\psi}(x'), & t > t' \\ -\bar{\psi}(x')\psi(x), & t' > t \end{cases}$$
$$= \theta(t-t')\psi(x)\bar{\psi}(x') - \theta(t'-t)\bar{\psi}(x')\psi(x).$$

The Feynman fermion propagator is then the vacuum expectation value of this time ordered product,

$$iS_F(x-x') = \langle 0|\, T(\psi(x)\bar{\psi}(x'))\, |0\rangle\,. \tag{5.61}$$

Remembering the destroying effect of annihilation operators on the vacuum, we can skip some trivial steps of the calculation. We look separately at both time ordering cases, getting,

$$\langle 0|\,\psi(x)\bar{\psi}(x')\,|0\rangle = \langle 0|\,\psi^+(x)\bar{\psi}^{\,-}(x')\,|0\rangle = \langle 0|\,\{\psi^+(x),\bar{\psi}^{\,-}(x')\}\,|0\rangle = iS^+(x-x'),$$
$$\langle 0|\,\bar{\psi}(x')\psi(x)\,|0\rangle = \langle 0|\,\bar{\psi}^{\,+}(x')\psi^-(x)\,|0\rangle = \langle 0|\,\{\bar{\psi}^{\,+}(x'),\psi^-(x)\}\,|0\rangle = iS^-(x-x'),$$

yielding,

$$S_F(x) = \theta(t)S^+(x) - \theta(-t)S^-(x) = (i\not{\partial}+m)\Delta_F(x), \tag{5.62}$$

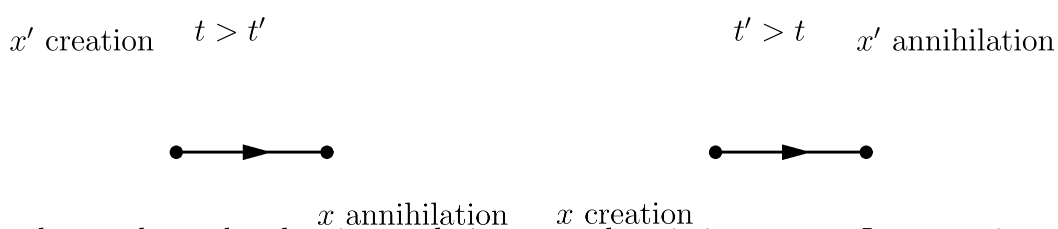
or, as a contour integral,

$$S_F(x) = \int\limits_{C_F} \frac{d^4p}{(2\pi)^4}\mathrm{e}^{-ip\cdot x}\frac{1}{\not{p}-m} = \lim_{\varepsilon\to 0+}\int \frac{d^4p}{(2\pi)^4}\mathrm{e}^{-ip\cdot x}\frac{\not{p}+m}{p^2-m^2+i\varepsilon}. \tag{5.63}$$

We then see that the fermion propagator is nothing else than the Green's function of the Dirac equation,

$$(i\not{\partial}-m)S_F(x) = \int \frac{d^4p}{(2\pi)^4}\mathrm{e}^{-ip\cdot x}\frac{(\not{p}-m)(\not{p}+m)}{p^2-m^2} = \delta^{(4)}(x)\mathbb{1}. \tag{5.64}$$

The interpretation of S_F is then similar to the one of the Green's function in classical electrodynamics:



We can ask ourselves why the time ordering procedure is important. In scattering processes both orderings are not distinguishable (see Fig. 5.4) in experiments, so that we can understand as a sum over both time ordering possibilities.

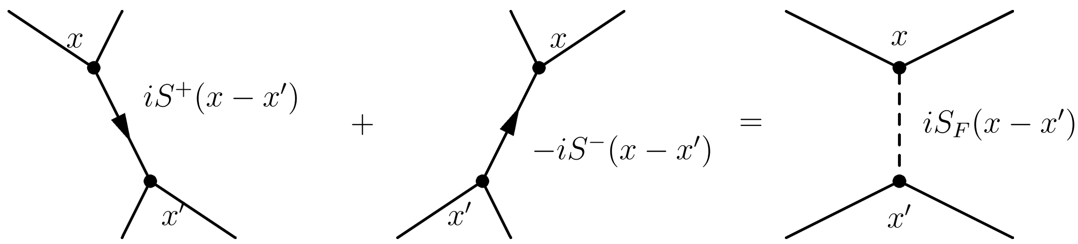


Figure 5.4: Sum of both time orderings

5.5 Photon field operator

After being able to describe free scalar fields (Klein-Gordon, spin 0) and free fermion fields (Dirac, spin 1/2), we go on to vector fields (spin 1) like the one describing the photon. The photon field will be shown to have a fundamental importance in QED since it is the interaction field between fermions.

To start, we recall the photon field operator of advanced quantum mechanics, which reads in Coulomb gauge,

$$\vec{A}(x) = \sum_{\alpha=1,2} \int \frac{d^3k}{(2\pi)^3} \frac{1}{\sqrt{2\omega_k}} \Big(a_\alpha(\vec{k})\, \vec{\varepsilon}_\alpha(\vec{k}) \mathrm{e}^{-ik\cdot x} + a_\alpha^\dagger(\vec{k})\, \vec{\varepsilon}_\alpha^{\,*}(\vec{k}) \mathrm{e}^{ik\cdot x} \Big). \tag{5.65}$$

In Eq. (5.65), $a_\alpha^\dagger(\vec{k})$ creates a photon of momentum $\vec{k}$ and polarization α, and $a_\alpha(\vec{k})$ destroys the same.

Since we are dealing with a bosonic field, we impose the commutation relations,

$$[a_\alpha(\vec{k}), a_\beta^\dagger(\vec{k}')] = -g_{\alpha\beta}(2\pi)^3 \delta^{(3)}(\vec{k} - \vec{k}'), \tag{5.66}$$

$$[a_\alpha(\vec{k}), a_\beta(\vec{k}')] = [a_\alpha^\dagger(\vec{k}), a_\beta^\dagger(\vec{k}')] = 0. \tag{5.67}$$

Supposing that the photon propagates in the z-direction ($k^\mu = (k, 0, 0, k)^\top$), we have the following possibilites for the polarization vectors :

- linear : $\varepsilon_1^\mu = (0, 1, 0, 0)^\top, \varepsilon_2^\mu = (0, 0, 1, 0)^\top$,
- circular : $\varepsilon_+^\mu = \frac{1}{\sqrt{2}}(\varepsilon_1^\mu + i\varepsilon_2^\mu) = \frac{1}{\sqrt{2}}(0, 1, i, 0)^\top, \varepsilon_-^\mu = \frac{1}{\sqrt{2}}(\varepsilon_1^\mu - i\varepsilon_2^\mu) = \frac{1}{\sqrt{2}}(0, 1, -i, 0)^\top$.

These vector sets satisfy the completness relation,

$$\Pi^{\mu\nu} = \sum_{\substack{\lambda = \pm \\ (\text{or } \lambda = 1,2)}} \varepsilon_\lambda^{*\mu} \varepsilon_\lambda^\nu = \begin{pmatrix} 0 & & & \\ & 1 & & \\ & & 1 & \\ & & & 0 \end{pmatrix}. \tag{5.68}$$

By applying a well chosen boost to $\Pi^{\mu\nu}$ we can easily check that it is in general not Lorentz invariant. We have to choose a specific gauge depending on the reference frame, parametrized by a real number n.

To do so we define a auxiliary vector $n^\mu = n(1,0,0,-1)^\top$ satisfying $n_\sigma k^\sigma = 2kn$ and get the "axial gauge",

$$\Pi^{\mu\nu} = -g^{\mu\nu} + \frac{n^\mu k^\nu + k^\mu n^\nu}{n_\sigma k^\sigma}. \tag{5.69}$$

For $n = 1$, we recover the Coulomb gauge,

$$\Pi^{\mu\nu} = \begin{pmatrix} -1 & & & \\ & 1 & & \\ & & 1 & \\ & & & 1 \end{pmatrix} + \begin{pmatrix} 1 & & & \\ & 0 & & \\ & & 0 & \\ & & & -1 \end{pmatrix} = \begin{pmatrix} 0 & & & \\ & 1 & & \\ & & 1 & \\ & & & 0 \end{pmatrix}.$$

In physical processes, the photon field couples to an external current,

$$j^\mu(x) = j^\mu(k)\mathrm{e}^{ik\cdot x},$$

and we have the current conservation,

$$\partial_\mu j^\mu = 0,$$

which yields in Fourier space,

$$k_\mu j^\mu = 0,$$

and thus,

$$j_\mu \Pi^{\mu\nu} = j_\nu \Pi^{\mu\nu} = 0,$$

i.e. the $n^\mu k^\nu + k^\mu n^\nu$ term vanishes when contracted with external currents, such that we are left with an effective polarization sum,

$$p^{\mu\nu}_{\mathit{eff}} = -g^{\mu\nu}. \tag{5.70}$$

We now look at the time ordered product of photon field operators,

$$T(A_\mu(x)A_\nu(x')) = \begin{cases} A_\mu(x)A_\nu(x'), & t > t' \\ A_\nu(x')A_\mu(x), & t' > t \end{cases}. \tag{5.71}$$

Repeating the same steps as in the fermion case, we get the **photon propagator**,

$$\begin{aligned} iD_{F,\mu\nu}(x-x') &= \langle 0|\, T(A_\mu(x)A_\nu(x'))\, |0\rangle && (5.72)\\ &= -ig_{\mu\nu}\Delta_F(x-x') && (5.73)\\ &= -ig_{\mu\nu} \lim_{\varepsilon\to 0+} \int \frac{d^4k}{(2\pi)^4} \frac{\mathrm{e}^{-ik\cdot x}}{k^2 + i\varepsilon}. && (5.74) \end{aligned}$$

Finally, we see that the photon propagator is the Green's function of the wave equation,

$$\Box D_{F,\mu\nu}(x) = g_{\mu\nu}\delta^{(4)}(x). \tag{5.75}$$

5.6 Interaction representation

In the previous sections, we have gained an understanding of the free fields occurring in QED. The next step is to introduce a way to handle interactions between those fields.

Idea: decompose the Hamiltonian in the Schrödinger representation,

$$H_S = H_{0,S} + H'_S,$$

and define states and operators in the free Heisenberg representation,

$$\begin{aligned} \psi_I &= \mathrm{e}^{iH_{0,S}t}\psi_S \\ O_I &= \mathrm{e}^{iH_{0,S}t}O_S\mathrm{e}^{-iH_{0,S}t}, \end{aligned}$$

and you get the **interaction representation** (also called **Dirac representation**).

We have, in particular,

$$H_{0,I} = H_{0,S} = H_0, \tag{5.76}$$

and the time evolution of ψ_I respectively O_I becomes,

$$i\partial_t\psi_I = H'_I\psi_I, \tag{5.77}$$
$$i\partial_t O_I = -H_0 O_I + O_I H_0 = [O_I, H_0], \tag{5.78}$$

i.e. ψ_I is influenced only by the “true” interaction part; the “trivial” time evolution (free part) has been absorbed in the operators O_I.

Comparison The Schrödinger, Heisenberg, and interaction representations differ in the way they describe time evolution:

- Schrödinger representation: states contain time evolution, operators are time independent;
- Heisenberg representation: states are time independent, operators contain time evolution;
- Interaction representation: time dependence of states only due to interactions, free (also called “trivial”) time evolution for operators.

This comparison shows that the interaction representation is a mixture of both other representations.

5.6.1 Time evolution operator

In preparation for time-dependent perturbation theory, we consider the time evolution operator $U(t, t_0)$ in the interaction representation:

$$\psi_I(t) = U(t, t_0)\psi_I(t_0). \tag{5.79}$$

The time evolution operator in Eq. (5.79) can be written in terms of the free and interaction Hamiltonians, Eq. (5.76), in the Schrödinger representation by using the time evolution properties:

$$\psi_I(t) = e^{iH_0t}\psi_S(t) = e^{iH_0t}e^{-iH_S(t-t_0)}\psi_S(t_0) = e^{iH_0t}e^{-iH_S(t-t_0)}e^{-iH_0t_0}\psi_I(t_0).$$

Comparing this result with Eq. (5.79) yields

$$U(t, t_0) = e^{iH_0t}e^{-iH_S(t-t_0)}e^{-iH_0t_0}. \tag{5.80}$$

An interaction picture operator is related by

$$O_H(t) = U^\dagger(t, t_0)O_IU(t, t_0)$$

to its Heisenberg picture equivalent.

Because of Eq. (5.80) the time evolution operator has the following properties:

- $U(t_0, t_0) = \mathbb{1}$,
- $U(t_2, t_1)U(t_1, t_0) = U(t_2, t_0)$,
- $U^{-1}(t_0, t_1) = U(t_1, t_0)$, and
- $U^\dagger(t_1, t_0) = U^{-1}(t_1, t_0) = U(t_0, t_1)$.

5.6.2 Time ordering

To find the time evolution operator, the time evolution (Schrödinger) equation

$$i\frac{\partial}{\partial t}U(t, t_0) = H_I'U(t, t_0) \tag{5.81}$$

has to be solved. This is equivalent to the integral equation

$$U(t, t_0) = \mathbb{1} + (-i)\int_{t_0}^{t} dt_1 H_I'(t_1)U(t_1, t_0)$$

which can be iterated to give the Neumann series

$$U(t,t_0) = \mathbb{1} + (-i)\int_{t_0}^{t} dt_1 H_I'(t_1) \tag{5.82}$$

$$+(-i)^2 \int_{t_0}^{t} dt_1 \int_{t_0}^{t_1} dt_2 H_I'(t_1) H_I'(t_2) \tag{5.83}$$

$$+\dots \tag{5.84}$$

$$+(-i)^n \int_{t_0}^{t} dt_1 \int_{t_0}^{t_1} dt_2 \dots \int_{t_0}^{t_{n-1}} dt_n H_I'(t_1)\dots H_I'(t_n). \tag{5.85}$$

This is not yet satisfactory since the boundary of every integral but the first depends on the foregoing integration. To solve this problem, one uses time ordering. Let us first consider the following identities:

$$\begin{aligned}\int_{t_0}^{t} dt_1 \int_{t_0}^{t_1} dt_2 H_I'(t_1) H_I'(t_2) &= \int_{t_0}^{t} dt_2 \int_{t_2}^{t} dt_1 H_I'(t_1) H_I'(t_2) \\ &= \int_{t_0}^{t} dt_1 \int_{t_1}^{t} dt_2 H_I'(t_2) H_I'(t_1)\end{aligned}$$

where in the first line the integration domains are identical (see Fig. 5.5) and in going to the second line the variable labels are exchanged. We can combine these terms in a more compact expression:

$$\begin{aligned}&2\int_{t_0}^{t} dt_1 \int_{t_0}^{t_1} dt_2 H_I'(t_1) H_I'(t_2) \\ &= \int_{t_0}^{t} dt_2 \int_{t_2}^{t} dt_1 H_I'(t_1) H_I'(t_2) + \int_{t_0}^{t} dt_1 \int_{t_1}^{t} dt_2 H_I'(t_2) H_I'(t_1) \\ &= \int_{t_0}^{t} dt_1 \int_{t_0}^{t} dt_2 \Big(H_I'(t_1) H_I'(t_2)\theta(t_1 - t_2) + H_I'(t_2) H_I'(t_1)\theta(t_2 - t_1)\Big) \\ &= \int_{t_0}^{t} dt_1 \int_{t_0}^{t} dt_2 T\Big(H_I'(t_1) H_I'(t_2)\Big).\end{aligned}$$

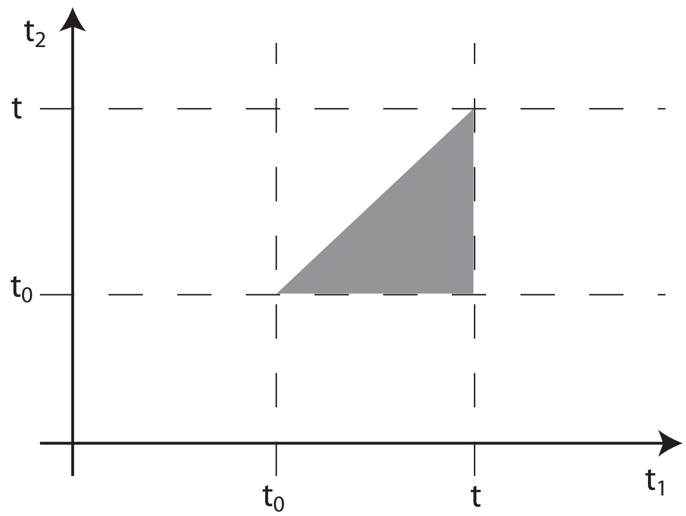


Figure 5.5: *Identical integration domains.*

All terms of the Neumann series can be rewritten in this way. For the n-th term in Eq. (5.85) we have

$$\begin{aligned} n! \int_{t_0}^{t} dt_1 \ldots \int_{t_0}^{t_{n-1}} dt_n H_I'(t_1) \ldots H_I'(t_n) \\ = \int_{t_0}^{t} dt_1 \ldots \int_{t_0}^{t} dt_n T\Big(H_I'(t_1) \ldots H_I'(t_n) \Big). \end{aligned}$$

We therefore obtain the following perturbation series[4] for the time evolution operator:

$$U(t,t_0) = \sum_{n=0}^{\infty} \frac{1}{n!} (-i)^n \int_{t_0}^{t} dt_1 \ldots \int_{t_0}^{t} dt_n T\Big(H'(t_1) \ldots H'(t_n) \Big). \tag{5.86}$$

Defining the time ordered exponential, Eq. (5.86) can be written as

$$\boxed{U(t,t_0) = T \exp\left(-i \int_{t_0}^{t} dt' H'(t') \right)}. \tag{5.87}$$

[4]We are working in the interaction picture and drop the index I for simplicity.

We check that this result indeed solves the time evolution equation (5.81):

$$
\begin{aligned}
i\frac{\partial}{\partial t}U(t,t_0) &= i\sum_{n=1}^{\infty}\frac{1}{n!}(-i)^n n\int_{t_0}^{t} dt_1 \ldots \int_{t_0}^{t} dt_{n-1} \\
&\qquad\qquad T\Big(H'(t_0)\ldots H'(t_{n-1})H'(t)\Big) \\
&= H'(t)\sum_{n=1}^{\infty}\frac{1}{(n-1)!}(-i)^{n-1}\int_{t_0}^{t} dt_1 \ldots \int_{t_0}^{t} dt_{n-1} \\
&\qquad\qquad T\Big(H'(t_0)\ldots H'(t_{n-1})\Big) \\
&= H'(t)U(t,t_0).
\end{aligned}
$$

5.7 Scattering matrix

Our overall aim is to develop a formalism to compute scattering matrix elements which describe the transition from initial states defined at $t \to -\infty$ to final states observed at $t \to +\infty$. To this end, we split up the Hamiltonian into a solvable free part which determines the operators' time evolution and an interaction part responsible for the time evolution of the states. Now we investigate how the time ordered exponential that is the time evolution operator, see Eq. (5.87), relates to the $\mathcal{S}$-matrix.

The scattering matrix element $\langle f|\,\mathcal{S}\,|i\rangle$ is the transition amplitude for $|i\rangle \to |f\rangle$ caused by interactions. The state of the system is described by the time dependent state vector $|\psi(t)\rangle$. The above statement about asymptotically large times can now be recast in a more explicit form: The initial state is given by

$$\lim_{t\to-\infty} |\psi(t)\rangle = |\phi_i\rangle$$

where $|\phi_i\rangle$ is an eigenstate of the free Hamilton operator and $t \to -\infty$ is justified since the interaction timescale is about 10^{-15} s. The scattering matrix element $\mathcal{S}_{fi}$ is given by the projection of the state vector $|\psi(t)\rangle$ onto a final state $|\phi_f\rangle$:

$$\mathcal{S}_{fi} = \lim_{t\to+\infty}\langle\phi_f|\psi(t)\rangle = \langle\phi_f|\,\mathcal{S}\,|\phi_i\rangle\,.$$

Using the time evolution operator (and its action on a state, see Eq. (5.79)), this can be expressed as

$$\mathcal{S}_{fi} = \lim_{t_2\to+\infty}\lim_{t_1\to-\infty}\langle\phi_f|\,U(t_2,t_1)\,|\phi_i\rangle\,.$$

We can therefore conclude that

$$\boxed{\mathcal{S} = U(+\infty,-\infty)} = \sum_{n=0}^{\infty}\frac{1}{n!}(-i)^n\int_{-\infty}^{\infty} dt_1 \ldots \int_{-\infty}^{\infty} dt_n T\Big(H'(t_1)\ldots H'(t_n)\Big). \tag{5.88}$$

As an instructive example, we consider $2 \to 2$ scattering:

$$k_1 + k_2 \to k_3 + k_4.$$

The scattering matrix element is given by

$$\mathcal{S}_{fi} = \langle f|\, \mathcal{S}\, |i\rangle = \underbrace{\langle 0|\, a(k_4)a(k_3)}_{\langle \phi_f|}\, |\mathcal{S}|\, \underbrace{a^\dagger(k_1)a^\dagger(k_2)\, |0\rangle}_{|\phi_i\rangle}.$$

The $\mathcal{S}$-operator itself consists of further creation and annihilation operators belonging to further quantum fields. By evaluation of the creators and annihilators in $\mathcal{S}$ (using commutation or anticommutation relations), it follows that there is only one single non-vanishing contribution to $\mathcal{S}_{fi}$ being of the ("normally ordered") form

$$f(k_1, k_2, k_3, k_4)a^\dagger(k_3)a^\dagger(k_4)a(k_2)a(k_1).$$

Note that in the above expression, the annihilation operators stand on the right hand side, while the creation operators are on the left. Such expressions are said to be in normal order and are denoted by colons, $: ABC :$. Since the aim is to find the non-vanishing contributions, a way has to be found how time ordered products can be related to products in normal order. For instance, consider the time ordered product of two Boson field operators (where $A^+,\ B^+$ are annihilators and $A^-,\ B^-$ creators)[5]

$$\begin{aligned} T\Big(A(x_1)B(x_2)\Big)\Big|_{t_1>t_2} &= A(x_1)B(x_2) \\ &= A^+(x_1)B^+(x_2) + A^-(x_1)B^+(x_2) \\ &+ \underbrace{A^+(x_1)B^-(x_2)}_{\text{not in normal order}} + A^-(x_1)B^-(x_2). \end{aligned}$$

[5]The $\pm$ sign is motivated by the decomposition of field operators in positive and negative frequency parts:

$$\phi(x) = \phi^+(x) + \phi^-(x).$$

Consider for example the Klein-Gordon field where

$$\phi(x) = \int \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{2p^0}} \Big(a(p)e^{+i\vec{p}\cdot\vec{x}} + a^\dagger(p)e^{-i\vec{p}\cdot\vec{x}}\Big)$$

and therefore

$$\begin{aligned} \phi^+(x) &= \int \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{2p^0}} a(p)e^{+i\vec{p}\cdot\vec{x}} \\ \phi^-(x) &= \int \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{2p^0}} a^\dagger(p)e^{-i\vec{p}\cdot\vec{x}}. \end{aligned}$$

One can observe that only one of the above terms is not in normal order while the other three would vanish upon evaluation in $\langle 0|\cdot|0\rangle$. Using

$$\underbrace{A^+(x_1)B^-(x_2)}_{\text{not in normal order}} = \underbrace{B^-(x_2)A^+(x_1)}_{\text{in normal order}} + \underbrace{[A^+(x_1), B^-(x_2)]}_{\text{c-number}},$$

we rewrite

$$\begin{aligned}
[A^+(x_1), B^-(x_2)] &= \langle 0|\,[A^+(x_1), B^-(x_2)]\,|0\rangle \\
&= \langle 0|\,A^+(x_1)B^-(x_2)\,|0\rangle \\
&= \langle 0|\,T(A(x_1)B(x_2))\,|0\rangle \,.
\end{aligned}$$

Since the same holds for $t_1 < t_2$, we draw the conclusion

$$T\Big(A(x_1)B(x_2)\Big) = \,:A(x_1)B(x_2):\, + \langle 0|\,T(A(x_1)B(x_2))\,|0\rangle \,.$$

An analogous calculation for fermion operators yields the same result.

The next step towards Feynman diagrams is to formalize this connection between time and normal ordered products. We first define the following shorthand

$$\underbracket{\phi_A(x_1)\phi_B}(x_2) = \langle 0|\,T(\phi_A(x_1)\phi_B(x_2))\,|0\rangle$$

which is called contraction of operators. This allows to state the following in compact notation.

Wick's theorem: The time ordered product of a set of operators can be decomposed into the sum of all corresponding contracted products in normal order. All combinatorially allowed contributions appear:

$$\begin{aligned}
&T(ABC\dots XYZ) = \,:ABC\dots XYZ: \\
&+ :\underbracket{AB}\,C\dots XYZ: + \dots + :\underbracket{ABC\dots XYZ}: + \dots + :ABC\dots X\,\underbracket{YZ}: \\
&+ :\underbracket{AB}\,\underbracket{CD}\dots XYZ: + :\underbracket{A\underbracket{BC}D}\dots XYZ: + \dots \\
&+ :\text{threefold contractions}: + \dots \;.
\end{aligned}$$

5.8 Feynman rules of quantum electrodynamics

The Lagrangian density of QED is given by

$$\mathcal{L} = \mathcal{L}_0^{\text{Dirac}} + \mathcal{L}_0^{\text{photon}} + \mathcal{L}'$$

where the subscript 0 denotes the free Lagrangian densities and $'$ denotes the interaction part. In particular, we have

$$\mathcal{L}_0^{\text{Dirac}} = \bar{\psi}(i\partial\!\!\!/ - m)\psi \tag{5.89}$$

$$\mathcal{L}_0^{\text{photon}} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} \tag{5.90}$$

$$\mathcal{L}' = -e\bar{\psi}\gamma_\mu\psi A^\mu = -j_\mu A^\mu \tag{5.91}$$

where $F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu$. Note that from $\mathcal{L}_0^{\text{photon}}$ the free Maxwell's equations can be derived using the Euler-Lagrange equations. Using $\mathcal{L}_0^{\text{photon}} + \mathcal{L}'$ yields Maxwell's equations in the presence of sources and $\mathcal{L}_0^{\text{Dirac}} + \mathcal{L}'$ does the same for the Dirac equation. The interaction term $\mathcal{L}'$ describes current-field interactions and therefore couples the fermions described by the Dirac equation to photons described by Maxwell's equations.

Using Eq. (5.91), one finds the quantized interaction Hamiltonian density

$$\mathcal{H}' = -\mathcal{L}' = e\bar{\psi}\gamma_\mu\psi A^\mu.$$

Integrating the interaction Hamiltonian density over all space yields the interaction Hamiltonian,

$$H' = \int d^3\vec{x}\,\mathcal{H}',$$

and, in the integral representation of $\mathcal{S}$ given in Eq. (5.88), this leads to integrations over space-time:

$$\mathcal{S} = \sum_{n=0}^{\infty}\frac{1}{n!}(-ie)^n\int d^4x_1\ldots d^4x_n T\Big(\bar{\psi}(x_1)\gamma_{\mu_1}\psi(x_1)A^{\mu_1}\ldots\bar{\psi}(x_n)\gamma_{\mu_n}\psi(x_n)A^{\mu_n}\Big). \tag{5.92}$$

Since $e = \sqrt{4\pi\alpha}$ (see Eq. (1.9)), the coupling constant appears in the interaction term and n-th order terms are suppressed with e^n. This means that we found an expansion of $\mathcal{S}$ in the small parameter e which is the starting point for perturbation theory. The structure of the n-th term in the perturbation series in Eq. (5.92) is

$$\mathcal{S}^{(n)} = \frac{1}{n!}\int d^4x_1\ldots d^4x_n\mathcal{S}_n \tag{5.93}$$

where

$$\mathcal{S}_n = \sum_{\text{contractions}} K(x_1,\ldots,x_n)\ :\ldots\bar{\psi}(x_i)\ldots\psi(x_j)\ldots A(x_n):. \tag{5.94}$$

For a specific scattering process, the relevant matrix element is

$$\mathcal{S}_{fi} = \underbrace{\langle f|}_{\sim a}\,\mathcal{S}\,\underbrace{|i\rangle}_{\sim a^\dagger}$$

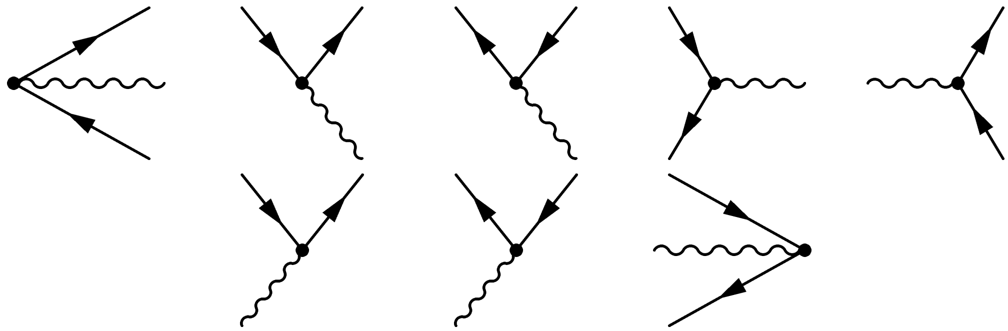


Figure 5.6: *First order contributions* $\mathcal{S}^{(1)}$. These processes violate energy-momentum conservation and are therefore unphysical.

which means that only terms in $\mathcal{S}$ matching $\langle f| \cdot |i\rangle$ yield contributions to the transition amplitude. The following field operators, which constitute the Feynman rules in position space, are contained in $\mathcal{S}$ ($\overrightarrow{\text{time}}$).

$\psi^+(x)$	absorption of electron at x	x
$\bar{\psi}^+(x)$	absorption of positron at x	x
$\bar{\psi}^-(x)$	emission of electron at x	x
$\psi^-(x)$	emission of positron at x	x
$A^+(x)$	absorption of photon at x	x
$A^-(x)$	emission of photon at x	x
$\underbrace{\psi(x_2)\bar{\psi}}(x_1)$ $= iS_F(x_2 - x_1)$	Fermion propagator	x_1 x_2
$\underbrace{A^\mu(x_2)A^\nu}(x_1)$ $= iD_F^{\mu\nu}(x_2 - x_1)$	photon propagator	x_1 x_2
$-ie\bar{\psi}(x)\gamma_\mu\psi(x)A^\mu(x)$ $= -ie\gamma_\mu \cdot$ vertex at x	vertex at x	x

The $\mathcal{S}$-operator at order n is examined using Wick's theorem. At fist order, this yields (remembering Eq. (5.92) while ignoring disconnected contributions from Wick's theorem) the following $2^3 = 8$ contributions:

$$\mathcal{S}^{(1)} = -ie \int d^4x T(\bar{\psi}(x)\gamma_\mu\psi(x)A^\mu(x)) = -ie \int d^4x :\bar{\psi}(x)\gamma_\mu\psi(x)A^\mu(x): .$$

There is a total of 8 possible combinations, since A^μ creates or annihilates a photon, $\bar{\psi}$ creates an electron or annihilates a positron, and ψ creates a position or annihilates an electron. Fig. 5.6 shows the corresponding Feynman diagrams.

However, all these processes are unphysical because they violate energy-momentum conservation:

$$\pm p_{e^+} \pm p_{e^-} \pm p_\gamma \neq 0$$

which is because free particles fulfill

$$p_{e^+}^2 = m_e^2 \qquad p_{e^-}^2 = m_e^2 \qquad p_\gamma^2 = 0.$$

To find physical contributions to the interaction Hamiltonian, we turn to the second order contributions to $\mathcal{S}$ (see Eq. (5.92)):

$$\mathcal{S}^{(2)} = \frac{1}{2!}(-ie)^2 \int d^4x_1 d^4x_2 T\Big(\bar{\psi}(x_1)\gamma_{\mu_1}\psi(x_1)A^{\mu_1}(x_1)\bar{\psi}(x_2)\gamma_{\mu_2}\psi(x_2)A^{\mu_2}(x_2)\Big).$$

Application of Wick's theorem yields contraction terms. We first note that contractions of the form

$$\psi(x_1)\psi(x_2) \qquad \overline{\psi}(x_1)\overline{\psi}(x_2)$$

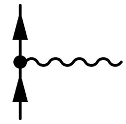
vanish because they contain creators and annihilators, respectively, for different particles and thus

$$\langle 0| T(\psi(x_1)\psi(x_2)) |0\rangle = 0.$$

The remaining terms read, using shorthands like $\bar{\psi}(x_1) = \bar{\psi}_1$,

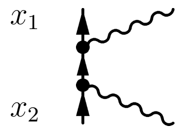
$$\mathcal{S}^{(2)} = \frac{(-ie)^2}{2!}\int d^4x_1 d^4x_2 \Big\{$$

$$:\bar\psi_1\gamma_{\mu_1}\psi_1\bar\psi_2\gamma_{\mu_2}\psi_2 A_1^{\mu_1}A_2^{\mu_2}:$$



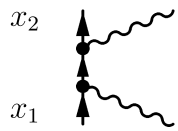
(a)

$$+:\bar\psi_1\gamma_{\mu_1}\underbrace{\psi_1\bar\psi_2}_{}\gamma_{\mu_2}\psi_2 A_1^{\mu_1}A_2^{\mu_2}:$$



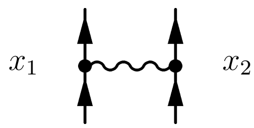
(b)

$$+:\underbrace{\bar\psi_1\gamma_{\mu_1}\psi_1\bar\psi_2\gamma_{\mu_2}\psi_2}_{} A_1^{\mu_1}A_2^{\mu_2}:$$



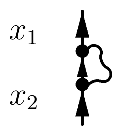
(c)

$$+:\bar\psi_1\gamma_{\mu_1}\psi_1\bar\psi_2\gamma_{\mu_2}\psi_2 \underbrace{A_1^{\mu_1}A_2^{\mu_2}}_{}:$$



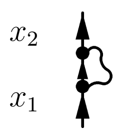
(d)

$$+:\bar\psi_1\gamma_{\mu_1}\underbrace{\psi_1\bar\psi_2}_{}\gamma_{\mu_2}\psi_2 \underbrace{A_1^{\mu_1}A_2^{\mu_2}}_{}:$$



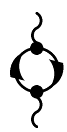
(e)

$$+:\underbrace{\bar\psi_1\gamma_{\mu_1}\psi_1\bar\psi_2\gamma_{\mu_2}\psi_2}_{} \underbrace{A_1^{\mu_1}A_2^{\mu_2}}_{}:$$



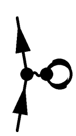
(f)

$$+:\bar\psi_1\gamma_{\mu_1}\psi_1\bar\psi_2\gamma_{\mu_2}\psi_2 A_1^{\mu_1}A_2^{\mu_2}:$$



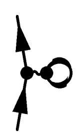
(g)

$$+:\bar\psi_1\gamma_{\mu_1}\psi_1 \underbrace{\bar\psi_2\gamma_{\mu_2}\psi_2}_{} \underbrace{A_1^{\mu_1}A_2^{\mu_2}}_{}:$$



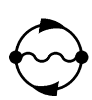
(h)

$$+:\underbrace{\bar\psi_1\gamma_{\mu_1}\psi_1}_{} \bar\psi_2\gamma_{\mu_2}\psi_2 \underbrace{A_1^{\mu_1}A_2^{\mu_2}}_{}:$$



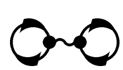
(i)

$$+:\underbrace{\bar\psi_1\gamma_{\mu_1}\underbrace{\psi_1\bar\psi_2}_{}\gamma_{\mu_2}\psi_2}_{} \underbrace{A_1^{\mu_1}A_2^{\mu_2}}_{}:$$



(j)

$$+:\underbrace{\bar\psi_1\gamma_{\mu_1}\psi_1}_{} \underbrace{\bar\psi_2\gamma_{\mu_2}\psi_2}_{} \underbrace{A_1^{\mu_1}A_2^{\mu_2}}_{}:\Big\}.$$



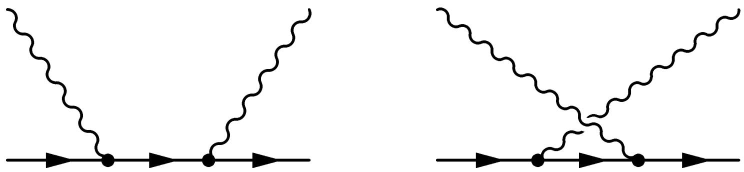
(k)

It follows a discussion of the contributions (a) through (k).

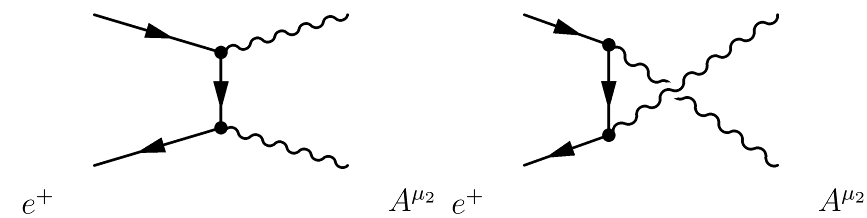
(a) *Independent emission or absorption.* These diagrams violate energy-momentum conservation.

(b)&(c) *Processes involving two electrons or positrons and two photons.*

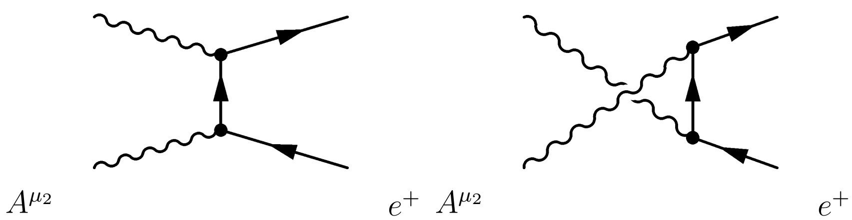
1. Compton scattering: $\gamma e^- \to \gamma e^-$, $\gamma e^+ \to \gamma e^+$



2. Electron-positron pair annihilation: $e^+e^- \to \gamma\gamma$

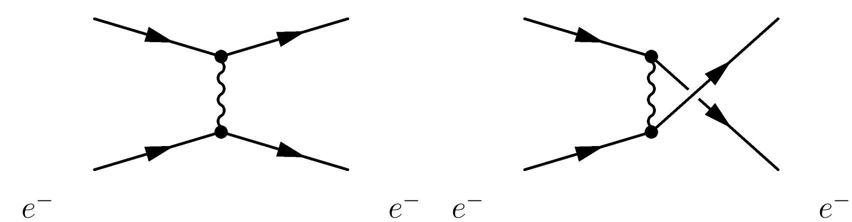


3. Electron-positron pair creation: $\gamma\gamma \to e^+e^-$

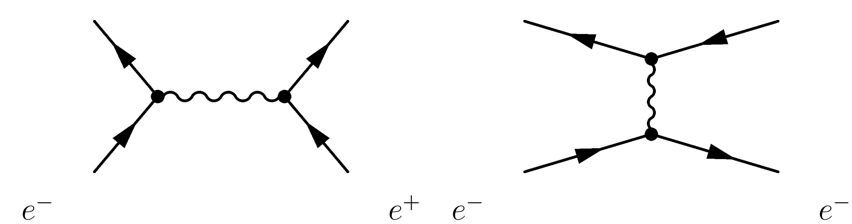


(d) *Processes involving four electrons or positions.*

1. Møller scattering: $e^-e^- \to e^-e^-$, $e^+e^+ \to e^+e^+$



2. Bhabha scattering: $e^+e^- \to e^+e^-$



(e)&(f) *No interaction between external particles.* No scattering takes place, these terms are corrections to the fermion propagator.

(g) *Correction to photon propagator.*

(h)&(i) *Corrections to fermion propagator, vanishing.*

(j)&(k) *Vacuum → vacuum transitions, disconnected graphs.*

This constitutes a list of all known processes (for practical purposes) in $\mathcal{S}^{(2)}$; in general, we can find all processes by examining all orders of the scattering matrix operator $\mathcal{S}$.

The $\mathcal{S}$-matrix elements are defined as matrix elements between single-particle states. Consequently, we need to apply the norm (5.49) repectively (5.50) to external states. The invariant amplitudes $\mathcal{M}_{fi}$, which are derived from the $\mathcal{S}$-matrix elements according to Eq. (3.11) properly account for this normalization factor, and are evaluated for continuum states as defined in Eq. (5.47) and Eq. (5.48).

The contractions of the field operators (see Eq. (5.43) and (5.44)) with external momentum eigenstates (as given in Eq. (5.49) and (5.50)) are for electrons

$$\begin{aligned}
\psi(x)\left|e^-(p,s)\right\rangle_{\text{single-particle}} &= \frac{1}{\sqrt{2E_pV}}\int\frac{d^3k}{(2\pi)^3}\frac{1}{\sqrt{2E_k}}\sum_r a_r(k)u_r(k)\mathrm{e}^{-ik\cdot x}\sqrt{2E_p}a_s^\dagger(p)\left|0\right\rangle \\
&= \frac{1}{\sqrt{2E_pV}}e^{-ip\cdot x}u_s(p)\left|0\right\rangle \\
\left\langle e^-(p,s)\right|_{\text{single-particle}}\bar{\psi}(x) &= \frac{1}{\sqrt{2E_pV}}e^{+ip\cdot x}\left\langle 0\right|\bar{u}_s(p),
\end{aligned}$$

for positrons

$$\begin{aligned}
\bar{\psi}(x)\left|e^+(p,s)\right\rangle_{\text{single-particle}} &= \frac{1}{\sqrt{2E_pV}}e^{-ip\cdot x}\bar{v}_s(p)\left|0\right\rangle \\
\left\langle e^+(p,s)\right|_{\text{single-particle}}\psi(x) &= \frac{1}{\sqrt{2E_pV}}e^{+ip\cdot x}\left\langle 0\right|v_s(p),
\end{aligned}$$

and for photons

$$\begin{aligned}
A_\mu(x)\left|\gamma(k,\lambda)\right\rangle &= \frac{1}{\sqrt{2E_kV}}e^{-ik\cdot x}\varepsilon_\mu^\lambda(k) \\
\left\langle\gamma(k,\lambda)\right|A_\mu(x) &= \frac{1}{\sqrt{2E_kV}}e^{+ik\cdot x}\varepsilon_\mu^{*\lambda}(k).
\end{aligned}$$

Example We treat the case of Møller scattering $e^-e^- \to e^-e^-$ as a typical example for the application of the Feynman rules.

We first define our initial and final states,

$$
\begin{aligned}
|i\rangle &= \left|e^-(p_1,s_1)\right\rangle_{\text{single-particle}} \otimes \left|e^-(p_2,s_2)\right\rangle_{\text{single-particle}} = \sqrt{2E_1 2E_2}\frac{1}{\sqrt{2E_1V2E_2V}} a^\dagger_{s_1}(p_1)a^\dagger_{s_2}(p_2)\,|0\rangle\,,\\
\langle f| &= \left\langle e^-(p_3,s_3)\right|_{\text{single-particle}} \otimes \left\langle e^-(p_4,s_4)\right|_{\text{single-particle}} = \sqrt{2E_3 2E_4}\frac{1}{\sqrt{2E_3V2E_4V}}\,\langle 0|\, a_{s_4}(p_4)a_{s_3}(p_3).
\end{aligned}
$$

The transition matrix element $\mathcal{S}_{fi}$ is then,

$$
\begin{aligned}
\mathcal{S}_{fi} = \langle f|\,\mathcal{S}\,|i\rangle = \frac{(-ie)^2}{2!}\int d^4x_1 d^4x_2 \sqrt{16E_1E_2E_3E_4}\,\langle 0|\,\underbrace{a_{s_4}(p_4)}_{E}\,\underbrace{a_{s_3}(p_3)}_{D}\\
:\underbrace{\bar\psi(x_1)}_{D}\,\gamma_\mu\,\underbrace{\underbrace{\psi(x_1)}_{C}\,\underbrace{\bar\psi(x_2)}_{E}}_{\overset{B}{=}-\bar\psi(x_2)\psi(x_1)}\,\gamma_\nu\,\underbrace{\psi(x_2)}_{A}:\,A^\mu(x_1)A^\nu(x_2)\,\underbrace{a^\dagger_{s_1}(p_1)}_{A}\,\underbrace{a^\dagger_{s_2}(p_2)}_{C}\,|0\rangle\,,
\end{aligned}
\tag{5.95}
$$

yielding $2\times 2 = 4$ Feynman graphs in position space (of which 2! are topologically identical). In Fig. 5.7, we labeled the last Feynman graph according to Eq. (5.95).

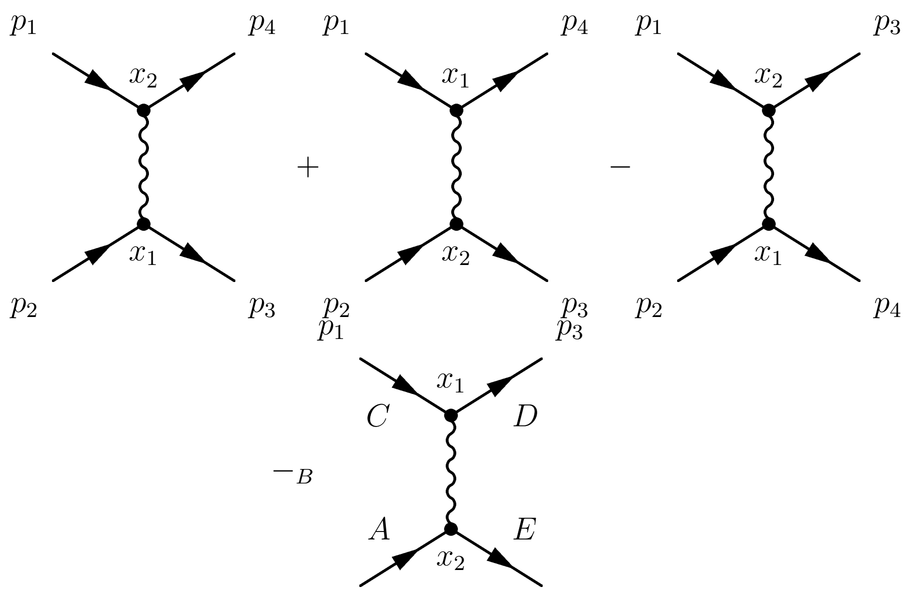


Figure 5.7: Feynman graphs associated with the Møller scattering.

We recall that each ordering of ψ, $\bar\psi$ corresponds to a Feynman diagram. The anticommutation relations are responsible for the relative sign changes.

With the photon propagator in momentum space,

$$iD_F^{\mu\nu}(q) = -\frac{ig^{\mu\nu}}{q^2}, \tag{5.96}$$

we get,

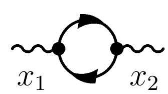
$$\begin{aligned}\mathcal{S}_{fi} =& (-ie)^2(2\pi)^4\delta^{(4)}(p_3+p_4-p_1-p_2)\frac{1}{\sqrt{16E_1E_2E_3E_4V^2}} \\ &\Big[\bar{u}_{s_4}(p_4)\gamma_\mu u_{s_2}(p_2) iD_F^{\mu\nu}(p_3-p_1)\bar{u}_{s_3}(p_3)\gamma_\mu u_{s_1}(p_1) \\ &- \bar{u}_{s_4}(p_4)\gamma_\mu u_{s_1}(p_1) iD_F^{\mu\nu}(p_3-p_2)\bar{u}_{s_3}(p_3)\gamma_\mu u_{s_2}(p_2)\Big].\end{aligned} \tag{5.97}$$

We now define the **invariant amplitude** $\mathcal{M}_{fi}$ (see Eq. (3.11)) via,

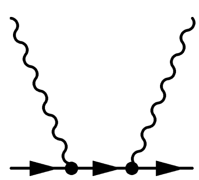
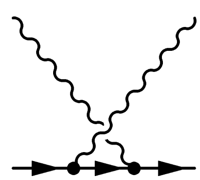
$$\mathcal{S}_{fi} = \delta_{fi} + i(2\pi)^4\delta^{(4)}(p_3+p_4-p_1-p_2)\frac{1}{\sqrt{16E_1E_2E_3E_4V^2}}\mathcal{M}_{fi}. \tag{5.98}$$

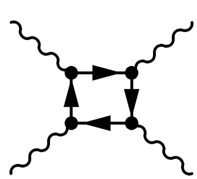
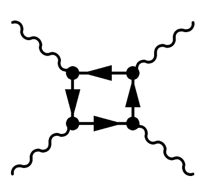
$\mathcal{M}_{fi}$ can then be computed using the Feynman rules in momentum space.

Application of the Feynman rules

- Momentum conservation at each vertex
- Fermion number conservation at each vertex (indicated by the direction of the arrows)
- All topologically allowed graphs contribute
- Exchange factor (-1) when interchanging two external fermions with each other
- Each closed fermion loop yields a factor (-1), e.g. 

coming from the contraction $:\overline{\psi}(x_1)\psi(x_1)\overline{\psi}(x_2)\psi(x_2):$
- graphs in which the ordering of the vertices along a fermion line is different are not topologically equivalent, and must be summed, eg.

$\mathcal{S}^{(2)}$:  $\neq$ 

$\mathcal{S}^{(4)}$:  $\neq$ 

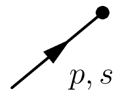
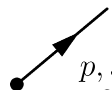
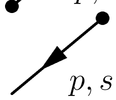
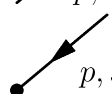
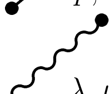
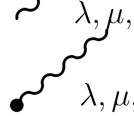
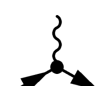
External states			
incoming electron		p, s	$u_s(p)$
outgoing electron		p, s	$\bar{u}_s(p)$
incoming positron		p, s	$\bar{v}_s(p)$
outgoing positron		p, s	$v_s(p)$
incoming photon		λ, μ, p	$(\varepsilon_\lambda)^\mu(p)$
outgoing photon		λ, μ, p	$(\varepsilon^*_\lambda)^\mu(p)$
Propagators			
electron		p	$\frac{i(\not{p}+m)}{p^2-m^2+i\varepsilon}$
photon	μ  ν	k	$\frac{-ig^{\mu\nu}}{k^2+i\varepsilon}$
Vertex		μ	
electron-photon-electron		q_e	$-ieq_e\gamma^\mu$

Table 5.2: Feynman rules in momentum space.

5.9 Trace techniques for γ-matrices

Cross sections are proportinal to $|\mathcal{M}_{fi}|^2 \propto |\bar{u}_{s_f}(p_f)\Gamma u_{s_i}(p_i)|^2$, where Γ denotes an arbitrary product of γ-matrices.

In many experiments – but not all! –, the spin states of the initial and final states are not observed. This is for example the case at the CMS and ATLAS experiments of LHC. We then need to follow the following procedure :

- If the spin state of the final state particles cannot be measured, one must sum over the final state spins : $\sum_{s_f} |\cdots|^2$,
- If the initial states particles are unpolarized, one must average over the initial state spins : $\frac{1}{2}\sum_{s_i} |\cdots|^2$.

Then, remembering that $\bar{u} = u^\dagger\gamma^0$, we can write,

$$\begin{aligned}
\frac{1}{2}\sum_{s_i,s_f} |\bar{u}_{s_f}(p_f)\Gamma u_{s_i}(p_i)|^2 &= \frac{1}{2}\sum_{s_i,s_f} \bar{u}_{s_f}(p_f)\Gamma u_{s_i}(p_i)u^\dagger_{s_i}(p_i)\gamma^0\gamma^0\Gamma^\dagger\gamma^0 u_{s_f}(p_f) \\
&= \frac{1}{2}\sum_{s_i,s_f} \bar{u}_{s_f}(p_f)\Gamma u_{s_i}(p_i)\bar{u}_{s_i}(p_i)\bar{\Gamma} u_{s_f}(p_f) \\
&= \frac{1}{2}\sum_{s_i,s_f} (\bar{u}_{s_f}(p_f))_\alpha \Gamma_{\alpha\beta}(u_{s_i}(p_i))_\beta(\bar{u}_{s_i}(p_i))_\gamma\bar{\Gamma}_{\gamma\delta}(u_{s_f}(p_f))_\delta \\
&\overset{(5.41)}{=} \frac{1}{2}\Gamma_{\alpha\beta}(\not{p}_i + m)_{\beta\gamma}\bar{\Gamma}_{\gamma\delta}(\not{p}_f + m)_{\delta\alpha} \\
&= \frac{1}{2}\Big(\Gamma(\not{p}_i + m)\bar{\Gamma}(\not{p}_f + m)\Big)_{\alpha\alpha} \\
&= \frac{1}{2}\mathrm{Tr}\Big(\Gamma(\not{p}_i + m)\bar{\Gamma}(\not{p}_f + m)\Big),
\end{aligned}$$

where the indices α, β, γ and δ label the matrix element, and $\bar{\Gamma} := \gamma^0\Gamma^\dagger\gamma^0$.

We thus get the important result,

$$\boxed{\frac{1}{2}\sum_{s_i,s_f} |\bar{u}_{s_f}(p_f)\Gamma u_{s_i}(p_i)|^2 = \frac{1}{2}\mathrm{Tr}\Big(\Gamma(\not{p}_i + m)\bar{\Gamma}(\not{p}_f + m)\Big)}, \tag{5.99}$$

and its analogon for antiparticles,

$$\boxed{\frac{1}{2}\sum_{s_i,s_f} |\bar{v}_{s_f}(p_f)\Gamma v_{s_i}(p_i)|^2 = \frac{1}{2}\mathrm{Tr}\Big(\Gamma(\not{p}_i - m)\bar{\Gamma}(\not{p}_f - m)\Big)}, \tag{5.100}$$

i.e. the Clifford algebra of γ-matrices is taking care of the spin summation for us.

We now compute $\bar{\Gamma}$ for an arbitrary number of γ-matrices.

- For $\Gamma = \gamma^\mu$, $(\gamma^0)^\dagger = \gamma^0$, $(\gamma^i)^\dagger = -\gamma_i$ hence $(\gamma^\mu)^\dagger = \gamma^0\gamma^\mu\gamma^0 \Rightarrow \bar{\gamma^\mu} = \gamma^\mu$. For later use, note that $\bar{\gamma^5} = -\gamma^5$.

- For $\Gamma = \gamma^{\mu_1}\cdots\gamma^{\mu_n}$, $\Gamma^\dagger = (\gamma^{\mu_1}\cdots\gamma^{\mu_n})^\dagger = \gamma^0\gamma^{\mu_n}\cdots\gamma^{\mu_1}\gamma^0 \Rightarrow \bar{\Gamma} = \gamma^{\mu_n}\cdots\gamma^{\mu_1}$. In other words, to get $\bar{\Gamma}$, we just need to read Γ in the inverse ordering.

We finally want to compute some traces for products of γ-matrices, since they appear explicitly $(\Gamma, \bar{\Gamma})$ and implicitely $(\not{p} = \gamma^\mu p_\mu)$ in the formulas (5.99) and (5.100). In doing this, one should remember that the trace is cyclic $(\mathrm{Tr}(ABC) = \mathrm{Tr}(BCA))$ and the Clifford algebra of γ-matrices.

- **0 γ-matrix :** $\mathrm{Tr}\,\mathbb{1} = 4$.

- **1 γ-matrix :** $\mathrm{Tr}\,\gamma^\mu = 0$, $\mathrm{Tr}\,\gamma^5 = 0$. The last one is shown using the fact that $\{\gamma^5, \gamma^\mu\} = 0$.

- **2 γ-matrices :** $\mathrm{Tr}(\gamma^\mu\gamma^\nu) = \frac{1}{2}\mathrm{Tr}(\gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu) = 4g^{\mu\nu} \Rightarrow \mathrm{Tr}(\not{a}\not{b}) = 4a\cdot b$, where $\cdot$ is the scalar product of 4-vectors.

- **4 γ-matrices :**

$$\begin{aligned}
\mathrm{Tr}(\gamma^\mu\gamma^\nu\gamma^\rho\gamma^\sigma) &= \mathrm{Tr}(\gamma^\nu\gamma^\rho\gamma^\sigma\gamma^\mu) = -\mathrm{Tr}(\gamma^\nu\gamma^\rho\gamma^\mu\gamma^\sigma) + 2g^{\mu\sigma}\mathrm{Tr}(\gamma^\nu\gamma^\rho)\\
&= \mathrm{Tr}(\gamma^\nu\gamma^\mu\gamma^\rho\gamma^\sigma) + 8g^{\mu\sigma}g^{\nu\rho} - 8g^{\mu\rho}g^{\nu\sigma}\\
&= -\mathrm{Tr}(\gamma^\mu\gamma^\nu\gamma^\rho\gamma^\sigma) + 8g^{\mu\sigma}g^{\nu\rho} - 8g^{\mu\rho}g^{\nu\sigma} + 8g^{\mu\nu}g^{\rho\sigma}\\
\Rightarrow \mathrm{Tr}(\gamma^\mu\gamma^\nu\gamma^\rho\gamma^\sigma) &= 4(g^{\mu\nu}g^{\rho\sigma} + g^{\mu\sigma}g^{\nu\rho} - g^{\mu\rho}g^{\nu\sigma})\\
\Rightarrow \mathrm{Tr}(\not{a}_1\not{a}_2\not{a}_3\not{a}_4) &= 4\left[(a_1\cdot a_2)(a_3\cdot a_4) + (a_1\cdot a_4)(a_2\cdot a_3) - (a_1\cdot a_3)(a_2\cdot a_4)\right].
\end{aligned}$$

  and in general,

$$\begin{aligned}
\mathrm{Tr}(\not{a}_1\cdots\not{a}_n) =&(a_1\cdot a_2)\mathrm{Tr}(\not{a}_3\cdots\not{a}_n) - (a_1\cdot a_3)\mathrm{Tr}(\not{a}_3\not{a}_4\cdots\not{a}_n)\\
&+\cdots\pm(a_1\cdot a_n)\mathrm{Tr}(\not{a}_2\cdots\not{a}_{n-1}),
\end{aligned}$$

  which implies inductivly that the trace of a string of γ-matrices is a real number.

- **n γ-matrices (n odd) :**

$$\begin{aligned}
\mathrm{Tr}(\gamma^{\mu_1}\cdots\gamma^{\mu_n}) &= \mathrm{Tr}(\gamma^{\mu_1}\cdots\gamma^{\mu_n}\underbrace{\gamma^5\gamma^5}_{=\mathbb{1}}) = \mathrm{Tr}(\gamma^5\gamma^{\mu_1}\cdots\gamma^{\mu_n}\gamma^5)\\
&= (-1)^n\mathrm{Tr}(\gamma^{\mu_1}\cdots\gamma^{\mu_n}) \Rightarrow \mathrm{Tr}(\gamma^{\mu_1}\cdots\gamma^{\mu_n}) = 0.
\end{aligned}$$

- **n γ-matrices (n even) :**

$$\mathrm{Tr}(\gamma^{\mu_1}\cdots\gamma^{\mu_n}) = \mathrm{Tr}((\gamma^{\mu_1}\cdots\gamma^{\mu_n})^\dagger) = \mathrm{Tr}(\gamma^0\gamma^{\mu_n}\cdots\gamma^{\mu_1}\gamma^0) = \mathrm{Tr}(\gamma^{\mu_n}\cdots\gamma^{\mu_1}).$$

- γ^5 **and 2 γ-matrices :** $\mathrm{Tr}(\gamma^5\gamma^\mu\gamma^\nu) = 0$. To show this identity, we remark that $\gamma^5\gamma^\mu\gamma^\nu$ is a rank-2 tensor, which does not depend on any 4-momenta. Therefore, $\mathrm{Tr}(\gamma^5\gamma^\mu\gamma^\nu) = cg^{\mu\nu}$. We contract with $g_{\mu\nu}$ to get $\mathrm{Tr}(\gamma^5\gamma^\mu\gamma_\mu) = cg^{\mu\nu}g_{\mu\nu} = 4c$, but since $\gamma^\mu\gamma_\mu = 4\mathbb{1}$ we get $c = \mathrm{Tr}\,\gamma^5 = 0$.
- γ^5 **and 4 γ-matrices :** $\mathrm{Tr}(\gamma^5\gamma^\mu\gamma^\nu\gamma^\rho\gamma^\sigma) = -4i\varepsilon^{\mu\nu\rho\sigma}$.
- **Contractions :**

$$\gamma^\mu\gamma_\mu = 4\mathbb{1} \tag{5.101}$$
$$\gamma^\mu \not{a}\gamma_\mu = -2\not{a} \tag{5.102}$$
$$\gamma^\mu \not{a}\not{b}\gamma_\mu = 4(a\cdot b)\mathbb{1} \tag{5.103}$$
$$\gamma^\mu \not{a}\not{b}\not{c}\gamma_\mu = -2\not{c}\not{b}\not{a} \tag{5.104}$$

5.10 Annihilation process : $e^+e^- \to \mu^+\mu^-$

In this section, we compute the differential cross section of the simplest of all QED process, the reaction

$$e^-(p_1)e^+(p_2) \to \mu^-(p_3)\mu^+(p_4),$$

illustrated on Fig. 5.8. The simplicity arises frome the fact that $e^- \neq \mu^-$, and hence only one diagram contributes (the e^+e^--pair *must* be annihilated).

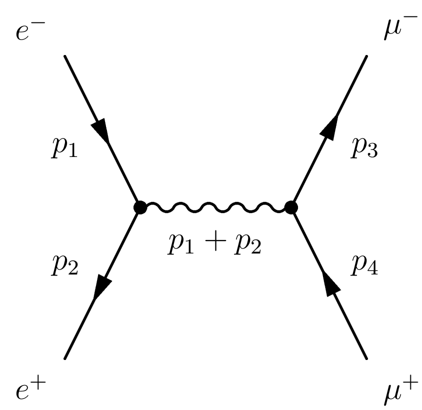


Figure 5.8: Annihilation process $e^+e^- \to \mu^+\mu^-$

We recall the Mandelstam variables for this process,

$$s = (p_1+p_2)^2$$
$$t = (p_1-p_3)^2$$
$$u = (p_1-p_4)^2.$$

We make the following assumptions (very common for QED processes),

- Unpolarized leptons : $\frac{1}{2}\sum_{s_i,s_f}$,
- High energy limit : $m_e, m_\mu = 0 \Leftrightarrow \sqrt{s} \gg m_e, m_\mu$.

Using the Feynman rules of Table 5.2 for the diagram depicted in Fig. 5.8, we get,

$$\begin{aligned}
-i\mathcal{M}_{fi} &= \bar{u}_{s_3}(p_3)ie\gamma^\mu v_{s_4}(p_4)\frac{-ig_{\mu\nu}}{(p_1+p_2)^2}\bar{v}_{s_2}(p_2)ie\gamma^\nu u_{s_1}(p_1) \\
\overline{|\mathcal{M}_{fi}|^2} &= \frac{1}{2}\sum_{s_1}\frac{1}{2}\sum_{s_2}\sum_{s_3}\sum_{s_4}|\mathcal{M}_{fi}|^2 \\
&= \frac{1}{4}\frac{e^4}{s^2}\mathrm{Tr}(\gamma^\mu \not{p}_4\gamma^\nu \not{p}_3)\mathrm{Tr}(\gamma_\mu \not{p}_1\gamma_\nu \not{p}_2) \\
&= \frac{1}{4}\frac{e^4}{s^2}16\left[2(p_1\cdot p_3)(p_2\cdot p_4)+2(p_1\cdot p_4)(p_2\cdot p_3)\right].
\end{aligned}$$

Since we are working in the high energy limit, we have $p_i^2 = 0$ and hence $t = -2p_1 \cdot p_3 = -2p_2 \cdot p_4$ and $u = -2p_1 \cdot p_4 = -2p_2 \cdot p_3$. Using the identity $s+t+u = 2m_e^2 + 2m_\mu^2 = 0$ to get rid of the Mandelstam u-variable and with $\alpha = \frac{e^2}{4\pi}$ we have,

$$\overline{|\mathcal{M}_{fi}|^2} = 32\pi^2\alpha^2\frac{t^2+(s+t)^2}{s^2}. \tag{5.105}$$

Considering the center of mass frame, we have $s = 4(E^*)^2$, $t = -\frac{s}{2}(1-\cos\Theta^*)$ and with the help of Eq. (3.34), this yields,

$$\frac{d\sigma}{dt} = \frac{1}{16\pi s^2}\overline{|\mathcal{M}_{fi}|^2} = \frac{\pi\alpha^2}{s^2}(1+\cos^2\Theta^*), \tag{5.106}$$

or using,

$$\frac{d\sigma}{dt} = \frac{d\Omega^*}{dt}\frac{d\sigma}{d\Omega^*} = \frac{4\pi}{s}\frac{d\sigma}{d\Omega^*},$$

we get the differential cross section for $e^+e^- \to \mu^+\mu^-$ in the center of mass frame,

$$\boxed{\frac{d\sigma^{e^+e^-\to\mu^+\mu^-}}{d\Omega^*} = \frac{\alpha^2}{4s}(1+\cos^2\Theta^*)}. \tag{5.107}$$

This differential cross section (see Fig. 5.9) has been very well measured and is one of the best tests of QED at high energies.

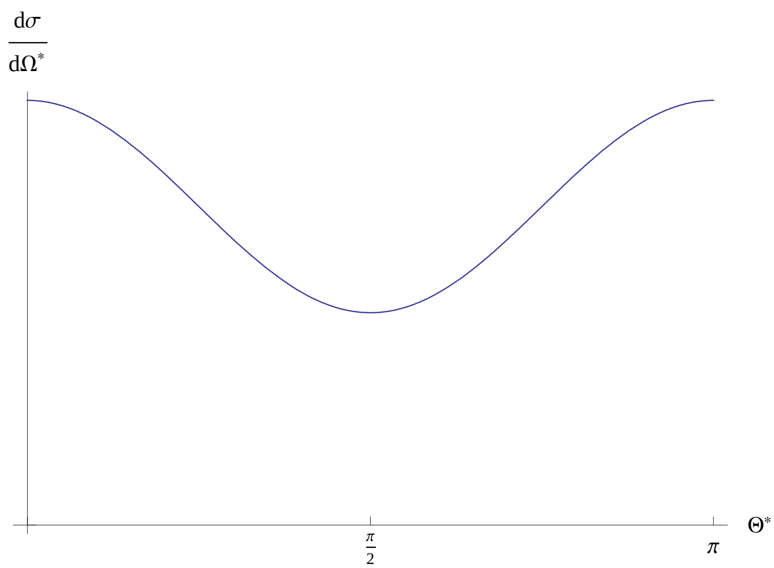


Figure 5.9: Differential cross section for $e^+e^- \to \mu^+\mu^-$ in the center of mass frame.

Using this result, we can calculate the total cross-section by integration over the solid angle:

$$\sigma = \int \frac{d\sigma}{d\Omega^*} d\Omega^* = \frac{\alpha^2}{4s} \int_0^\pi (1+\cos^2\Theta^*) \underbrace{\sin\Theta^* d\Theta^*}_{d\cos\Theta^*} \underbrace{\int_0^{2\pi} d\phi}_{2\pi} \tag{5.108}$$

$$= \frac{\alpha^2}{4s} 2\pi \frac{8}{3} \tag{5.109}$$

$$\Rightarrow \boxed{\sigma^{e^+e^-\to\mu^+\mu^-} = \frac{4\pi\alpha^2}{3s}} = \frac{86.9\,\text{nb}}{s\,[\text{GeV}^2]} \tag{5.110}$$

where $1\,\text{nb} = 10^{-33}\,\text{cm}^2$. If one considers non-asymptotic energies, $s \simeq m_\mu^2$ (but $s \gg m_e^2$), one finds a result which reduces to Eq. (5.110) for $m_\mu^2 = 0$:

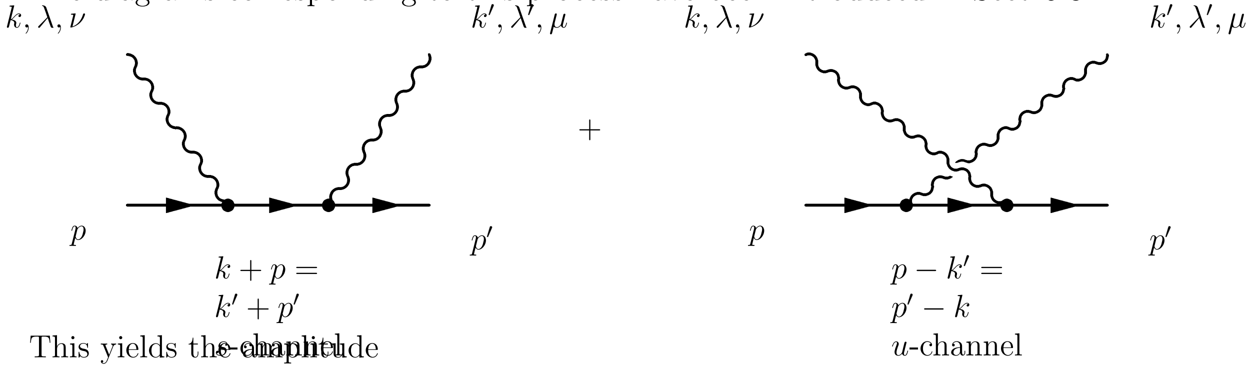
$$\sigma^{e^+e^-\to\mu^+\mu^-} = \frac{4\pi\alpha^2}{3s}\left(1 + 2\frac{m_\mu^2}{s}\right)\sqrt{1 - \frac{4m_\mu^2}{s}}.$$

5.11 Compton scattering

Let us now consider Compton scattering:

$$\gamma(k) + e^-(p) \to \gamma(k') + e^-(p').$$

The diagrams corresponding to this process have been introduced in Sect. 5.8.



This yields the amplitude

$$-i\mathcal{M}_{fi} = \varepsilon^*_\mu(k', \lambda')\varepsilon_\nu(k, \lambda)\bar{u}(p') \left[\underbrace{ie\gamma^\mu \frac{i}{\not{p} + \not{k} - m} ie\gamma^\nu}_{\text{LHS diagram}} + \underbrace{ie\gamma^\nu \frac{i}{\not{p} - \not{k}' - m} ie\gamma^\mu}_{\text{RHS diagram}} \right] u(p)$$

where the on-shell conditions read

$$k^2 = k'^2 = 0 \qquad\qquad p^2 = p'^2 = m^2$$

and the photons are transversal:

$$k \cdot \varepsilon(k) = k' \cdot \varepsilon(k') = 0.$$

It is instructive to check that the invariant amplitude is indeed also gauge invariant. Consider the gauge transformation

$$A_\nu(x) \to A_\nu(x) + \partial_\nu \Lambda(x)$$

which leaves Maxwell's equations unaltered. In the photon field operator this can be implemented by

$$\varepsilon_\nu(k, \lambda) \to \varepsilon_\nu(k, \lambda) + \beta k_\nu, \;\; \beta \in \mathbb{R} \text{ arbitrary.}$$

We observe the change of the matrix element for transformation of one of the photons:

$$-i\mathcal{M}_{fi}(\varepsilon_\nu \to k_\nu) = -ie^2 \varepsilon^*_\mu(k', \lambda')\bar{u}(p') \left[\gamma^\mu \frac{1}{\not{p} + \not{k} - m} \not{k} + \not{k} \frac{1}{\not{p} - \not{k}' - m} \gamma^\mu \right] u(p).$$

In simplifying this expression, we use

$$\frac{1}{\not{p} + \not{k} - m} \not{k} u(p) = \frac{1}{\not{p} + \not{k} - m} (\not{k} + \not{p} - m) u(p) = \mathbb{1} u(p)$$

where we added a zero since $(\not{p} - m)u(p) = 0$ and analogously

$$\bar{u}(p')\not{k}\frac{1}{\not{p} - \not{k}' - m} = \bar{u}(p')(\not{k} - \not{p}' + m)\frac{1}{\not{p}' - \not{k} - m} = -\bar{u}(p')\mathbb{1}.$$

Putting the terms together, we therefore find

$$-i\mathcal{M}_{fi}(\varepsilon_\nu \to k_\nu) = -ie^2\varepsilon^*_\mu(k', \lambda')\bar{u}(p')(\gamma^\mu\mathbb{1} - \mathbb{1}\gamma^\mu)u(p) = 0.$$

The result is the same for the transformation $\varepsilon^*_\mu \to \varepsilon^*_\mu + \beta k'_\mu$.

It is generally true that only the sum of the contributing diagrams is gauge invariant. Individual diagrams are not gauge invariant and thus without physical meaning.

Recall that the aim is to find the differential cross section and therefore the squared matrix element. Since there are two contributing diagrams, one has to watch out for interference terms. Applying the trace technology developed in Sect. 5.9 yields

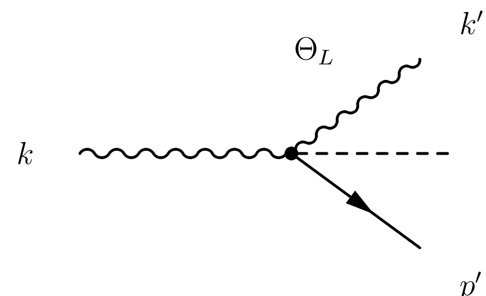
$$\begin{aligned}\overline{|\mathcal{M}_{fi}|^2} &= \frac{1}{2}\sum_\lambda \frac{1}{2}\sum_s \sum_{\lambda'} \sum_{s'} |\mathcal{M}_{fi}|^2 \\ &= 2e^4\left[\frac{m^2 - u}{s - m^2} + \frac{m^2 - s}{u - m^2} + 4\left(\frac{m^2}{s - m^2} + \frac{m^2}{u - m^2}\right) + 4\left(\frac{m^2}{s - m^2} + \frac{m^2}{u - m^2}\right)^2\right].\end{aligned} \tag{5.111}$$

Bearing in mind that $s + t + u = 2m^2$, this yields the unpolarized Compton cross-section

$$\frac{d\sigma}{dt} = \frac{1}{16\pi(s - m^2)^2}\overline{|\mathcal{M}_{fi}|^2} \tag{5.112}$$

which is a frame independent statement.

Head-on electron-photon collision is rather uncommon; usually photons are hitting on a target. Therefore it is useful to consider the electron's rest frame (laboratory frame):



With $\omega = |\vec{k}| = E^L_\gamma$, $\omega' = |\vec{k}'| = E'^L_\gamma$, and $p = (m, \vec{0})^T$ one finds

$$s - m^2 = 2m\omega \tag{5.113}$$

$$u - m^2 = -2p \cdot k' = -2m\omega' \tag{5.114}$$

$$t = -2\omega\omega'(1 - \cos\Theta_L). \tag{5.115}$$

One of the three variables can be eliminated using $s+t+u=2m^2$:

$$\omega' = \frac{1}{2m}(s+t-m^2) = \omega - \frac{\omega\omega'}{m}(1-\cos\Theta_L) \tag{5.116}$$

$$\Rightarrow \frac{1}{\omega'} - \frac{1}{\omega} = \frac{1}{m}(1-\cos\Theta_L) \tag{5.117}$$

$$\Rightarrow \omega' = \frac{\omega}{1+\frac{\omega}{m}(1-\cos\Theta_L)}. \tag{5.118}$$

We continue calculating the differential cross-section. Eq. (5.115) yields

$$dt = \frac{\omega'^2}{\pi} 2\pi d\cos\Theta_L = \frac{\omega'^2}{\pi} d\Omega_L.$$

Furthermore, we can use Eq. (5.117) to simplify Eq. (5.111):

$$\frac{m^2}{s-m^2} + \frac{m^2}{u-m^2} = \frac{m^2}{2m\omega} + \frac{m^2}{-2m\omega'} = \frac{m}{2}\left(\frac{1}{\omega} - \frac{1}{\omega'}\right) = -\frac{1}{2}(1-\cos\Theta_L).$$

Using these results and remembering Eq. (5.112), we obtain

$$\frac{d\sigma^{\gamma e\to\gamma e}}{d\Omega_L} = \frac{dt}{d\Omega_L}\frac{d\sigma}{dt} = \frac{\omega'^2}{\pi}\frac{1}{16\pi(2m\omega)^2}2e^2\left[\frac{2m\omega'}{2m\omega} + \frac{-2m\omega}{-2m\omega'} - \sin^2\Theta_L\right] \tag{5.119}$$

$$= \frac{\omega'^2}{\pi}\frac{2\cdot 16\pi^2\alpha^2}{16\pi 4m^2\omega^2}\left[\frac{\omega'}{\omega} + \frac{\omega}{\omega'} - \sin^2\Theta_L\right] \tag{5.120}$$

$$\Rightarrow \boxed{\frac{d\sigma^{\gamma e\to\gamma e}}{d\Omega_L} = \frac{\alpha^2}{2m^2}\left(\frac{\omega'}{\omega}\right)^2\left[\frac{\omega'}{\omega} + \frac{\omega}{\omega'} - \sin^2\Theta_L\right]} \tag{5.121}$$

which is called the Klein-Nishima formula.

It follows a discussion of important limiting cases.

- *Classical limit:* $\omega \ll m \Rightarrow \omega' \simeq \omega$
  In the classical limit, Eq. (5.121) simplifies to the classical Thomson cross-section (which was used to measure α)

$$\boxed{\frac{d\sigma^{\gamma e\to\gamma e}}{d\Omega_L} = \frac{\alpha^2}{2m^2}\left[1+\cos^2\Theta_L\right]},$$

  yielding the total cross-section

$$\boxed{\sigma^{\gamma e\to\gamma e} = \frac{\alpha^2}{2m^2}\frac{16\pi}{3}}.$$

- *Asymptotic limit:* $s \gg m^2 \Rightarrow \omega \gg m$
  In this case, the so-called leading log approximation holds:
$$\sigma^{\gamma e \to \gamma e} = \frac{2\pi\alpha^2}{m^2}\frac{m^2}{s}\left[\ln\frac{s}{m^2} + \frac{1}{2} + \mathcal{O}\left(\frac{m^2}{s}\right)\right] \simeq \frac{2\pi\alpha^2}{s}\ln\frac{s}{m^2}.$$

- *In general* we can conclude that
$$\sigma^{\gamma e \to \gamma e} \sim \frac{\alpha^2}{m^2} \simeq 10^{-25}\,\mathrm{cm}^2$$
  from which one can infer the "classical electron radius"
$$r_e^{\text{classical}} \sim \sqrt{\sigma_{\text{Thomson}}} \sim \frac{\alpha}{m} = 2.8 \cdot 10^{-13}\,\mathrm{cm}.$$

5.12 QED as a gauge theory

Recall the QED Lagrangian

$$\begin{aligned}\mathcal{L}^{\text{QED}} &= \bar{\psi}(i\gamma^\mu\partial_\mu - m)\psi - eq_e\bar{\psi}\gamma^\mu\psi A_\mu - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} \\ &= \mathcal{L}_0^{\text{Dirac}} + \mathcal{L}' + \mathcal{L}_0^{\text{photon}}\end{aligned}$$

introduced in Sect. 5.8 which includes the following observables:

- *Fermions:* components of $\bar{\psi}\gamma^\mu\psi = j^\mu$
- *Photons:* components of $F^{\mu\nu}$: $\vec{E}$ and $\vec{B}$ field.

Neither ψ nor A_μ as such are observables. In particular, the phase of ψ cannot be observed. This means that QED must be invariant under phase transformations of ψ:

$$\psi(x) \to \psi'(x) = e^{ieq_e\chi(x)}\psi(x)$$

which is a unitary one-dimensional i. e. $U(1)$ transformation. Observe first the action on the Dirac Lagrangian:

$$\begin{aligned}\mathcal{L}_0^{\text{Dirac}} &\to \bar{\psi}'(i\gamma^\mu\partial_\mu - m)\psi' \\ &= \bar{\psi}e^{-ieq_e\chi(x)}e^{ieq_e\chi(x)}(i\partial\!\!\!/ - m)\psi - \bar{\psi}\gamma^\mu(\partial_\mu eq_e\chi(x))\psi \\ &= \mathcal{L}_0^{\text{Dirac}} - eq_e\bar{\psi}\gamma^\mu\psi(\partial_\mu\chi(x)).\end{aligned}$$

Therefore, the free Dirac field Lagrangian alone is not invariant under this transformation. In order for the extra term to vanish, A_μ has to be transformed, too:

$$A_\mu \to A'_\mu = A_\mu - \partial_\mu\chi(x)$$

QED	photon	$U(1)$
Weak interaction	$W^{\pm},\ Z^0$	$SU(2)$
QCD	gluon	$SU(3)$

Table 5.3: Summary of gauge theories.

such that $F^{\mu\nu} = F'^{\mu\nu}$ since $F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu$. This means that we are dealing with the gauge transformation known from classical electrodynamics. Because we have

$$-eq_e\bar{\psi}\gamma^\mu\psi A_\mu \to -eq_e\bar{\psi}\gamma^\mu\psi A_\mu + eq_e\bar{\psi}\gamma^\mu\psi(\partial_\mu\chi)$$

the complete Lagrangian $\mathcal{L}^{\text{QED}}$ is invariant under $U(1)$ gauge transformations. This motivates the definition of the gauge covariant derivative

$$\boxed{D_\mu = \partial_\mu + ieq_eA_\mu}$$

which contains the photon-electron interaction.

In summary, the requirement of gauge invariance uniquely determines the photon-electron interaction and QED is a $U(1)$ gauge theory.

This suggests a new approach on theory building: start from symmetries instead of finding them in the final Lagrangian:

$$\boxed{\begin{array}{c}\text{local symmeties}\\ \text{(gauge invariance)}\end{array}} \to \boxed{\begin{array}{c}\text{existence of vector fields}\\ \text{(gauge fields)}\end{array}} \to \boxed{\text{gauge interactions}}.$$

A summary of gauge theories with the corresponding gauge fields and gauge groups is given in Tab. 5.3.

Chapter 6

Tests of QED

In the previous chapter elements of the quantum electrodynamics theory are discussed. We now turn to precision tests of the theory which usually consist in the measurement of the electromagnetic fine structure constant α in different systems. Experimental results are compared with theoretical predictions. The validation process requires very high precision in both measurements and theoretical calculations. QED is then confirmed to the extent that these measurements of α from different physical sources agree with each other. The most stringent test of QED is given by the measurement of the electron magnetic moment. However, several other experimental tests have been performed in different energy ranges and systems:

- Low energy range, accessible with small experiments;
- High energy range, accessible with particle colliders (e.g. e^+e^- colliders);
- Condensed matter systems (quantum Hall effect, Josephson effect).

As we will see, the achieved precision makes QED one of the most accurate physical theories constructed so far.

6.1 Measurement of the electron anomalous magnetic moment

6.1.1 Electron magnetic moment

A rotating electrically charged body creates a magnetic dipole. In classical analogy, this is also the case for the spinning electron. External magnetic fields exert a torque on the electron magnetic moment. Electrons have an *intrinsic magnetic moment* μ, related to their spin s:

$$\mu = -g\frac{e}{2m}s = -\frac{g}{2}\frac{e}{2m} \tag{6.1}$$

where e is the unit charge and m the electron mass. In the case of electrons the magnetic moment is anti-parallel to the spin. The g-factor is equal to 2, as calculated from Dirac's equation:

$$a \equiv \frac{g-2}{2} = 0.$$

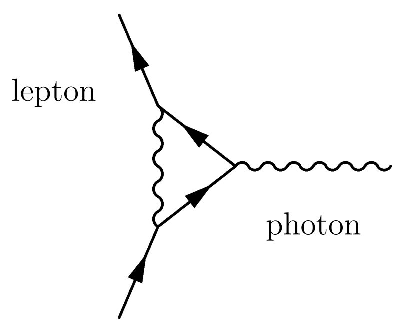
Corrections to the g-factor are given by higher order QED contributions as well as hadronic and weak interactions. There could be additional contributions from physics beyond the Standard Model (SM):

$$\frac{g}{2} = 1 + a_{\mathrm{QED}}(\alpha) + a_{\mathrm{hadronic}} + a_{\mathrm{weak}} + a_{\mathrm{new}}.$$

When adding the corrections we usually talk of the *anomalous magnetic moment* of the electron.

6.1.2 QED: higher order corrections

The one-loop corrections to the magnetic moment are due to vacuum fluctuation and polarization effects. A corresponding diagram is for example



The textbook calculation of the one-loop corrections gives corrections $\sim 10^{-3}$ (see [14, pp. 189]):

$$a = \frac{\alpha}{2\pi} \approx 0.0011614.$$

Hadronic and weak interactions are calculated (within the SM) to be very small and negligible, respectively.

As we will see, the precision achieved by experimental results needs QED predictions with α^4 precision.

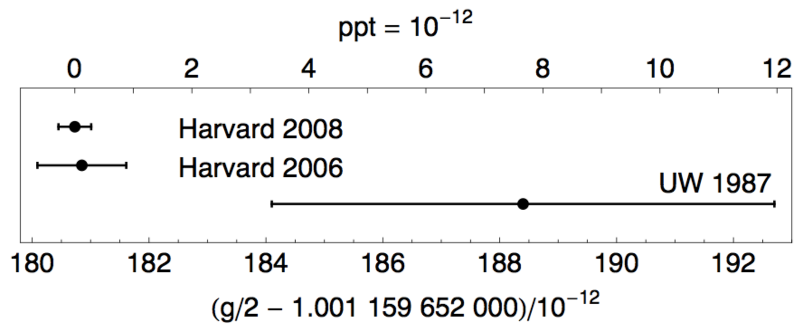


Figure 6.1: *Most accurate measurements of the electron* $g/2$. Source: [16, p. 177].

6.1.3 $g/2$ measurements

Nowadays the precision of the $g/2$ measurements is below 10^{-12} as is shown in Fig. 6.1. The latest measurements are 15 times more precise than the previous result which stood for about 20 years. As one can see in Fig. 6.1, the latest value is shifted by 1.7 standard deviations with respect to the previous result from 1987.

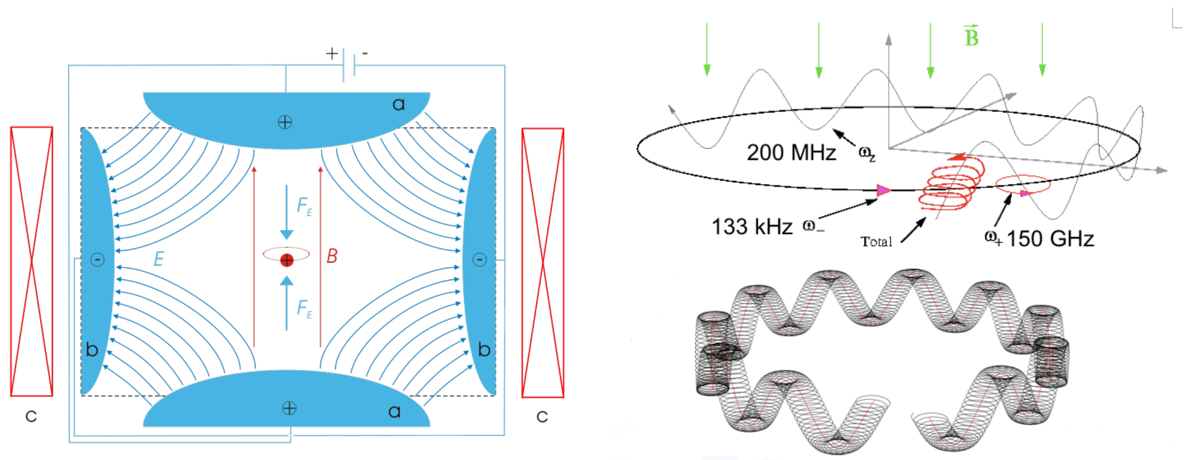
So, how did we get to this astonishing precision?

6.1.3.1 Experiment

The main ingredients of the experiment are:

- *Single-electron quantum cyclotron*
  A Penning trap suspends and confines the electron in an atom-like state.
- *Fully resolved cyclotron and spin energy levels*
  Accurate measurements of the resonant frequencies of driven transitions between the energy levels of this homemade atom—an electron bound to the trap—reveals the electron magnetic moment in units of Bohr magnetons, $g/2$.
- *Detection sensitivity sufficient to detect one quantum transitions*
  Frequency detection sensitivity in the radio and microwave region.

The Penning trap confines electrons by using a strong vertical magnetic field B for radial confinement and a quadrupole electric field for axial confinement (see Fig. 6.2(a)). The magnetic field is produced by a solenoid while the electric field is produced by three electrodes: one ring and two endcaps. A sketch of the electron trajectory is shown in Fig. 6.2(b). The trajectory in the radial plane is characterized by two frequencies: The magneton frequency ω_- and the modified cyclotron frequency ω_+. The cyclotron frequency is then $\omega = \omega_+ + \omega_-$. Since there is also a low-frequency oscillation in the z-direction, the overall trajectory has the shown form.



(a) (b)

Figure 6.2: *Sketch of the fields and the electron trajectory in a Penning trap.* Confinement is achieved by a vertical magnetic field and a quadrupole electric field. Source: [17]. (a) The magneton frequency ω_- and the modified cyclotron frequency ω_+ contribute to the electron trajectory as well as a low-frequency oscillation in z-direction. (b)

A non-relativistic electron in a magnetic field has the following energy levels:[1]

$$E(n, m_s) = \frac{g}{2} h\nu_c m_s + \left(n + \frac{1}{2}\right) h\nu_c \tag{6.2}$$

depending on the cyclotron frequency

$$\nu_c = \frac{eB}{2\pi m} \tag{6.3}$$

and on the spin frequency

$$\nu_s = \frac{g}{2}\nu_c = \frac{g}{2}\frac{eB}{2\pi m}. \tag{6.4}$$

Here n is the principal quantum number and m_s the spin quantum number. Eq. (6.4) yields

$$\frac{g}{2} = \frac{\nu_s}{\nu_c} = 1 + \frac{\nu_s - \nu_c}{\nu_c} \equiv 1 + \frac{\nu_a}{\nu_c}.$$

Since ν_s and ν_c differ only by one part per 10^3, measuring ν_a and ν_c to a precision of one part per 10^{10} gives $g/2$ to one part per 10^{13}.

This technique of measuring $g/2$ has two main advantages:

[1]See e. g. [18, § 112].

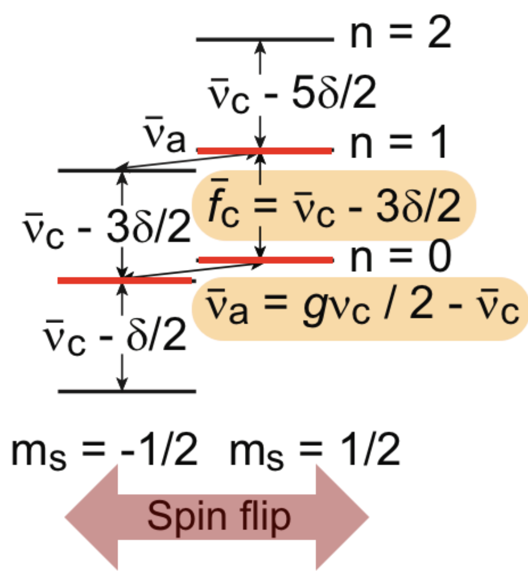


Figure 6.3: *Lowest cyclotron and spin levels of an electron in a Penning trap.* Source: [16, p. 180, modified].

1. One can measure the ratio of two frequencies to very high precision.
2. Since the B field appears in both numerator and denominator (see Eq. (6.4)), the dependence on the magnetic field cancels in the ratio.

Including the relativistic corrections, Eq. (6.2) is modified and the energy levels are given by:

$$E(n, m_s) = \frac{g}{2} h\nu_c m_s + \left(n + \frac{1}{2}\right) h\bar{\nu}_c - \underbrace{\frac{1}{2} h\delta \left(n + \frac{1}{2} + m_s\right)^2}_{\text{relativistic correction term}}$$

where $\bar{\nu}_c$ denotes the cyclotron frequency, shifted due to the Penning trap. Higher states are excited via microwave radiation. The experiment measures the following transition frequencies (see Fig. 6.3):

$$\bar{f}_c \equiv \bar{\nu}_c - \frac{3}{2}\delta, \text{ corresponding to } (n, m_s) = (1, 1/2) \to (0, 1/2) \text{ and}$$
$$\bar{\nu}_a \equiv \frac{g}{2}\nu_c - \bar{\nu}_c, \text{ corresponding to } (0, 1/2) \to (0, -1/2)$$

with the cyclotron frequency $\nu_c \sim 150\,\text{GHz}$.

A sketch of the experimental setup is shown in Fig. 6.4(a) and 6.4(b). A Penning trap is used to artificially bind the electron in an orbital state. For confinement, a high voltage

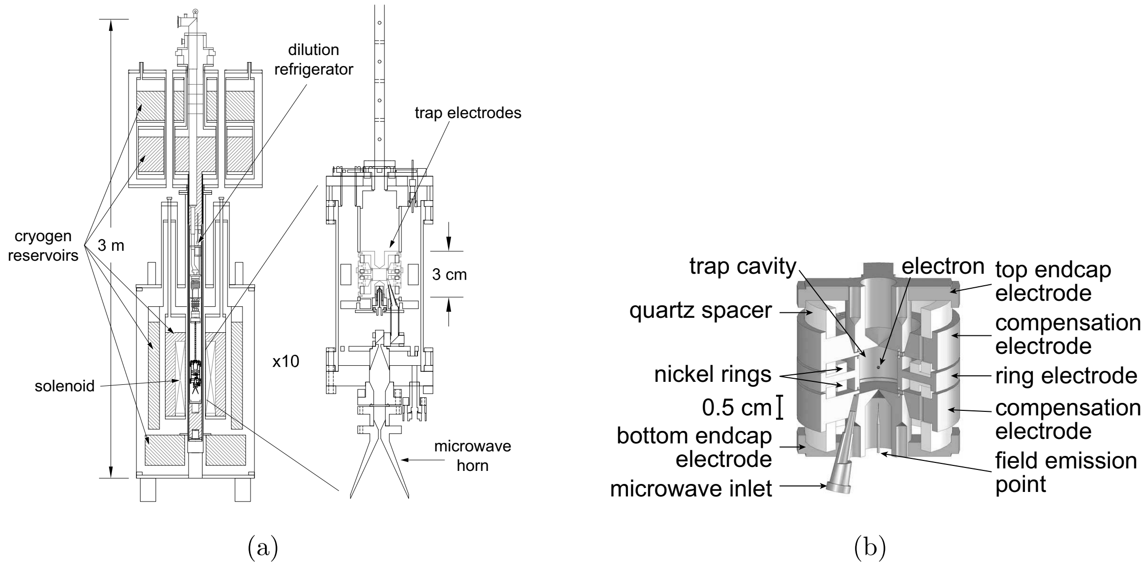


Figure 6.4: *Sketch of the experimental setup.* Overview of experimental apparatus. Source: [16, p. 185]. (a) The Penning trap cavity is used to confine a single electron and to inhibit spontaneous emission. Source: [16, p. 182]. (b)

(100 V) is applied between the cylindric and endcap contacts. Since $\nu_c \propto B$ (see Eq. (6.3)), a high magnetic field (5 T) is necessary to increase the spacing between the cyclotron energy levels. And finally, because the probability to occupy the orbital ground state is proportional to the Boltzmann factor,

$$\exp\left(-\frac{h\bar{\nu}_c}{k_\mathrm{B}T}\right),$$

very low temperatures (100 mK) are needed.

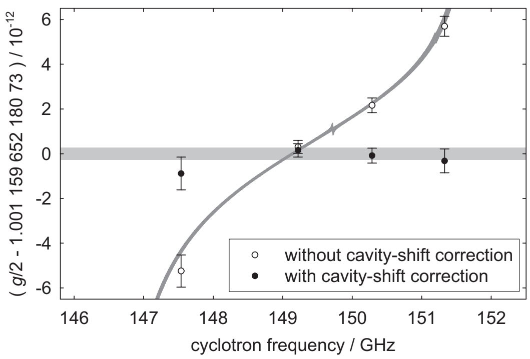
In analyzing the results of Penning trap measurements, one has to correct for the frequency shifts due to the cavity. This can be done by measuring at various frequencies (see Fig. 6.5(a)). The result for $g/2$ given in [16] is

$$g/2 = 1.001\ 159\ 652\ 180\ 73\ (28) \quad [0.28\ \text{ppt}]. \tag{6.5}$$

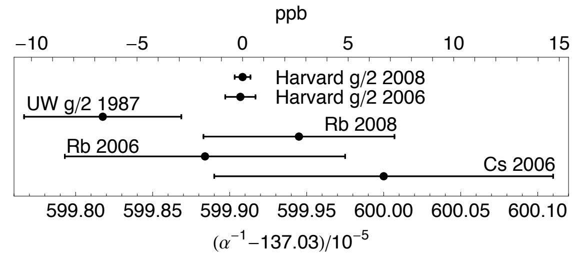
6.1.3.2 Theoretical predictions

The QED calculations provide the prediction for $g/2$ up to the fifth power of α:

$$\frac{g}{2} = 1 + C_2\left(\frac{\alpha}{\pi}\right) + C_4\left(\frac{\alpha}{\pi}\right)^2 + C_6\left(\frac{\alpha}{\pi}\right)^3 + C_8\left(\frac{\alpha}{\pi}\right)^4 + C_{10}\left(\frac{\alpha}{\pi}\right)^5 + \cdots + a_\text{hadronic} + a_\text{weak} \tag{6.6}$$



(a)



(b)

Figure 6.5: $g/2$ *and fine structure constant.* Four measurements of $g/2$ without (open) and with (filled) cavity-shift corrections. The light gray uncertainty band shows the average of the corrected data. The dark gray band indicates the expected location of the uncorrected data given the result in Eq. (6.5) and including only the cavity-shift uncertainty. Source: [16, p. 201]. (a) The most precise determinations of α. Source: [19, p. 264]. (b)

where

$$\begin{aligned}
C_2 &= 0.500\ 000\ 000\ 000\ 00\ \text{(exact)}\\
C_4 &= -0.328\ 478\ 444\ 002\ 90\ (60)\\
C_6 &= 1.181\ 234\ 016\ 827\ (19)\\
C_8 &= -1.914\ 4\ (35)\\
C_{10} &= 0.0\ (4.6)\\
a_{\text{hadronic}} &= 1.682(20) \cdot 10^{-12}.
\end{aligned}$$

From Eq. (6.6) and the theoretical predictions we can on the one hand measure the coupling constant α (see Fig. 6.5(b)):

$$\begin{aligned}
\alpha^{-1} &= 137.035\ 999\ 084\ (33)\ (39) \quad [0.24\ \text{ppb}][0.28\ \text{ppb}]\\
&= 137.035\ 999\ 084\ (51) \quad [0.37\ \text{ppb}]
\end{aligned}$$

and on the other hand, we can compare the measured $g/2$ with the expectation using α from other measurements

$$\begin{aligned}
g/2 &= 1.001\ 159\ 652\ 180\ 73\ (28) \quad [0.28\ \text{ppt}]\ \text{(measured)}\\
g(\alpha)/2 &= 1.001\ 159\ 652\ 177\ 60\ (520) \quad [5.2\ \text{ppt}]\ \text{(predicted)}.
\end{aligned}$$

6.2 High energy tests

6.2.1 e^+e^- colliders

In addition to the low-energy experiments, QED has been tested also in high energy e^+e^- collisions [20, 21, 22].

We discuss here the following reactions:

- Bhabha scattering : $e^+e^- \to e^+e^-$
- Lepton pair production : $e^+e^- \to \mu^+\mu^-, \tau^+\tau^-$
- Hadronic processes : $e^+e^- \to q\bar{q} \to$ hadrons

The energy range $12\,\text{GeV} \le \sqrt{s} \le 47\,\text{GeV}$ was investigated with the PETRA accelerator at DESY (Hamburg). High energy ranges ($90\,\text{GeV} \le \sqrt{s} \le 200\,\text{GeV}$) were covered by the LEP collider at CERN (Geneva). However, electroweak contributions to the cross-sections, like the one shown in Fig. 6.6, become considerable at these energies. Intermediate energies were covered by TRISTAN and SLC. Table 6.1 gives an overview of the e^+e^- colliders.

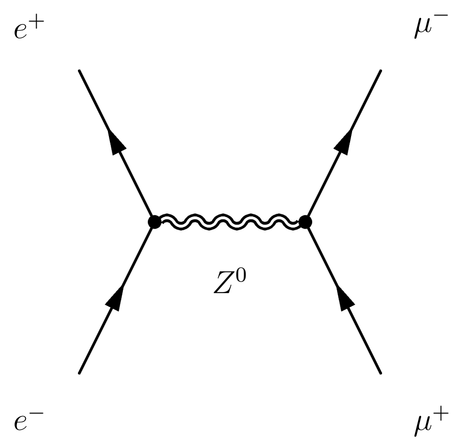


Figure 6.6: Electroweak contribution to $\mathcal{M}_{fi}(e^+e^- \to \mu^+\mu^-)$ at high energies.

The PETRA collider is shown in Fig. 6.7 as an example.

As an example for a typical detector we take a look at JADE (Figs. 6.8 and 6.9), using the same numbering as in the figure.

1. *Beam pipes counters.*
2. *End plug lead glass counters.*
3. *Pressure tank.*

Accelerator	Experiment(s)	$\sqrt{s}$ [GeV]	$\mathcal{L}_{\text{int}}$ [pb^{-1}]
SPEAR	SPEAR	2-8	-
PEP	ASP, DELCO, HRS, MARK II, MAC	0-29	300
PETRA	JADE, MARK J, PLUTO, TASSO, CELLO	12-47	20
TRISTAN	TRISTAN	50-60	20
SLC	MARK II, SLD	90	25
LEP	ALEPH, DELPHI, OPAL, L3	90-200	200 700

Table 6.1: Table of e^+e^- colliders

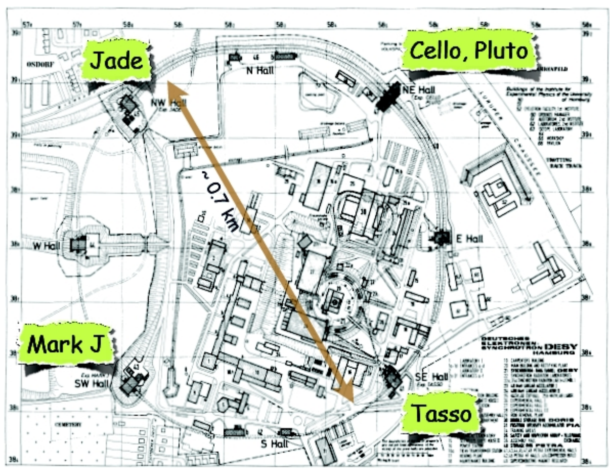


Figure 6.7: PETRA storage ring

4. *Muon chambers.* Detect muons.

5. *Jet chambers.* Records the trajectories of the produced particles.

6. *Time of flight counters.* Measure the time necessary for the particle to get from the collision center and thus its velocity.

7. *Coil.* Produces a magnetic field of 0.5 [T] parallel to the beam in the central region to measure the momentum of the particles by providing the curvature of their trajectories.

8. *Central lead glass counters.*

9. *Magnet yoke.*

10. *Muon filter.*

11. *Removable end plug.*

12. *Beam pipe.*

13. *Tagging counter.*

14. *Mini beta quadrupole.* Focus the beam to increase the luminosity of the beam in the experiment.

15. *Moving devices.*

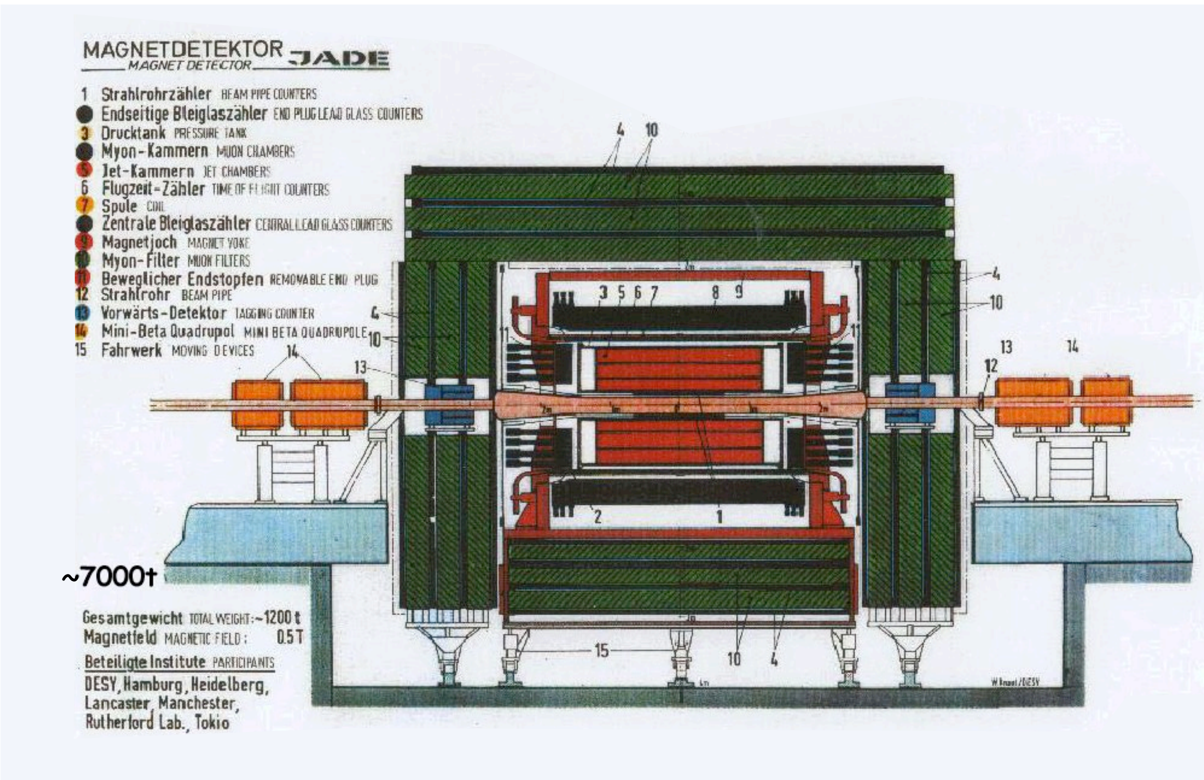


Figure 6.8: JADE detector : schematics

6.2.2 Detector elements

In order to help identify the particles produced in a collision (or their decay product) we can determine their charge and invariant mass using the methods presented in chapter 4. This measurement proceeds mostly in the inner part of the detector, see Fig. 4.15, by means of drift chambers or silicon trackers. If some of the produced particles are long living (i.e. are stable or decay weakly), this setup gives also the possibility to detect a decay vertex.

Further away from the beam axis are the **calorimeters**, whose function is to stop the particles and measure the energy they deposit. There are mostly two types of calorimeters:

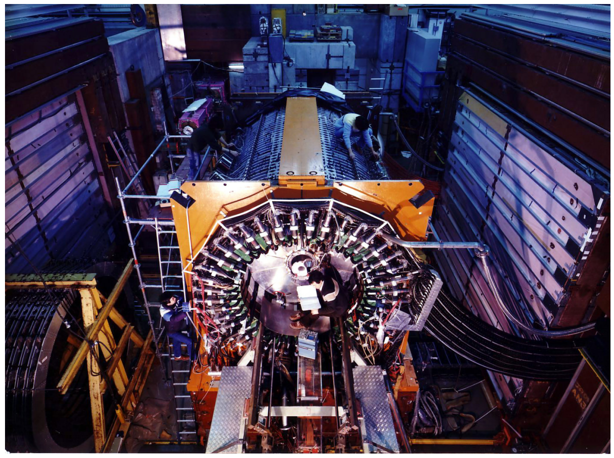


Figure 6.9: JADE detector

electromagnetic and hadronic. The angular resolution is limited by the size of each detector cell. Calorimeters are also able to measure neutral particles while the tracking devices described above can only detect charged particles.

Electromagnetic calorimeters stop and measure the energy of electrons, positrons and photons. All electromagnetically interacting particles leave at least a part of their energy in this detector part.

Hadronic calorimeters stop and measure the energy of hadrons, e.g. protons, neutrons and pions. Muons and antimuons are not stopped but leave some energy. Most modern experiments are also surrounded by **muon detectors** in order to distinguish the energy deposit of low energetic hadrons from the one of muons. Since it is practically impossible to stop muons, this last detector records the direction of passage of muons and, eventually, their momentum.

Fig. 6.10 shows the schematic view of the different signal hits for different types of particles. The energy deposit is usually depicted by a histogram.

- *Electron signature.* Eletrons leave a curved trace in the inner tracking detector and deposit all their energy in the electromagnetic calorimeter, where they are completely stopped. There is hence no signal stemming from electrons in detectors further away from the collision point.

- *Hadron signature.* Charged hadrons leave a trace in the inner detector (curved by the magnetic field), – whereas uncharged hadrons do not –, deposit a part of their energy in the electromagnetic calorimeter and the rest of their energy in the hadronic calorimeter.

- *Muon signature.* Muons leave a curved trace in the inner detector and deposit some energy in the electromagnetic and hadronic calorimeters whitout being stopped, and then leave a signal in the muon detector.

- *Photon signature.* Photons do not leave a trace in the inner detector and are stopped in the electromagnetic calorimeter.

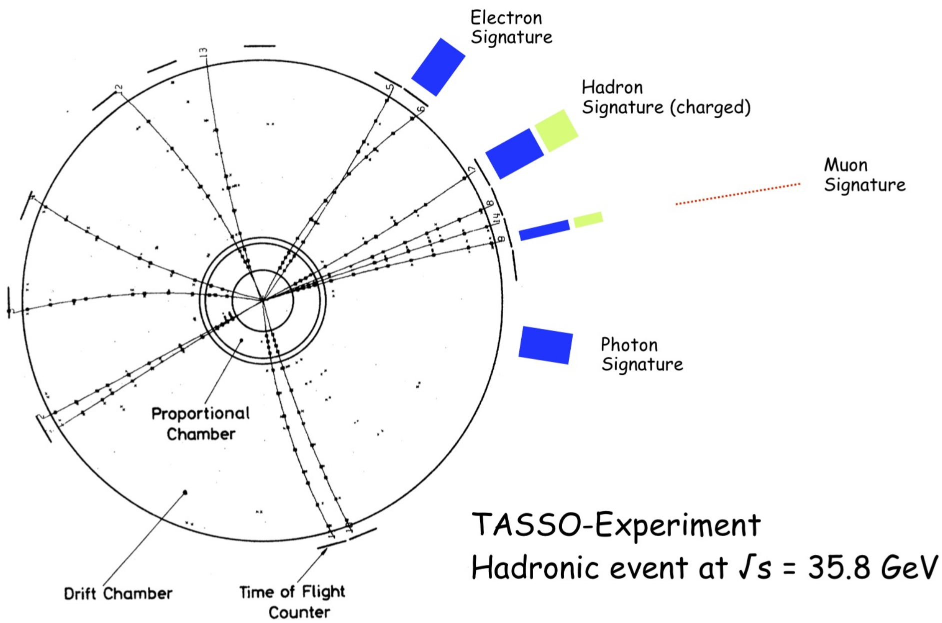


Figure 6.10: Event reconstruction principle

6.2.3 Cross section measurement

To measure a cross section we divide the measured number of events N by the integrated luminosity at that energy $\mathcal{L}(s)$,

$$\sigma(s) = \frac{N}{\mathcal{L}(s)}. \tag{6.7}$$

The last one is measured by counting the events occurring at small scattering angles and using the relation,

$$\sigma^{\text{theo}}_{ee,\gamma\gamma} = \frac{N(1-b)}{(\varepsilon A)\cdot\mathcal{L}}, \tag{6.8}$$

where A and b depend on the detector geometry, while ε is the efficiency (the probability to measure a particle, if it hits the detector).

Fig. 6.11 shows a typical integrated luminosity spectrum over the energy range $0-47\,\text{GeV}$.

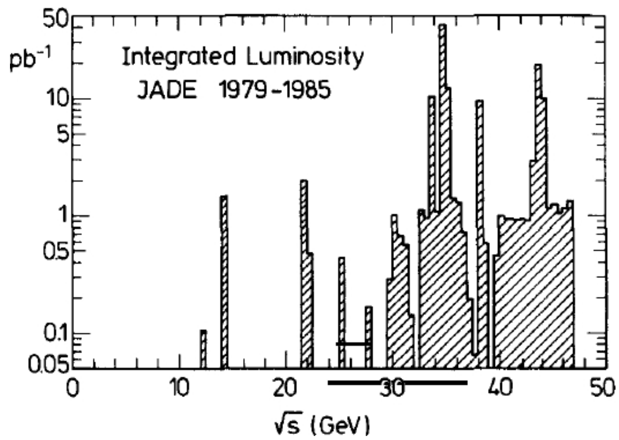


Figure 6.11: Integrated luminosity for the JADE experiment at PETRA

Reminder : e^+e^- kinematics One can write the differential cross section as,

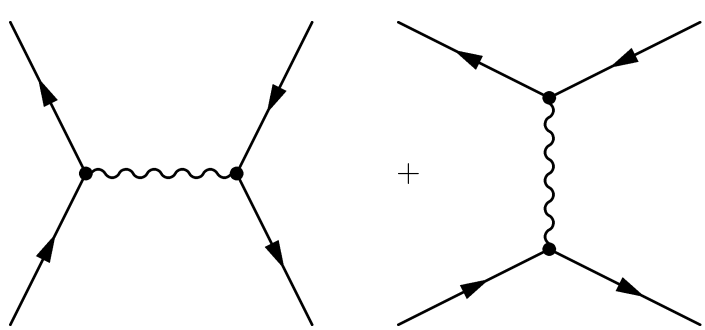
$$\frac{d\sigma_{\text{QED}}}{d\Omega} = \frac{d\sigma_0}{d\Omega}(1+\delta_{\text{rad}}), \tag{6.9}$$

where δ_{rad} stands for the radiative corrections, i.e. terms coming form diagrams with more vertices (proportional to α in the case of QED). These include emission of further low energy exchange bosons and loop corrections.

6.2.4 Bhabha scattering

Leading order We first treat the leading order term, the one yielding $d\sigma_0/d\Omega$.

The following two diagrams contribute to the invariant amplitude :



s-channel t-channel

Using Eq. (3.32) and the trace theorems of section 5.9, we get,

$$\begin{aligned}\frac{d\sigma_0}{d\Omega} &= \frac{\alpha^2}{4s}\left(\underbrace{\frac{t^2+s^2}{u^2}}_{t\text{-channel}} + \underbrace{\frac{2t^2}{us}}_{\text{interference}} + \underbrace{\frac{t^2+u^2}{s^2}}_{s\text{-channel}}\right) \\ &= \frac{\alpha^2}{4s}\left(\frac{3+\cos^2\vartheta}{1-\cos\vartheta}\right)^2 .\end{aligned} \tag{6.10}$$

Note that it is divergent for $\vartheta \to 0$. Fig. 6.12 shows the $\cos\vartheta$-dependence of each component in Eq. (6.10). We remark that the differential cross section is dominated by the t-channel component at all angles, and that the s-channel is almost constant, when compared to the last. The interference term is always negative. It is small in magnitude for large scattering angles ($\vartheta \sim \pi \Leftrightarrow \cos\vartheta \sim -1$) and diverges in the case of forward scattering ($\vartheta = 0 \Leftrightarrow \cos\vartheta = 1$).

Fig. 6.13 shows the typical trace left in the electronic calorimeter by a scattered e^+e^--pair. Fig. 6.14 shows $\sigma^{e^+e^-\to e^+e^-}$ measured as a function of $\cos\vartheta$ for different center of mass

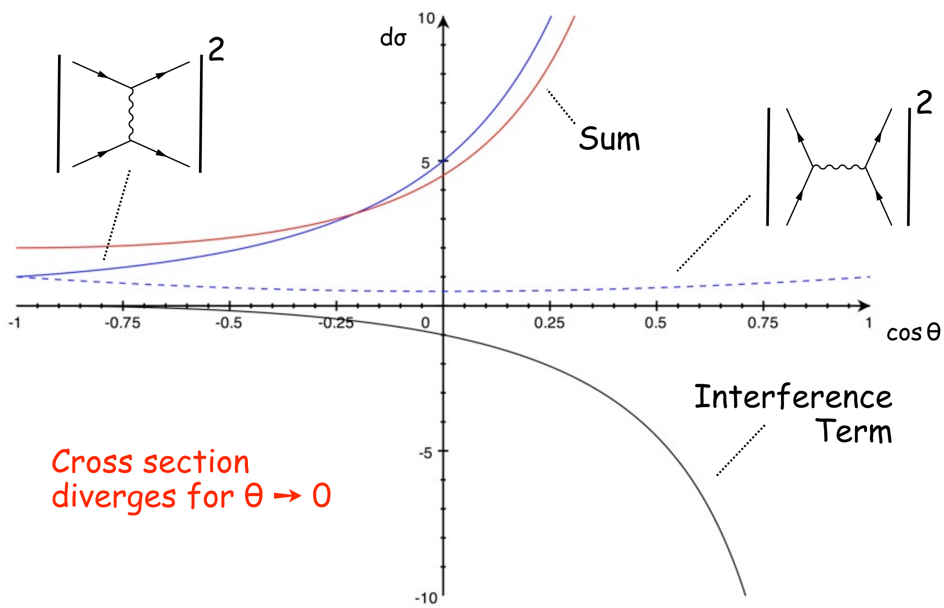


Figure 6.12: Relative magnitude of the different terms in $d\sigma_0/d\Omega$.

energies. It decreases following a $1/s$-dependence.

Radiative corrections The diagrams contributing to the cross section and proportional to higher powers of α (or e) are shown in Table 6.2.

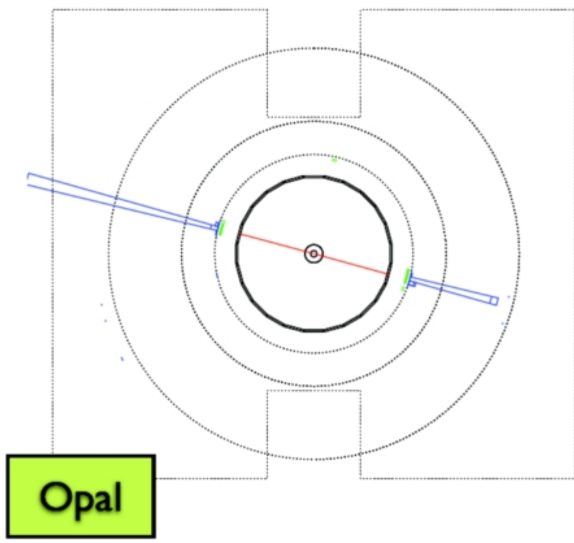


Figure 6.13: Typical event display of a Bhabha scattering event recorded by the Opal experiment. The length of the blue histogram corresponds to the amount of energy deposited in the electromagnetic calorimeter.

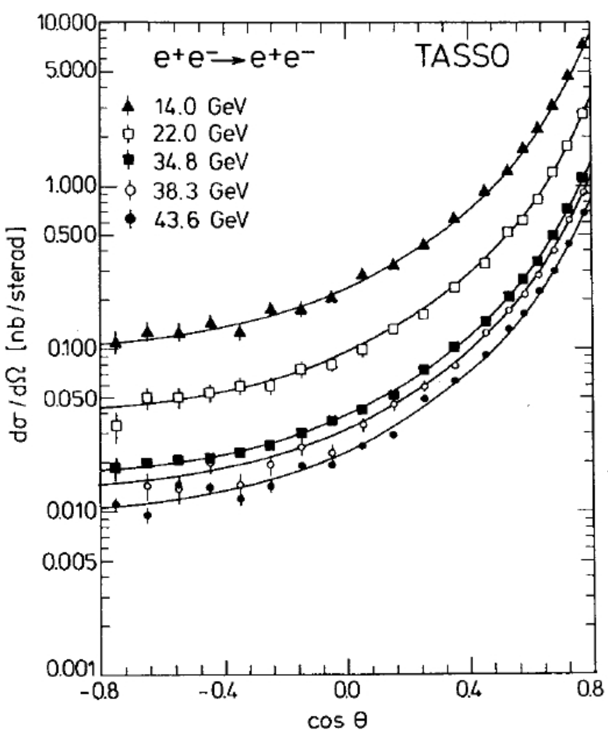


Figure 6.14: Energy and angle dependence of the cross section measured at TASSO and compared to leading order calculations.

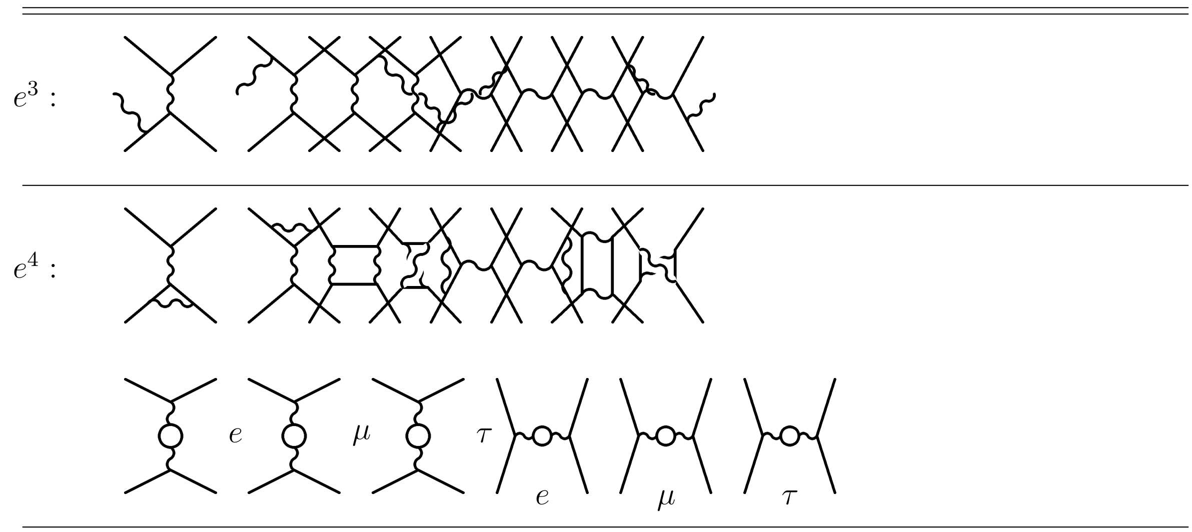


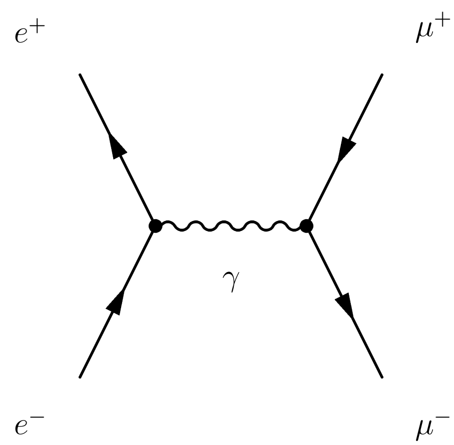
Table 6.2: Diagrams of radiative and loop corrections up to e^4

Because of momentum conservation, the diagrams of the e^3-order imply that the electron-positron pair is no longer back-to-back after the collision. This effect is called **acollinearity**. The acollinearity angle is the angle $\xi = \pi - \phi$, where ϕ is the angle between the direction of the scattered electron and the scattered positron; for a back-to-back flight there is no acollinearity, thus $\xi = 0$. This angle has been measured at the JADE experiment and confirms higher order QED corrections in a very impressive way (see Fig. 6.15).

6.2.5 Lepton pair production

Muon pair production Looking at different final states gives also different results. We illustrate this by looking at the process $e^+e^- \to \mu^+\mu^-$. This is the simplest process of QED and is often used to normalize cross sections of other processes.

There is only one leading order Feynman diagram, namely,



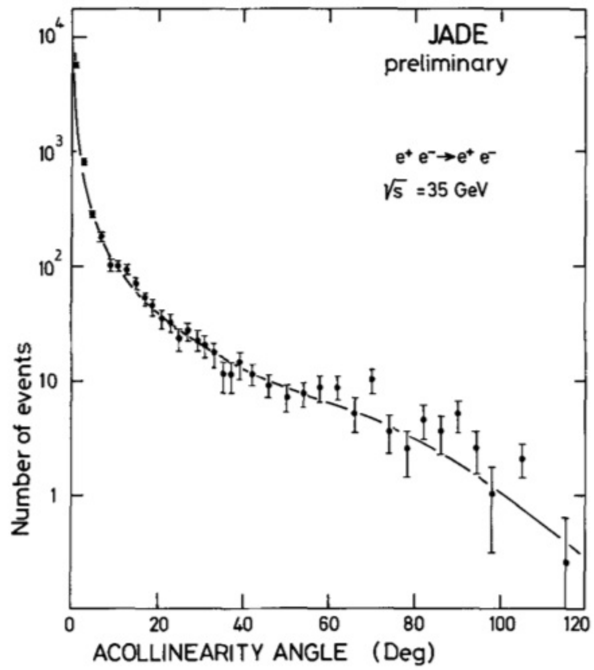


Figure 6.15: Comparison of measured acollinearity at JADE with the QED prediction.

and the leading order differential cross section is,

$$\frac{d\sigma_0}{d\Omega} = \frac{\alpha^2}{4s}\left(\frac{t^2+u^2}{s^2}\right) = \frac{\alpha^2}{4s}(1+\cos^2\vartheta), \tag{6.11}$$

which is shown in Fig. 5.9.

Fig. 6.16 shows an event candidate: low energy deposits in the electromagnetic calorimeter and hits in the muon chambers.

Muon pair production : Z^0 exchange Since only s-channel contributes to the muon pair production, the diagram containing a Z^0 boson instead of a photon [2],

[2] This contribution is also present in the case of Bhabha scattering, yet since the t-channel dominates over the s-channel, the effect is virtually invisible.

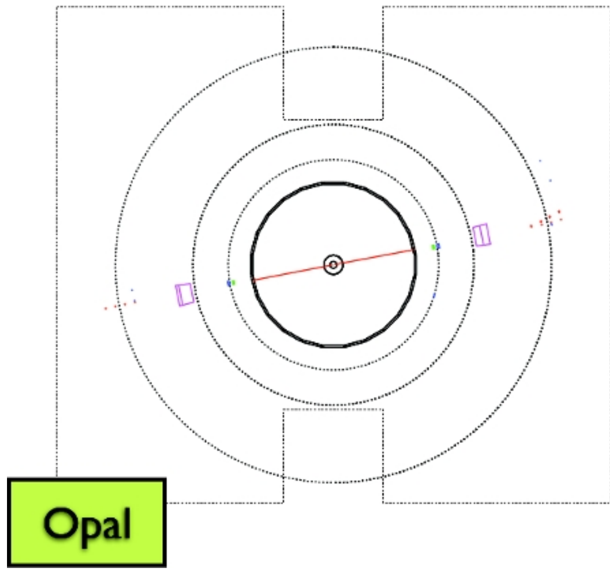
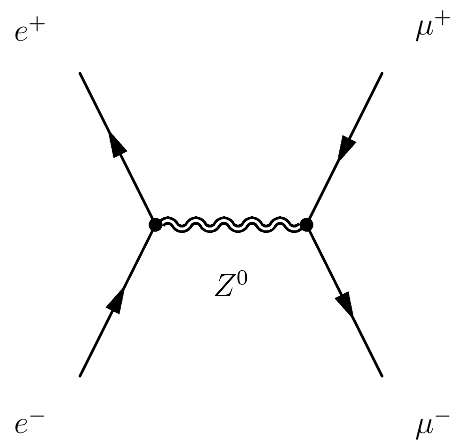


Figure 6.16: Typical event display of a muon pair production event recorded by the Opal experiment.



becomes comparable with the photon term (approx. 10%), even at leading order. This leads to a the modified cross section,

$$\frac{d\sigma_0^{\mathrm{EW}}}{d\Omega} = \frac{\alpha^2}{4s}(1 + \cos^2\vartheta + A\cos\vartheta). \tag{6.12}$$

This is illustrated in Fig. 6.17 comparing the QED and electroweak predictions to the data.

As an easy integration of Eq. (6.12) shows, the *total* cross section is not sensitive to the effects of electroweak interaction and we have a very good agreement with the QED value (Fig. 6.18).

For the $\sqrt{s}$-range measured at PETRA, electroweak corrections are small. In the case of LEP they are instead quite important, especally in the range around the Z^0 resonance,

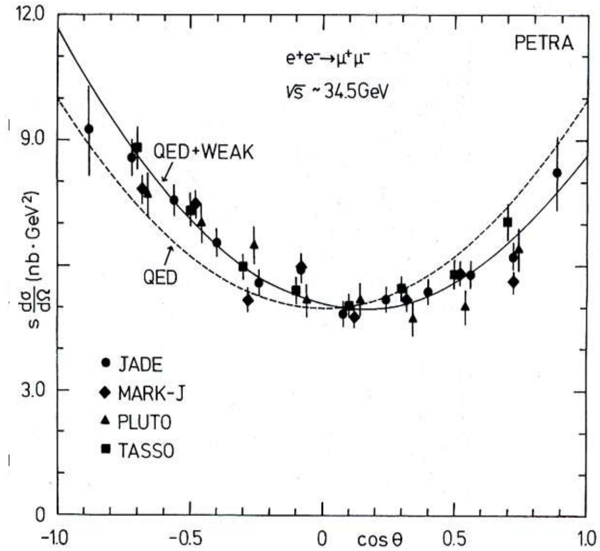


Figure 6.17: Comparison of $e^+e^- \to \mu^+\mu^-$ differential cross section measured at PETRA with the QED and electroweak predictions.

$\sqrt{s} \sim m_Z \approx 90\,[\text{GeV}]$.

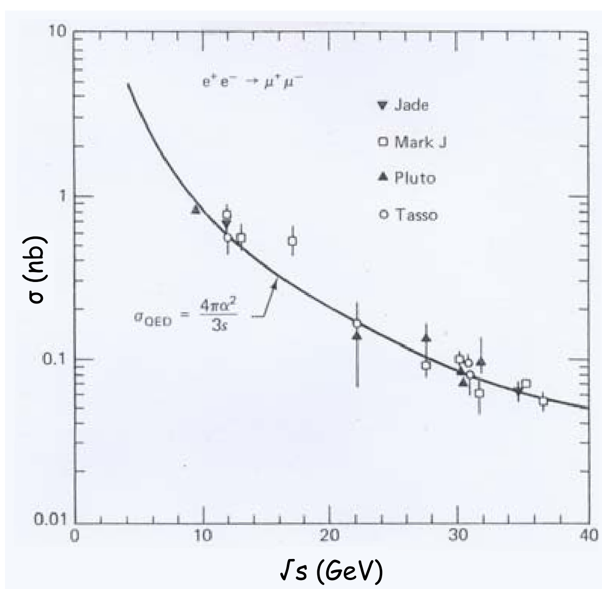
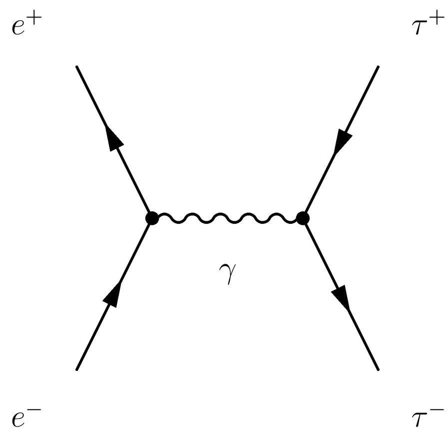


Figure 6.18: Comparison of measured total cross section at PETRA with the QED prediction.

Tau pair production At high enough energy ($\sqrt{s} \geq 2m_\tau \approx 3.6\,[\text{GeV}]$) the production of $\tau^+\tau^-$-pair – which is very similar to the case of muon pair production – is possible:



The final state of a tau pair production event observed in the detector can contain hadrons, since the lifetime of τ is very short ($\tau_\tau = 2.9 \cdot 10^{-13}$ [s]) and it is the only lepton with sufficiently high mass to produce $q\bar{q}$-pairs.

Fig. 6.19 shows an event where one of the two tau survived long enough, e.g. because of a large energy and thus a longer lifetime $\gamma\tau_\tau$ in the laboratory frame, to hit the electromagnetic calorimeter, while the other one decayed in three pions which then left traces in the electromagnetic and hadronic calorimeters.

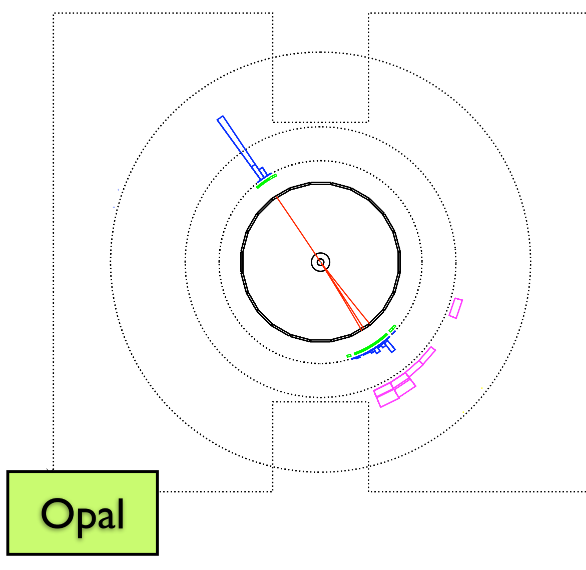


Figure 6.19: Typical event display of a tau pair production event recorded by the Opal experiment.

6.2.6 Hadronic processes

The production of quark-antiquark $q\bar{q}$ pair is another possible final state in e^+e^- annihilation.

When a bound $q\bar{q}$ state is produced, we speak of a **resonance** because the e^+e^- cross section looks like the amplification curve of a periodic system such as a pendulum or an RLC circuit near the resonant frequency. A famous resonance is the J/ψ resonance corresponding to a bound state of $c\bar{c}$.

Away from the resonances, there is in general no visible bound state, and the produced quarks hadronize in **jets** due to the confinement of the strong interaction : quarks cannot be seen as free particles.

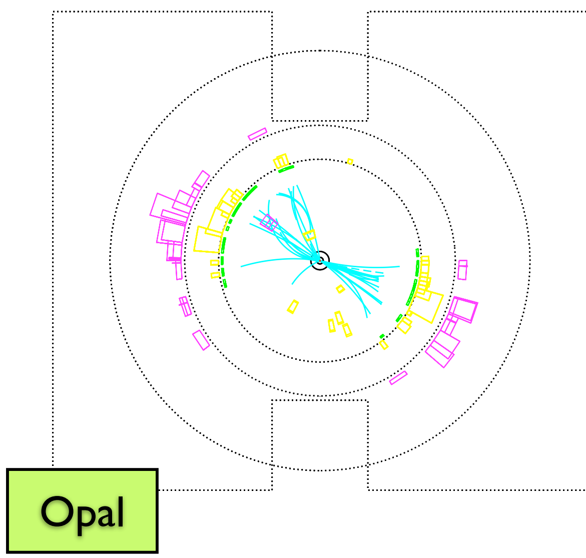
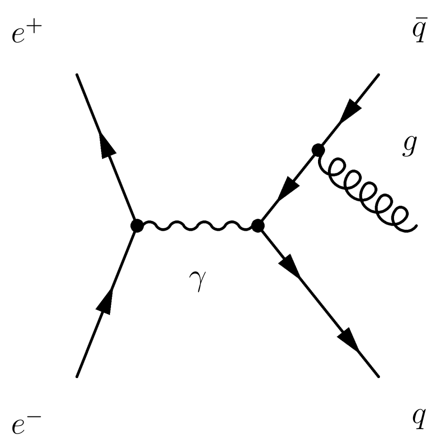


Figure 6.20: Typical display of jet production event recorded by the Opal experiment.

Due to the strength of strong interaction at low energy, the radiative effects (this time the radiated bosons are gluons),



take a much more dramatic form than in QED : Since gluons also have a color charge, they hadronize and for each emitted gluon one observes one more jet (Fig. 6.21).

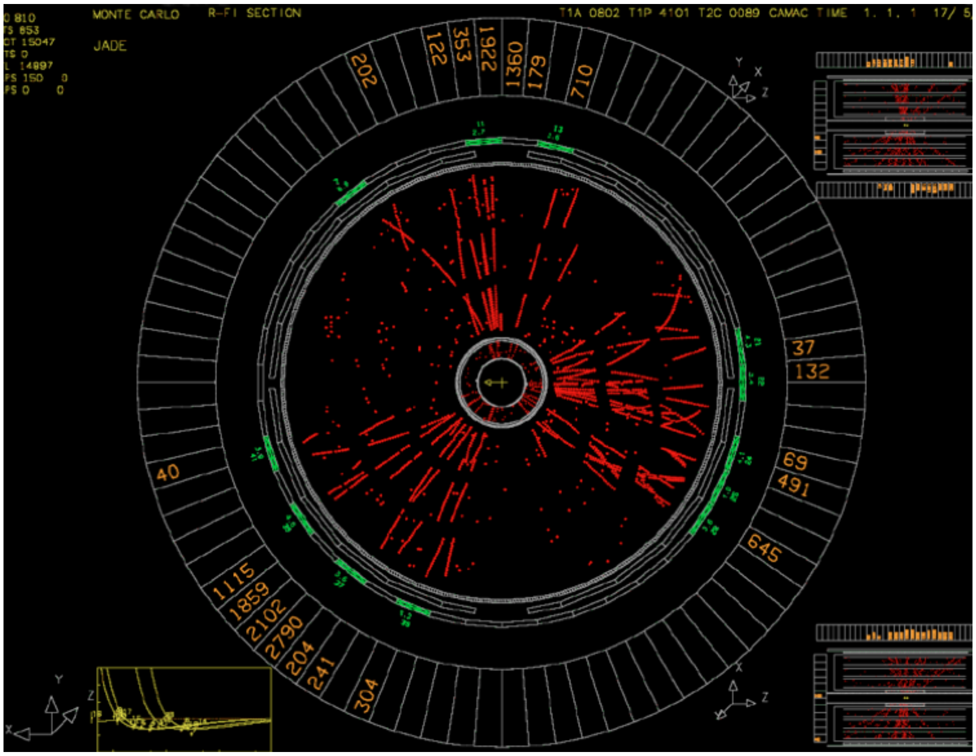


Figure 6.21: Typical event display of a 3-jets production.

6.2.7 Limits of QED

In this section, one addresses the question : what do we expect if QED is not the only theoretical model involved in the scattering processes discussed so far?

Suppose there is an energy scale Λ (equivalent to a length scale Λ^{-1}) at which QED does not describe the data anymore.

We would have changes of the various quantities, for instance, the potential, photon propagator and total cross section would be modified as follows :

$$\frac{1}{r} \to \frac{1}{r}\left(1 - \mathrm{e}^{-\Lambda r}\right) \qquad \text{(potential)}$$

$$-\frac{1}{q^2} \to -\frac{1}{q^2}\left(1 + \frac{q^2}{\Lambda^2}\right) \qquad \text{(propagator)}$$

$$\sigma^{e^+e^- \to \mu^+\mu^-} \to \frac{4\pi\alpha^2}{3s}\left(1 \pm \frac{s}{\Lambda^2 - s}\right)^2 \qquad \text{(cross section).}$$

The form of the potential is typical of a Yukawa coupling of a fermion with a massive spin 0 field. Since this particle is imagined as heavy – the energy available is smaller or similar to the production threshold Λ – we can treat thie particle as spinless since spin effect are only significant in the relativistic case. This type of ansatz is thus standard in the sense that any new heavy particle that can be produced from an e^+e^--annihilation will have the same effect on the potential, regardless of it being a scalar or a vector particle. The other quantities are then directly related to the change in the potential.

We have seen the electroweak effects to the QED cross section at the end of the previous subsection. This corresponds to $\Lambda \approx m_{Z^0}$ (Fig. 6.22).

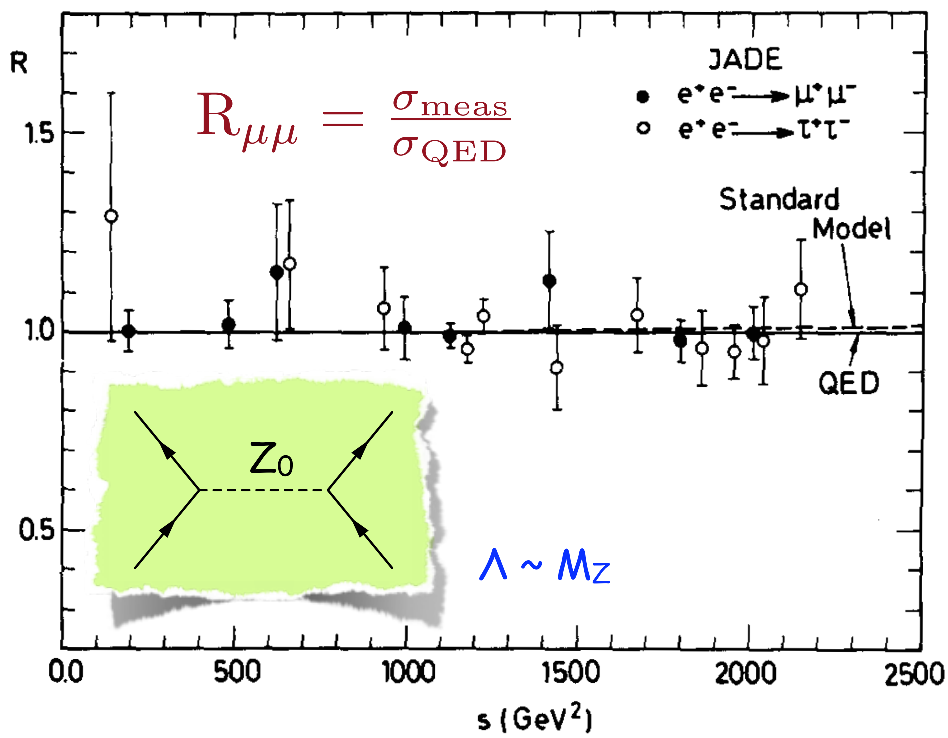


Figure 6.22: Comparison of measured total cross section at PETRA with the QED prediction for muon and tau pair production.

Fig. 6.23 shows the ratio,

$$R_{\mu\mu} = \frac{\sigma_{\rm meas}^{e^+e^-\to\mu^+\mu^-}}{\sigma_{\rm QED}^{e^+e^-\to\mu^+\mu^-}},$$

as measured at PETRA and TRISTAN. By comparing data and theory and varying Λ within the experimental error one can infer that – if any – new physics can only be brought in with a mass scale $\Lambda \geq 200$ [GeV].

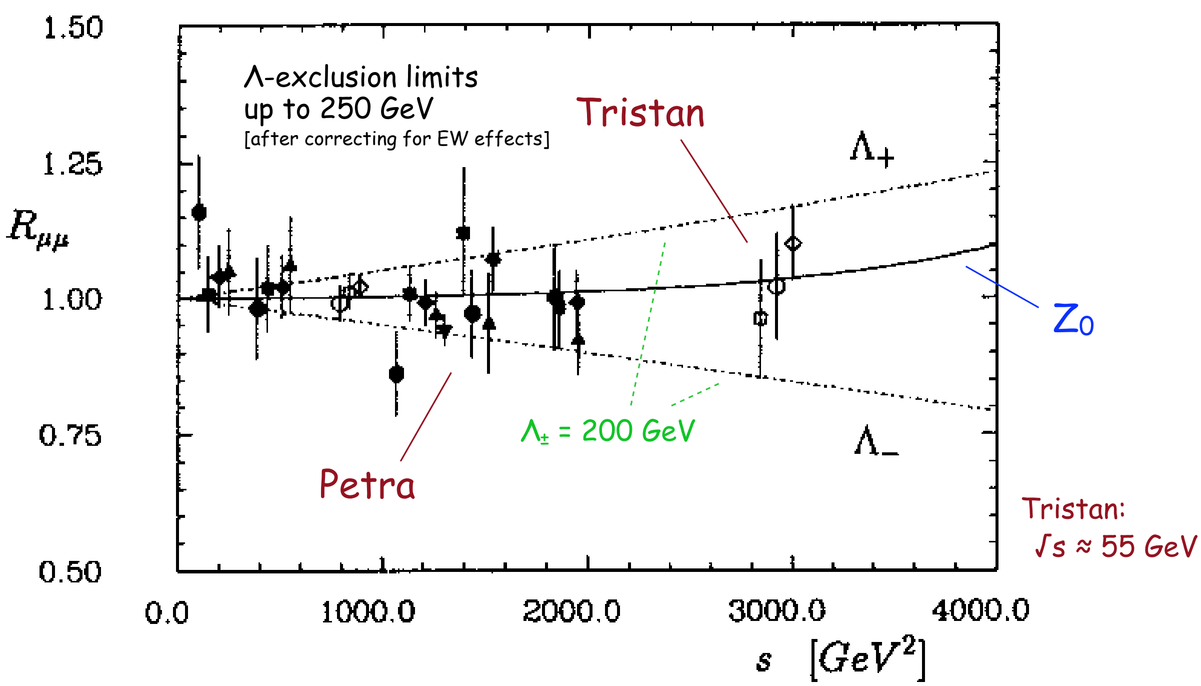


Figure 6.23: Comparison of measured total cross section at PETRA and TRISTAN with the QED prediction for muon pair production.

Chapter 7

Unitary symmetries and QCD as a gauge theory

Literature:

- Lipkin [23] (group theory concepts from a physicist's point of view)
- Lee [24], chapter 20 (extensive treatment of Lie groups and Lie algebras in the context of differential geometry)

Interactions between particles should respect some observed symmetry. Often, the procedure of postulating a specific symmetry leads to a unique theory. This way of approach is the one of **gauge theories**. The usual example of a gauge theory is QED, which corresponds to a local $U(1)$-symmetry of the Lagrangian :

$$\psi \to \psi' = \mathrm{e}^{ieq_e\chi(x)}\psi, \tag{7.1}$$

$$A_\mu \to A'_\mu = A_\mu - \partial_\mu\chi(x). \tag{7.2}$$

We can code this complicated transformation behavior by replacing in the QED Lagrangian ∂_μ by the **covariant derivative** $D_\mu = \partial_\mu + ieq_e A_\mu$.

7.1 Isospin $SU(2)$

For this section we consider only the strong interaction and ignore the electromagnetic and weak interactions. In this regard, isobaric nuclei (with the same mass number A) are very similar. Heisenberg proposed to interpret protons and neutrons as two states of the same object : the **nucleon**:

$$|p\rangle = \psi(x)\begin{pmatrix}1\\0\end{pmatrix},$$

$$|n\rangle = \psi(x)\begin{pmatrix}0\\1\end{pmatrix}.$$

We note the analogy to the spin formalism of nonrelativistic quantum mechanics, which originated the name *isospin*.

In isospin-space, $|p\rangle$ and $|n\rangle$ can be represented as a two-component spinor with $I = \frac{1}{2}$. $|p\rangle$ has then $I_3 = +\frac{1}{2}$ and $|n\rangle$ has $I_3 = -\frac{1}{2}$.

Since the strong interaction is blind to other charges (electromagnetic charge, weak hypercharge), the (strong) physics must be the same for any linear combinations of $|p\rangle$ and $|n\rangle$. In other words, for,

$$\begin{aligned}|p\rangle \to |p'\rangle &= \alpha\,|p\rangle + \beta\,|n\rangle\,,\\ |n\rangle \to |n'\rangle &= \gamma\,|p\rangle + \delta\,|n\rangle\,,\end{aligned}$$

for some $\alpha, \beta, \gamma, \delta \in \mathbb{C}$, or,

$$|N\rangle = \begin{pmatrix}\psi_p\\ \psi_n\end{pmatrix} \to |N'\rangle = U\,|N\rangle\,, \tag{7.3}$$

for some $2{\times}2$ matrix U with complex entries, the (strong) physics does not change if we switch from $|N\rangle$ to $|N'\rangle$ to describe the system.

We remark at this point that this symmetry is only an *approximate* symmetry since it is violated by the other interactions, and is hence *not* a symmetry of nature.

First we require the conservation of the norm $\langle N|N\rangle$ which we interpret as the number of particles like in quantum mechanics. This yields,

$$\begin{aligned}\langle N|N\rangle \to \langle N'|N'\rangle = \langle N|\,U^\dagger U|N\rangle &\stackrel{!}{=} \langle N|N\rangle\\ \Rightarrow U^\dagger U = UU^\dagger = \mathbb{1} &\Rightarrow U \in U(2).\end{aligned} \tag{7.4}$$

A general unitary matrix has 4 real parameters. Since the effect of U and $\mathrm{e}^{i\varphi}U$ are the same, we fix one more parameter by imposing,

$$\det U \stackrel{!}{=} 1 \Rightarrow U \in SU(2), \tag{7.5}$$

the special unitary group in 2 dimensions. This group is a **Lie group** (a group which is at the same time a manifold). We use the representation,

$$U = \mathrm{e}^{i\alpha_j \hat{I}_j}, \tag{7.6}$$

where the α_j's are arbitrary group parameters (constant, or depending on the spacetime coordinate x), and the $\hat{I}_j$'s are the generators of the Lie group.

We concentrate on infinitesimal transformations, for which $\alpha_j \ll 1$. In this approximation we can write

$$U \approx \mathbb{1} + i\alpha_j \hat{I}_j. \tag{7.7}$$

The two defining conditions of $SU(2)$, Eq. (7.4) and (7.5), imply then for the generators,

$$\hat{I}_j^\dagger = \hat{I}_j \qquad (hermitian), \tag{7.8}$$

$$\mathrm{Tr}\,\hat{I}_j = 0 \qquad (traceless). \tag{7.9}$$

In order for the exponentiation procedure to converge for noninfinitesimal α_j's, the generators must satisfy a comutation relation, thus defining the **Lie algebra** $su(2)$ of the group $SU(2)$.

Quite in general, the commutator of two generators must be expressible as a linear combination of the other generators [1]. In the case of $su(2)$ we have,

$$[\hat{I}_i, \hat{I}_j] = i\varepsilon_{ijk}\hat{I}_k, \tag{7.10}$$

where ε_{ijk} is the totally antisymmetric tensor with $\varepsilon_{123} = +1$. They are characteristic of the (universal covering group of the) Lie group (but independent of the chosen representation) and called structure constants of the Lie group.

The representations can be characterized according to their total isospin. Consider now $I = 1/2$, where the generators are given by

$$\hat{I}_i = \frac{1}{2}\tau_i$$

with $\tau_i = \sigma_i$ the Pauli spin matrices (this notation is chosen to prevent confusion with ordinary spin):

$$\tau_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad \tau_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \qquad \tau_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

which fulfill $[\sigma_i, \sigma_j] = 2i\varepsilon_{ijk}\sigma_k$. The action of the matrices of the representation (see Eq. (7.6)) is a non-abelian phase transformation:

$$|N'\rangle = e^{i\vec{\alpha}\cdot\frac{\vec{\tau}}{2}}\,|N\rangle\,.$$

For $SU(2)$, there exists only one diagonal matrix (τ_3). In general, for $SU(N)$, the following holds true:

- *Rank* $r = N - 1$: There are r simultaneously diagonal operators.
- *Dimension* of the Lie algebra $o = N^2 - 1$: There are o generators of the group and therefore o group parameters. E. g. in the case of $SU(2)/\{\pm\mathbb{1}\} \cong SO(3)$ this means that there are three rotations/generators and three angles as parameters.

[1] Since we are working in a matrix representation of $SU(2)$ this statement makes sense. The difference between the abstract group and its matrix representation is often neglected.

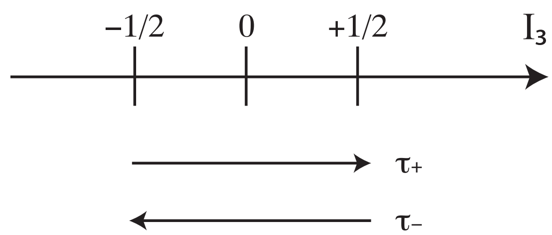


Figure 7.1: *The nucleons* $|n\rangle$ *and* $|p\rangle$ *form an isospin doublet.*

Isospin particle multiplets (representations) can be characterized by their quantum numbers I and I_3: There are $2I+1$ states. Consider for example once again the case $I = 1/2$. There are two states, characterized by their I_3 quantum number:

$$\begin{pmatrix} \left|I = \frac{1}{2}, I_3 = +\frac{1}{2}\right\rangle \\ \left|I = \frac{1}{2}, I_3 = -\frac{1}{2}\right\rangle \end{pmatrix} = \begin{pmatrix} |p\rangle \\ |n\rangle \end{pmatrix}.$$

This is visualized in Fig. 7.1, along with the action of the operators $\tau_\pm = 1/2(\tau_1 \pm i\tau_2)$:

$$\begin{aligned} \tau_- |p\rangle &= \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} = |n\rangle \\ \tau_+ |n\rangle &= |p\rangle \\ \tau_- |n\rangle &= \tau_+ |p\rangle = 0. \end{aligned}$$

This is the smallest non-trivial representation of $SU(2)$ and therefore its fundamental representation.

Further examples for isospin multiplets are

I	multiplets	I_3
$\frac{1}{2}$	$\begin{pmatrix} p \\ n \end{pmatrix}$ $\begin{pmatrix} K^+ \\ K^0 \end{pmatrix}$ $\begin{pmatrix} {}^3_2\mathrm{He} \\ {}^3_1\mathrm{H} \end{pmatrix}$	$+\frac{1}{2}$, $-\frac{1}{2}$
1	$\begin{pmatrix} \pi^+ \\ \pi^0 \\ \pi^- \end{pmatrix}$	$+1$, 0, -1
$\frac{3}{2}$	$\begin{pmatrix} \Delta^{++} \\ \Delta^+ \\ \Delta^0 \\ \Delta^- \end{pmatrix}$	$+\frac{3}{2}$, $+\frac{1}{2}$, $-\frac{1}{2}$, $-\frac{3}{2}$

where $m_\Delta \approx 1232\,\mathrm{MeV}$ and $m_{p,n} \approx 938\,\mathrm{MeV}$.

All $I \geqslant 1$ representations can be obtained from direct products out of the fundamental $I = 1/2$ representation $\underline{2}$ where "2" denotes the number of states. In analogy to the addition of two electron spins where the Clebsch-Gordan decomposition reads rep.$_{1/2}$ $\otimes$

rep.$_{1/2}$ = rep.$_0 \oplus$ rep.$_1$ and where there are two states for the spin-1/2 representation, one state for the spin-0 representation, and three states for the spin-1 representation, we have

$$\underbrace{\underline{2} \otimes \underline{2}}_{I=|\frac{1}{2}\pm\frac{1}{2}|=0,1} = \underbrace{\underline{1}}_{\text{isosinglet, } I=0} \oplus \underbrace{\underline{3}}_{\text{isotriplet, } I=1} . \tag{7.11}$$

However, there is an important difference between isospin and spin multiplets. In the latter case, we are considering a bound system and the constituents carrying the spin have the same mass. On the other hand, pions are not simple bound states. Their structure will be described by the quark model.

7.1.1 Isospin invariant interactions

Isospin invariant interactions can be constructed by choosing $SU(2)$ invariant interaction terms $\mathcal{L}'$. For instance, consider the Yukawa model, describing nucleon-pion coupling, where

$$\mathcal{L}'_{\pi N} = ig\bar{N}\vec{\tau}N \cdot \vec{\pi} = ig\bar{N}'\vec{\tau}N' \cdot \vec{\pi}' \tag{7.12}$$

which is an isovector and where the second identity is due to $SU(2)$ invariance. Infinitesimally, the transformation looks as follows:

$$N' = UN \qquad U = \mathbb{1} + \frac{i}{2}\vec{\alpha}\cdot\vec{\tau} \tag{7.13}$$

$$\bar{N}' = \bar{N}U^\dagger \qquad U^\dagger = \mathbb{1} - \frac{i}{2}\vec{\alpha}\cdot\vec{\tau} = U^{-1} \tag{7.14}$$

$$\vec{\pi}' = V\vec{\pi} \qquad V = \mathbb{1} + i\vec{\alpha}\cdot\vec{t}\,. \tag{7.15}$$

The parameters $\vec{t}$ can be determined from the isospin invariance condition in Eq. (7.12):

$$\bar{N}\tau_j N\pi_j = \bar{N}U^{-1}\tau_i UNV_{ij}\pi_j .$$

With $V_{ij} = \delta_{ij} + i\alpha_k(t_k)_{ij}$ (cp. Eq. (7.15)) and inserting the expressions for U and $U^\dagger$, this yields

$$\begin{aligned}
\tau_j &= \underbrace{\left(\mathbb{1} - \frac{i}{2}\alpha_k\tau_k\right)\tau_i\left(\mathbb{1} + \frac{i}{2}\alpha_k\tau_k\right)}_{\substack{=\tau_i + \frac{i}{2}\alpha_k[\tau_i,\tau_k] + \mathcal{O}(\alpha_k^2) \\ =\tau_i + \frac{i}{2}\alpha_k 2i\varepsilon_{ikl}\tau_l + \mathcal{O}(\alpha_k^2)}}\left(\delta_{ij} + i\alpha_k(t_k)_{ij}\right) \\
&= \tau_j + i\alpha_k\left\{i\varepsilon_{jkl}\tau_l + \tau_i(t_k)_{ij}\right\} \\
&= \tau_j + i\alpha_k\tau_i\underbrace{\left\{i\varepsilon_{jki} + (t_k)_{ij}\right\}}_{\overset{!}{=}0} \\
&\Rightarrow (t_k)_{ij} = -i\varepsilon_{kij}.
\end{aligned}$$

This means that the 3×3 matrices t_k, $k = 1, 2, 3$, are given by the structure constants (see Eq. (7.10)). For the commutator we therefore have

$$[t_k, t_l]_{ij} = -\varepsilon_{kim}\varepsilon_{lmj} + \varepsilon_{lim}\varepsilon_{kmj} = \varepsilon_{klm}\varepsilon_{mij} = i\varepsilon_{klm}(-i\varepsilon_{mij}) = i\varepsilon_{klm}(t_m)_{ij} \tag{7.16}$$

where the second identity follows using the Jacobi identity. This means that the matrices t_k fulfill the Lie algebra

$$[t_k, t_l] = i\varepsilon_{klm}t_m.$$

The t_ks form the adjoint representation of $SU(2)$.

7.2 Quark model of hadrons

It is experimentally well established that the proton and the neutron have inner structure. The evidence is:

- Finite electromagnetic charge radius

  $$\langle r_{p,n}\rangle = 0.8 \cdot 10^{-15}\,\mathrm{m}$$

  (The neutron is to be thought of as a neutral cloud of electromagnetically interacting constituents.)

- Anomalous magnetic moment

  $$\vec{\mu} = g\frac{q}{2m}\vec{s} \qquad g_p = 5.59 \qquad g_n = -3.83$$

- Proliferation of strongly interacting hadronic states (particle zoo)

  $$p,\ n,\ \Lambda,\ \Delta^-,\ \Xi,\ \Sigma,\ \Omega,\ \ldots$$

The explanation for these phenomena is that protons and neutrons (and the other hadrons) are bound states of quarks:

$$\left.\begin{aligned} |p\rangle &= |uud\rangle \\ |n\rangle &= |udd\rangle \end{aligned}\right\} \text{ 3 quark states.}$$

The up quark and the down quark have the following properties

$$\begin{aligned} |u\rangle &: \ q = +\frac{2}{3},\ I = \frac{1}{2},\ I_3 = +\frac{1}{2},\ S = \frac{1}{2}; \\ |d\rangle &: \ q = -\frac{1}{3},\ I = \frac{1}{2},\ I_3 = -\frac{1}{2},\ S = \frac{1}{2}. \end{aligned}$$

Quarks			Charge	Baryon number
Up 1.5 − 3 Mev	Charm 1270 MeV	Top 171 000 MeV	+2/3 e	1/3
Down 3.5 − 6 MeV	Strange 105 MeV	Bottom 4200 MeV	−1/3 e	1/3
Leptons			Charge	Lepton number
e^-	μ^-	τ^-	− e	1
ν_e	ν_μ	ν_τ	0	1

Table 7.1: Quarks and leptons.

Thus, $|u\rangle$ and $|d\rangle$ form an isospin doublet and combining them yields the correct quantum numbers for $|p\rangle$ and $|n\rangle$. There are also quark-antiquark bound states: The pions form an isospin triplet while the $|\eta\rangle$ is the corresponding singlet state (see Eq. (7.11)):

$$\left.\begin{array}{l} |\pi^+\rangle = |u\bar{d}\rangle \\ |\pi^0\rangle = \frac{1}{\sqrt{2}}\left(|u\bar{u}\rangle - |d\bar{d}\rangle\right) \\ |\pi^-\rangle = |d\bar{u}\rangle \end{array}\right\} \text{triplet states, } I = 1$$

$$\left.|\eta\rangle = \frac{1}{\sqrt{2}}\left(|u\bar{u}\rangle + |d\bar{d}\rangle\right)\right\} \text{singlet state, } I = 0.$$

There are in total three known quark doublets:

$$\underbrace{\begin{pmatrix}|u\rangle\\|d\rangle\end{pmatrix}}_{\text{up/down}} \qquad \underbrace{\begin{pmatrix}|c\rangle\\|s\rangle\end{pmatrix}}_{\text{charm/strange}} \qquad \underbrace{\begin{pmatrix}|t\rangle\\|b\rangle\end{pmatrix}}_{\text{top/bottom}} \qquad \begin{pmatrix} q = +\frac{2}{3},\ I_3 = +\frac{1}{2} \\ q = -\frac{1}{3},\ I_3 = -\frac{1}{2}\end{pmatrix}.$$

These quarks can be combined to give states like, e. g., $|\Lambda\rangle = |uds\rangle$.

7.3 Hadron spectroscopy

7.3.1 Quarks and leptons

Experimental evidence shows that, in addition to the three quark isospin doublets, there are also three families of leptons, the second type of elementary fermions (see Tab. 7.1). The lepton families are built out of an electron (or μ or τ) and the corresponding neutrino. The summary also shows the large mass differences between the six known quarks. All of the listed particles have a corresponding antiparticle, carrying opposite charge and baryon or lepton number, respectively.

Stable matter is built out of quarks and leptons listed in the first column of the family table. Until now, there is no evidence for quark substructure and they are therefore considered to be elementary. Hadrons, on the other hand, are composite particles. They are divided in two main categories as shown in the following table:

Quarks	Flavor	Other numbers
Up, Down	—	$S = C = B = T = 0$
Charm	$C = +1$	$S = B = T = 0$
Strange	$S = -1$	$C = B = T = 0$
Top	$T = +1$	$S = C = B = 0$
Bottom	$B = -1$	$S = C = T = 0$

Table 7.2: Additional quantum numbers for the characterization of unstable hadronic matter. Antiquarks have opposite values for these quantum numbers.

Type	Matter	Antimatter
Baryons	qqq	$\bar{q}\bar{q}\bar{q}$
Mesons	$q\bar{q}$	

Bound states such as $|qq\rangle$ or $|qq\bar{q}\rangle$ are excluded by the theory of quantum chromodynamics (see Sect. 7.4).

Unstable hadronic matter is characterized by the following additional flavor quantum numbers: Charm (C), Strangeness (S), Beauty (B), and Topness (T) (see Tab. 7.2). It is important to remember that in strong and electromagnetic interactions both baryon and flavor quantum numbers are conserved while in weak interactions only baryon quantum numbers are conserved. Therefore, weak interactions allow heavy quarks to decay into the stable quark family. The quark decay channels are shown in the following table:

Quark $\rightarrow$	Decay products
$u,\ d$	stable
s	uW^-
c	sW^+
b	cW^-
t	bW^+

As we have seen, protons and neutrons are prominent examples of baryons. Their general properties can be summarized as follows:

	Proton	**Neutron**
Quarks	$\vert uud\rangle$	$\vert udd\rangle$
Mass	0.9383 GeV	0.9396 GeV
Spin	1/2	1/2
Charge	$\mathrm{e} = 1.6 \cdot 10^{-19}\,\mathrm{C}$	0 C
Baryon number	1	1
Lifetime	stable: $\tau \geq 10^{32}$ years	unstable: $\tau_{n\rightarrow pe^-\bar{\nu}_e} = 887 \pm 2\,\mathrm{s}$
Production	gaseous hydrogen: ionization through electric field	under 1 MeV: nuclear reactors; $1-10$ MeV: nuclear reactions
Target for experiments	liquid hydrogen	liquid deuterium

The respective antiparticles can be produced in high-energy collisions, e. g.

$$
\begin{aligned}
pp &\to pp\bar{p}p \quad \text{with } |\bar{p}\rangle = \left|\bar{u}\bar{u}\bar{d}\right\rangle \quad \text{or} \\
pp &\to pp\bar{n}n \quad \text{with } |\bar{n}\rangle = \left|\bar{u}\bar{d}\bar{d}\right\rangle .
\end{aligned}
$$

Recall that in Sect. 4.1 we calculate the energy threshold for the reaction $pp \to pp\bar{p}p$ and find that a proton beam colliding against a proton target must have at least $|\vec{p}| = 6.5\,\text{GeV}$ for the reaction to take place.

7.3.2 Strangeness

We now take a more detailed look at the strangeness quantum number. In 1947, a new neutral particle, K^0, was discovered from interactions of cosmic rays:

$$\pi^- p \xrightarrow{\text{s}} K^0\Lambda, \text{ with consequent decays: } K^0 \xrightarrow{\text{w}} \pi^+\pi^-,\ \Lambda \xrightarrow{\text{w}} \pi^- p. \tag{7.17}$$

This discovery was later confirmed in accelerator experiments. The processes in Eq. (7.17) is puzzling because the *production* cross section is characterized by the strong interaction while the long lifetime ($\tau \sim 90\,\text{ps}$) indicates a weak *decay*. In this seemingly paradoxical situation, a new quantum number called "strangeness" is introduced. A sketch of production and decay of the K^0 is shown in Fig. 7.2. As stated before, the strong interaction conserves flavor which requires for the production $\Delta S = 0$. The decay, on the other hand, proceeds through the weak interaction: The s quark decays via $s \to uW^-$.

Baryons containing one or more strange quarks are called hyperons. With three constituting quarks we can have, depending on the spin alignment, spin-1/2 ($|\uparrow\downarrow\uparrow\rangle$) or spin-3/2 ($|\uparrow\uparrow\uparrow\rangle$) baryons (see Tab. 7.3).[2] There are 8 spin-1/2 baryons (octet) and 10 spin-3/2 baryons (decuplet). Octet and decuplet are part of the $SU(3)$ multiplet structure (see Sect. 7.4).[3] All hyperons in the octet decay weakly (except for the Σ^0). They therefore have a long lifetime of about $10^{-10}\,\text{s}$ and decay with $|\Delta S| = 1$, e. g.

$$
\begin{aligned}
\Sigma^+ &\to p\pi^0,\ n\pi^+ \\
\Xi^0 &\to \Lambda\pi^0.
\end{aligned}
$$

The members of the decuplet, on the other hand, all decay strongly (except for the Ω^-) with $|\Delta S| = 0$. They therefore have short lifetimes of about $10^{-24}\,\text{s}$, e. g.

$$
\begin{aligned}
\Delta^{++}(1230) &\to \pi^+ p \\
\Sigma^+(1383) &\to \Lambda\pi^+.
\end{aligned}
$$

[2]The problem that putting three fermions into one symmetric state violates the Pauli exclusion principle is discussed in Sect. 7.4.

[3]However, this "flavor $SU(3)$" is only a sorting symmetry and has nothing to do with "color $SU(3)$" discussed in Sect. 7.4.

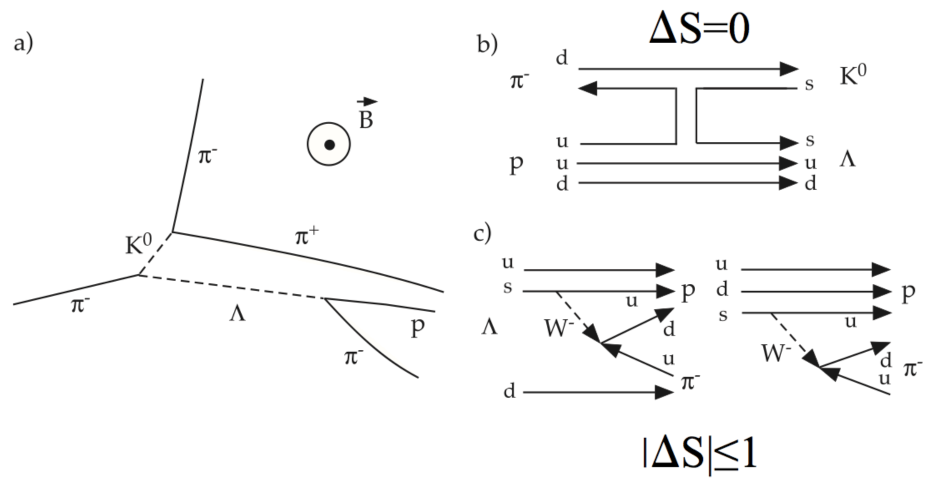


Figure 7.2: *Sketch of the reaction $\pi^- p \to K^0 \Lambda$ and the decays of the neutral K^0 and Λ.* Tracks detected in a bubble chamber (a). Feynman diagrams for the production and the Λ decay (b). Notice that $S(K^0) = 1,\ |K^0\rangle = |d\bar{s}\rangle$ and $S(\Lambda) = -1,\ |\Lambda\rangle = |uds\rangle$. Source: [8, p. 140].

Spin-1/2: **Octet**			Spin-3/2: **Decuplet**		
Baryon	State	Strangeness	Baryon	State	Strangeness
$p(938)$	$\|uud\rangle$	0	$\Delta^{++}(1230)$	$\|uuu\rangle$	0
$n(940)$	$\|udd\rangle$	0	$\Delta^{+}(1231)$	$\|uud\rangle$	0
$\Lambda(1115)$	$\|(ud-du)s\rangle$	-1	$\Delta^{0}(1232)$	$\|udd\rangle$	0
$\Sigma^+(1189)$	$\|uus\rangle$	-1	$\Delta^{-}(1233)$	$\|ddd\rangle$	0
$\Sigma^0(1192)$	$\|(ud+du)s\rangle$	-1	$\Sigma^{+}(1383)$	$\|uus\rangle$	-1
$\Sigma^-(1197)$	$\|dds\rangle$	-1	$\Sigma^{0}(1384)$	$\|uds\rangle$	-1
$\Xi^0(1315)$	$\|uss\rangle$	-2	$\Sigma^{-}(1387)$	$\|dds\rangle$	-1
$\Xi^-(1321)$	$\|dss\rangle$	-2	$\Xi^{0}(1532)$	$\|uss\rangle$	-2
			$\Xi^{-}(1535)$	$\|dss\rangle$	-2
			$\Omega^{-}(1672)$	$\|sss\rangle$	-3

Table 7.3: Summary of the baryon octet and decuplet.

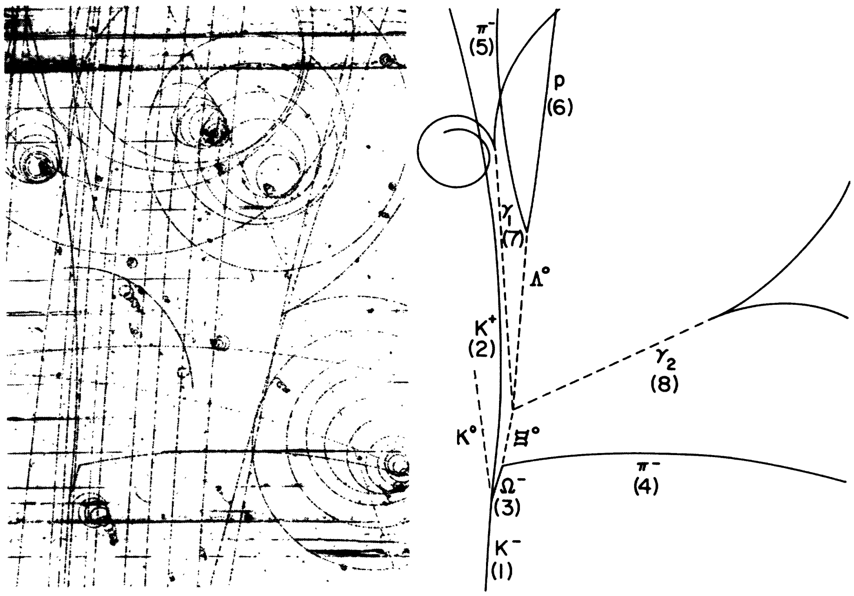


Figure 7.3: *Bubble chamber photograph (LHS) and line diagram (RHS) of an event showing the production and decay of* Ω^-. Source: [25, p. 205].

The quark model, as outlined so far, predicts the hyperon $|\Omega^-\rangle = |sss\rangle$ as a member of the spin-3/2 decuplet. Therefore, the observation of the production,

$$K^-p \to \Omega^- K^+ K^0,$$

and decay,

$$\Omega^- \to \Xi^0 \pi^-, \ \Xi^0 \to \Lambda \pi^0, \ \Lambda \to p\pi^-,$$

of the Ω^- at Brookhaven in 1964 is a remarkable success for the quark model. A sketch of the processes is given in Fig. 7.3. Note that the production occurs via a strong process, $\Delta S = 0$, while the decay is weak: $|\Delta S| = 1$.

7.3.3 Strong vs. weak decays

Generally speaking, strong processes yield considerably shorter lifetimes than weak processes. Consider, for instance, the following two decays,

$$\begin{aligned}
&\Delta^+ \to p + \pi^0 && \Sigma^+ \to p + \pi^0 \\
&\tau_\Delta = 6 \cdot 10^{-24}\,\mathrm{s} && \tau_\Sigma = 8 \cdot 10^{-11}\,\mathrm{s} \\
&|uud\rangle \to |uud\rangle + \frac{1}{\sqrt{2}}\left(|u\bar{u}\rangle + |d\bar{d}\rangle\right) && |uus\rangle \to |uud\rangle + \frac{1}{\sqrt{2}}\left(|u\bar{u}\rangle - |d\bar{d}\rangle\right) \\
&\text{(strong)} && \text{(weak).}
\end{aligned}$$

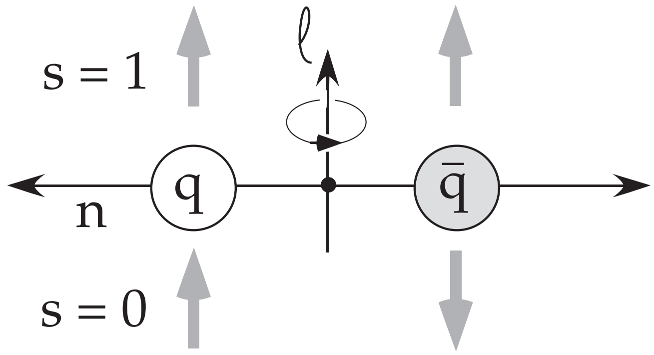


Figure 7.4: *Sketch of the possible spin configurations for quark-antiquark bound states.* The $q\bar{q}$ pair is characterized by orbital excitations l (rotation) and radial excitations n (vibration). Source: [8, p. 141].

The final state is identical in both decays but the lifetime is much longer for the weak process. Since the final state is equal, this difference in lifetime must come from a difference in the coupling constants. For $\tau \sim 1/\alpha^2$ where α is a coupling constant:

$$\frac{\alpha_{\text{weak}}}{\alpha_{\text{strong}}} \sim \sqrt{\frac{\tau_\Delta}{\tau_\Sigma}} = 2.7 \cdot 10^{-7}.$$

7.3.4 Mesons

Mesons are quark-antiquark bound states: $|q\bar{q}\rangle$. In analogy to the spin states of a two-electron system (and not to be confused with the isospin multiplets discussed on p. 130), the $|q\bar{q}\rangle$ bound state can have either spin 0 (singlet) or spin 1 (triplet) (see Fig. 7.4). Radial vibrations are characterized by the quantum number n while orbital angular momentum is characterized by the quantum number l. The states are represented in spectroscopic notation:

$$n^{2s+1}l_J$$

where $l = 0$ is labeled by S, $l = 1$ by P and so on. A summary of the $n = 1$, $l = 0$ meson states is shown in Tab. 7.4. A summary of the states with $l \leq 2$ can be found in Fig. 7.5.

7.3.5 Gell-Mann-Nishijima formula

Isospin is introduced in Sect. 7.1. The hadron isospin multiplets for $n = 1$, $l = 0$ are shown in Fig. 7.6. This summary leads to the conclusion that the charge Q of an hadron with baryon number B and strangeness S is given by

$$\boxed{Q = I_3 + \frac{B + S}{2}}$$

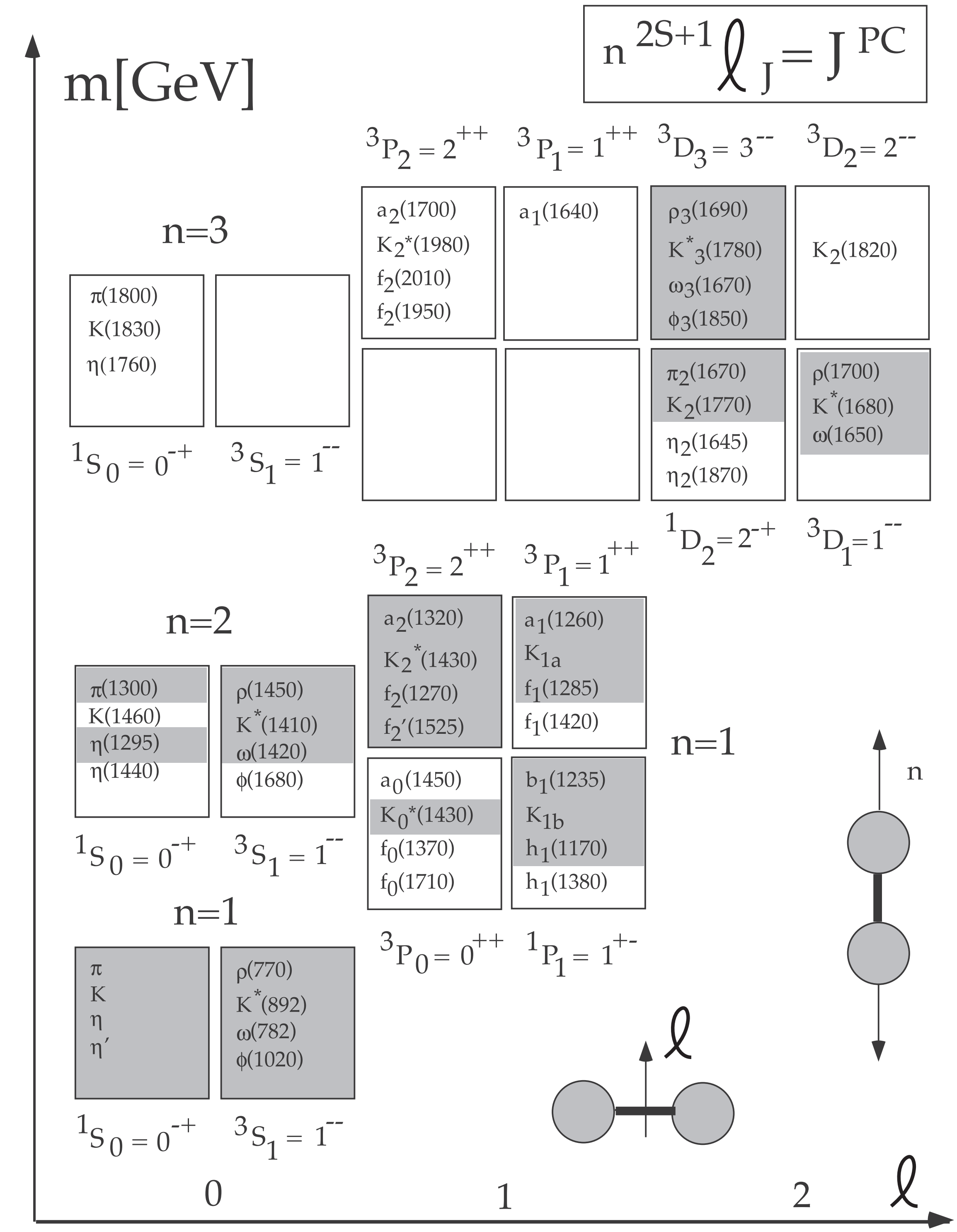


Figure 7.5: *Summary of mesons from u, d s quarks for $l \leq 2$.* Cells shaded in grey are well established states. Source: [8, p. 143].

Mesons ($n = 1$, $l = 0$)			
1^1S_0 (spin 0)		1^3S_1 (spin 1)	
$\pi^+(140)$	$\|u\bar{d}\rangle$	$\rho^+(770)$	$\|u\bar{d}\rangle$
$\pi^-(140)$	$\|\bar{u}d\rangle$	$\rho^-(770)$	$\|\bar{u}d\rangle$
$\pi^0(135)$	$1/\sqrt{2}\,\|d\bar{d} - u\bar{u}\rangle$	$\rho^0(770)$	$1/\sqrt{2}\,\|d\bar{d} - u\bar{u}\rangle$
$K^+(494)$	$\|u\bar{s}\rangle$	$K^{*+}(892)$	$\|u\bar{s}\rangle$
$K^-(494)$	$\|\bar{u}s\rangle$	$K^{*-}(892)$	$\|\bar{u}s\rangle$
$K^0(498)$	$\|d\bar{s}\rangle$	$K^{*0}(896)$	$\|d\bar{s}\rangle$
$\bar{K}^0(498)$	$\|\bar{d}s\rangle$	$\bar{K}^{*0}(896)$	$\|\bar{d}s\rangle$
$\eta(547)$	$\sim 1/\sqrt{6}\,\|u\bar{u} + d\bar{d} - 2s\bar{s}\rangle$	$\phi(1020) = \psi_1$	$-\|s\bar{s}\rangle$
$\eta'(958)$	$\sim 1/\sqrt{3}\,\|u\bar{u} + d\bar{d} + s\bar{s}\rangle$	$\omega(782) = \psi_2$	$1/\sqrt{2}\,\|u\bar{u} + d\bar{d}\rangle$

Table 7.4: Summary of $n = 1$, $l = 0$ meson states.

which is called Gell-Mann-Nishijima formula. As an example, consider the Ω^- hyperon where $0 + (1-3)/2 = -1$.

7.4 Quantum chromodynamics and color $SU(3)$

The quark model, as discussed so far, runs into a serious problem: Since the quarks have half-integer spin, they are fermions and therefore obey Fermi-Dirac statistics. This means that states like

$$\Delta^{++} = \left|u^\uparrow u^\uparrow u^\uparrow\right\rangle, \; S = \frac{3}{2}$$

where three quarks are in a symmetric state (have identical quantum numbers) are forbidden by the Pauli exclusion principle.

The way out is to introduce a new quantum number that allows for one extra degree of freedom which enables us to antisymmetrize the wave function as required for fermions:

$$\Delta^{++} = \mathcal{N} \sum_{ijk} \varepsilon_{ijk} \left|u_i^\uparrow u_j^\uparrow u_k^\uparrow\right\rangle$$

where $\mathcal{N}$ is some normalization constant and the quarks come in three different "colors":[4]

$$|q\rangle \to |q_{1,2,3}\rangle = \begin{pmatrix} |q_1\rangle \\ |q_2\rangle \\ |q_3\rangle \end{pmatrix} .$$

Since color cannot be observed, there has to be a corresponding new symmetry in the Lagrangian due to the fact that the colors can be transformed without the observables

[4] The new charge is named "color" because of the similarities to optics: There are three fundamental colors, complementary colors and the usual combinations are perceived as white.

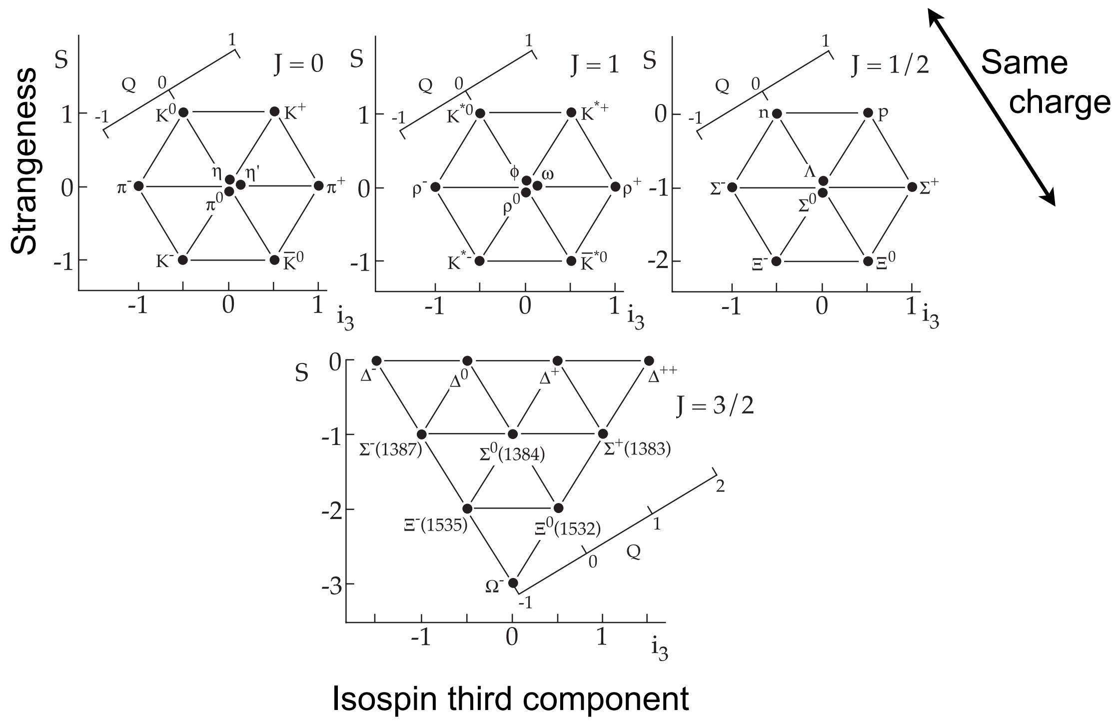


Figure 7.6: *Summary of hadron isospin multiplets.* $n = 1$, $l = 0$. Source: [8, p. 147].

being affected. In the case of our new charge in three colors the symmetry group is $SU(3)$, the group of the special unitary transformations in three dimensions. The Lie algebra of $SU(3)$ is

$$\left[T^a, T^b\right] = i f^{abc} T^c$$

where, in analogy to Eq. (7.10), f^{abc} denotes the structure constants and where there are 8 generators T^a (recall that $o = N^2 - 1 = 8$, see p. 129) out of which $r = N - 1 = 2$ are diagonal.

The fundamental representation is given by the 3×3 matrices $T^a = \frac{1}{2}\lambda^a$ with the Gell-Mann matrices

$$\lambda_1 = \begin{pmatrix} \overbrace{\boxed{\begin{matrix} 0 & 1 \\ 1 & 0 \end{matrix}}}^{\tau_1} & \begin{matrix} 0 \\ 0 \end{matrix} \\ \begin{matrix} 0 & 0 \end{matrix} & 0 \end{pmatrix} \quad \lambda_2 = \begin{pmatrix} \overbrace{\boxed{\begin{matrix} 0 & -i \\ i & 0 \end{matrix}}}^{\tau_2} & \begin{matrix} 0 \\ 0 \end{matrix} \\ \begin{matrix} 0 & 0 \end{matrix} & 0 \end{pmatrix} \quad \lambda_3 = \begin{pmatrix} \overbrace{\boxed{\begin{matrix} 1 & 0 \\ 0 & -1 \end{matrix}}}^{\tau_3} & \begin{matrix} 0 \\ 0 \end{matrix} \\ \begin{matrix} 0 & 0 \end{matrix} & 0 \end{pmatrix} \quad \lambda_4 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

$$\lambda_5 = \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix} \quad \lambda_6 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \quad \lambda_7 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix} \quad \lambda_8 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix}.$$

One can observe that these matrices are hermitian and traceless,

$$\lambda_a^\dagger = \lambda_a \qquad\qquad \mathrm{Tr}\,\lambda^a = 0.$$

Furthermore, one can show that

$$\mathrm{Tr}\left(\lambda^a \lambda^b\right) = 2\delta^{ab}$$

and

$$\lambda^a_{ij}\lambda^a_{kl} = 2\left(\delta_{il}\delta_{kj} - \frac{1}{3}\delta_{ij}\delta_{kl}\right) \text{ (Fierz identity).}$$

The structure constants of $SU(3)$ are given by

$$f_{abc} = \frac{1}{4i}\mathrm{Tr}\left(\left[\lambda_a, \lambda_b\right]\lambda_c\right)$$

and are antisymmetric in a, b, and c. The numerical values are

$$\begin{aligned}
&f_{123} = 1\\
&f_{458} = f_{678} = \frac{\sqrt{3}}{2}\\
&f_{147} = f_{156} = f_{246} = f_{257} = f_{345} = f_{367} = \frac{1}{2}\\
&f_{abc} = 0 \quad \text{else.}
\end{aligned}$$

As in the case of $SU(2)$, the adjoint representation is given by the structure constants which, in this case, are 8×8 matrices:

$$(t^a)_{bc} = -if_{abc}.$$

The multiplets (again built out of the fundamental representations) are given by the direct sums

$$\underline{3} \otimes \underline{\bar{3}} = \underline{1} \oplus \underline{8} \tag{7.18}$$

where the bar denotes antiparticle states and

$$\underline{3} \otimes \underline{3} \otimes \underline{3} = \underline{1} \oplus \underline{8} \oplus \underline{8} \oplus \underline{10}. \tag{7.19}$$

The singlet in Eq. (7.18) corresponds to the $|q\bar{q}\rangle$ states, the mesons (e. g. π), while the singlet in Eq. (7.19) is the $|qqq\rangle$ baryon (e. g. p, n). The other multiplets are colored and can thus not be observed.'Working out the$SU(3)$ potential structure, one finds that an attractive QCD potential exists only for the singlet states, while the potential is repulsive for all other multiplets.

The development of QCD outlined so far can be summarized as follows: Starting from the observation that the nucleons have similar properties, we considered isospin and $SU(2)$ symmetry. We found that the nucleons n and p correspond to the fundamental representations of $SU(2)$ while the π is given by the adjoint representation. To satisfy the Pauli exclusion principle, we had to introduce a new quantum number and with it a new $SU(3)$ symmetry of the Lagrangian. This in turn led us to multiplet structures where the colorless singlet states correspond to mesons and baryons.

Construction of QCD Lagrangian We now take a closer look at this $SU(3)$ transformation of a color triplet,

$$|q\rangle = \begin{pmatrix} q_1 \\ q_2 \\ q_3 \end{pmatrix} \rightarrow |q'\rangle = \begin{pmatrix} q'_1 \\ q'_2 \\ q'_3 \end{pmatrix} = \mathrm{e}^{ig_s \alpha_a T^a} \begin{pmatrix} q_1 \\ q_2 \\ q_3 \end{pmatrix} = U\,|q\rangle\,, \tag{7.20}$$

where $g_s \in \mathbb{R}$ is used as a rescaling (and will be used for the perturbative expansion) of the group parameter α introduced previously. The reason of introducing it becomes clear in the context of gauge theories.

In analogy to the QED current,

$$j^{\mu}_{\mathrm{QED}} = eq_e \bar{q}\gamma^{\mu} q,$$

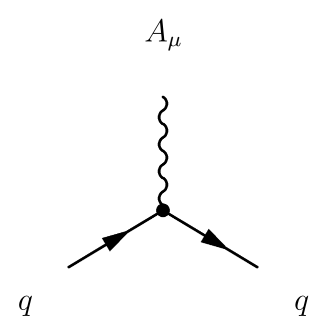
we introduce the **color current** [5], which is the conserved current associated with the $SU(3)$ symmetry,

$$j^{\mu}_a = g_s \bar{q}_i \gamma^{\mu} T^a_{ij} q_j \qquad\qquad a = 1 \cdots 8. \tag{7.21}$$

In the same spirit, by looking at the QED interaction,

$$\mathcal{L}^{\mathrm{int}}_{\mathrm{QED}} = -j^{\mu}_{\mathrm{QED}} A_{\mu} = eq_e \bar{q}\gamma^{\mu} q A_{\mu},$$

yielding the vertex,

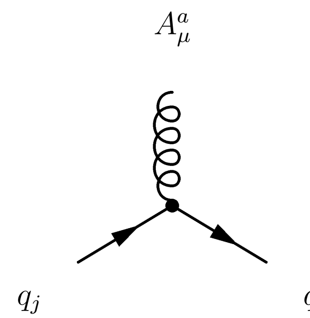


where we can see the photon – the electrically uncharged $U(1)$ gauge boson of QED –, we postulate an interaction part of the QCD Lagrangian of the form,

$$\mathcal{L}^{\mathrm{int}}_{\mathrm{QCD}} = -j^{\mu}_a A^a_{\mu} = g_s \bar{q}_i \gamma^{\mu} T^a_{ij} q_j A^a_{\mu}, \tag{7.22}$$

which translates in the vertex (which is not the only one of QCD as we shall see),

[5] The Einstein summation convention still applies, even if the color index i and j are not in an upper and lower position. This exception extends also to the color indices $a, b, \ldots$ of the gauge fields to be introduced. There is no standard convention in the literature, and since there is no metric tensor involved, the position of a color index, is merely an esthetic/readability problem.



Now there are 8 $SU(3)$ gauge bosons A_μ^a for QCD : one for each possible value of a. They are called **gluons** and are themselves colored.

Continuing with our analogy, we define the **covariant derivative of QCD**[6],

$$D_\mu = \partial_\mu \mathbb{1} + i g_s T^a A_\mu^a, \tag{7.23}$$

and state that the QCD Lagrangian should have a term of the form,

$$\tilde{\mathcal{L}}_{\mathrm{QCD}} = \bar{q}(i\not{D} - m)q. \tag{7.24}$$

Up to this point, both QED and QCD look nearly identical. Their differences become crucial when we look at local gauge symmetries. Such a transformation can be written,

$$|q(x)\rangle \to |q'(x)\rangle = \mathrm{e}^{i g_s \alpha_a(x) T^a} |q(x)\rangle\,, \tag{7.25}$$

and we impose as before that the Lagrangian must be invariant under any such transformation. This is equivalent of imposing,

$$\begin{aligned} D'_\mu |q'(x)\rangle &\stackrel{!}{=} \mathrm{e}^{i g_s \alpha_a(x) T^a} D_\mu |q(x)\rangle \\ &\Leftrightarrow \langle \bar{q}'(x)| \, i\not{D}' \, |q'(x)\rangle = \langle \bar{q}(x)| \, i\not{D} \, |q(x)\rangle\,. \end{aligned}$$

For $\alpha_a(x) \ll 1$, we can expand the exponential and keep only the first order term,

$$\begin{aligned} D'_\mu |q'(x)\rangle &= \left(\partial_\mu + i g_s T^c A'^c_\mu\right) \left(\mathbb{1} + i g_s \alpha_a(x) T^a\right) |q(x)\rangle \\ &\stackrel{!}{=} \left(\mathbb{1} + i g_s \alpha_a(x) T^a\right) \underbrace{\left(\partial_\mu + i g_s T^c A^c_\mu\right)}_{D_\mu} |q(x)\rangle\,. \end{aligned}$$

Making the ansatz $A'^c_\mu = A^c_\mu + \delta A^c_\mu$ where $|\delta A^c_\mu| \ll |A^c_\mu|$ and expanding the former equation to first order in δA^c_μ (the term proportional to $\alpha_a(x)\delta A^c_\mu$ has also been ignored), we get,

$$\begin{aligned} & i g_s T^c \delta A^c_\mu + i g_s (\partial_\mu \alpha_a(x)) T^a + i^2 g_s^2 T^c A^c_\mu \alpha_a(x) T^a \stackrel{!}{=} i^2 g_s^2 \alpha_a(x) T^a T^c A^c_\mu \\ & \Rightarrow T^c \delta A^c_\mu \stackrel{!}{=} -(\partial_\mu \alpha_a(x)) T^a + i g_s [T^a, T^c] \alpha_a(x) A^c_\mu, \end{aligned}$$

[6]Note that D_μ acts on color triplet and gives back a color triplet; ∂_μ does not mix the colors, whereas the other summand does (T^a is a 3×3 matrix).

or, renaming the dummy indices and using the Lie algebra $su(3)$,

$$\begin{aligned} T^a \delta A^a_\mu &= -(\partial_\mu \alpha_a(x))T^a - g_s f_{abc} T^a \alpha_b(x) A^c_\mu \qquad \forall T^a \\ \Rightarrow A'^a_\mu &= A^a_\mu - \underbrace{\partial_\mu \alpha_a(x)}_{\text{like in QED}} - \underbrace{g_s f_{abc} \alpha_b(x) A^c_\mu}_{\text{non-abelian part}}. \end{aligned} \tag{7.26}$$

Eq. (7.26) describes the (infinitesimal) gauge transformation of the gluon field.

In order for the gluon field to become physical, we need to include a kinematical term (depending on the derivatives of the field). Remember the photon term of QED,

$$\mathcal{L}^{\text{photon}}_{\text{QED}} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu} \qquad\qquad F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu,$$

where the last is gauge invariant. As we might expect from Eq. (7.26), the non-abelian part will get us into trouble. Let's look at,

$$\begin{aligned} \delta(\partial_\mu A^c_\nu - \partial_\nu A^c_\mu) = -\partial_\mu \partial_\nu \alpha_a + \partial_\nu \partial_\mu \alpha_a - g_s f_{abc} \alpha_b (\partial_\mu A^c_\nu - \partial_\nu A^c_\mu) \\ - g_s f_{abc} \left((\partial_\mu \alpha_b) A^c_\nu - (\partial_\nu \alpha_b) A^c_\mu \right). \end{aligned}$$

We remark that the two first summands cancel each other and that the third looks like the $SU(3)$ transformation under the adjoint representation.

We recall that,

$$\begin{aligned} q_i \to q'_i &= (\delta_{ij} + i g_s \alpha_a T^a_{ij}) q_j \qquad &\text{(fundamental representation)} \\ B_a \to B'_a &= (\delta_{ac} + i g_s \alpha_b t^b_{ac}) B_c \qquad &\text{(adjoint representation)} \end{aligned}$$

respectively, where,

$$t^b_{ac} = -i f_{bac} = i f_{abc}.$$

Hence, if $F^a_{\mu\nu}$ transforms in the adjoint representation of $SU(3)$, we should have,

$$\delta F^a_{\mu\nu} \stackrel{!}{=} -g_s f_{abc} \alpha_b F^c_{\mu\nu}.$$

We now make the ansatz,

$$F^a_{\mu\nu} = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu - g_s f_{abc} A^b_\mu A^c_\nu, \tag{7.27}$$

and prove that it fulfills the above constraint.

$$\begin{aligned} \delta F^a_{\mu\nu} =& \delta(\partial_\mu A^a_\nu - \partial_\nu A^a_\mu) - g_s f_{abc} \delta(A^b_\mu A^c_\nu) \\ =& - g_s f_{abc} \alpha_b (\partial_\mu A^c_\nu - \partial_\nu A^c_\mu) - g_s f_{abc} \left((\partial_\mu \alpha_b) A^c_\nu - (\partial_\nu \alpha_b) A^c_\mu \right) \\ & - g_s f_{abc} \left(-(\partial_\mu \alpha_b) A^c_\nu + (\partial_\nu \alpha_b) A^c_\mu \right) - g_s f_{abc} \left(-g_s f_{bde} \alpha_d A^e_\mu A^c_\nu - g_s f_{cde} \alpha_d A^b_\mu A^e_\nu \right), \end{aligned}$$

Using,

$$f_{abc}f_{bde}\alpha_d A^e_\mu A^c_\nu = f_{abe}f_{bdc}\alpha_d A^c_\mu A^e_\nu = f_{ace}f_{cdb}\alpha_d A^b_\mu A^e_\nu,$$

and

$$f_{aec}f_{dbc} - f_{acb}f_{dec} = (iT^a_{ec})(iT^d_{cb}) - (iT^d_{ec})(iT^a_{cb}) = \left[T^a, T^d\right]_{eb} = if_{adc}T^c_{eb},$$

we get the desired result.

We check finally that a kinematic term based on the above definition of $F^a_{\mu\nu}$ is gauge invariant :

$$\delta\left(F^a_{\mu\nu}F^{\mu\nu}_a\right) = 2F^{\mu\nu}_a \delta F^a_{\mu\nu} = -2g_s \underbrace{f_{abc}}_{=-f_{cba}} \alpha_b \underbrace{F^{\mu\nu}_a F^c_{\mu\nu}}_{=F^{\mu\nu}_c F^a_{\mu\nu}} = 0.$$

Finally, we get the full **QCD Lagrangian**,

$$\mathcal{L}_{\mathrm{QCD}} = -\frac{1}{4}F^a_{\mu\nu}F^{\mu\nu}_a + \bar{q}(i\not{D} - m_q)q, \tag{7.28}$$

with $\not{D}$ and $F^a_{\mu\nu}$ definded by Eqs. (7.23) and (7.27) respectively.

This Lagrangian is per construction invariant under local $SU(3)$ gauge transformations. It is our first example of a non-abelian gauge theory, a so-called **Yang-Mills theory**.

Structure of the kinematic term From the definition of $F^a_{\mu\nu}$, Eq. (7.27), we see that,

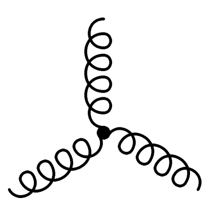
$$F^a_{\mu\nu}F^{\mu\nu}_a = \left(\partial_\mu A^a_\nu - \partial_\nu A^a_\mu - g_s f_{abc}A^b_\mu A^c_\nu\right)\left(\partial^\mu A^\nu_a - \partial^\nu A^\mu_a - g_s f_{ade}A^\mu_d A^\nu_e\right),$$

will have a much richer structure than in the case of QED.

First, we have – as in QED – a 2-gluon term $\left(\partial_\mu A^a_\nu - \partial_\nu A^a_\mu\right)\left(\partial^\mu A^\nu_a - \partial^\nu A^\mu_a\right)$ corresponding to the **gluon propagator**,

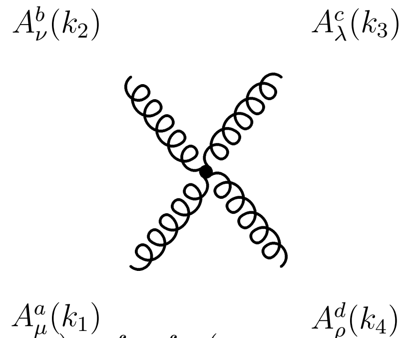
$$= -\frac{g^{\mu\nu}}{k^2}\delta^{ab}. \tag{7.29}$$

Then we have a 3-gluon term $\left(-g_s f_{abc}A^b_\mu A^c_\nu\right)\left(\partial^\mu A^\nu_a - \partial^\nu A^\mu_a\right)$ yielding a **3-gluon vertex**



$$= g_s f_{abc}\left[g_{\mu\nu}(k_1 - k_2)_\lambda + g_{\nu\lambda}(k_2 - k_3)_\mu + g_{\lambda\mu}(k_3 - k_1)_\nu\right]. \tag{7.30}$$

Finally we have also a 4-gluon term $\left(-g_s f_{abc} A^b_\mu A^c_\nu\right)\left(-g_s f_{ade} A^\mu_d A^\nu_e\right)$ yielding the **4-gluon vertex**

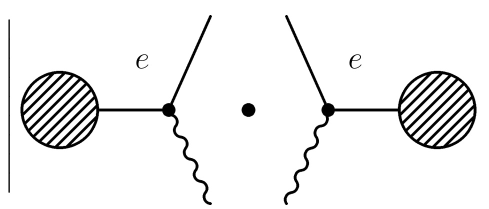


$$= -ig_s^2\left[f_{abe}f_{cde}(g_{\mu\lambda}g_{\nu\rho} - g_{\mu\rho}g_{\nu\lambda}) + f_{ade}f_{bce}(g_{\mu\nu}g_{\lambda\rho} - g_{\mu\lambda}g_{\nu\rho}) + f_{ace}f_{bde}(g_{\mu\rho}g_{\nu\lambda} - g_{\mu\nu}g_{\rho\lambda})\right] \tag{7.31}$$

Unlike in QED, gluons are able to interact with themselves. This comes from the fact that the theory is non-abelian. As a consequence, there is no superposition principle for QCD: the field of a system of strongly interacting particles is *not* the sum of the individual fields. Thence, there is no plane wave solution to QCD problems, and we cannot make use of the usual machinery of Green's functions and Fourier decomposition. Up to now there is no known solution.

7.4.1 Strength of QCD interaction

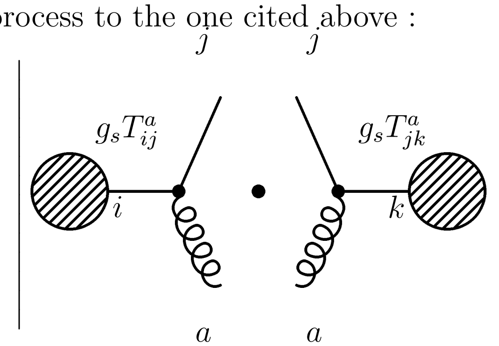
In QED, when we take a term of the form,



where the  denotes some other part of the Feynman diagram, the expression is proportional to $e^2 = 4\pi\alpha$.

In the case of QCD, we have a few more possibilities. We look at the general $SU(n)$ case. The QCD result can be found by setting $n = 3$.

First, for the analogous process to the one cited above :



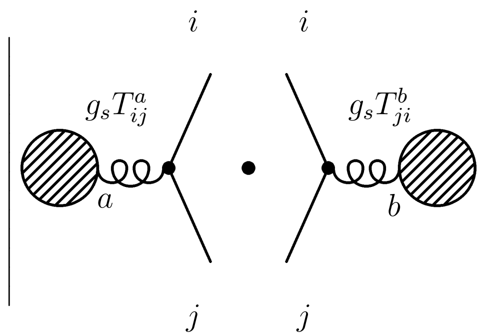
which is proportional to $g_s^2 T^a_{ij} T^a_{jk} = 4\pi\alpha_s C_F \delta_{ik}$, where

$$C_F = \frac{n^2 - 1}{2n}, \tag{7.32}$$

is the color factor, the Casimir operator of $SU(n)$. To find it, we used one of the Fierz identities (see exercises), namely,

$$\begin{aligned} T^a_{ij} T^a_{jk} &= \frac{1}{2}\left(\delta_{ik}\delta_{jj} - \frac{1}{n}\delta_{ij}\delta_{jk}\right) \\ &= \frac{1}{2}\left(n\delta_{ik} - \frac{1}{n}\delta_{ik}\right) = \frac{n^2-1}{2n}\delta_{ik}. \end{aligned}$$

Next we look at,



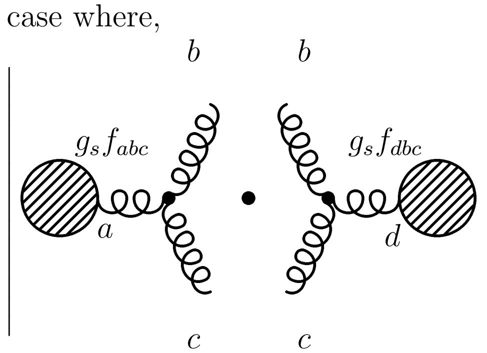
which is proportional to $g_s^2 T^a_{ij} T^b_{ji} = 4\pi\alpha_s T_F \delta^{ab}$, where

$$T_F = \frac{1}{2}. \tag{7.33}$$

To find it, we used the fact that,

$$\mathrm{Tr}\left(T^a T^b\right) = \frac{1}{2}\delta^{ab}.$$

Finally we investigate the case where,



which is proportional to $g_s^2 f_{abc} f_{dbc} = 4\pi\alpha_s C_A \delta^{ad}$, where

$$C_A = n. \tag{7.34}$$

To find it, we used the relation,

$$f_{abc} = -2i\mathrm{Tr}\left(\left[T^a, T^b\right]T^c\right),$$

that we have shown in the beginning of this section.

In the case of QCD, $C_F = \frac{4}{3}, T_F = \frac{1}{2}, C_A = 3$. From the discussion above, we can heuristically draw the conclusion that gluons tend to couple more to other gluons, than to quarks.

At this stage, we note two features specific to the strong interaction, which we are going to handle in more detail in a moment :

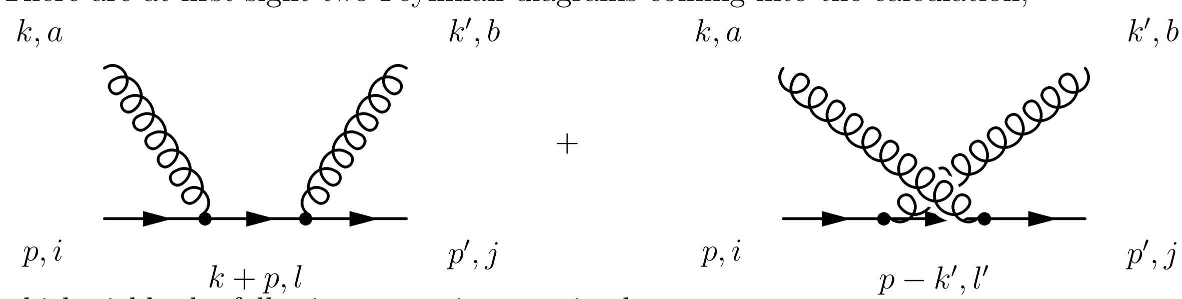
- **Confinement** : At low energies (large distances), the coupling becomes very large, so that the perturbative treatment is no longer valid, an the process of hadronization becomes inportant. This is the reason why we cannot observe color directly.
- **Asymptotic freedom** : At high energies (small distances) the coupling becomes negligible, and the quarks and gluons can move almost freely.

As an example, of typical QCD calculation, we sketch the calculation of the

Gluon Compton scattering

$$g(k) + q(p) \to g(k') + q(p').$$

There are at first sight two Feynman diagrams coming into the calculation,



which yields the following scattering matrix element,

$$\begin{aligned}-i\mathcal{M}_{fi} = &- ig_s^2\left[\bar{u}(p')\not{\varepsilon}^*(k')\frac{1}{\not{p}+\not{k}-m}\not{\varepsilon}(k)u(p)T^b_{jl}T^a_{li}\right.\\ &\left.+\bar{u}(p')\not{\varepsilon}(k)\frac{1}{\not{p}-\not{k}'-m}\not{\varepsilon}^*(k')u(p)T^a_{jl'}T^b_{l'i}\right].\end{aligned} \tag{7.35}$$

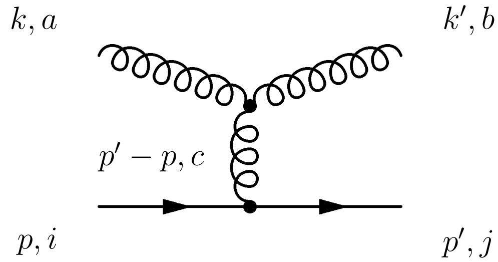
We start by checking the gauge invariance ($\mathcal{M}_{fi}$ must vanish under the substitution $\varepsilon_\mu(k) \to k_\mu$):

$$-i\mathcal{M}'_{fi} = -ig_s^2\bar{u}(p')\not{\varepsilon}^*(k')u(p)\left(T^b_{ji}T^a_{li} - T^a_{jl'}T^b_{l'i}\right),$$

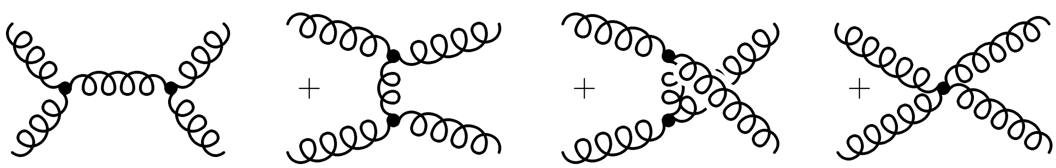
where

$$T^b_{ji}T^a_{li} - T^a_{jl'}T^b_{l'i} = \left[T^b, T^a\right]_{ji} = if_{bac}T^c_{ji} \neq 0!$$

So we need another term, which turns out to be the one corresponding to the Feynman diagram,



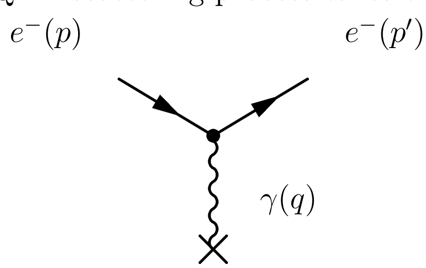
The calculation of the gluon-gluon scattering goes analogously. We need to consider the graphs,



.

7.4.2 QCD coupling constant

To leading order, a typical QED scattering process takes the form,

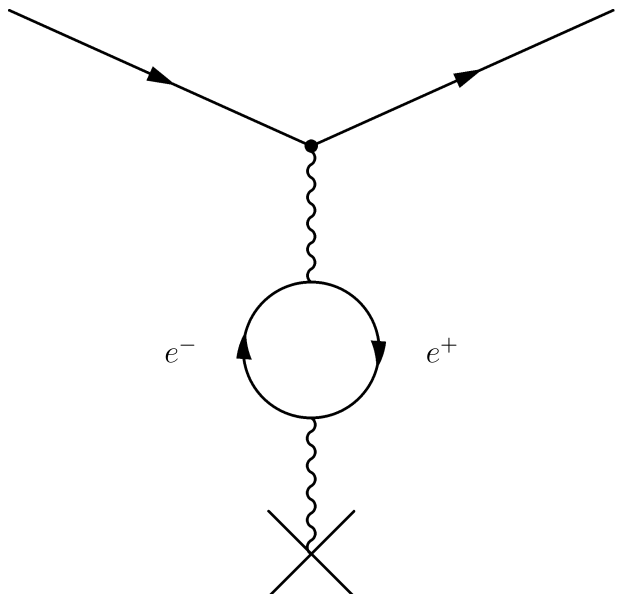


with $q^2 = (p' - p)^2 \leq 0$.

In the Coulomb limit (long distance, low momentum transfer), the potential takes the form,

$$V(R) = -\frac{\alpha}{R} \qquad\qquad R \gtrsim \frac{1}{m_e} \approx 10^{-11}\,[\mathrm{cm}]. \tag{7.36}$$

When $R \leq m_e^{-1}$, quantum effects become important (loop corrections, also known as vacuum polarization), since the next to leading order (NLO) diagram,



starts to play a significant (measurable) role. This results in a change of the potential to,

$$V(R) = -\frac{\alpha}{R}\left[1 + \frac{2\alpha}{3\pi}\ln\frac{1}{m_e R} + \mathcal{O}(\alpha^2)\right] = -\frac{\bar{\alpha}(R)}{R}, \tag{7.37}$$

where $\bar{\alpha}(R)$ is called the effective coupling.

We can understand the effective coupling in analogy to a solid state physics example : in an insulator, an excess of charge gets screened by the polarization of the nearby atoms. Here we create e^+e^- pairs out of the vacuum, hence the name **vacuum polarization**.

As we can see from Eq. (7.37), the smaller the distance $R \leq m_e^{-1}$, the bigger the observed "charge" $\bar{\alpha}(R)$. What we call the electron charge e (or the fine structure constant α) is the limiting value for very large distances or low momentum transfer as shown in Fig. 7.4.2.

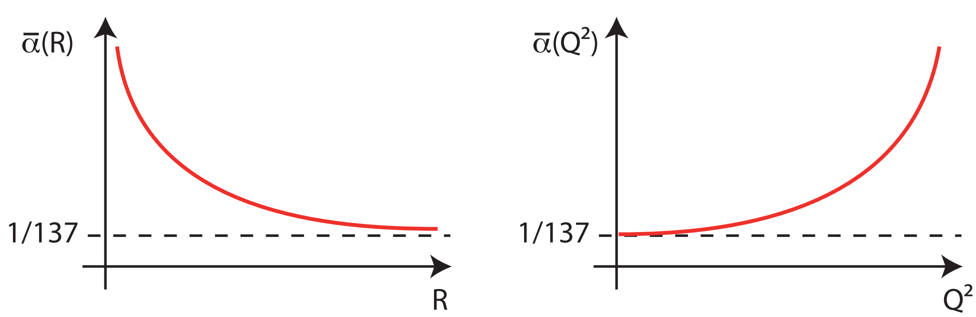
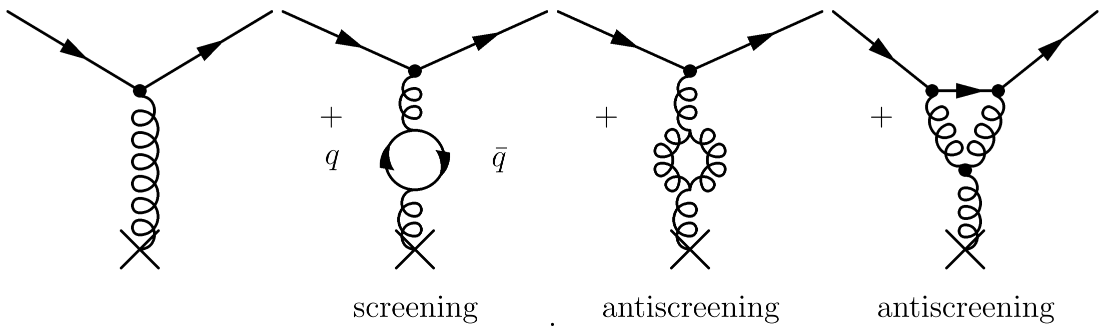


Figure 7.7: Evolution of the effective electromagnetic coupling with distance and energy ($Q^2 = -q^2$).

For example the measurements done at LEP show that, $\bar{\alpha}(Q^2 = m_Z^2) \approx \frac{1}{128} > \alpha$.

In the case of QCD, we have at NLO, the following diagrams,



We can picture the screening/antiscreening phenomenon as follows,

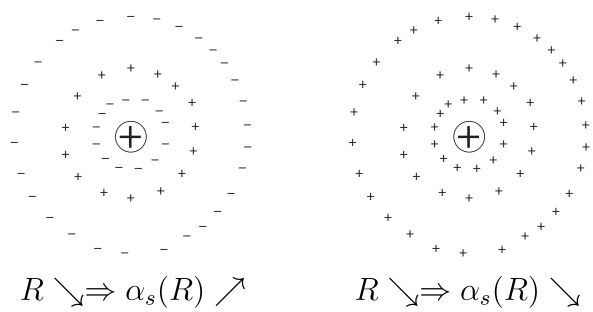


Figure 7.8: Screening and antiscreening.

For QCD, the smaller the distance R (or the bigger the energy Q^2), the smaller the observed coupling $\bar{\alpha}_s(R)$. At large distances, $\bar{\alpha}_s(R)$ becomes comparable with unity, and the perturbative approach breaks down as we can see in Fig. 7.4.2. The region concerning confinement and asymptotic freedom are also shown.

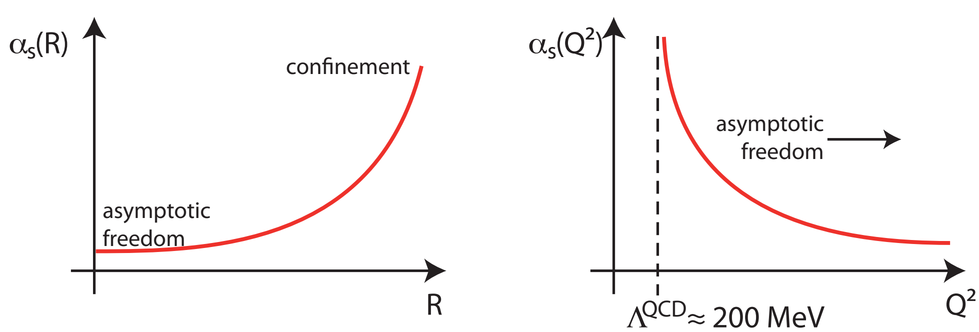


Figure 7.9: Evolution of the effective strong coupling with distance and energy ($Q^2 = -q^2$).

The β-function of QCD In the renormalization procedure of QCD, we get a differential equation for $\alpha_s(\mu^2)$ where μ is the renormalization scale,

$$\mu^2 \frac{\partial \alpha_s}{\partial(\mu^2)} = \beta(\alpha_s) \tag{7.38}$$

$$\beta(\alpha_s) = -\alpha_s \left[\beta_0 \frac{\alpha_s}{4\pi} + \beta_1 \left(\frac{\alpha_s}{4\pi}\right)^2 + \beta_2 \left(\frac{\alpha_s}{4\pi}\right)^3 + \cdots\right], \tag{7.39}$$

with

$$\beta_0 = \frac{11}{3} n_c - \frac{2}{3} n_f = 11 - \frac{2}{3} n_f \qquad \text{(NLO)} \tag{7.40}$$

$$\beta_1 = \frac{17}{12} n_c^2 - \frac{5}{12} n_c n_f - \frac{1}{4} \left(\frac{n_c^2 - 1}{2n_c}\right) n_f, \qquad \text{(NNLO)} \tag{7.41}$$

where n_c is the number of colors and n_f is the number of quark flavors. These two numbers enter into the calculation through gluon respectively quark loop corrections to the propagators.

We remark at this stage that unless[7] $n_f \geq 17$, we have $\beta_0 > 0$, whereas in the case of QED, we get,

$$\beta_0^{\text{QED}} = -\frac{4}{3} < 0. \tag{7.42}$$

This fact explains the completely different behavior of the effective couplings of QCD and QED.

To end this chapter, we will solve Eq. (7.38) retaining only the first term of the power expansion of β.

$$\begin{aligned}
\mu^2 \frac{\partial \alpha_s}{\partial(\mu^2)} &= -\frac{\beta_0}{4\pi} \alpha_s^2 \\
\frac{\partial \alpha_s}{\alpha_s^2} &= -\frac{\beta_0}{4\pi} \partial(\ln \mu^2) \\
\int\limits_{\alpha_s(Q_0^2)}^{\alpha_s(Q^2)} \frac{d\alpha_s}{\alpha_s^2} &= -\frac{\beta_0}{4\pi} \int\limits_{\ln Q_0^2}^{\ln Q^2} d(\ln \mu^2),
\end{aligned}$$

and hence,

$$\boxed{\frac{1}{\alpha_s(Q^2)} = \frac{1}{\alpha_s(Q_0^2)} + \frac{\beta_0}{4\pi} \ln \frac{Q^2}{Q_0^2}}. \tag{7.43}$$

[7] As of 2009, only 6 quark flavors are known and there is experimental evidence (decay witdth of the Z^0 boson) that there are no more than 3 generations with light neutrinos.

We thus have a relation between $\alpha_s(Q^2)$ and $\alpha_s(Q_0^2)$, giving the evolution of the effective coupling.

A mass scale is also generated, if we set,

$$\frac{1}{\alpha_s(Q^2 = \Lambda^2)} = 0 \Rightarrow \alpha_s(\Lambda^2) = \infty.$$

Choosing $\Lambda = Q_0$, we can rewrite Eq. (7.43) as,

$$\boxed{\alpha_s(Q^2) = \frac{4\pi}{\beta_0 \ln \frac{Q^2}{\Lambda^2}}}. \tag{7.44}$$

Chapter 8

QCD in e^+e^- annihilations

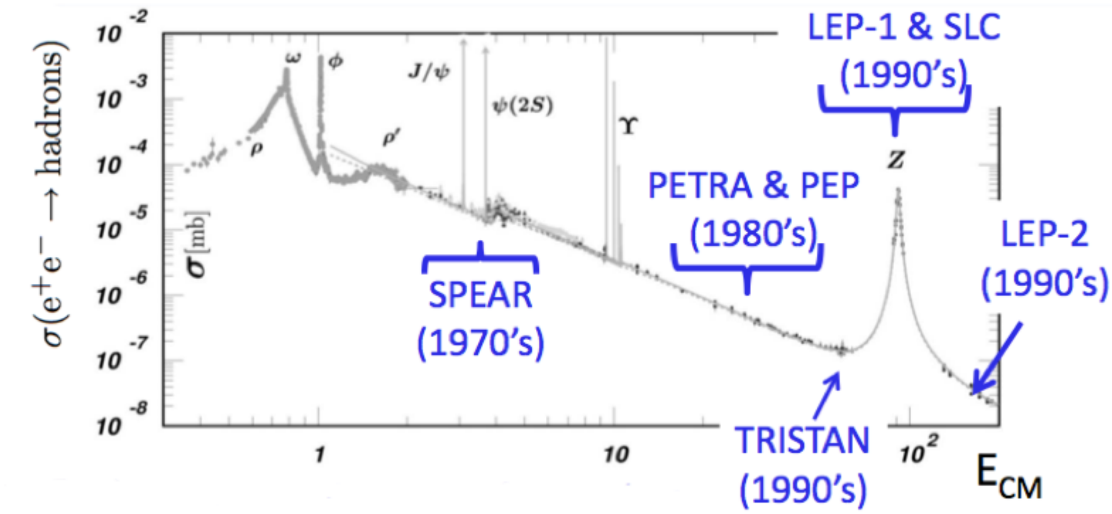
Literature:

- Dissertori/Knowles/Schmelling [27]
- Ellis/Stirling/Webber [28]
- Bethke [29, 30]
- Particle Data Group [26]
- JADE, Durham, and Cambridge jet algorithms [31, 32, 33, 34]
- FastJet Package, Fast k_T, SISCone [35, 36, 37]

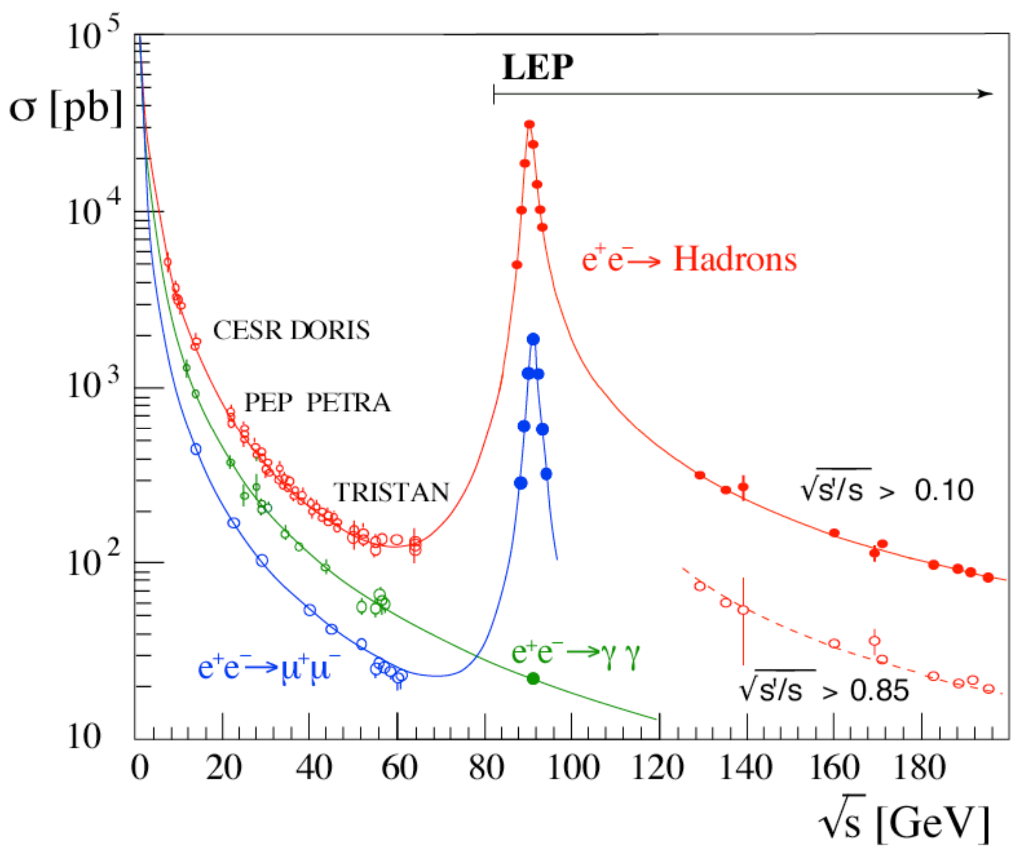
In Chap. 7, QCD is introduced as an $SU(3)$ gauge theory. Here we continue this discussion and consider QCD processes following e^+e^- annihilations. The main focus is on the definition and application of observables linking theoretical predictions with measurable quantities: Jets and event shapes are discussed; the applications include measurements of the parton spins, the strong coupling constant, and the QCD color factors. The chapter is concluded by an outlook to hadronization and non-perturbative QCD.

Some examples of e^+e^- colliders and their energies are given in Tab. 6.1. Fig. 8.1(a) maps the corresponding eras onto the available center of mass energies. A half-logarithmic plot comparing $\sigma^{e^+e^-\to\text{hadrons}}$ to $\sigma^{e^+e^-\to\mu^-\mu^+}$ is given in Fig. 8.1(b). Experimental milestones include:

- *SPEAR (SLAC):* Discovery of quark jets.
- *PETRA (DESY) & PEP (SLAC):* First high energy ($> 10\,\text{GeV}$) jets; discovery of gluon jets (at the PETRA collider, see Fig. 8.2); many pioneering QCD studies.

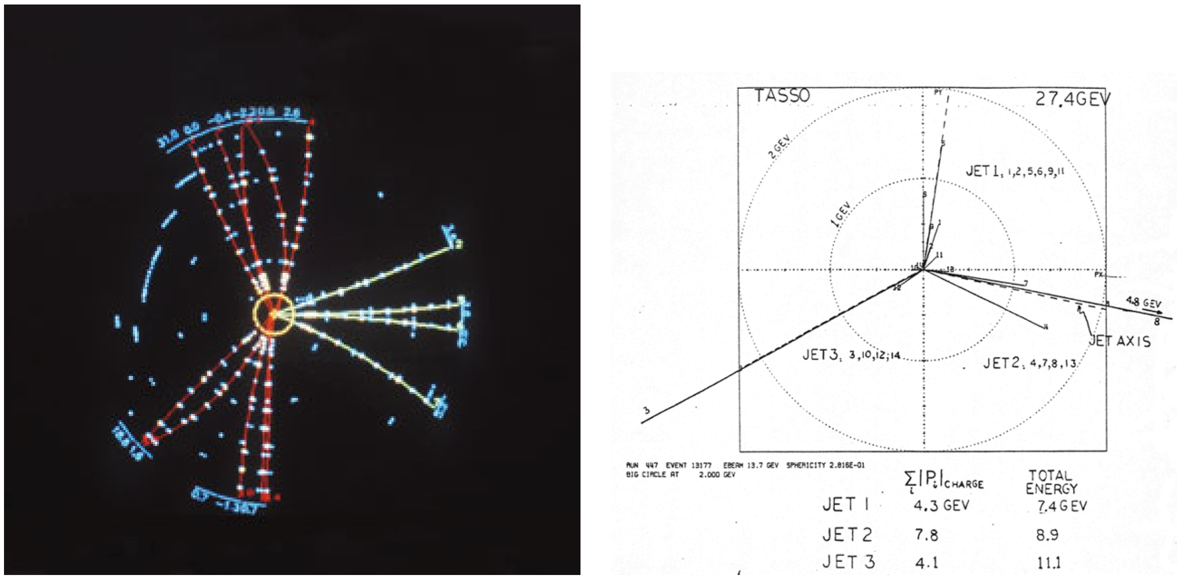


(a)



(b)

Figure 8.1: *Cross sections in e^+e^- annihilations.* (a) Cross section for $e^+e^- \to$ hadrons as a function of the center of mass energy. The E_{CM} dependence is linear because the plot is double-logarithmic. Source: [38]. (b) Comparison of cross sections for $e^+e^- \to$ hadrons and for $e^+e^- \to \mu^-\mu^+$. Both cross sections show the same $1/s$ dependence on the center of mass energy squared, except at the Z resonance.



(a) (b)

Figure 8.2: *Gluon discovery at the PETRA collider at DESY, Hamburg.* Event display (a) and reconstruction (b).

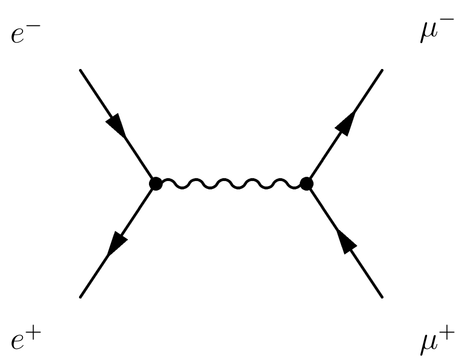
- *LEP (CERN) & SLC (SLAC):* Large energies (small α_s, see later) mean more reliable calculations and smaller hadronization uncertainties. Large data samples are collected: $\sim 3 \cdot 10^6$ hadronic Z decays per experiment. This allows for precision tests of QCD.

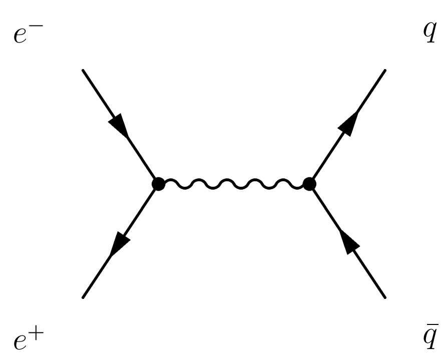
8.1 The basic process: $e^+e^- \to q\bar{q}$

In Sect. 5.10 we calculated the cross section for $e^+e^- \to \mu^+\mu^-$ and found

$$\sigma^{e^+e^- \to \mu^+\mu^-} = \frac{4\pi\alpha_{\mathrm{em}}^2}{3s} = \frac{86.9\,\mathrm{nbGeV}^2}{s} \tag{8.1}$$

where the finite electron and muon masses have been neglected. Here, we consider the basic process $e^+e^- \to q\bar{q}$. In principle, the same Feynman diagram contributes:





The only differences are the fractional electric charges of the quarks and the fact that the quarks appear in $N_c = 3$ different colors which cannot be distinguished by measurement. Therefore, the cross section is increased by a factor N_c. For the quark-antiquark case one thus finds (for $m_q = 0$)

$$\sigma_0^{e^+e^-\to q\bar{q}} = \frac{4\pi\alpha_{\text{em}}^2}{3s} e_q^2 N_c = \frac{86.9\,\text{nbGeV}^2}{s} e_q^2 N_c. \tag{8.2}$$

We assume $\sum_q \sigma^{e^+e^-\to q\bar{q}} = \sigma^{e^+e^-\to\text{hadrons}}$, i. e. the produced quark-antiquark pair will always hadronize.

With Eq. (8.1) and (8.2), neglecting mass effects and gluon as well as photon radiation, we find the following ratio:

$$R = \frac{\sigma^{e^+e^-\to\text{hadrons}}}{\sigma^{e^+e^-\to\mu^+\mu^-}} = N_c \sum_q e_q^2. \tag{8.3}$$

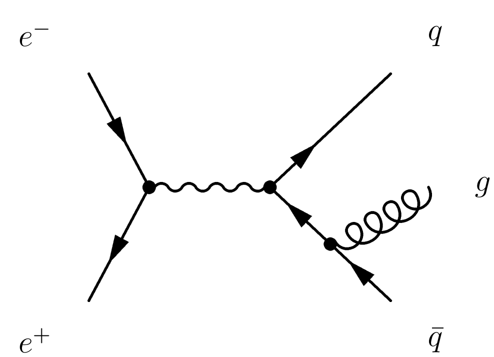
The sum runs over all flavors that can be produced at the available energy. For E_{CM} below the Z peak and above the Υ resonance (see Fig. 8.3), we expect[1]

$$R = N_c \sum_q e_q^2 = N_c \left[\underbrace{\left(\frac{2}{3}\right)^2}_{u} + \underbrace{\left(-\frac{1}{3}\right)^2}_{d} + \underbrace{\left(-\frac{1}{3}\right)^2}_{s} + \underbrace{\left(\frac{2}{3}\right)^2}_{c} + \underbrace{\left(-\frac{1}{3}\right)^2}_{b} \right] = N_c \frac{11}{9}.$$

This is in good agreement with the data for $N_c = 3$ which confirms that there are three colors. At the Z peak one also has to include coupling to the Z boson which can be created from the e^+e^- pair instead of a photon. The small remaining difference visible in the plot is because of QCD corrections for gluon radiation (see later).

8.1.1 Singularities

In order to achieve a better prediction, we have to go beyond the basic QED prediction by including QCD dynamics: Consider the production of a quark-antiquark pair along with a gluon:



[1] Recall that the top quark mass is $m_t \approx 171\,\text{GeV}$.

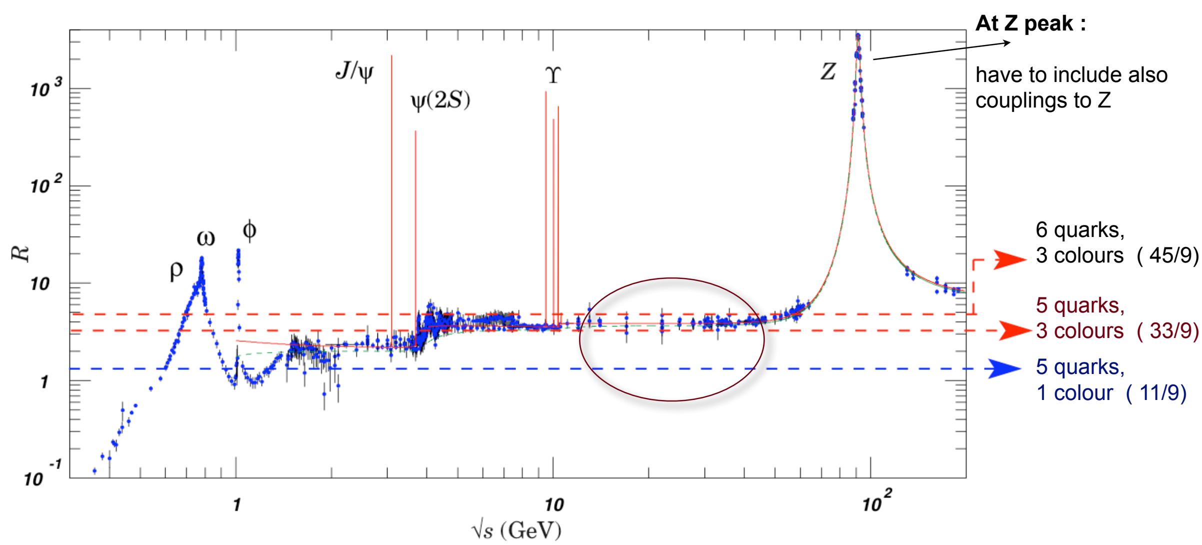


Figure 8.3: *Ratio* $R = \sigma^{e^+e^-\to hadrons}/\sigma^{e^+e^-\to\mu^+\mu^-}$ *as a function of the center of mass energy.* As expected by Eq. (8.3), there is roughly no energy dependence besides various resonances. The data confirm that there are three quark colors.

We define the kinematic variables

$$x_i = 2\frac{p_i \cdot Q}{Q^2} = \frac{E_i^*}{E_{\text{beam}}} \tag{8.4}$$

where $Q = p_{e^+} + p_{e^-} = p_{\gamma/Z}$ and $Q^2 = s$. Energy-momentum conservation ($\sum_i p_i = Q$) requires that, in this case,

$$x_q + x_{\bar{q}} + x_g = 2 \tag{8.5}$$

$$x_i \leq 1. \tag{8.6}$$

One can calculate the differential cross section

$$\frac{d^2\sigma}{dx_q dx_{\bar{q}}} = \sigma_0 \frac{\alpha_s}{2\pi} C_F \frac{x_q^2 + x_{\bar{q}}^2}{(1 - x_{\bar{q}})(1 - x_q)} \tag{8.7}$$

where $C_F = 4/3$ is the color factor of the fundamental representation. Note that this expression is singular for

- $x_q \to 1$, e. g. $\bar{q}||g$,
- $x_{\bar{q}} \to 1$, e. g. $q||g$, and for
- $(x_q, x_{\bar{q}}) \to (1,1)$, e. g. $x_g \to 0$.

Because of the kinematic constrains imposed by energy-momentum conservation (Eq. (8.5) and (8.6)), the allowed region (part of which we have to integrate Eq. (8.7) over to find a cross section) for a $\gamma^\star \to q\bar{q}g$ event is of the form shown in Fig. 8.4.

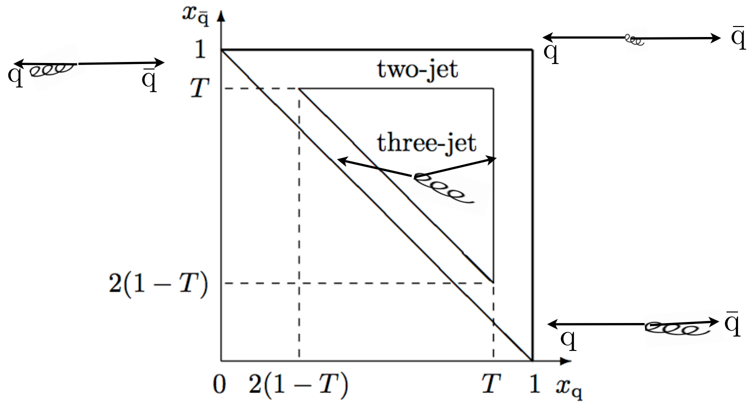
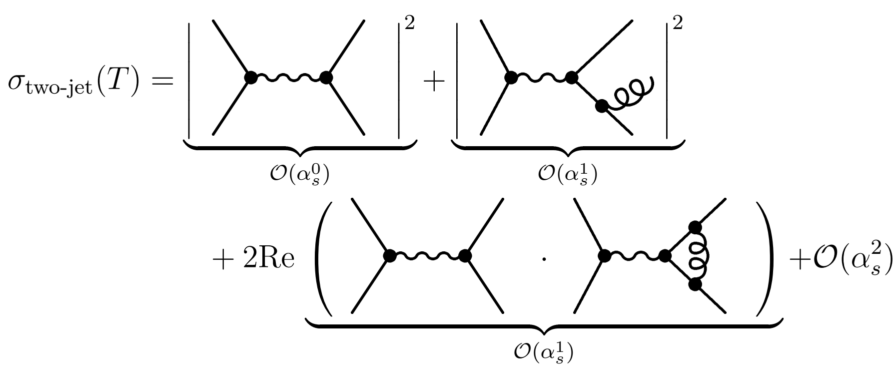


Figure 8.4: *A Dalitz plot showing the allowed region of the x_q-$x_{\bar{q}}$ plane for a $\gamma^\star \to q\bar{q}g$ event with massless partons.* The thick lines indicate the singularities where $x_q = 1$ and $x_{\bar{q}} = 1$. Their intersection marks the position of the soft gluon singularity: $x_g = 0$. The concept of jets will be introduced later, but it is clear that there has to be at least a certain angle between the gluon and the quarks if the jet in gluon direction is to be detected separately. Source: [27, p. 74].

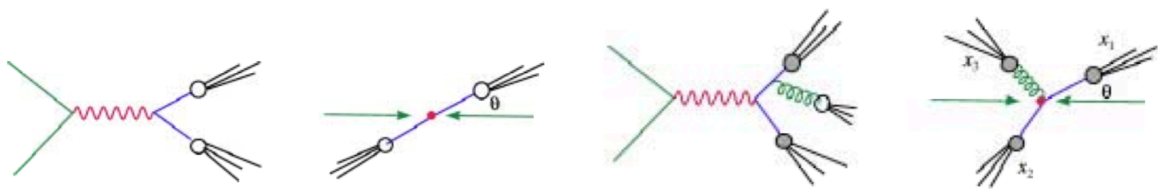
So, how does one deal with these singularities to find a meaningful expression for the cross section to first order? Consider first the two-jet cross section. Two jets are detected when the gluon is either very soft or almost parallel to the quarks such that only two energy flows back-to-back can be measured. Including interference terms, the cross section in the case of an unresolved gluon is given by (integration over two-jet region, see Fig. 8.4)



$$= \sigma_0 \left(1 + \alpha_s f(T) + \mathcal{O}(\alpha_s^2)\right)$$

where T stems from the criterion separating the two- and three-jet regions of the Dalitz plot: $\max\{x_q, x_{\bar{q}}, x_g\} < T$. The singularities of the second and third term cancel and the result is a function of the parameter T. However, our problem is not yet resolved, since $\lim_{T\to 1} f(T) = -\infty$.

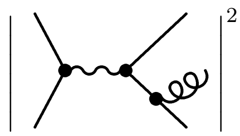
If the gluon can be resolved, a three-jet event is detected and the integration is over the



(a) (b)

Figure 8.5: *Hadronization of quarks and gluons.* Diagrams of the processes $e^+e^- \to q\bar{q} \to$ hadrons (a) and $e^+e^- \to q\bar{q}g \to$ hadrons (b). The RHS shows the situation in the center of mass frame. Source: [39, p. 5 and 6].

three-jet region of Fig. 8.4:

$\sigma_{\text{three-jet}}(T) =$  $+ \mathcal{O}(\alpha_s^2) = \sigma_0 \alpha_s g(T) + \mathcal{O}(\alpha_s^2)$

where $\lim_{T\to 1} g(T) = +\infty$. Combining the two-jet and three-jet cross sections, one finds that the dependence on T cancels yielding a finite result for the total cross section:

$$\begin{aligned}\sigma_{\text{tot}} &= \sigma_{\text{two-jet}} + \sigma_{\text{three-jet}} + \cdots = \sigma_0 \left(1 + \alpha_s \left[f(T) + g(T)\right] + \mathcal{O}(\alpha_s^2)\right) \\ &= \boxed{\sigma_0 \left(1 + \frac{3}{4} C_F \frac{\alpha_s}{\pi} + \mathcal{O}(\alpha_s^2)\right)}.\end{aligned}$$

8.2 Jets and other observables

We now focus on entities actually observable in experiment. We do not observe free quarks but only colorless hadrons produced by the "hadronization/fragmentation" of quarks and gluons. For instance, the processes discussed so far can be visualized as in Fig. 8.5.

The anatomy of the process $e^+e^- \to Z \to$ hadrons is sketched in Fig. 8.6. The things that we can do based on such a process include

- Measure α_s,
- Measure the masses of (heavy) quarks,
- Measure gluon self-coupling,
- Study hadronization and particle correlations, and
- Study the transition between the non-perturbative and perturbative regime and the properties of quark or gluon jets.

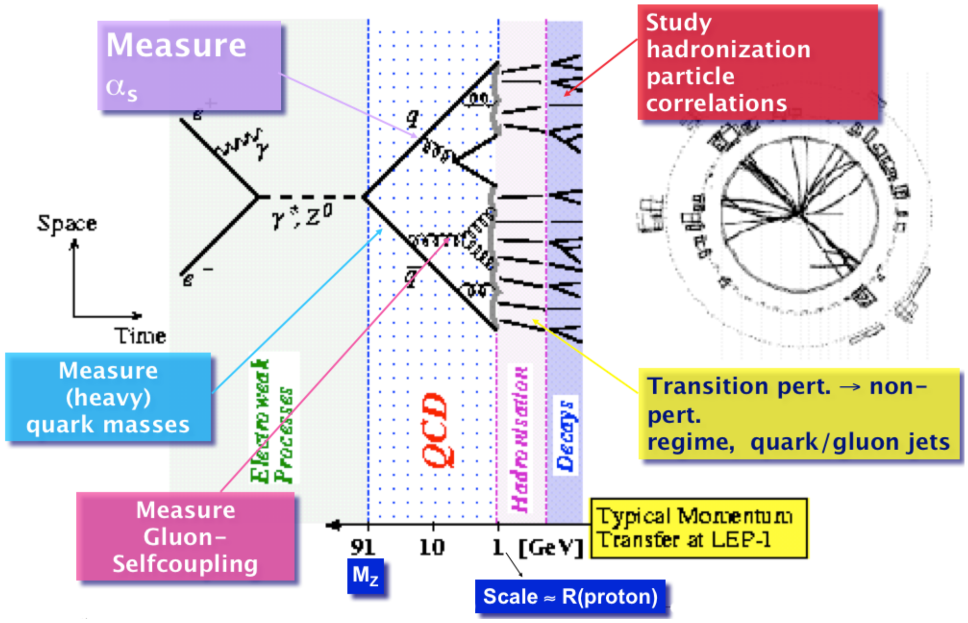


Figure 8.6: *The anatomy of the process $e^+e^- \to \gamma^\star, Z \to$ hadrons.* Source: [40, p. 13, modified].

8.2.1 Jet algorithms

Let us turn to the question: What is a jet?

Fig. 8.7 shows a multi-jet event recorded by the ALEPH detector. A possible verbal definition of "jet" would be "cluster/spray of particles (tracks, calorimeter deposits) or flow of energy in a restricted angular region". Jets are the connection between the quarks and gluons of QCD and the signals actually measured in the detectors. If we are to extract this information from the data, we clearly need some kind of algorithmic definition of this concept: In the "final state" of many interesting interactions there are quarks and gluons. These are the fundamental particles of QCD. Confinement (see p. 150) means that in the detector we see hadrons (together with leptons and photons), but not single quarks or gluons. At energies much larger than $\Lambda^{\mathrm{QCD}}(\sim 1\,\mathrm{GeV})$ these hadrons appear confined into jets. Our aim is to compare the predictions based on partons (quarks and gluons) with the measurements on hadrons. Therefore, we need an algorithmic definition of a jet which

1. can be applied both to data *and* predictions and
2. gives a close relationship between partons and jets of hadrons.

The basic requirement for such an algorithm is applicability at all relevant levels of theory and experiment: partons, stable particles, measured objects (calorimeter objects, tracks, etc.) while always finding the same jet. Furthermore, the algorithm has to be independent of the very details of the detector, e. g. the granularity of the calorimeter, the energy

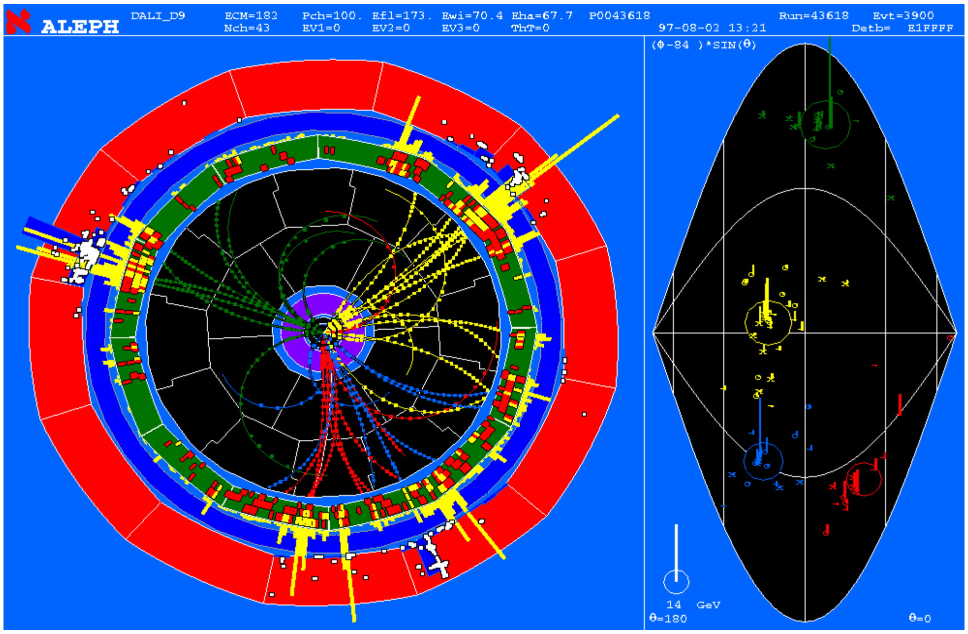


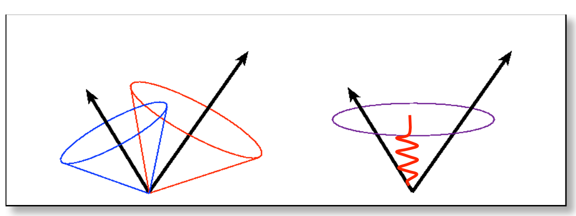
Figure 8.7: *Multi-jet event in the ALEPH detector.*

response, etc. Finally, it should also be easy to implement. In order that we can test QCD predictions, there has to be a close correspondence between the jet momentum (i. e. energy, momentum, and angle) at the parton level and at the hadron level.

NB: Other requirements might strongly depend on the specific application/measurement being performed: For a precision test of QCD there may be requirements which for an analysis of W decays or searches for new physics might not be necessary (e. g. infrared safety).

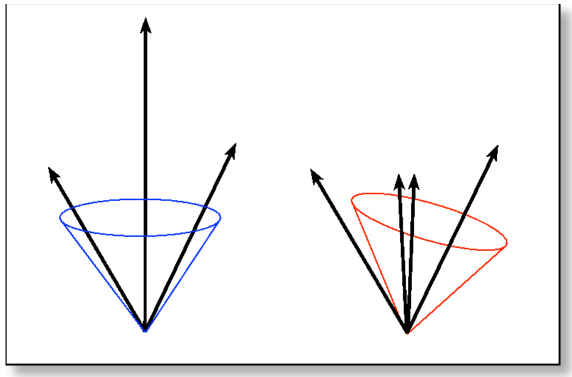
Further requirements come from QCD: We want to compare perturbative calculations with the data. Therefore, the algorithm has to be insensitive to "soft physics" which requires infrared safety and collinear safety.

Infrared safety requires that the configuration must not change when adding a further soft particle. This would be violated by the following behavior[2]:



Collinear safety means that the configuration does not change when substituting one particle with two collinear particles. The problem is visualized in this figure:

[2]Source: [41, pp. 4].



Infrared and collinear safety yield algorithms with the required insensitivity to soft physics: They guarantee the cancellation (between real and virtual emission diagrams) of the infrared and collinear divergencies in *every order* of perturbation theory.

8.2.1.1 Examples of jet algorithms

There are two classes of jet algorithms in use. Algorithms of the class "JADE" are used mainly for e^+e^- annihilations (i. e. for the analysis of events with purely leptonic initial states), but more recently, this class of jet algorithms is also used at hadron colliders. We will concentrate on this class here. The second class of jet algorithms is called "CONE" and is mainly used at hadron colliders with some applications also at e^+e^- colliders.

JADE class algorithms are characterized by

- a "metric" y_{ij} (measure of distance in momentum space),
- a criterion of resolution y_{cut}, and a
- procedure of recombination.

The original definition of the metric from the JADE experiment at PETRA reads

$$y_{ij} = \frac{2E_iE_j(1-\cos\theta_{ij})}{E_{CM}^2} \approx \frac{m_{ij}^2}{E_{CM}^2} \tag{8.8}$$

where m_{ij} is the invariant mass of the particle pair (i,j), see Fig. 8.8(a) Given this metric and a pre-defined resolution y_{cut}, the algorithm is:

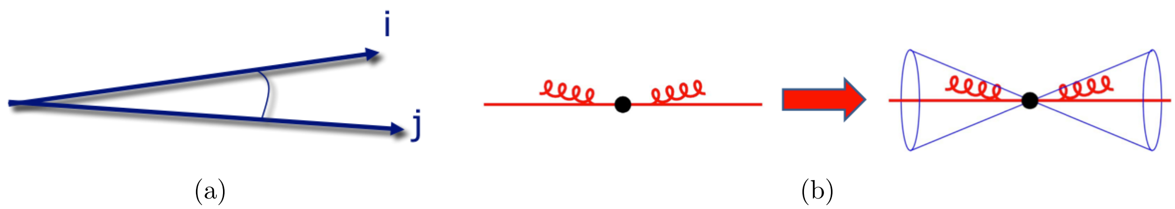
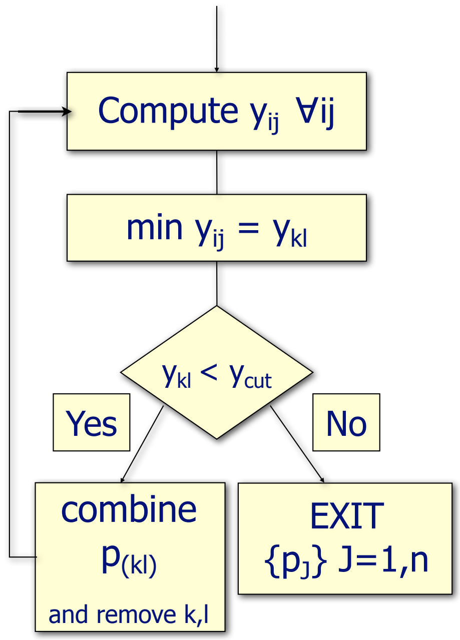


Figure 8.8: *Particle pair (a) and recombination of close particles (b).*



First, all distances y_{ij} between pairs (i, j) are calculated. Then we search for the smallest invariant mass: $\min_{(i,j)} y_{ij} = y_{kl}$. The fact that y_{kl} is the smallest distance in momentum space of all pairs (of particles or, in the subsequent steps, also pseudo-particles) means that the pair (k, l) is either nearly parallel, $\theta_{kl} = 0$, or one or both of the particles are very soft, see Fig. 8.8(b). If the distance cannot be resolved, $y_{kl} < y_{\text{cut}}$, the two particles (k, l) are combined (clustered) into one new pseudo-particle with the combined momentum $p_{(kl)} = p_k + p_l$ (i. e. momentum is conserved), see Fig. 8.9(a). This is the so-called E scheme. Applying this algorithm will reduce complex events until there is a certain number of jets left, as is sketched in Fig. 8.9(b).

The proposed algorithm has some very useful characteristics:

- Infrared safety,

$$y_{ij} \to 0 \text{ for } E_i \text{ or } E_j \to 0,$$

  and collinear safety,

$$y_{ij} \to 0 \text{ for } \theta_{ij} \to 0,$$

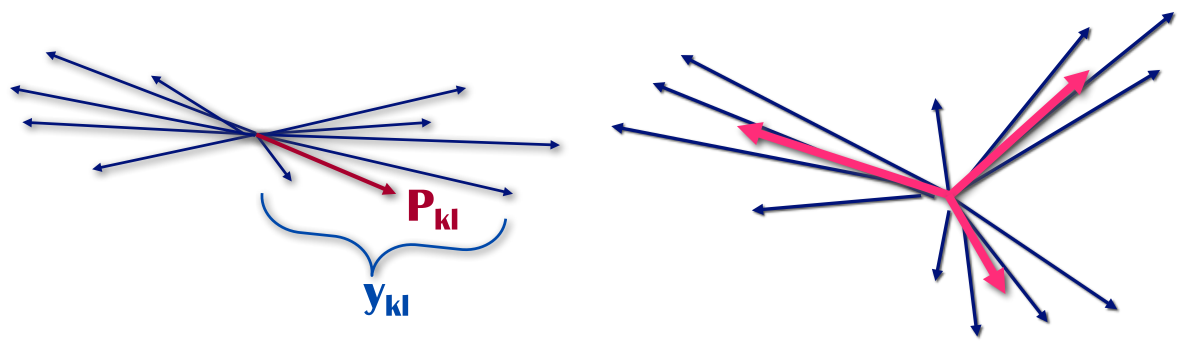


Figure 8.9: *Recombination of particle pair with small invariant mass (a) and reduction of particle pattern to jets (b).*

(in every order of perturbation theory, see p. 163 and Eq. (8.8)).

- All particles are assigned to one and only one jet.
- The algorithm's sequence does not depend on y_{cut}.
- The number of found jets is a monotonic function of y_{cut}.

For the discussed algorithm there is no need to stick to the JADE metric of Eq (8.8); alternative metrics can be introduced. For instance, the DURHAM metric is

$$y_{ij} = \frac{2\min\left(E_i^2, E_j^2\right)(1-\cos\theta_{ij})}{E_{CM}^2} \approx \frac{k_\perp^2}{E_{CM}^2} \tag{8.9}$$

where $k_\perp$ is the transverse momentum of the less energetic particle with respect to the more energetic one. The introduction of this metric was motivated by perturbative QCD calculations: It allows for the resummation of large logarithms of the type $\ln^m(y_{\text{cut}})$ in all orders of perturbation theory (see e. g. [27, pp. 139]). These logarithms appear order-by-order in the expressions for jet cross sections, jet rates, etc.

Now is a good time to recall the Dalitz plot of Fig. 8.4 where we separated a two-jet and a three-jet region. The algorithmic jet definition we have developed enables us to define the thee-jet region: Apply the jet algorithm until three jets are left. If the distance between the jets can be resolved, $\min_{(i,j)}(y_{ij}) > y_{\text{cut}}$, there are three jets, else it is a two-jet event. The shape of the found three-jet region is somewhat different, since y_{ij} also depends on the angle θ_{ij}, see Fig. 8.10.

In order to compare the analyzed data to the predictions of QCD, we need perturbative predictions for jet rates. For the reaction $e^+e^- \to$ hadrons the leading order predictions are as follows. For the JADE algorithm we have

$$\sigma^{LO}_{\text{three-jet}}(y_{\text{cut}}) = \sigma_0 C_F \frac{\alpha_s}{2\pi}\left[2\ln^2 y_{\text{cut}} + 3\ln y_{\text{cut}} - \frac{\pi^2}{3} + \frac{5}{2} - f(y_{\text{cut}})\right] \tag{8.10}$$

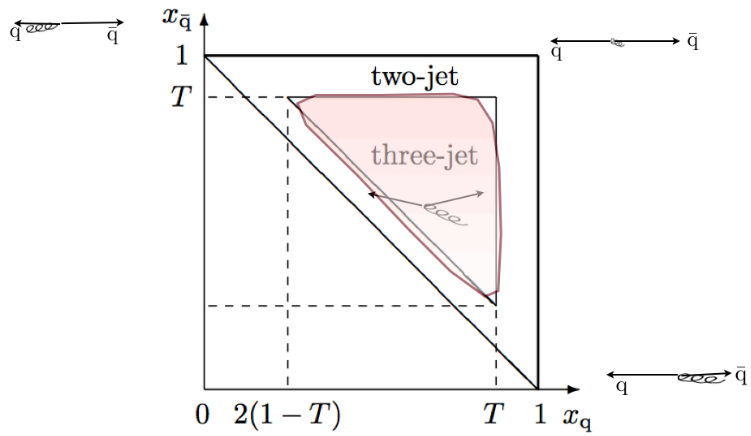


Figure 8.10: *A Dalitz plot showing the allowed region of the x_q-$x_{\bar{q}}$ plane for a $\gamma^\star \to q\bar{q}g$ event with massless partons.* The three-jet region is determined using an algorithmic jet definition.

where $f(y_{\text{cut}}) \to 0$ for $y_{\text{cut}} \to 0$. The prediction for the DURHAM algorithm is the same, except for the factor "2" in front of "$\ln^2 y_{\text{cut}}$". In simple terms, the logarithm terms arise because the vertex where the gluon is radiated off contributes a factor proportional to $\alpha_s/E_{\text{gluon}}$ to the integrand which upon integration yields $\int_{y_{\text{cut}}} dE/E$.

Resummation[3] with the DURHAM algorithm looks as follows. First, let

$$R_2(y_{\text{cut}}) = \frac{\sigma_{\text{two-jet}}}{\sigma_{\text{tot}}}.$$

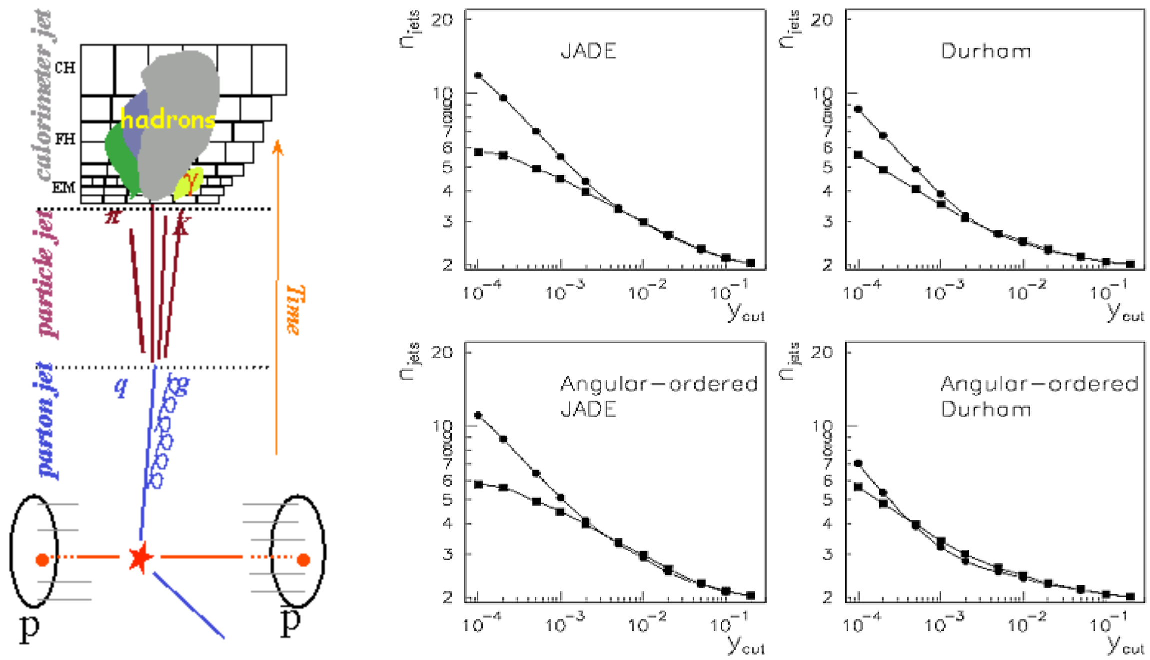
One can show that

$$\begin{aligned} R_2 &= \exp\left\{ - \int\limits_{sy_{\text{cut}}}^{s} \frac{dq^2}{q^2} \frac{C_F \alpha_s(q^2)}{2\pi} \left[\ln \frac{s}{q^2} - \frac{3}{2} \right] \right\} \\ &\approx 1 - \int\limits_{sy_{\text{cut}}}^{s} \frac{dq^2}{q^2} \frac{C_F \alpha_s(q^2)}{2\pi} \ln \cdots + \cdots \approx 1 - \frac{C_F \alpha_s}{2\pi} \ln^2 y_{\text{cut}} + \ldots \end{aligned}$$

where $R_2(y_{\text{cut}} \to 0) = 0$. This is an example of the characteristics an algorithm has to have if you want to perform "high-precision" perturbative QCD calculations. Now there also exists an algorithm of the k_t (DURHAM) type for hadron colliders, see later.

To conclude this section, we turn to the comparison of jet algorithms. There is no such thing as the best "benchmark" variable which allows to compare algorithms in a general manner. The suitability and performance of an algorithm depends very strongly on the performed analysis. Usually we would like to have a good resolution of energies and angles

[3]Resummation in QCD is analogous to the treatment of infrared divergencies in QED, see e. g. [14, pp. 202]



(a) (b)

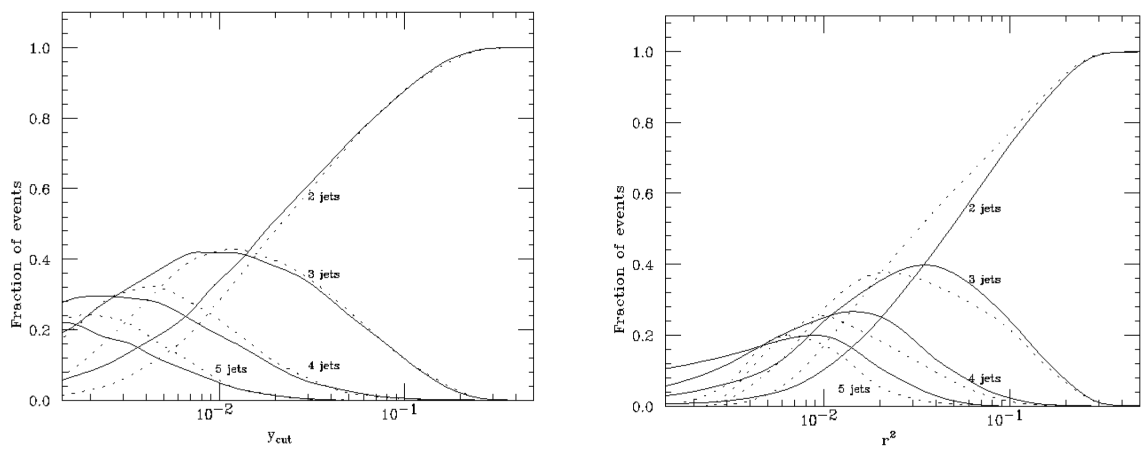
Figure 8.11: *Visualization of levels at which the algorithms have to deliver good resolution (a) and comparison of jet algorithms (b).* The mean number of jets is displayed as a function of y_{cut}. The parton level is denoted by squares and the hadron level by circles. The results were obtained by HERWIG Monte Carlo simulation at $E_{CM} = M_z$. Source: [34, p. 28]. For details compare [34, pp. 7].

of the jets at the parton, hadron, and detector levels (see Fig. 8.11(a) for a visualization), as well as a good efficiency and purity to find a certain number of jets at a certain level. For some jet algorithms, the mean number of jets as a function of y_{cut} at the hadron and parton levels, as obtained by HERWIG (Hadron Emission Reactions With Interfering Gluons) Monte Carlo simulation at $E_{CM} = M_Z$, is compared in Fig. 8.11(b). Another comparison[4] is shown in Fig. 8.12. The fraction of events with 2 jets which have 2, 3, 4, and 5 sub-jets is given as a function of y_{cut} or r^2, the radius fraction sqared, respectively. The data stem from HERWIG Monte Carlo simulations at $E_{CM} = 1.8\,\text{TeV}$ with $75\,\text{GeV} < E_t(\text{jet } 2) < 100\,\text{GeV}$. Data from a k_t algorithm are shown in Fig. 8.12(a) while the results in Fig. 8.12(b) come from a CONE algorithm with radius $R = 0.7$.

8.2.2 Event shape variables

The introduced jet algorithms can be used as a starting point to define more refined observables that capture the event topologies.

[4]More on k_t and CONE algorithms can be found in [41].

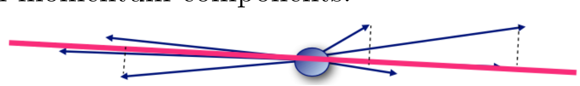


(a) (b)

Figure 8.12: *Comparison of k_t (a) and CONE (b) algorithms.* Legend: —parton level, $\cdots$ calorimeter level. The fraction of two-jet events with 2, 3, 4, and 5 sub-jets is given as a function of y_{cut} or r^2. The data is generated by HERWIG Monte Carlo simulations at $E_{CM} = 1.8\,\text{TeV}$ with $75\,\text{GeV} < E_t(\text{jet 2}) < 100\,\text{GeV}$.

An example for an event shape variable is the *differential two-jet rate*. The definition goes as follows: Apply the DURHAM algorithm until exactly three jets are left (in contrast to the possibility to run the algorithm until a certain resolution is reached). Then take the minimal distance y_{ij} of all pairs (i,j) and call it y_{23} (or y_3): $\min_{(i,j)} y_{ij} = y_{23} = y_3$. This gives one value for each event. The distribution of these values for all events is an "event-shape distribution". Therefore, one can plot the differential cross section as in Fig. 8.13. There is one histogram entry for each event. The data come from hadronic Z decays at LEP. Observe that two-jet events are more likely than three-jet events. The perturbative regime is limited to high gluon energies. Hadronization effects that have to be phenomenologically modeled spoil the perturbative calculations at low y_3 values.

As another example for an event-shape variable, let us consider *thrust*. It was invented around 1978 and first used at PETRA. The idea is to select the axis that maximizes the sum of the longitudinal momentum components:



The thrust of an event is then defined as

$$T = \max_{\vec{n}} \frac{\sum_i |\vec{p}_i \cdot \vec{n}|}{\sum_i |\vec{p}_i|}$$

where $|\vec{n}| = 1$ and the sum runs over the three-momenta of all final states. The thrust axis is defined by the vector $\vec{n}_T$ for which the maximum is obtained. This definition means that for $T = 1$ the event is perfectly back-to-back while for $T = 1/2$ the event is spherically symmetric:

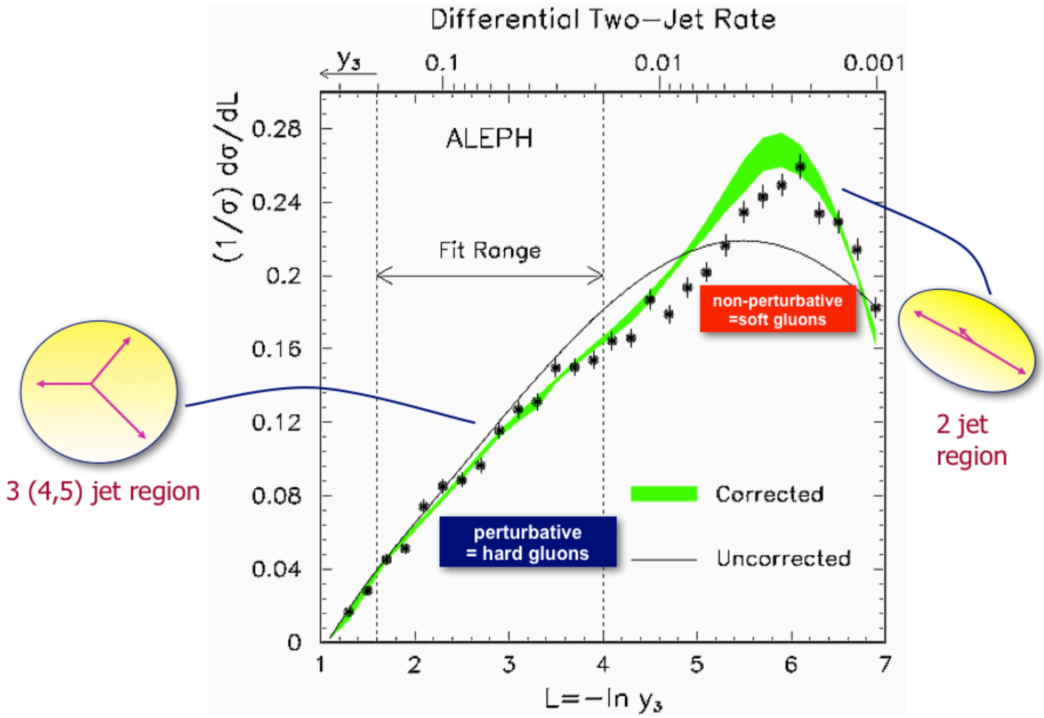
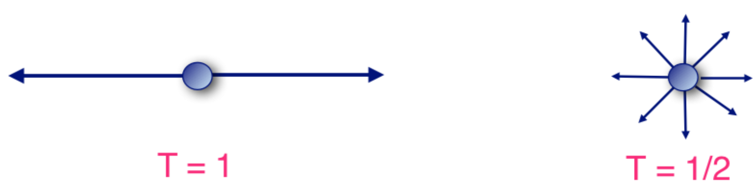


Figure 8.13: *Differential two-jet rate for hadronic Z decays at LEP.*



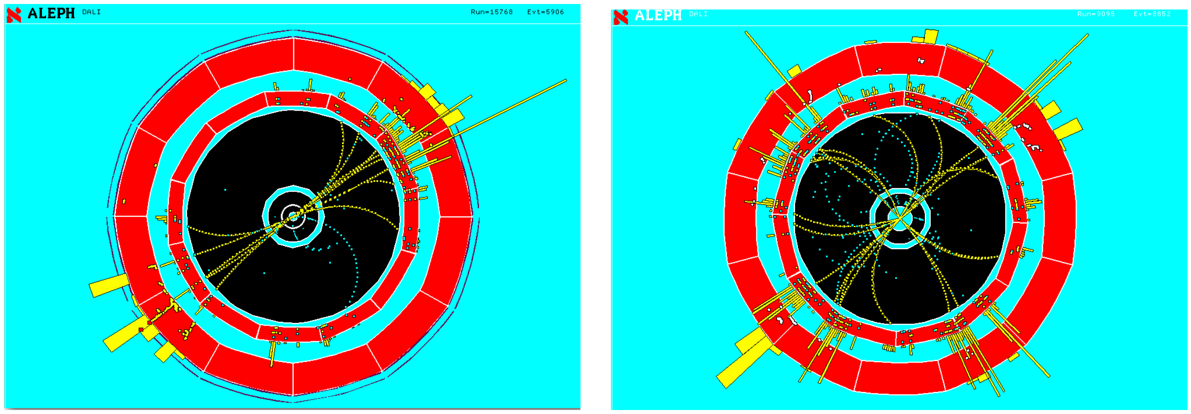
This point is also illustrated with ALEPH data of Z decays in Fig. 8.14 where Fig. 8.14(a) corresponds to $T \to 1$ and Fig. 8.14(b) to $T \to 1/2$. The corresponding event-shape distribution is shown in Fig.8.15 (compare also the differential two-jet rate event-shape distribution in Fig. 8.13).

There are further event-shape variables suitable for different purposes. Some examples are given in the following.

- *Thrust major* T_{major}: The thrust major vector $\vec{n}_{\text{Ma}}$ is defined in the same way as the thrust vector $\vec{n}_T$, but with the additional condition that $\vec{n}_{Ma}$ must lie in the plane perpendicular to $\vec{n}_T$:

$$T_{\text{major}} = \max_{\vec{n}_{\text{Ma}} \perp \vec{n}_T} \frac{\sum_i |\vec{p}_i \cdot \vec{n}_{\text{Ma}}|}{\sum_i |\vec{p}_i|}.$$

- *Thrust minor* T_{minor}: The minor axis is perpendicular to both the thrust axis and



(a) (b)

Figure 8.14: *Event displays of Z decays recorded at ALEPH.* The thrust is nearly 1 for (a) and close to 1/2 for (b).

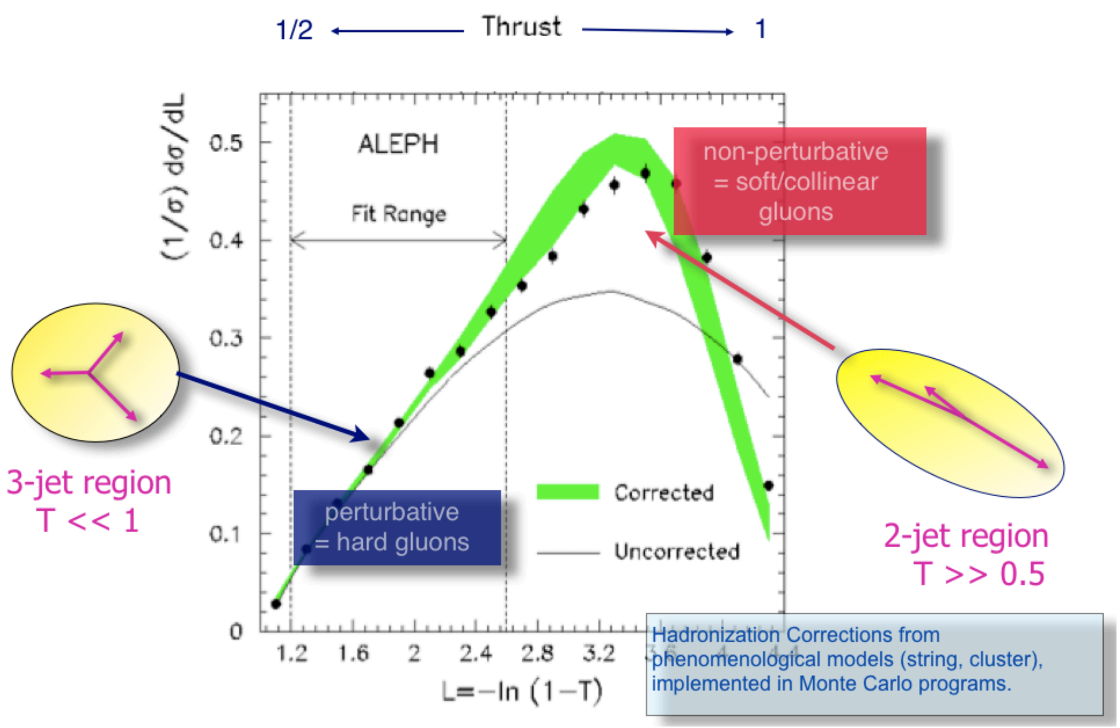


Figure 8.15: *Thrust for hadronic Z decays at LEP.* Observe that the two- and three-jet events are indicated by thrust values close to 1 and 1/2, respectively. Again, in the non-perturbative regime hadronization corrections from phenomenological models are needed.

the major axis: $\vec{n}_{\rm Mi} = \vec{n}_T \times \vec{n}_{Ma}$. The value of the thrust minor is given by

$$T_{\rm minor} = \frac{\sum_i |\vec{p}_i \cdot \vec{n}_{\rm Mi}|}{\sum_i |\vec{p}_i|}.$$

- *Oblateness* O: The oblateness is defined as the difference between thrust major and thrust minor:

$$O = T_{\rm major} - T_{\rm minor}.$$

- *Sphericity* S: The sphericity is calculated from the ordered eigenvalues $\lambda_{i=1,2,3}$ of the quadratic momentum tensor:

$$\begin{aligned} M^{\alpha\beta} &= \frac{\sum_i p_i^\alpha p_i^\beta}{\sum_i |\vec{p}_i|^2},\ \alpha,\beta = 1,2,3 \\ \lambda_1 &\geq \lambda_2 \geq \lambda_3,\ \lambda_1 + \lambda_2 + \lambda_3 = 1 \\ S &= \frac{3}{2}(\lambda_2 + \lambda_3). \end{aligned}$$

  The sphericity axis $\vec{n}_S$ is defined along the direction of the eigenvector associated to λ_1, the semi-major axis $\vec{n}_{\rm sMa}$ is along the eigenvector associated to λ_2.

- *Aplanarity* A: The aplanarity is calculated from the third eigenvalue of the quadratic momentum tensor:

$$A = \frac{3}{2}\lambda_3.$$

- *Planarity* P: The planarity is a linear combination of the second and third eigenvalue of the quadratic momentum tensor:

$$P = \lambda_2 - \lambda_3.$$

- *Heavy jet mass* ρ: A plane through the origin and perpendicular to $\vec{n}_T$ divides the event into two hemispheres, H_1 and H_2 from which the corresponding normalized hemisphere invariant masses are obtained:

$$M_i^2 = \frac{1}{E_{CM}^2}\left(\sum_{k\in H_i} p_k\right)^2,\ i = 1,2.$$

  The larger of the two hemisphere masses is called the heavy jet mass,

$$\rho = \max(M_1^2, M_2^2),$$

  and the smaller is the light jet mass M_L,

$$M_L = \min(M_1^2, M_2^2).$$

- *Jet mass difference* M_D: The difference between ρ and M_L is called the jet mass difference:

$$M_D = \rho - M_L.$$

- *Wide jet broadening* B_W: A measure of the broadening of particles in transverse momentum with respect to the thrust axis can be calculated for each hemisphere H_i using the relation

$$B_i = \frac{\sum_{k \in H_i} |\vec{p}_k \times \vec{n}_T|}{2 \sum_j |\vec{p}_j|}, \ i = 1, 2$$

where j runs over all particles in the event. The wide jet broadening is the larger of the two hemisphere broadenings,

$$B_W = \max(B_1, B_2),$$

and the smaller is called the narrow jet broadening B_N,

$$B_N = \min(B_1, B_2).$$

- *Total jet broadening* B_T: The total jet broadening is the sum of the wide and the narrow jet broadenings:

$$B_T = B_W + B_N.$$

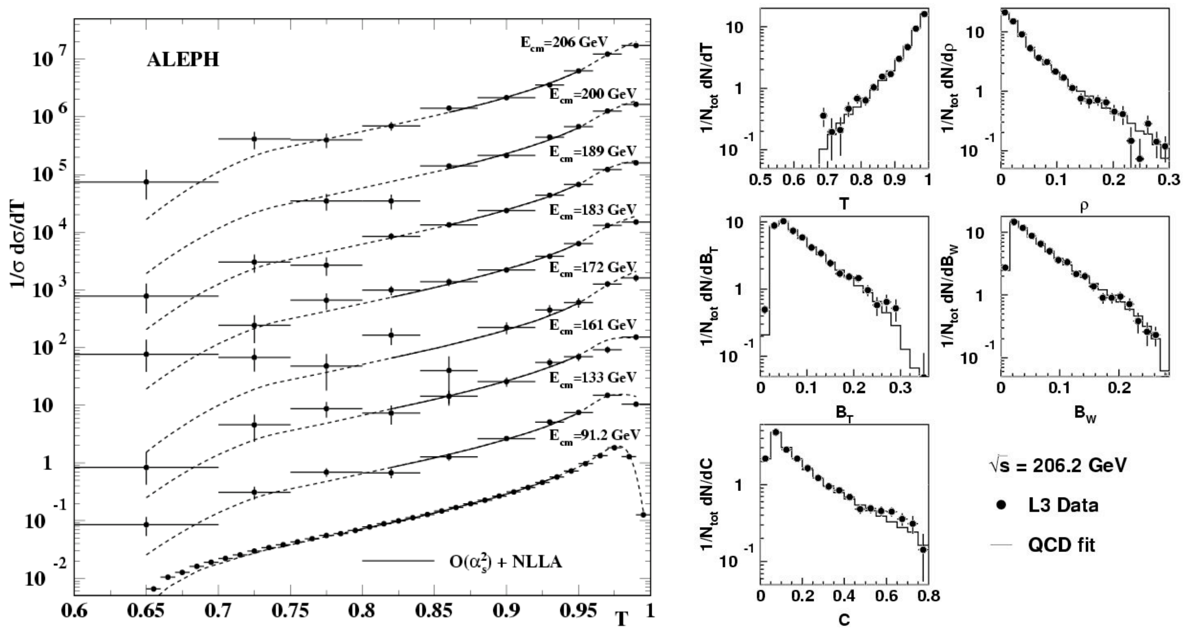
- *C-parameter* C: The C-parameter is derived from the eigenvalues of the linearized momentum tensor $\Theta^{\alpha\beta}$:

$$\Theta^{\alpha\beta} = \frac{1}{\sum_i |\vec{p}_i|} \sum_i \frac{p_i^\alpha p_i^\beta}{|\vec{p}_i|}, \ \alpha, \beta = 1, 2, 3.$$

The eigenvalues λ_j of this tensor define C by

$$C = 3(\lambda_1\lambda_2 + \lambda_2\lambda_3 + \lambda_3\lambda_1).$$

The discussed event-shape variables have been extensively used to analyze LEP data. Examples are given in Fig. 8.16: Fig. 8.16(a) shows thrust predictions and measurements; predictions and data for thrust, heavy jet mass, total jet broadening, wide jet broadening, and the C-parameter are shown in Fig. 8.16(b).



(a) (b)

Figure 8.16: *Comparison of predictions and LEP data for some event-shape variables.* Thrust data are shown for several center of mass energies (a). The other analyses deal with heavy jet mass, total jet broadening, wide jet broadening, and the C-parameter (b).

8.2.3 Applications

Examples for applications of the observables discussed above in this section are measurements of the strong coupling constant α_s (see later, Sect. 8.3), the discovery of quark and gluon jets, measurements of the quark and gluon spin, the triple-gluon vertex, and jet rates or the analysis of differences between quark and gluon jets.

Quark jets were discovered at the SPEAR storage ring (SLAC) [42]. The data are shown in Fig. 8.17. For higher energies particles cluster around an axis and the Monte Carlo simulation based on a jet model fits the data better than the simulation based on an isotropic phase-space model. This is the first observation of a jet structure.

Gluon jets were discovered at PETRA (DESY) [43, 44, 45, 46]. Here, the relevant observable is oblateness (see p. 172). The first three-jet event seen by TASSO is shown in Fig. 8.18(a). In Fig. 8.18(b) one can observe that events at $E_{CM} \sim 30\,\mathrm{GeV}$ exhibit larger oblateness (planar structure) than predicted by models without hard gluon radiation.

When it comes to parton spins the question is: How do you measure the spin of unobservable particles? For spin-1/2 fermions annihilating into a vector boson, conservation of angular momentum predicts a distribution

$$\frac{d\sigma}{d\cos\Theta^*} \sim 1 + \cos^2\Theta^*$$

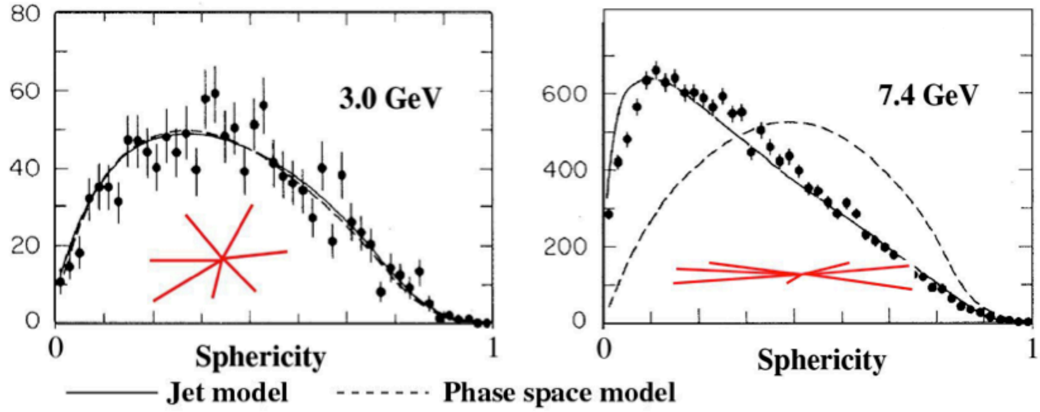
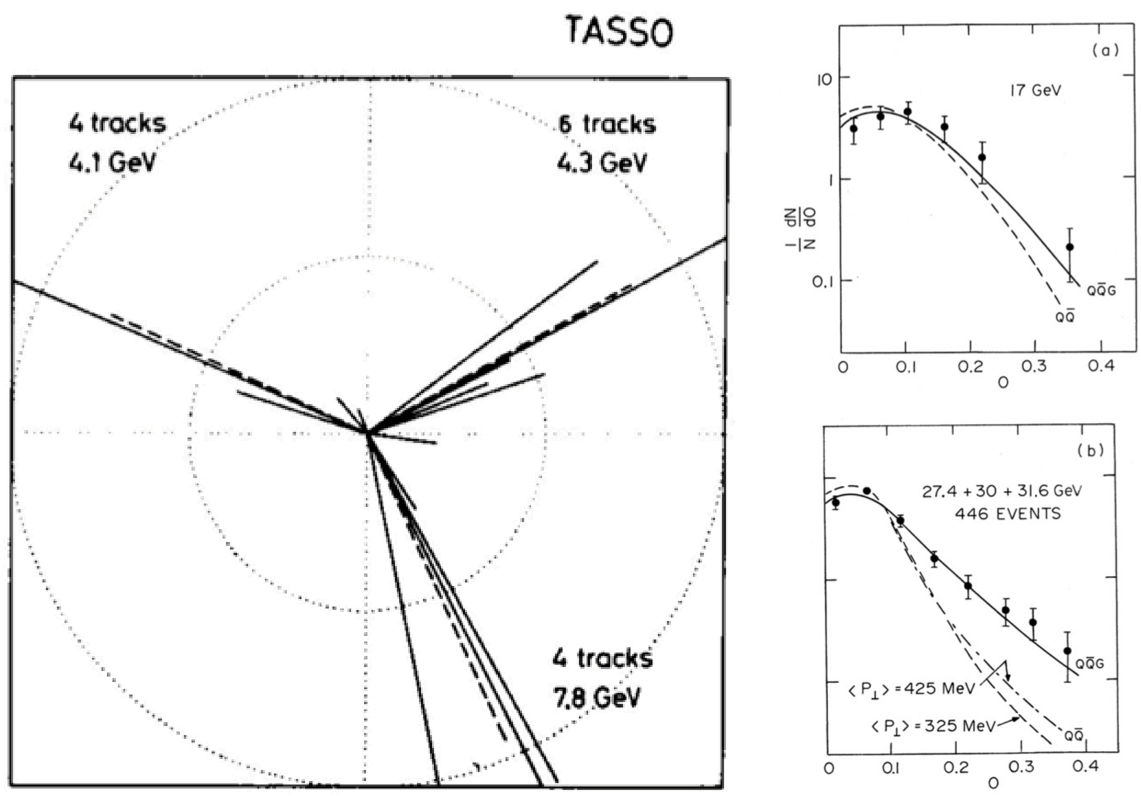
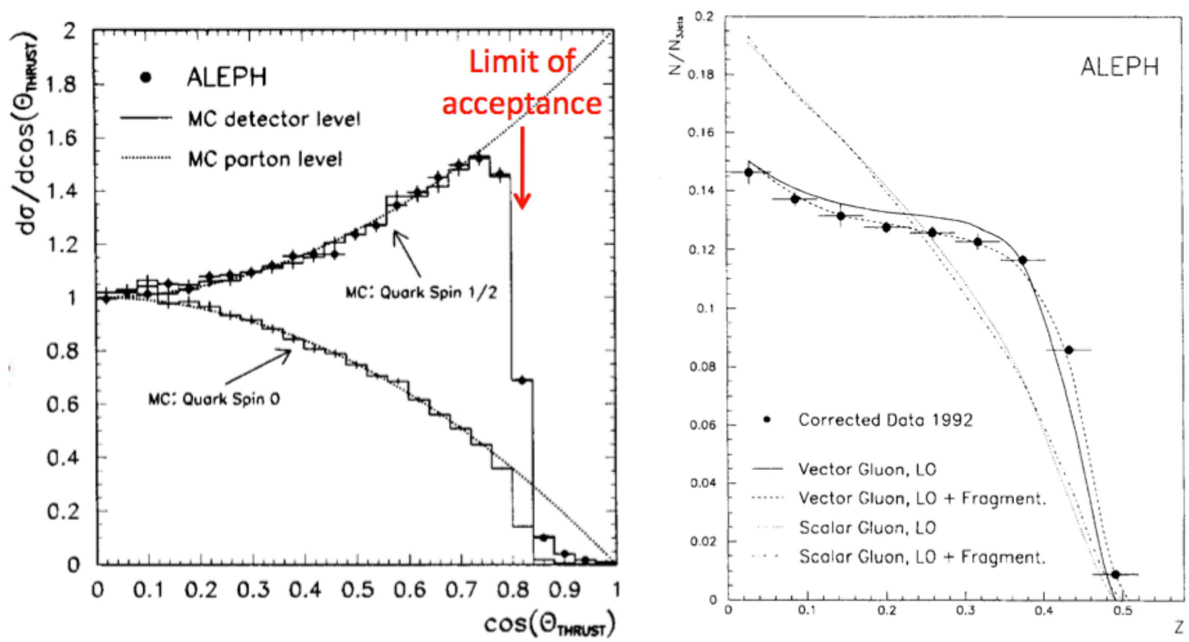


Figure 8.17: *Discovery of quark jets at SPEAR (SLAC).* Observed sphericity (see p. 172) distributions for data, jet model (solid curves) and phase-space model (dashed curves) for $E_{CM} = 3\,\mathrm{GeV}$ (LHS) and $7.4\,\mathrm{GeV}$ (RHS). Source: [42, 38, p. 1611].



(a) (b)

Figure 8.18: *The first three-jet event seen by TASSO (a) and the distribution $N^{-1}dN/dO$ as a function of oblateness, measured at MARK-J (b).* In both figures of (b) the solid curves are the predictions based on the $q\bar{q}g$ model and the dashed curve is based on the standard $q\bar{q}$ model. Source: [44, p. 832].



(a) (b)

Figure 8.19: *Measurements of quark (a) and gluon (b) spin by ALEPH.* Source: [47].

if the final state particles have spin 1/2 and

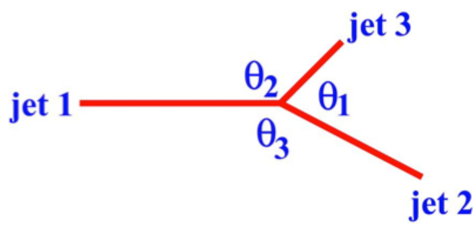
$$\frac{d\sigma}{d\cos\Theta^*} \sim 1 - \cos^2\Theta^* = \sin^2\Theta^*$$

for spin-0 particles in the final state. Therefore, the quark direction has to be measured to measure the quark spin. At LEP energies the thrust axis in two-jet events to a very good approximation aligns with the direction of the primary quarks. Thus, one can take the thrust direction in two-jet events. The exact expression for the spin-1/2 case reads

$$\frac{d\sigma}{d\cos\Theta^*} = \frac{\alpha_{\rm em}^2 e_q^2 \pi N_c}{2s}\left(2 - \beta_q^{*2} + \beta_q^{*2}\cos^2\Theta^*\right)\beta_q^*$$

where $\beta_q^* = \sqrt{1 - 4m_q^2/s} \to 1$ for $m_q = 0$. The resulting angular distribution found by ALEPH [47] is shown in Fig. 8.19(a).The experimental data are compared to a Monte Carlo simulation. The data are in perfect agreement with the spin-1/2 assignment for the quarks while a spin-0 assignment is clearly excluded. The sharp drop in the distribution around $\cos\Theta^* \sim 0.8$ is due to the finite detector acceptance.

Let us turn to the gluon spin. Hard gluon radiation leads to three-jet events. So, after applying a jet algorithm to select the three-jet events, how do we know which one is the gluon jet? Recall that the probability to radiate off a soft gluon is larger than to radiate off a hard gluon. Therefore, for three jets



with energies

$$E_i = E_{CM} \frac{\sin \theta_i}{\sum_j \sin \theta_j},$$

if ordered by energy, $E_1 > E_2 > E_3$, jet 3 is the gluon jet in 75% of the events. Defining the variable

$$Z = \frac{1}{\sqrt{3}}(x_2 - x_3)$$

(recall $x_i = 2E_i/E_{CM}$), the Dalitz plot looks like in Fig. 8.20. The arrow length is proportional to the jet energy. The following cases have to be compared: In the spin-1 case ("vector gluon") the prediction reads

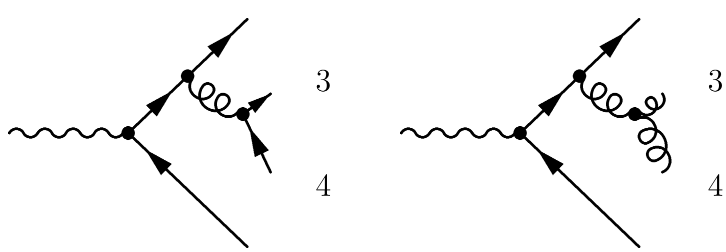
$$\frac{d^2\sigma^v}{dx_1 dx_2} \propto \left[\frac{x_1^2 + x_2^2}{(1-x_1)(1-x_2)} + \text{permutations}\,(1,2,3)\right]$$

while for spin-0 ("scalar gluon")

$$\frac{d^2\sigma^s}{dx_1 dx_2} \propto \left[\frac{x_3^2}{(1-x_1)(1-x_2)} + \text{permutations}\,(1,2,3) - 10\frac{\sum a_q^2}{\sum a_q^2 + v_q^2}\right]$$

where a_q and v_q are the axial-vector and vector couplings of the quarks to the intermediate photon or Z boson and the sums run over all contributing quark flavors. For e^+e^- annihilation via a photon only the vector coupling contributes, on the Z resonance both terms have to be taken into account. The ALEPH data shown in Fig. 8.19(b) clearly indicate that gluons have spin 1.

As we have seen before (see p. 146), the kinematic term of the QCD ($SU(3)$, non-abelian, gluon) Lagrangian contains a three-gluon term yielding a three-gluon vertex, a feature not present in QED ($U(1)$, abelian, photon). The splitting of a radiated gluon into two gluons will lead to a four-jet event, just like the splitting into a quark-antiquark pair:



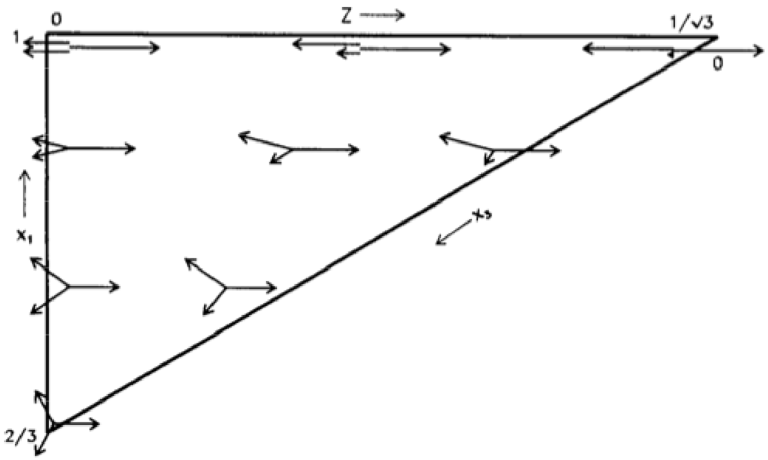


Figure 8.20: *Phase space as function of x_1 and Z for energy-ordered jet configurations,* $x_1 > x_2 > x_3$. The arrow length is proportional to the energy. Source: [47].

For gluon radiation off quarks one finds that the gluon is preferentially polarized in the plane of the splitting process. On the other hand, for a gluon splitting into two gluons there is a positive correlation between the plane spanned by the two new gluons and the polarization of the branching one. Finally, in case a gluon splits into two quarks, the plane defined by the momenta of the two quarks is anticorrelated with the polarization of the splitting gluon. So we conclude that for four-jet events induced by a gluon splitting into a $q\bar{q}$ pair, the distribution of the angle between the plane defined by the two primary quarks and the plane defined by the two secondary quarks should be enhanced around 90° (see Fig. 8.21). However, in a non-abelian theory we have contributions also from the triple-gluon interaction, and in this case the favored angle between the two planes spanned by the primary and secondary partons is rather small. Therefore, the shape of the distribution of this angle is sensitive to the color factors (see Sect. 7.4.1). Like in the three-jet case, it is difficult to distinguish between jets induced by the primary and the secondary partons. However, because of the $1/E$ characteristic of radiated gluons we expect the two secondary particles to be less energetic than the two primary quarks: If the jets are ordered by energy, $E_1 > E_2 > E_3 > E_4$, jets 3 and 4 are more likely to come from the radiated particles. So we arrive at the definition of the angular correlation variable called Bengtsson-Zerwas angle

$$\chi_{\mathrm{BZ}} = \angle\left[(\vec{p}_1 \times \vec{p}_2), (\vec{p}_3 \times \vec{p}_4)\right] = \frac{(\vec{p}_1 \times \vec{p}_2) \cdot (\vec{p}_3 \times \vec{p}_4)}{|(\vec{p}_1 \times \vec{p}_2)||(\vec{p}_3 \times \vec{p}_4)|}$$

where $\vec{p}_i$, $i = 1, \ldots, 4$ are the energy-ordered momenta of the four partons (jets). In Fig. 8.21 LEP measurements of χ_{BZ} are compared with the predictions by QCD on the one hand and an abelian model with three quark colors but no three-gluon coupling on the other. The data agree with QCD being an $SU(3)$ gauge theory rather than an abelian gauge theory.

At the end of our discussion of jet algorithms Fig. 8.12(a) is shown. It displays the fraction

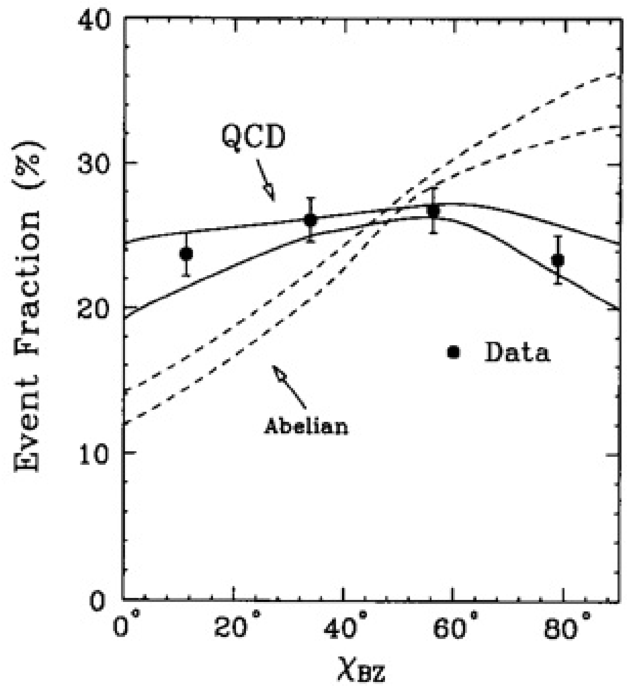


Figure 8.21: *Distribution of* χ_{BZ} *measured by L3.* The predictions for QCD and the abelian model are shown as bands indicating the theoretical uncertainties. Source: [48, p. 233].

of 2-jet events with 2, 3, 4, and 5 sub-jets as a function of y_{cut}. These predictions can be tested comparing measurements at highest LEP energies to Monte Carlo simulations which incorporate leading-order matrix elements for two-jet and three-jet production, plus approximations for multiple soft or collinear gluon radiation. Fig. 8.22 shows the n-jet rate according to the DURHAM (k_t) algorithm as a function of y_{cut}.

We conclude this section with a discussion of the differences between quark and gluon jets. Quark and gluon jets have different coupling strengths to emit gluons (see Sect. 7.4.1 and Fig. 8.23). Therefore, from couplings alone one expects a larger multiplicity in gluon jets of the order $C_A/C_F = 9/4$, and a softening of the momentum distributions for particles coming from the gluon jet. Thus gluon jets are more "soft" and "fat" than quark jets (see Fig. 8.24). Also the scaling violations, i. e. change of multiplicities with energy and momentum scale are different. In Fig. 8.24(d) the CONE algorithm is applied to data of OPAL (LEP) and compared to CDF data. The variable r denotes the radius of the considered cone fraction when R is the radius parameter of the cone algorithm:

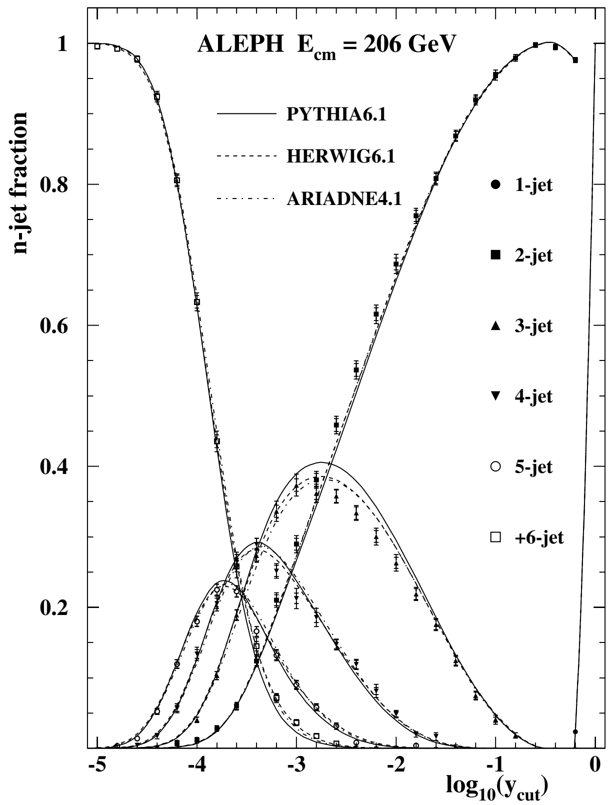


Figure 8.22: *ALEPH measurements of the n-jet rate (DURHAM) as a function of y_{cut}.*

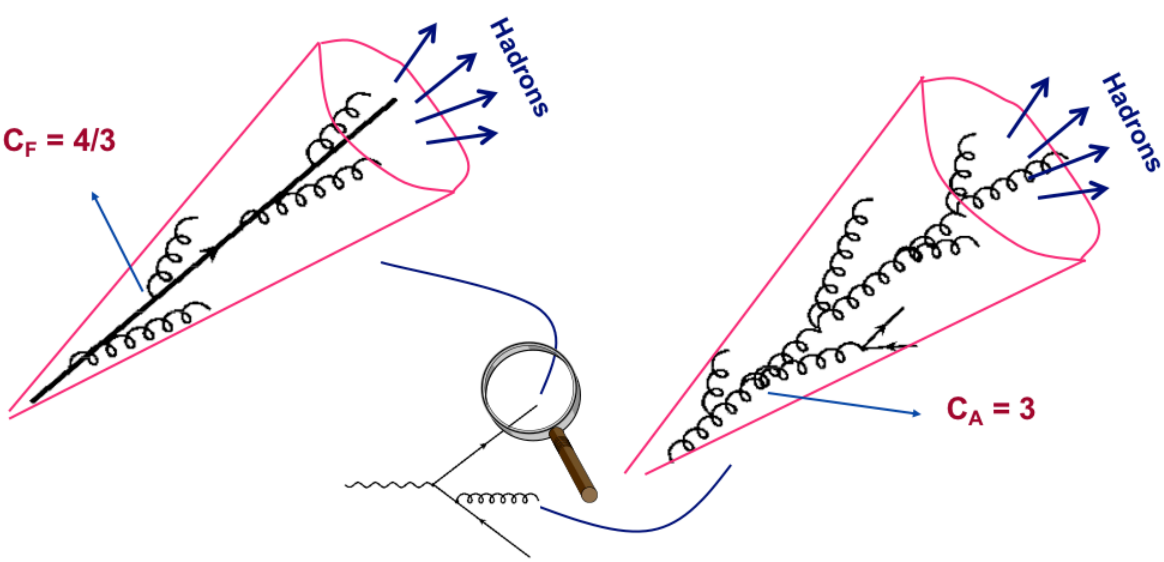
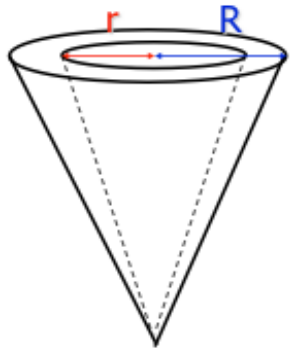


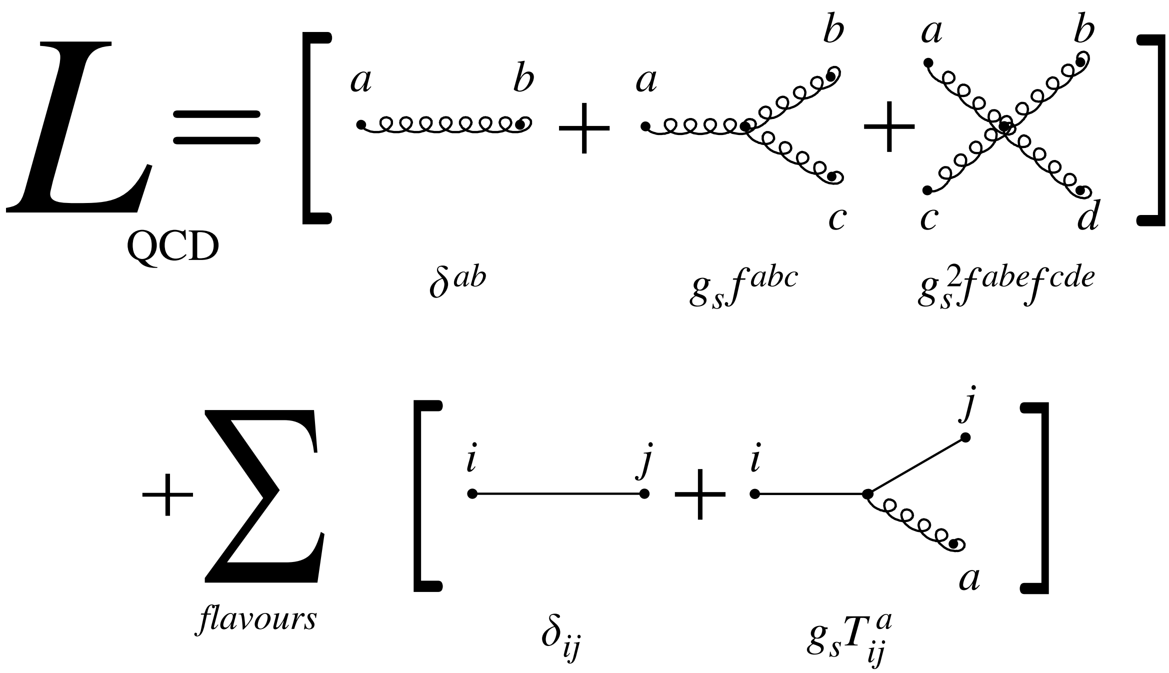
Figure 8.23: *Comparison of quark and gluon jets.* For a discussion of the difference in coupling strength see Sect. 7.4.1.



$\Psi(r)$ denotes the energy in a fraction of the cone and $\Phi(r)$ is defined by $\Phi(r) = d/dr\Psi(r)$.

8.3 Measurements of the strong coupling constant

The QCD Lagrangian is introduced in Sect. 7.4:



Except for the quark masses, there is only one free parameter in it: the strong coupling constant α_s which is discussed in Sect. 7.4.2. Recall that the differential equation for the strong coupling constant depending on the renormalization scale μ, $\alpha_s(\mu^2)$, is

$$\mu^2 \frac{\partial \alpha_s^2(\mu^2)}{\partial \mu^2} = \beta(\alpha_s(\mu^2))$$

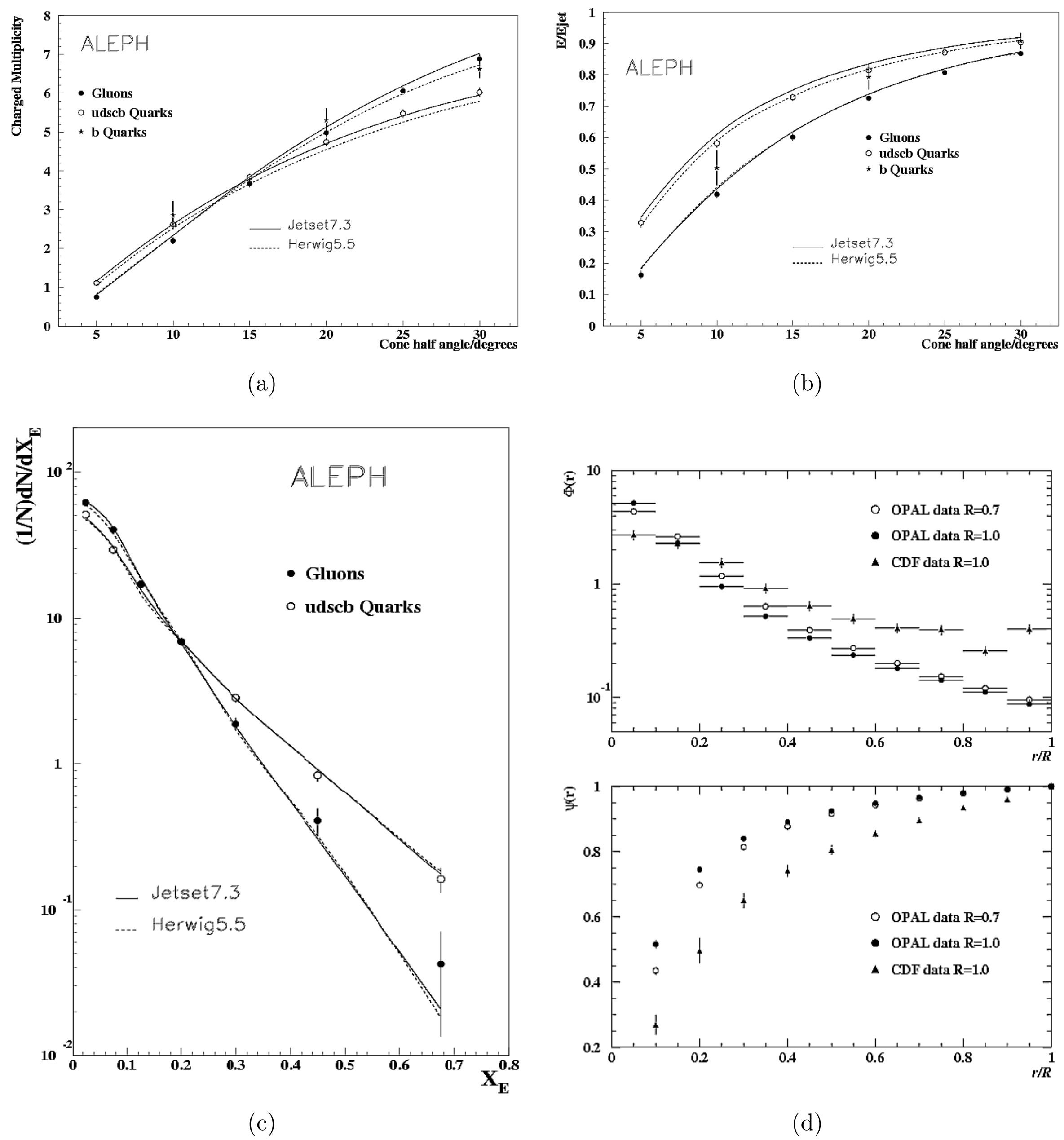


Figure 8.24: *Comparison of quark and gluon jets.* Note that gluon jets are more "soft" and "fat" than quark jets. The variable x_E is the energy fraction of the particles with respect to the jet energy (c). The variable r in (d) denotes the considered fraction of the cone.

which, retaining only the first term of the power expansion for β and absorbing the factor of 4π into the coefficient β_0, yields

$$\alpha_s(Q^2) \equiv \frac{g_s^2(Q^2)}{4\pi} = \frac{1}{\beta_0 \ln(Q^2/\Lambda_{\mathrm{QCD}}^2)}.$$

At that point we also stressed that

$$\beta_0 = \frac{1}{4\pi}\left(11 - \frac{2}{3}n_f\right) > 0 \text{ for (the likely case of) } n_f < 17$$

which makes the effective coupling constant behave like shown in Fig. 7.4.2. The following expansion holds for $\alpha_s(\mu^2)$ (see Eq. (7.43)):

$$\boxed{\alpha_s(\mu^2) \approx \alpha_s(Q^2)\left[1 - \alpha_s(Q^2)\beta_0 \ln\frac{\mu^2}{Q^2} + \alpha_s^2(Q^2)\beta_0^2 \ln^2\frac{\mu^2}{Q^2} + \mathcal{O}(\alpha_s^3)\right]}. \tag{8.11}$$

To measure the coupling strength one uses as many methods as possible in order to demonstrate that QCD really is the correct theory of strong interactions by showing that one universal coupling constant describes all strong interactions phenomena. Consider the perturbative expansion of the cross section for some QCD process:

$$\sigma^{\mathrm{pert}} = \alpha_s(\mu^2)A + \alpha_s^2(\mu^2)\left[B + \beta_0 A \ln\frac{\mu^2}{Q^2}\right] + \mathcal{O}(\alpha_s^3) \tag{8.12}$$

where the coefficients A and B depend on the specific process. So, if only the leading oder (LO) expansion is known, the following holds:

$$\sigma_{\mathrm{LO}}^{\mathrm{pert}} = \alpha_s(\mu^2)A = \alpha_s(Q^2)A - \alpha_s^2(Q^2)\beta_0 A \ln\frac{\mu^2}{Q^2} + \mathcal{O}(\alpha_s^2)$$

where in the second step we inserted the expansion from Eq. (8.11). This means that the result depends strongly on the choice of the renormalization scale μ. Since the corrections to the cross section can be relatively large, it is possible to find significantly different values for the measured effective coupling constant $\alpha_s^{\mathrm{meas,eff}}$ for two different processes: Consider two processes, where the LO calculations predict

$$\sigma_{\mathrm{LO;1}}^{\mathrm{pert}} = \alpha_s A_1$$
$$\sigma_{\mathrm{LO;2}}^{\mathrm{pert}} = \alpha_s A_2.$$

The predictions are compared to the cross sections σ_1^{exp} and σ_2^{exp} from experiment. Finally, because of the said strong scale dependence, the result may be $\alpha_{s;1}^{\mathrm{meas,eff}} \neq \alpha_{s;1}^{\mathrm{meas,eff}}$.

To solve the problem of the correction depending on the renormalization scale being too large, one has to take the calculation to next-to-leading order (NLO) to reduce the scale dependence of the prediction. For our example reaction $e^+e^- \to q\bar{q}g$ this means

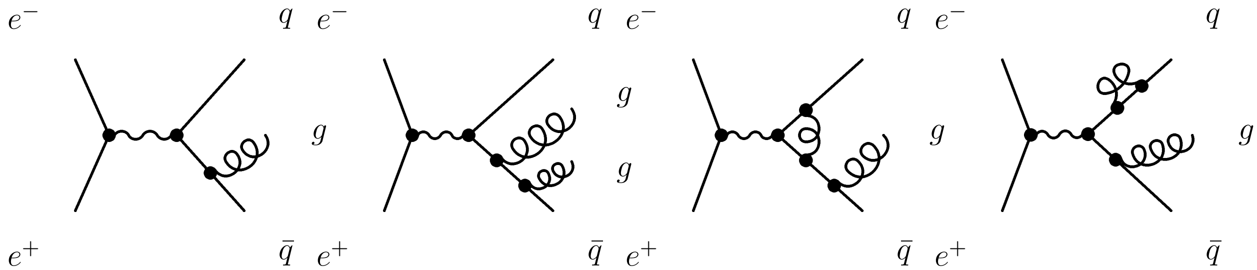


Figure 8.25: *Feynman diagrams for $e^+e^- \to q\bar{q}g$ to NLO.*

considering the diagrams shown in Fig. 8.25. The NLO expression is again obtained from the expansion in Eq. (8.12):

$$\begin{aligned}\sigma_{\mathrm{NLO}}^{\mathrm{pert}} &= \alpha_s(\mu^2)A + \alpha_s^2(\mu^2)\left[B + \beta_0 A \ln\frac{\mu^2}{Q^2}\right] + \mathcal{O}(\alpha_s^3)\\ &= \alpha_s(Q^2)A + \alpha_s^2(Q^2)B + \alpha_s^3(Q^2)\beta_0^2 A^2 \ln^2\frac{\mu^2}{Q^2} + \mathcal{O}(\alpha_s^4)\end{aligned}$$

where in the second line we inserted for $\alpha_s(\mu^2)$ the expansion from Eq. (8.11) and the dependence on $\ln(\mu^2/Q^2)$ cancels. Thus, the scale dependence of the prediction is much smaller than in the LO case. The scale dependence cancels completely at fully calculated order.

By comparing the NLO prediction for the cross section to experiment, one can extract $\alpha_s(Q^2)$, e. g. $\alpha_s(M_Z^2)$. This information can in turn be used to predict other process cross sections at NLO. Furthermore, by varying the scale μ^2 one can estimate the size of the NNLO contributions.

This procedure extends analogously to NNLO. Diagrams that have to be included at NNLO are shown in Fig. 8.26. The prediction reads

$$\sigma_{\mathrm{NNLO}}^{\mathrm{pert}} = \alpha_s(Q^2)A + \alpha_s^2(Q^2)B + \alpha_s^3(Q^2)C + \mathcal{O}\left(\alpha_s^4, \ln^3\frac{\mu^2}{Q^2}\right)$$

where the scale dependence is reduced even further. NNLO is the lowest order at which scale variations at NLO can be tested.

As an example for the scale dependence of the extracted strong coupling constant, see Fig. 8.27 where $\alpha_s(M_Z^2)$ from jet rates at LEP is shown as a function of $\ln(\mu^2/Q^2)$. Note that the scale dependence is reduced by the extension to NLO, as mentioned before. The theoretical error is taken to be the range of values covered by the projection of the bands over $-1 < \ln(\mu^2/Q^2) < 1$ on the abscissa. The right figure shows how the central values and errors obtained this way for three different shape variables converge with improvements in the theory.

There has been an enormous progress in the measurements of the strong coupling during the last 20 years. This is due to major improvements on the theoretical and also the

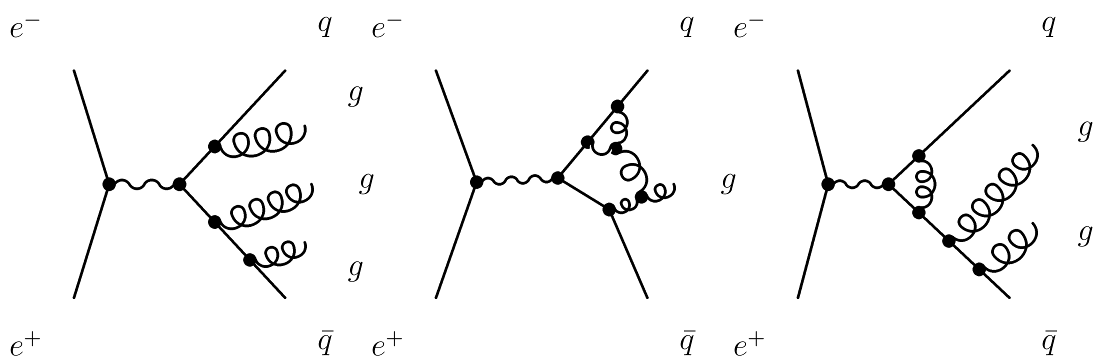


Figure 8.26: *Feynman diagrams for $e^+e^- \to q\bar{q}g$ at NNLO.*

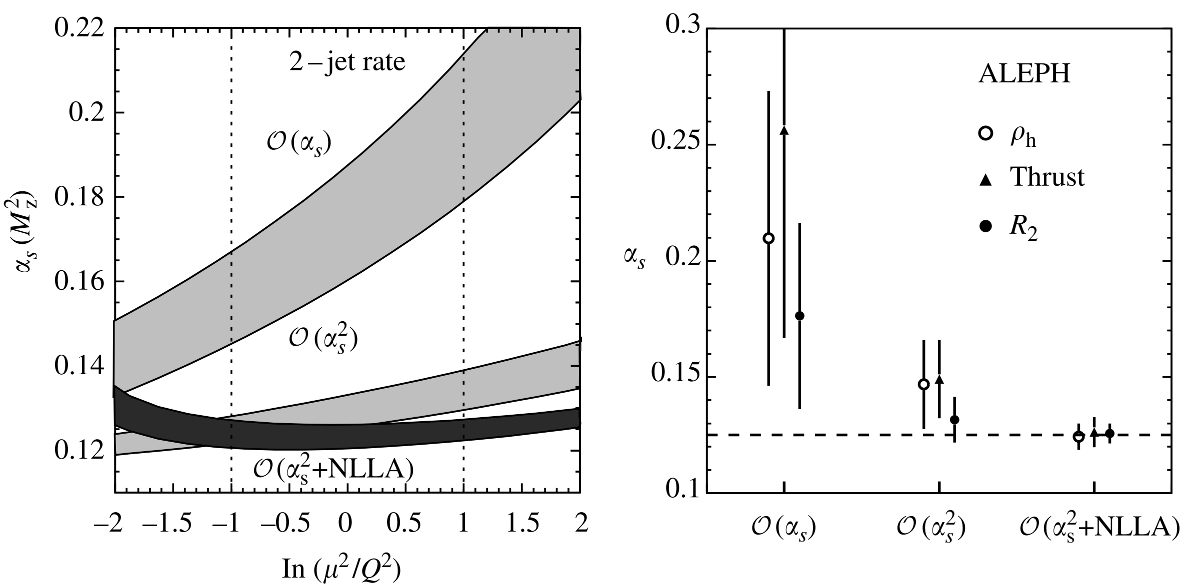


Figure 8.27: *Estimate of theoretical uncertainties for a measurement of the strong coupling constant from event shape variables.* NLLA refers to resummation of logarithms. Source: [27, p. 307].

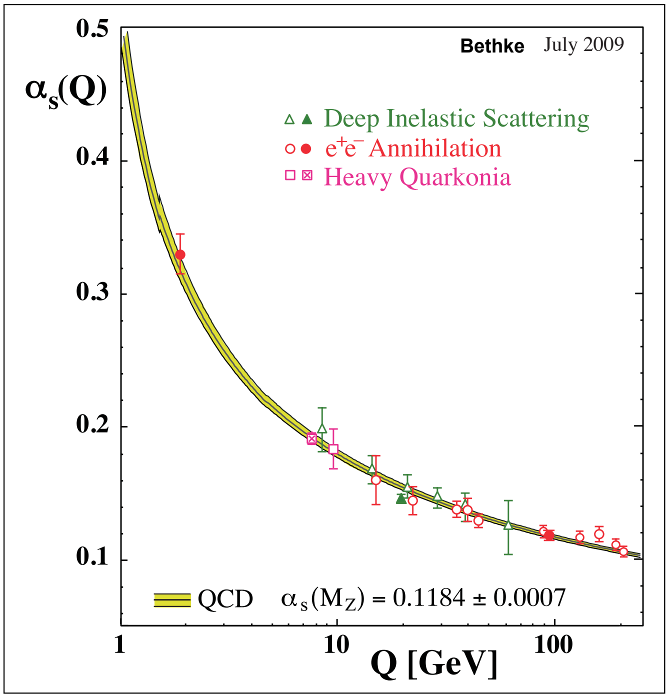


Figure 8.28: *Summary of measurements of α_s as a function of the respective energy scale Q*. The curves are QCD predictions. Source: [30, p. 12].

experimental side. A summary of measurements of α_s as a function of the respective energy scale Q is shown in Fig. 8.28.

In general, observables can be classified according to the influence the structure of the final state has on their value.

Inclusive observables do not look at the structure of the final state. Examples are total cross sections and ratios of cross sections (see e. g. Eq. (8.3)). Advantages of inclusive observables are that they do not (or only weakly) depend on non-perturbative corrections (hadronization) and that the perturbative series is now known to NNNLO. The disadvantage lies in the low sensitivity in some cases.

Non-inclusive (exclusive) observables, on the other hand, look at some structure in the final state depending on the momenta of the final state particles. Examples are jet rates and event shape distributions. Advantages of non-inclusive observables are high sensitivity and that the perturbative series is now known to NNLO (and resummation, see later). Disadvantages are that in some cases even the NNNLO corrections might be relevant and that hadronization (non-perturbative) corrections are needed.

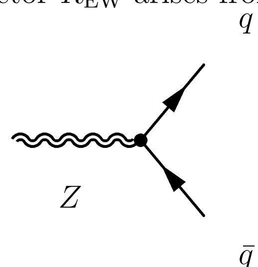
As an example for the usage of inclusive observables, consider the determination of α_s from inclusive Z or τ decays. In general, the prediction of the cross section ratio R reads

$$R = \frac{\sigma^{Z,\tau\to\text{ hadrons}}}{\sigma^{Z,\tau\to\text{ leptons}}} = R_{\text{EW}}(1 + \delta_{\text{QCD}} + \delta_{\text{mass}} + \delta_{\text{np}})$$

where the overall factor R_{EW} depends on the electroweak couplings of the quarks.[5] The corrections are dominated by the perturbative QCD correction δ_{QCD}. The other terms take into account the finite quark masses and the non-perturbative corrections. The perturbative QCD correction term is given by

$$\delta_{\mathrm{QCD}} = c_1 \frac{\alpha_s}{\pi} + c_2 \left(\frac{\alpha_s}{\pi}\right)^2 + c_3 \left(\frac{\alpha_s}{\pi}\right)^3 + \ldots .$$

Diagrammatically speaking, the factor R_{EW} arises from



while the perturbative QCD corrections come from diagrams like in Fig. 8.25 and 8.26. For the case of

$$R_Z = \frac{\sigma^{Z\to\ \text{hadrons}}}{\sigma^{Z\to\ \text{leptons}}}$$

the prediction reads $R_{\mathrm{EW}} = 19.934$, $c_1 = 1.045$, $c_2 = 0.94$, and $c_3 = -15$. The corresponding measurement is visualized in Fig. 8.29: Divide the number of hadronic decays by the number of leptonic decays to find $R_Z = 20.767 \pm 0.025$. From this ratio the following value of the strong coupling at the Z resonance can be extracted:

$$\begin{aligned}\alpha_s(M_Z) &= 0.1226 \pm \underbrace{0.0038}_{\text{exp., mostly statistical}} \pm \underbrace{0.0002}_{M_t:\,\pm 5\,\mathrm{GeV}} \pm \underbrace{0.0002}_{\text{renormalization shemes}} \\ &= 0.1226 \begin{array}{l} +0.0058 \\ -0.0038 \end{array} .\end{aligned}$$

Finally, we state a new result from 2009, obtained using NNNLO predictions:

$$\alpha_s(M_Z) = 0.1193 \begin{array}{l} +0.0028 \\ -0.0027 \end{array} \pm 0.0005.$$

We now turn to non-inclusive observables such as event-shapes and jet rates. We have already seen perturbative predictions for some examples of non-inclusive quantities in Sect. 8.2. There it is stated that the log terms in the predictions are because of the $\int dE/E$ integration arising from

$$\frac{d\sigma^{q\to qg}}{dE_{\mathrm{gluon}}} \propto \sigma_0 \frac{\alpha_s}{2\pi} \frac{1}{E_{\mathrm{gluon}}}$$

[5] R_{EW} is a modified version of the ratio $R = N_c 11/9$ of Sect. 8.1.

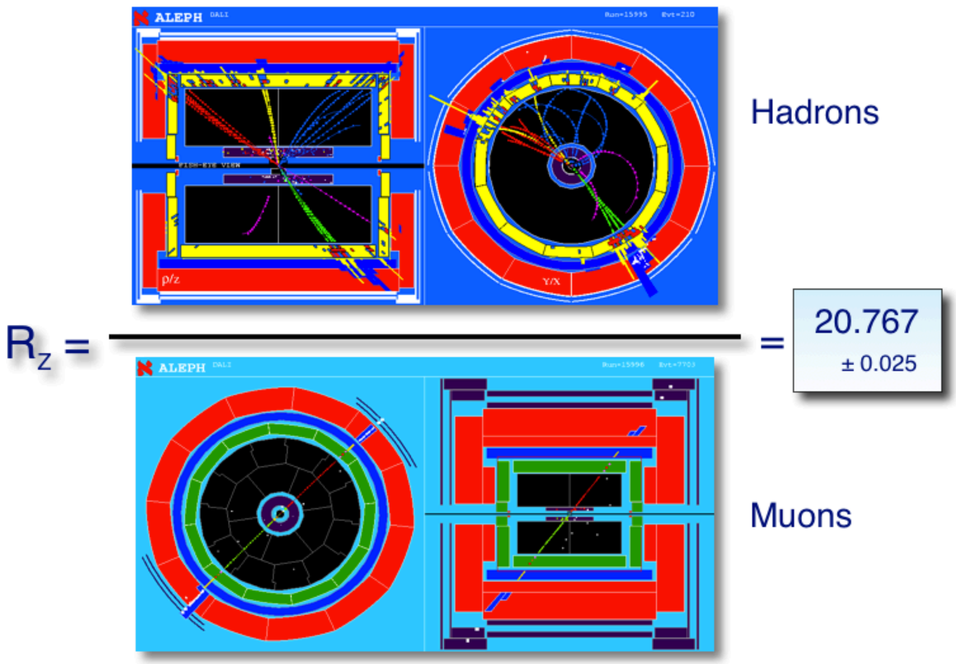


Figure 8.29: *Visualization of R_Z measurement.*

where σ_0 is the Born cross section for $Z \to q\bar{q}$ (see Sect. 8.1). Recall that the perturbative prediction is given by:

$$\frac{1}{\sigma_0}\frac{d\sigma}{dx} = \alpha_s(\mu^2)A(x) + \alpha_s^2(\mu^2)\left[B(x) + \beta_0 A(x)\ln\frac{\mu^2}{Q^2}\right] + \mathcal{O}(\alpha_s^3)$$

where the coefficients A and B are calculable for the class of observables x which are infrared and collinear safe, i. e. infrared singularities from real and virtual radiative corrections cancel (thrust, jet rates, C-parameter, etc.). To recall the important example of thrust, see Fig. 8.15.

Let us take a look at the results obtained by NLO fits. First measurements gave indications that the missing higher order terms are large: The coupling constant should be the same for all variables, but the results vary too much (see Fig. 8.30) which indicates that the expansion to NLO does not suffice. Typical results obtained by NLO fits are

$$\alpha_s(M_Z) = 0.120 \pm 0.010.$$

As we have seen before, to obtain perturbative corrections, we have to do integrals of the type $\int_{y_{\text{cut}}}^{s} dE_{\text{gluon}}/E_{\text{gluon}}$ which gives rise to the logarithm terms in $\sigma^{LO}_{\text{three-jet}}$ (see Eq. (8.10)):

$$\sigma^{LO}_{\text{three-jet}} = \sigma_0 C_F \frac{\alpha_s}{2\pi}\left[\ln^2 y_{\text{cut}} + \ldots\right]$$

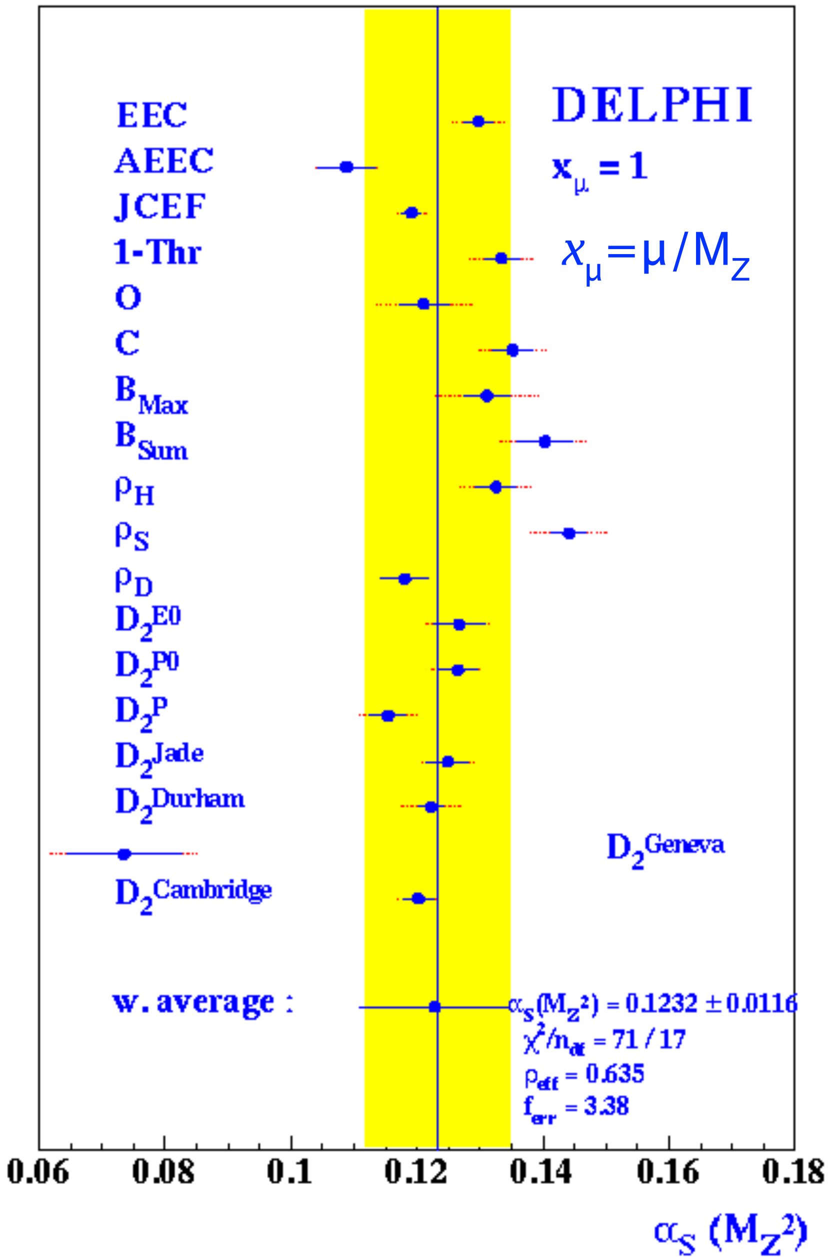


Figure 8.30: *NLO results for* $\alpha_s(M_Z)$*.* Source: [40, p. 29, modified].

where the color factor $C_F = 4/3$—the problem being that for $y_{\rm cut} \to 0$ the series does not converge.[6] The resummation procedure mentioned earlier (see p. 167) also works for the three-jet rate:

$$R_3 = \frac{C_F\alpha_s}{2\pi}\ln^2 y_{\rm cut} - \frac{C_F^2\alpha_s^2}{8\pi^2}\ln^4 y_{\rm cut} + \dots$$
$$= 1 - \exp\left\{ - \int\limits_{sy_{\rm cut}}^{s} \frac{dq^2}{q^2}\frac{C_F\alpha_s(q^2)}{2\pi}\left[\ln\frac{s}{q^2} - \frac{3}{2}\right]\right\}.$$

Combined (to avoid double counting of logarithmic terms in resummed expressions and in full fixed order prediction) with full NLO calculations this gives theoretically much improved predictions. Typical results are:

$$\alpha_s(M_Z) = 0.120 \pm 0.005.$$

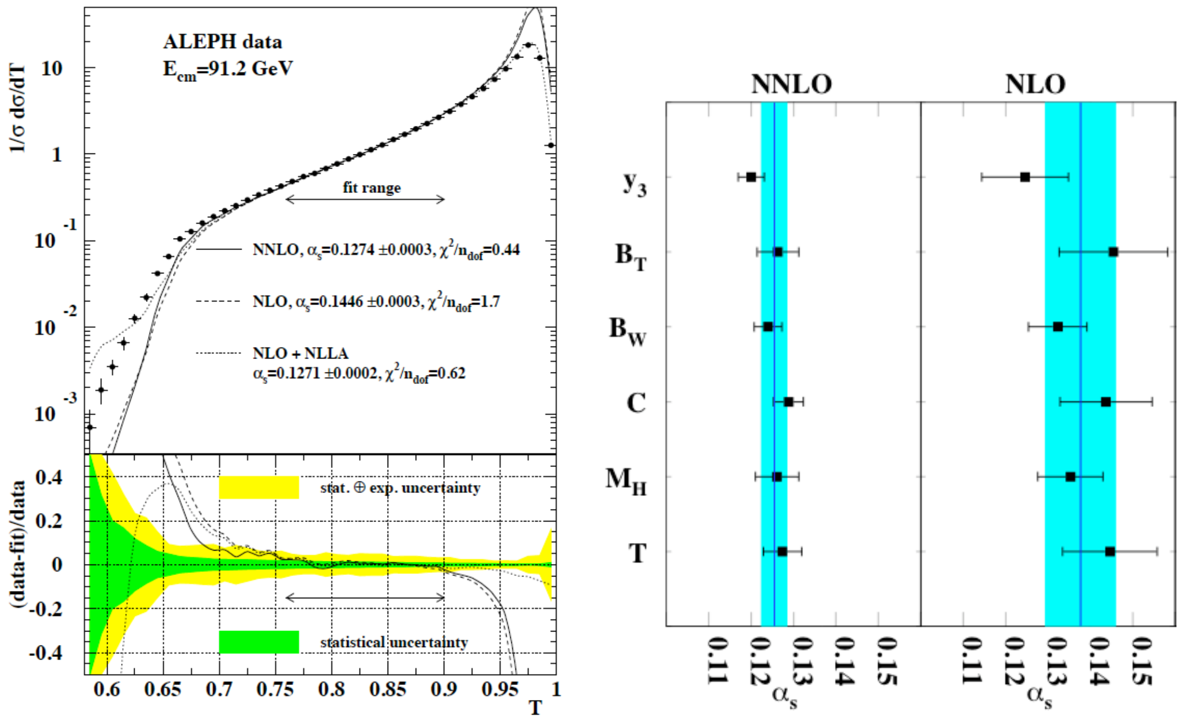
There are different sources of the remaining uncertainties. *Experimental uncertainties* include

- track reconstruction,
- event selection,
- detector corrections (via cut variations or different Monte Carlo generators),
- background subtraction (LEP2), and
- ISR corrections (LEP2).

They amount to about 1% uncertainty. Furthermore, there are *hadronization uncertainties* arising from the differences in behavior of various models for hadronization such as PYTHIA (string fragmentation), HERWIG (cluster fragmentation), or ARIADNE (dipole model and string fragmentation). Theses uncertainties are typically about 0.7 to 1.5%. Finally, there are also *theoretical uncertainties*, for instance

- renormalization scale variation,
- matching of NLO with resummed calculation, and
- quark mass effects.

[6] Recall that $y_{\rm cut}$ is the resolution parameter deciding if two particles are distinguished or seen as one pseudo-particle.



(a) (b)

Figure 8.31: *NNLO fit to ALEPH thrust data (a) and visualization of improvement in NNLO over NLO (b).* Source: [50, p. 11 and 17].

The corresponding uncertainty is typically 3.5 to 5 %.

As we have seen, the perturbative predictions have to be to sufficiently high order if we are to accurately determine the strong coupling constant: Now a NNLO prediction is available. Bearing in mind the foregoing, it has to be of the form

$$\frac{1}{\sigma_0}\frac{d\sigma}{dy}(y,Q,\mu) = \alpha_s(\mu)A(y) + \alpha_s^2(\mu)B(y,x_\mu) + \alpha_s^3(\mu)C(y,x_\mu) + \mathcal{O}(\alpha_s^4)$$

where y denotes an event shape variable and $x_\mu = \mu/Q$. At this level of precision, one has to take care of additional issues, such as quark mass effects and electro-weak effects which typically contribute around or below the per-cent range.

The first determination of $\alpha_s(M_Z)$ based on NNLO (and NLLA) calculations of event shape distributions [49, 50] yields

$$\alpha_s(M_Z) = 0.1224 \pm 0.0009 \text{ (stat) } \pm 0.0009 \text{ (exp) } \pm 0.0012 \text{ (hadr) } \pm 0.0035 \text{ (theo)}.$$

The fit to ALEPH thrust data is shown in Fig. 8.31(a). The largely reduced scatter of values for different variables at NNLO is visualized in Fig. 8.31(b). Note that the reduced perturbative uncertainty is 0.003.

The most precise determination of the strong coupling constant is obtained from jet observables at LEP. Precision at the 2% level is achieved from the three-jet rate [51]:

$$\alpha_s(M_Z) = 0.1175 \pm 0.0020 \text{ (exp) } \pm 0.0015 \text{ (theo)}.$$

The three-jet rate is known to have small non-perturbative corrections and to be very stable under scale variations (for a certain range of the jet resolution parameter). For a comparison of LO, NLO, and NNLO predictions to the corresponding ALEPH data, see Fig. 8.32(a).

The LEP results concerning the determination of the strong coupling constant (see Fig.8.32(b)) can be summarized as follows (combination by S. Bethke, a couple of years ago).

- Tau decays (NNLO)

$$\alpha_s(M_Z) = 0.1181 \pm 0.0030$$

- R_Z (NNLO)

$$\alpha_s(M_Z) = 0.1226 \begin{array}{l} +0.0058 \\ -0.0038 \end{array}$$

- Event shapes (NLO + NNLO)

$$\alpha_s(M_Z) = 0.1202 \pm 0.0050$$

- All (not including recent NNNLO results)

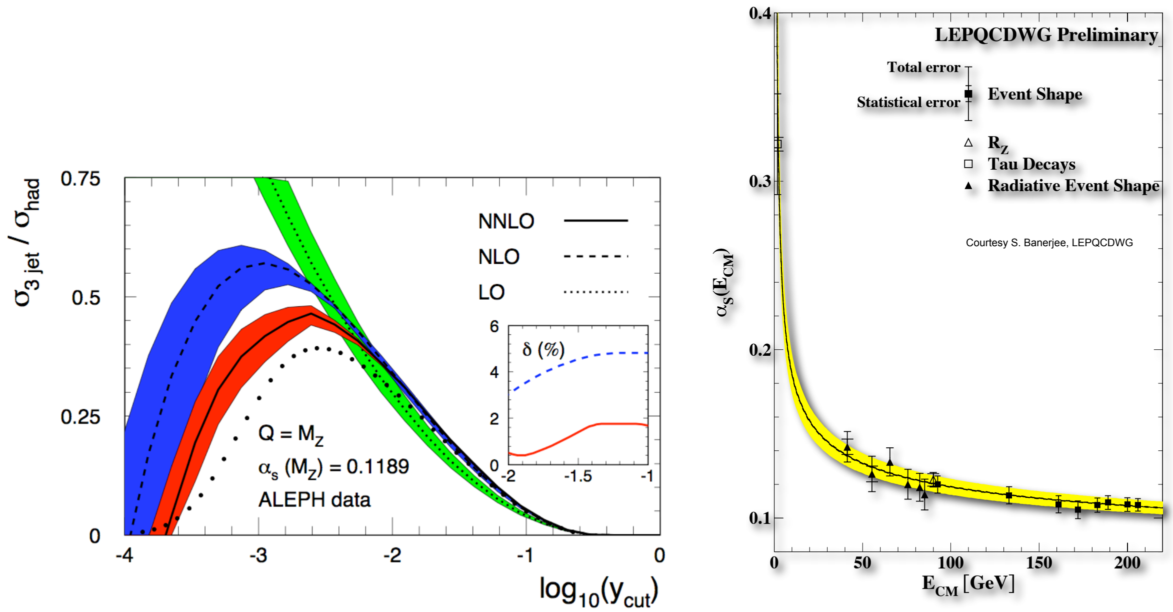
$$\alpha_s(M_Z) = 0.1195 \pm 0.0035$$

- Latest world average (S. Bethke, 2009 [30])

$$\boxed{\alpha_s(M_Z) = 0.1184 \pm 0.0007}$$

8.4 Measurements of the QCD color factors

Because they determine the gauge structure of strong interactions, the color factors are the most important numbers in QCD, besides α_s. Discussing the triple-gluon vertex we concluded that our observables also allow to test the gauge structure of QCD. We have already learned that the color factors (for $SU(3)$) $C_F = 4/3$, $C_A = 3$, and $T_F = 1/2$ measure the relative probabilities of gluon radiation ($q \to qg$), triple gluon vertex ($g \to gg$), and gluon splitting ($g \to q\bar{q}$).



(a) (b)

Figure 8.32: *NNLO, NLO, and LO fits to ALEPH data for the thee-jet rate (a) and summary of LEP results for α_s (b).*

The cross section prediction for four-jet events at order α_s^2 can be shown to be

$$\frac{1}{\sigma_0}\frac{d\sigma_{\text{four-jet}}}{dy} = \frac{\alpha_s^2 C_F^2}{\pi^2}\left[\sigma_A(y) + \left(1 - \frac{1}{2}\frac{C_A}{C_F}\right)\sigma_B(y) + \left(\frac{C_A}{C_F}\right)\sigma_C(y) + \left(\frac{T_F}{C_F}n_f\right)\sigma_D(y) + \left(1 - \frac{1}{2}\frac{C_A}{C_F}\right)\sigma_E(y)\right]$$

where σ_i, $i = A, \ldots, E$ are kinematic factors independent of the gauge group of QCD.

The combined measurements of the QCD color factors are summarized in Fig. 8.33: Four-jet and event shape results have been combined accounting for correlations between the measurements. In addition, constraints on C_A/C_F from differences between quark and gluon jets (see p. 179) are included. This yields

$$C_A = 2.89 \pm 0.21$$
$$C_F = 1.30 \pm 0.09$$

which is precise to 7% and agrees with the $SU(3)$ values of $C_A = 3$ and $C_F = 1.33$.

8.5 Hadronization

The trouble with hadronization is that perturbative calculations are no longer useful since α_s ceases to be comparatively small at length scales of about the proton radius.

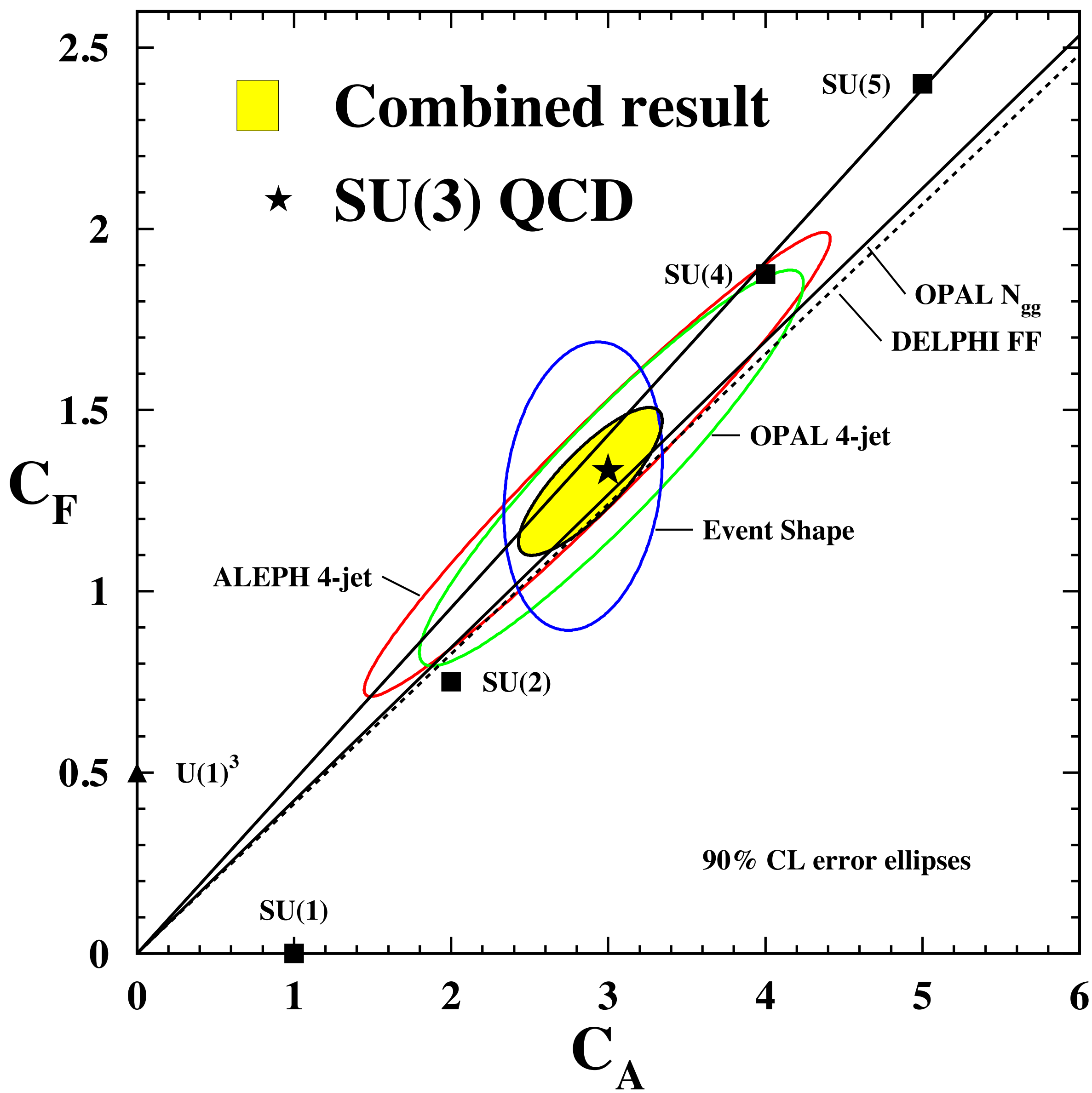


Figure 8.33: *Combined measurements of the color factors C_A and C_F.* The ellipses show the correlated measurements using four-jet events or event shape distributions while the lines represent the results of determinations of C_A/C_F from DELPHI (dashed) and OPAL (solid). Source: [52, p. 82].

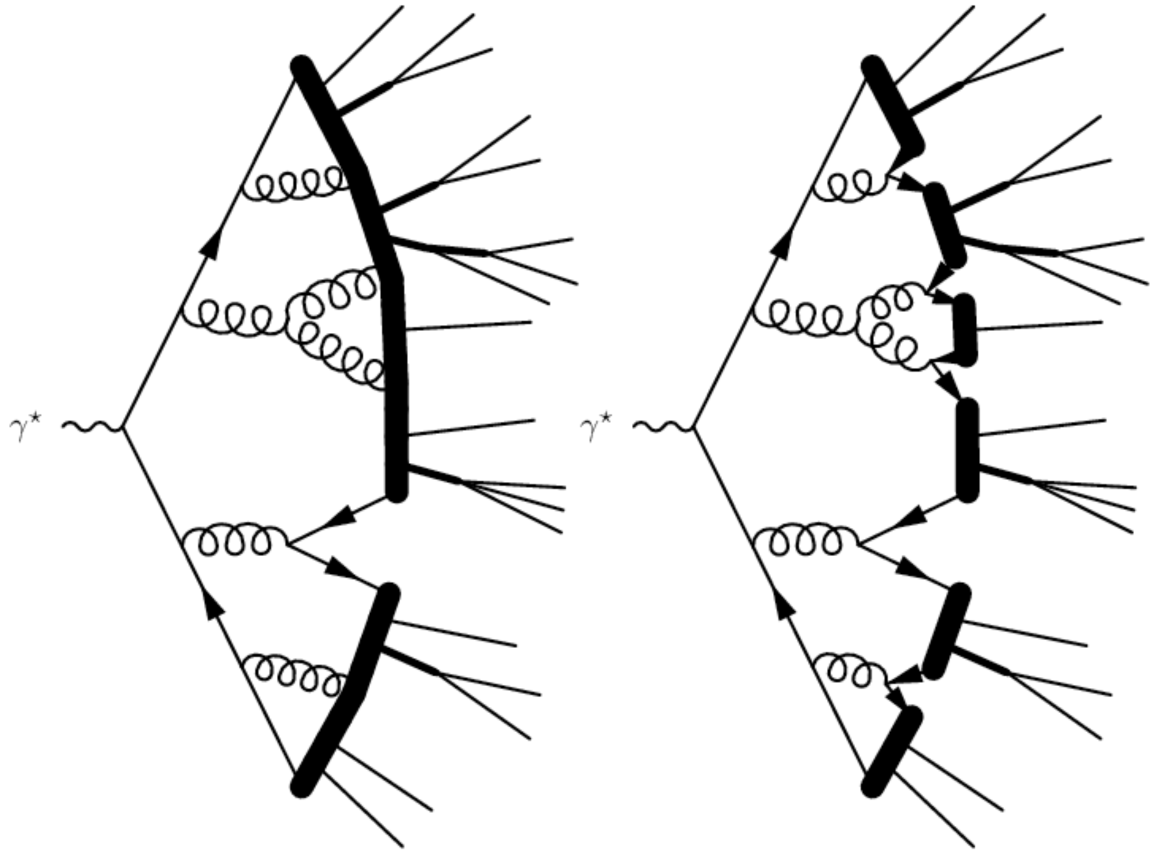
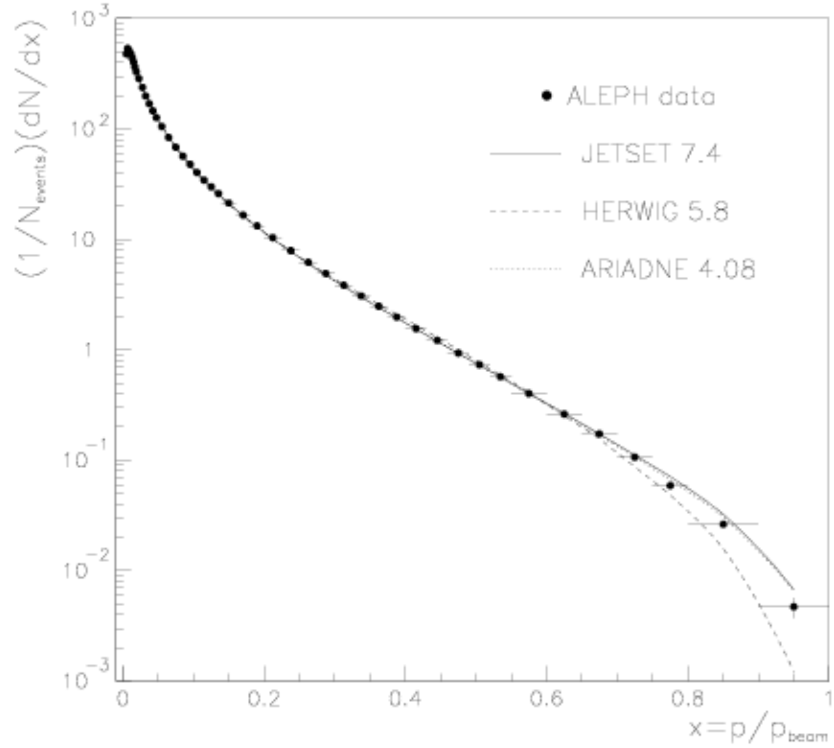


Figure 8.34: *Visualization of phenomenological models of hadronization.* (LHS) string fragmentation: JETSET/PYTHIA; (RHS) Cluster fragmentation: HERWIG. Source: [27, p. 164]

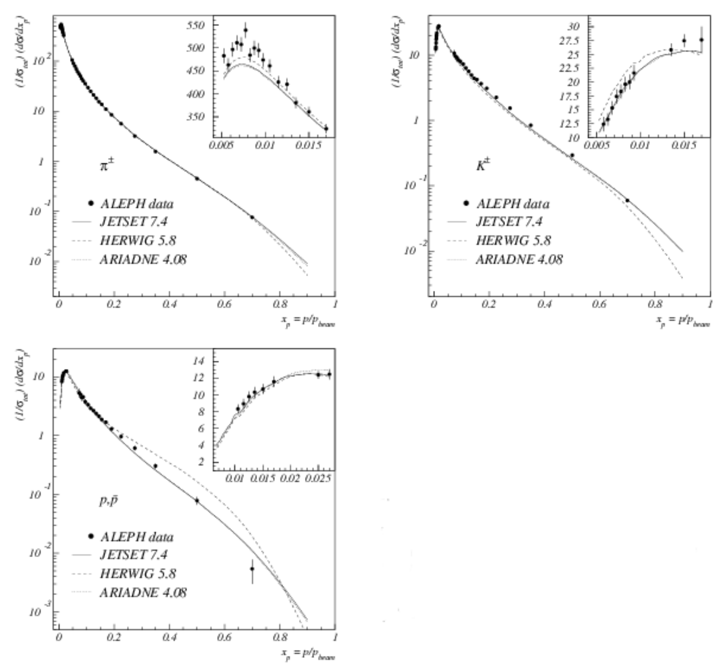
Perturbative QCD is applicable to the transition from the primary partons to a set of final state partons. This is pictured as a cascading process that is dominated by the collinear and soft emissions of gluons and mainly light quark-antiquark pairs. By contrast, phenomenological models are used to describe the non-perturbative transition from these final state partons to hadrons which then may decay according to further models (recall Fig. 8.6).

The parameters determining the behavior of the numerical models have to be adjusted using experimental data. Hadronization can be modeled by string fragmentation (JETSET/PYTHIA) or cluster fragmentation (HERWIG). For a visualization of this difference, see Fig. 8.34.

Fig. 8.35 shows comparisons of simulations to ALEPH data for hadron momentum distributions of the final state: Fig. 8.35(a) shows simulation and data for an inclusive variable and Fig. 8.35(b) deals with pions, kaons, and protons, respectively.



(a)



(b)

Figure 8.35: *Hadron momentum distributions, ALEPH data and simulation.* Inclusive measurement (a) and differential cross section for pions, kaons, and protons (b) compared with the predictions of JETSET, HERWIG, and ARIADNE. All observables are shown as functions of $x = p_{\mathrm{hadron}}/p_{\mathrm{beam}}$.

Part II

Spring semester

Chapter 9

Proton structure in QCD

Literature:

- Halzen/Martin [1], Chap. 8-10.

This chapter reviews the study of the proton structure, which lasted form after World War II to the closure of HERA (DESY) in 2007. The understanding gained from those results is of essential importance to predict cross-sections for the Tevatron (Fermilab) and the LHC (CERN), since both of them use hadrons as colliding particles.

First, the methods used to study the proton structure are presented and the relevant kinematic quantities are defined, starting from the similar case of $e^-\mu^-$-scattering. We then generalize to the case of a composite hadron. After that, the Bjorken scaling is introduced. Finally, the steps leading to the discovery of the uncharged parton – the gluon – are described.

One must remember that the link between the particle zoo and the results concerning the proton structure was not at all obvious, as the quark model had not yet imposed itself as a leading theory.

9.1 Probing a charge distribution & form factors

To probe a charge distribution in a target one can scatter electrons on it and measure their angular distribution (Fig. 9.1). The measurement of the cross-section can be compared with the expectation for a point charge distribution,

$$\frac{d\sigma}{d\Omega} = \left(\frac{d\sigma}{d\Omega}\right)_{\text{point}} |F(q)|^2, \tag{9.1}$$

where $F(q)$ is called the **form factor**, and $q := k_i - k_f$ is the momentum transfer from the probing particle to the target. The momentum transfer is also related to the resolution power of the probe.

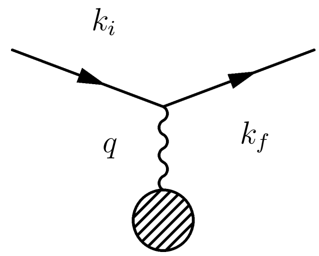


Figure 9.1: Probing a charge distribution

When probing a point ($\equiv$ spinless & structureless) target, $F(q) \equiv 1$ and one gets the **Mott cross section**,

$$\left(\frac{d\sigma}{d\Omega}\right)_{\text{point}} = \frac{(Z\alpha)^2 E^2}{4k^4 \sin^4(\theta/2)} \left(1 - \frac{k^2}{E^2} \sin^2(\theta/2)\right), \tag{9.2}$$

where Z is the electric charge measured in units of the elementary charge, E and $k = |k_i| = |k_f|$ are respectively the energy and the momentum of the probing particle, and θ is the scattering angle. One typically measures θ and E of the scattered electron.

Comparing the angular dependence of the differential cross-section of eletrons scattering off protons with the Mott cross sections, measurements show that the two distributions do not agree at large scattering angles as shown in Fig. 9.2.

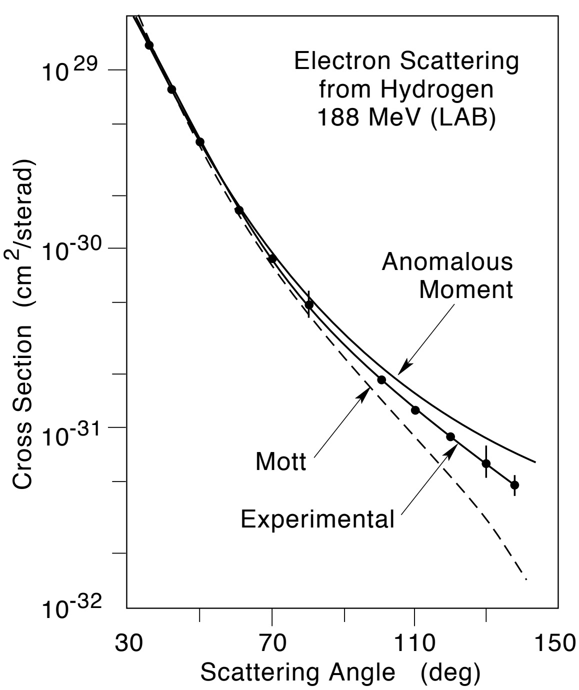


Figure 9.2: Mott cross section (dashed line) and compared to the experimental data form electron-hydrogen scattering. The measurement disagrees with the point-linke cross section at large scattering angles.

9.2 Structure functions

Starting from the example of scattering of two different elementary spin-$\frac{1}{2}$ particles, an ansatz is made for the general case.

9.2.1 $e^-\mu^-$-scattering in the laboratory frame

In the case of the $e^-\mu^-$-scattering in the laboratory frame at high energy ($s \gg M = m_\mu$), the matrix element is given by,

$$\begin{aligned}\overline{|\mathcal{M}_{fi}|^2} &= \frac{e^4}{q^4} L^{\mu\nu}_{e^-} L^{\mu^-}_{\mu\nu} \\ &= \frac{8e^4}{q^4} 2M^2 E'E \left(\cos^2(\theta/2) - \frac{q^2}{2M^2} \sin^2(\theta/2) \right) ,\end{aligned}$$

where E' is the energy of the scattered electron, and the transferred momentum,

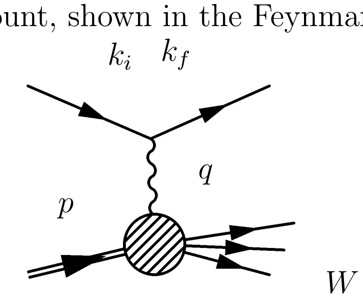
$$q^2 \approx -2k \cdot k' \approx -4EE' \sin^2(\theta/2),$$

yielding – upon inclusion of the flux factor and phase space – the differential cross section for $e^-\mu^-$ in the laboratory frame,

$$\frac{d\sigma}{d\Omega} = \frac{\alpha^2}{4E^2 \sin^4(\theta/2)} \frac{E}{E'} \left(\cos^2(\theta/2) - \frac{q^2}{2M^2} \sin^2(\theta/2) \right) . \tag{9.3}$$

9.2.2 e^-p-scattering & the hadronic tensor

When dealing with hadrons, the possibility of inelastic scattering, i.e. scattering where the final state contains excited states or other particles than the probe and the scattering particle, must be taken into account, shown in the Feynman diagram,



where W is the invariant mass of the particles in the final state (Sect. 4.4.4, p. 52). The scattering cross-section as a function of W is shown in Fig. 9.3. One notes the elastic peak at $W = m_p$ followed by a peak at 1232 MeV corresponding to the Δ^+ resonance and produced by the reaction,

$$e^-p \to e^-\Delta^+ \to e^-p\pi^0$$

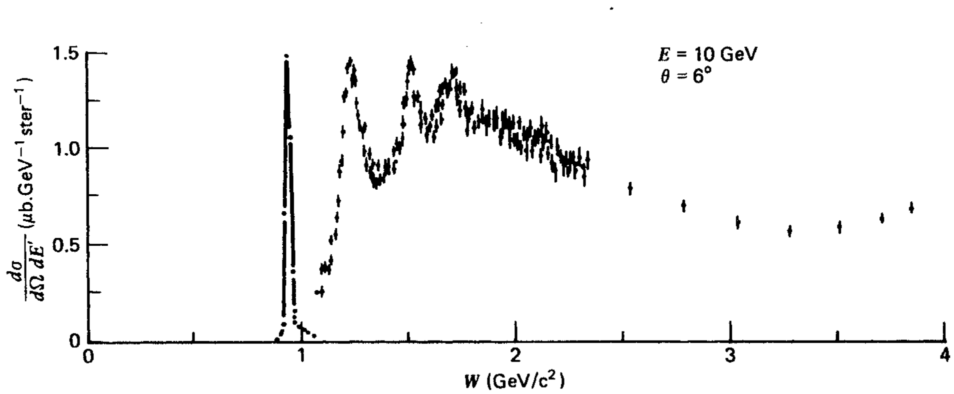


Figure 9.3: Differential cross section as a function of the invariant mass W.

To calculate the e^-p-scattering, one makes the substitution $L^{\mu\nu}_{\mu^-} \to W^{\mu\nu}_p$, where,

$$W^{\mu\nu}_p = -W_1 g^{\mu\nu} + \frac{W_2}{M^2} p^\mu p^\nu + \frac{W_4}{M^2} q^\mu q^\nu + \frac{W_5}{M^2}(p^\mu q^\nu + q^\mu p^\nu), \tag{9.4}$$

is the most general rank-2 tensor with functions $W_1, ..., W_5$ constructed from Lorentz scalars[1] depending on the internal structure of the proton, constructible from the 4-momentum of the proton (p) and the momentum transfer (q).

Imposing current conservation $\partial_\mu j^\mu_p = 0$, one can rewrite W_4 and W_5 in terms of W_1 and W_2 :

$$W_5 = -\frac{p \cdot q}{q^2} W_2$$
$$W_4 = \left(\frac{p \cdot q}{q^2}\right)^2 W_2 + \frac{M^2}{q^2} W_1,$$

Replacing W_4 and W_5 in Eq. (9.4) :

$$W^{\mu\nu}_p = W_1 \left(-g^{\mu\nu} + \frac{q^\mu q^\nu}{q^2}\right) + \frac{W_2}{M^2}\left(p^\mu - \frac{p \cdot q}{q^2} q^\mu\right)\left(p^\nu - \frac{p \cdot q}{q^2} q^\nu\right). \tag{9.5}$$

W_1 and W_2 are the so-called the **structure functions** of the proton. They depend on two independent variables,

$$Q^2 := -q^2 : \text{the 4-momentum transfer squared,}$$
$$\nu = \frac{p \cdot q}{M} : \text{the energy transferred to the nucleon by the scattering electron,}$$

[1] The "missing" W_3-term is related to the axial part of the current, and is relevant when considering the weak interaction. It is discarded in what follows.

or their dimensionless counterparts,

$$x = -\frac{q^2}{2p \cdot q} = \frac{Q^2}{2M\nu} \text{ : the Bjorken scaling } x\text{-variable,} \qquad 0 \leq x \leq 1,$$
$$y = \frac{p \cdot q}{p \cdot k_i}, \qquad 0 \leq y \leq 1.$$

With the variables defined above, we have the following expression for the invariant mass :

$$W^2 = (p+q)^2 = M^2 + 2M\nu - Q^2. \tag{9.6}$$

The elastic scattering case $W^2 = M^2$ corresponds to the value $x = 1$. Fig. 9.4 shows the

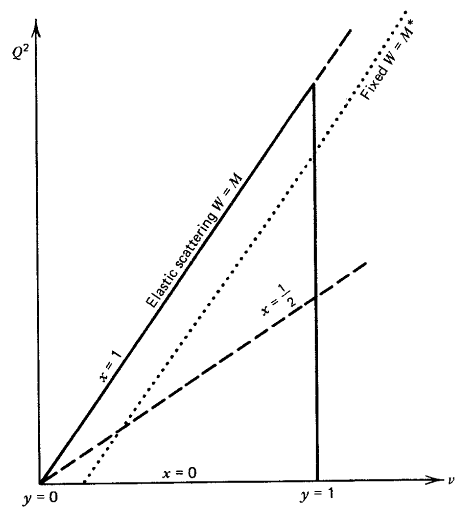


Figure 9.4: Allowed kinematical region of the Q^2-ν-plane.

kinematic region in the Q^2-ν-plane.
Using the hadron tensor, Eq. (9.5), the scattering matrix element is,

$$L^{e^-}_{\mu\nu} W^{\mu\nu}_p = 4EE' \left(W_2(Q^2,\nu)\cos^2(\theta/2) + W_1(Q^2,\nu)\sin^2(\theta/2) \right)$$

Including the flux and phase-space factors (Sect. 2.2.4, p. 15 & 3.2.3, p. 25) one finds the differential cross-section in the laboratory frame,

$$\frac{d\sigma}{dE'd\Omega} = \frac{\alpha^2}{4E^2\sin^4(\theta/2)} \left(W_2(Q^2,\nu)\cos^2(\theta/2) + W_1(Q^2,\nu)\sin^2(\theta/2) \right)$$

Integrating over the energy of the outgoing election E', one gets,

$$\frac{d\sigma}{d\Omega} = \frac{\alpha^2}{4E^2\sin^4(\theta/2)} \frac{E'}{E} \left(W_2(Q^2,\nu)\cos^2(\theta/2) + W_1(Q^2,\nu)\sin^2(\theta/2) \right).$$

9.3 Parton model

The key factor for investigating the proton substructure is the wavelength of the probing photon, which is related to the transferred momentum by,

$$\lambda \sim \frac{1}{\sqrt{Q^2}},$$

Therefore, large momentum transfer is equivalent to high resolution. As shown in Fig. 9.5, for $\lambda \approx 1\,\text{fm}$, one can "see" the proton as a single particle, whereas for, $\lambda \ll 1\,\text{fm}$, the

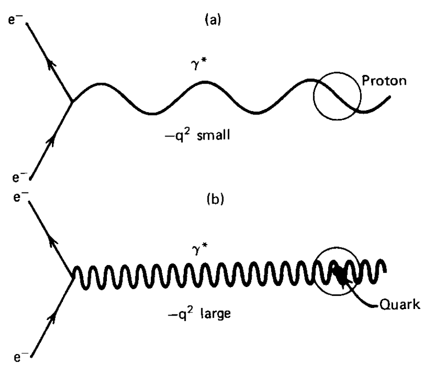
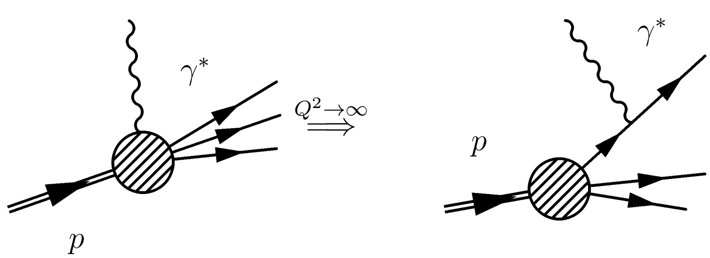


Figure 9.5: Relationship between resolution and transferred momentum.

probed particles are the constituents of the proton.

9.3.1 Bjorken scaling

J. Bjorken proposed in 1968 that, in the limit of infinite Q^2, the structure functions should only depend on the scaling variable x, and not on Q^2 and ν independently. This corresponds to postulating that at large Q^2 the inelastic e^-p-scattering is a sum of elastic scatterings of the electron on free **partons** within the proton, as illustrated below.



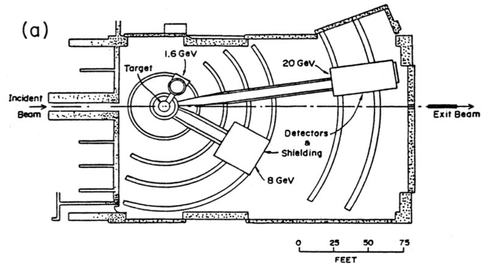
In this limit, one defines then the functions,

$$F_1(x) := \lim_{Q^2\to\infty} MW_1(Q^2,\nu), \tag{9.7}$$

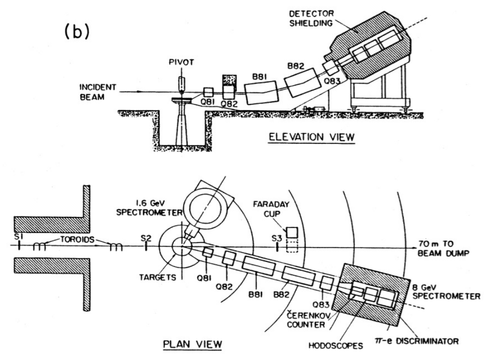
$$F_2(x) := \lim_{Q^2\to\infty} \nu W_2(Q^2,\nu). \tag{9.8}$$

9.3.2 SLAC-MIT experiment

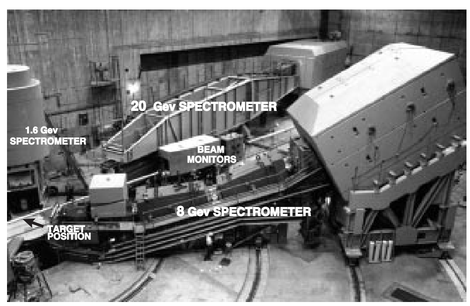
To test the hypothesis of Bjorken, a joint experiment of the SLAC and MIT groups was performed at the SLAC laboratory. Sketches and photographs of the experiment are shown in Fig. 9.6.



(a)



(b)



(c)

Figure 9.6: *SLAC-MIT experiment.* (a), (b) Sketches showing the 1.5 GeV, 8 GeV and 20 GeV spectrometers. (c) Photograph of the experiment.

The setup measured the scattering cross-section for fixed energies of the scattered electron and various angles. Fixing x (or $\omega = \frac{1}{x}$) one gets different values of Q^2 by varying the angle. The experimental result is shown in Fig. 9.7. This experiment confirmed the scaling hypothesis of Bjorken and gave a decisive piece of evidence in favour of the parton model introduced by Feynman in 1969. This model describes the proton as composed of **partons** which are the object one "sees" during an $e^- p$-scattering. One may describe the scattering process as shown in the following diagrams,

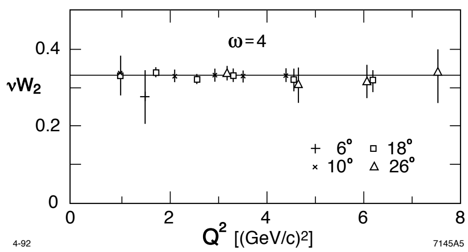
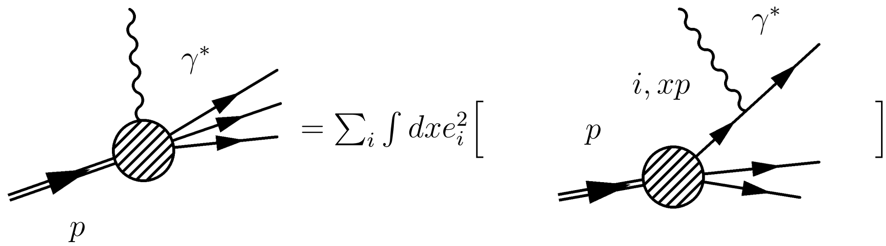


Figure 9.7: Experimental evidence for Bjorken scaling as measured at the SLAC-MIT experiment ($\omega = 1/x$).



The sum runs over all possible partons, each carrying an electric charge e_i (in units of the elementary charge) and a fraction x of the total momentum of the proton. This gives us a physical interpretation of the Bjorken scaling variable x. Since the fraction of proton momentum carried by the i-th parton is not known a priori, one needs to integrate over all possible values of x between zero (the parton carries no momentum) and one (the parton carries all the proton momentum).

The probability $f_i(x)$ that the struck parton carries a fraction x of the proton momentum is called **parton distribution function** (PDF). The total probability must be equal to 1, in order for the proton as a whole to carry all its momentum :

$$\sum_i \int_0^1 dx\, x f_i(x) = 1. \tag{9.9}$$

In Feynman's parton model the structure functions are sums of the parton densities constituting the proton,

$$\nu W_2(Q^2, \nu) \to F_2(x) = \sum_i e_i^2 x f_i(x) \tag{9.10}$$

$$M W_1(Q^2, \nu) \to F_1(x) = \frac{1}{2x} F_2(x) \tag{9.11}$$

9.3.3 Callan-Gross relation

The result,

$$\boxed{2xF_1 = F_2}, \tag{9.12}$$

is known as **Callan-Gross relation** and is a consequence of quarks being spin-$\frac{1}{2}$ particles. It can be derived by comparing the e^-p and $e^-\mu^-$ differential cross sections and setting the mass of the quark to be $m = xM$. Remembering the definitions of F_1 and F_2, Eqs. (9.7) and (9.8), one has,

$$\frac{F_1(x)}{F_2(x)} = \frac{W_1(Q^2,\nu)}{W_2(Q^2,\nu)}\frac{M}{\nu},$$

and since the scattering is elastic with a point particle (the parton),

$$\begin{aligned} 2W_1(Q^2,\nu) &= \frac{Q^2}{2m^2}\delta\left(\nu - \frac{Q^2}{2m}\right) \\ W_2(Q^2,\nu) &= \delta\left(\nu - \frac{Q^2}{2m}\right) \qquad \Rightarrow \frac{W_1(Q^2,\nu)}{W_2(Q^2,\nu)} = \frac{Q^2}{4m^2}, \end{aligned}$$

and one gets the desired result, by putting in the definition of x and $m = xM$,

$$\frac{F_1(x)}{F_2(x)} = \frac{Q^2}{4m^2}\frac{M}{\nu} = \frac{Q^2}{2M\nu}\frac{1}{2x^2} = \frac{1}{2x}$$

Fig. 9.8 shows the Q^2-independence of the Callan-Gross relation.

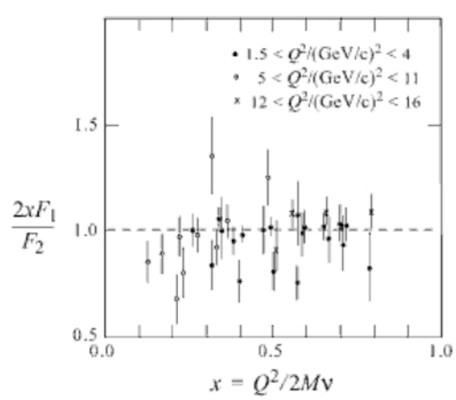


Figure 9.8: Experimental evidence for the Callan-Gross relation.

9.3.4 Parton density functions of protons and neutrons

The proton is know to be composed of two up and one down quarks (Sect. 7.3, p. 133). These quarks are known as valence quarks and are denoted q_v. They are the ones determining the properties of a hadron. It can however occur (in particular at high Q^2, corresponding to a high resolution) that a valence quark radiates a gluon which then splits in a quark-antiquark pair which is then probed by the virtual photon. These quarks are referred to as sea quarks and are denoted q_s.

In the case of e^-p-scattering and e^-n-scattering, writing q^N instead of $f_q^N(x)$ for convenience and using Eq. (9.10), we get respectively,

$$\frac{1}{x}F_2^{ep} = \left(\frac{2}{3}\right)^2 (u^p + \bar{u}^p) + \left(\frac{1}{3}\right)^2 (d^p + \bar{d}^p) + \left(\frac{1}{3}\right)^2 (s^p + \bar{s}^p) \tag{9.13}$$

$$\frac{1}{x}F_2^{en} = \left(\frac{2}{3}\right)^2 (u^n + \bar{u}^n) + \left(\frac{1}{3}\right)^2 (d^n + \bar{d}^n) + \left(\frac{1}{3}\right)^2 (s^n + \bar{s}^n), \tag{9.14}$$

where we have discarded the contributions of partons heavier than the strange quark.

One makes the assumption that these functions are not independent (exchanging an up quark for a down turns basically a proton into a neutron), and defines the total PDF of a given quark as the sum of its valence and sea components,

$$u := u_v + u_s = u^p = d^n$$
$$d := d_v + d_s = d^p = u^n.$$

Furthermore, we assume that the three lightest quark flavours (u,d,s) occur with equal probability in the sea:

$$S := u_s = \bar{u}_s = d_s = \bar{d}_s = s_s = \bar{s}_s.$$

Combining all definitions and assumptions one obtains,

$$\frac{1}{x}F_2^{ep} = \frac{1}{9}(4u_v + d_v) + \frac{4}{3}S \tag{9.15}$$

$$\frac{1}{x}F_2^{en} = \frac{1}{9}(4d_v + u_v) + \frac{4}{3}S. \tag{9.16}$$

At small momentum fractions ($x \approx 0$) the structure function is dominated by low-momentum $q\bar{q}$-pairs constituting the "sea", and hence

$$\frac{F_2^{en}}{F_2^{ep}} \to 1,$$

whereas for $x \approx 1$ the valence quarks dominate and,

$$\frac{F_2^{en}}{F_2^{ep}} \to \frac{1}{4}.$$

The experimental evidence is shown in Fig. 9.9.

Fig. 9.10 shows the distribution of F_2^{ep} that one would observe in different scenarios of proton structure.

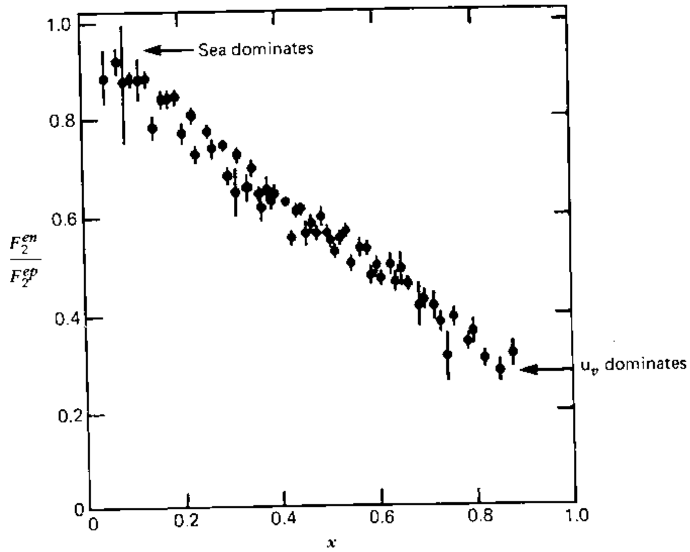


Figure 9.9: Ratio of the proton and neutron structure functions as a function of the Bjorken x-variable.

9.4 Gluons

9.4.1 Missing momentum

Summing the measured momenta of the partons cited above should give the proton momentum. However this is not the case.

$$\int_0^1 dx\ x(u+\bar{u}+d+\bar{d}+s+\bar{s}) = 1-\varepsilon_g,$$

where,

$$\varepsilon_q := \int_0^1 dx\ x(q+\bar{q}).$$

The experimental data, neglecting the contribution of strange quarks, show that,

$$\int_0^1 dx F_2^{ep} = \frac{4}{9}\varepsilon_u + \frac{1}{9}\varepsilon_d = 0.18,$$

$$\int_0^1 dx F_2^{en} = \frac{1}{9}\varepsilon_u + \frac{4}{9}\varepsilon_d = 0.12.$$

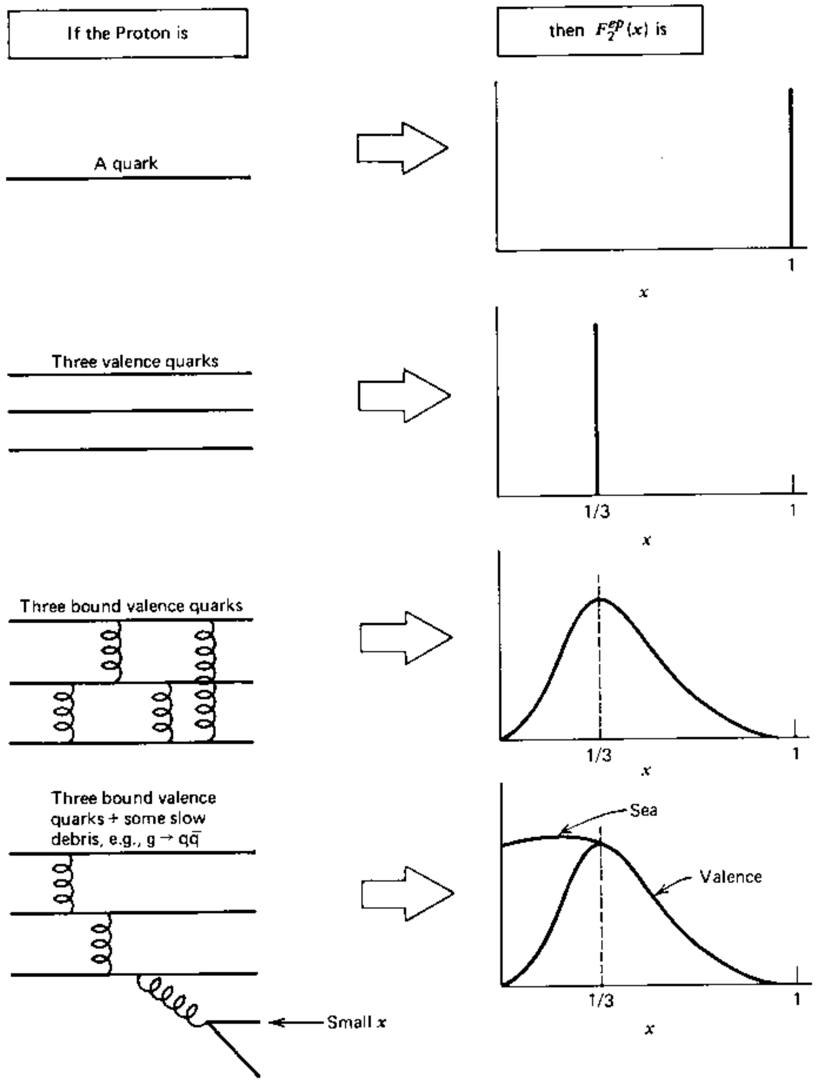


Figure 9.10: Structure functions F_2^{ep} in different scenarios of the proton structure.

Therefore,

$$\varepsilon_u = 0.36$$
$$\varepsilon_d = 0.18,$$

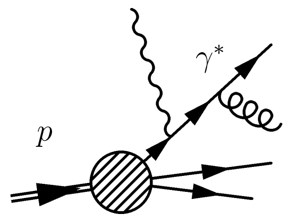
and the fraction of the proton momentum not carried by quarks is,

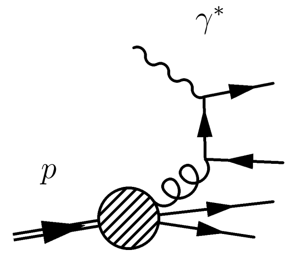
$$\varepsilon_g = 1 - \varepsilon_u - \varepsilon_d = 0.46.$$

Almost half of the proton momentum is carried by electrically uncharged partons. By repeating the scattering experiments with neutrinos instead of electrons, one observes that these uncharged partons do not interact weakly either. The parton carrying the missing momentum is now known as the **gluon**, the gauge boson of QCD.

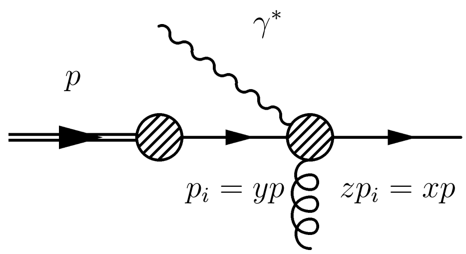
9.4.2 Gluons and the parton model at $\mathcal{O}(\alpha\alpha_s)$

By including the gluons into the parton model, the following diagrams need to be taken into account :





Looking specifically at the contribution of the first diagram, and using the kinematic variables defined in the following diagram,



one can show that the contribution to the proton structure function is of the form :

$$\frac{1}{x} F_2^{\gamma^* q \to qg} = \sum_i e_i^2 \int_x^1 \frac{dy}{y} f_i(y) \left[\frac{\alpha_s}{2\pi} P_{qq}(x/y) \log\left(\frac{Q^2}{\mu^2}\right) \right], \tag{9.17}$$

where μ is a cutoff to regularize soft gluon emission and,

$$P_{qq}(z) = \frac{4}{3}\left(\frac{1+z^2}{1-z}\right),$$

is called splitting function. It is the probability of a quark to emit a gluon and reduce momentum by a fraction z. It is obviously divergent for soft gluons ($z \to 1$).

From the form of Eq. (9.17), one sees that Q^2 appear explicitely, and not divided by $2M\nu$. This logarithmic term is responsible for the phenomenon of scaling violations wo be discussed in the next chapter.

Why did the SLAC-MIT experiment not see this violation? The effect of scaling violation is only visible at extremely small x-values which were not available at this time. The scaling violation was indeed observed in later experiments as we will discuss in the following sections.

9.5 Experimental techniques

The main site dedicated to the study of the proton structure is the HERA accelerator (DESY), shown in Fig. 9.11. It was the only e^-p-collider ever built and reached the beam energies $E_e = 30\,\mathrm{GeV}$ and $E_p = 900\,\mathrm{GeV}$ for electrons and protons respectively.

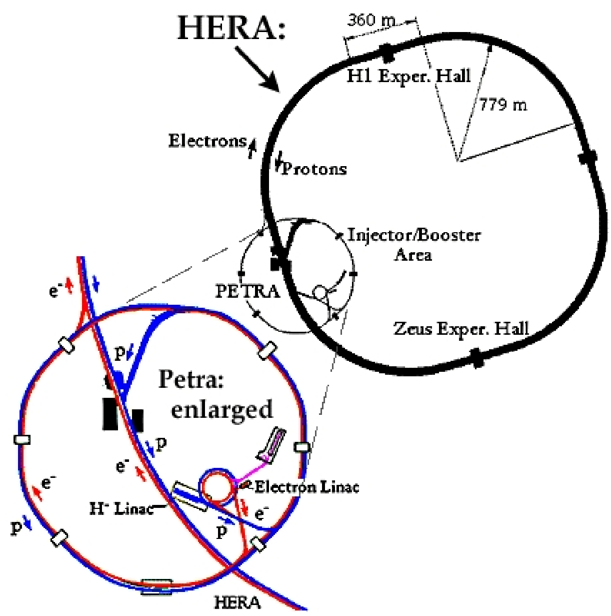


Figure 9.11: Schematics of the HERA accelerator at DESY.

Fig. 9.12 shows the coverage of the Q^2-x-kinematic region achieved at HERA and other experiments. The data at low Q^2 and low x allowed the observation of scaling violation and definitively confirmed the existence of the gluon as a constituent of the proton.

Fig. 9.13 shows the sketches of the H1 and ZEUS experiments at HERA, as well as the integrated luminosity collected by ZEUS. One can notice the asymmetrical configuration due to the different beam energies.

A typical deep inelastic scattering (DIS) event at ZEUS is shown in Fig. 9.14. One can observe the different properties of the final state : the quark jet deposits energy in the hadron calorimeter, while the electron is stopped in the electromagnetic section. The angles of the electron and hadronic system are measured in the central tracking chamber.

A "two jets" event, corresponding to the reaction,

$$e^- + p \to e^- + q + \bar{q} + X,$$

where X denotes the proton remnant (whose products are visible in the forward calorimeter), is shown in Fig. 9.15. An interesting feature of this event is the presence of a muon in correspondence of the jet. This muon may originate from the decay of a heavy quark.

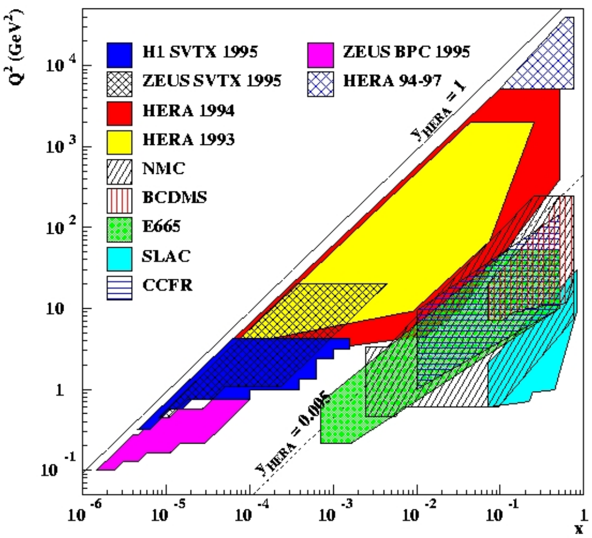


Figure 9.12: Coverage of the Q^2-x-kinematic region at HERA.

Since scaling is no longer preserved, both Q^2 and x (or $y = \frac{Q^2}{sx}$) have to be measured. Those can be obtained by measuring the energy E'_e and angle θ_e of the scattered eletron and using,

$$\begin{aligned} y_e &= 1 - \frac{E'_e}{2E_e}(1 - \cos\theta_e) \\ Q^2_e &= 2E_e E'_e(1 + \cos\theta_e). \end{aligned}$$

Fig. 9.16 shows the kinematic region measured at ZEUS while Fig. 9.17 shows the experimental results for the structure function F_2 as well as the NLO QCD fits. For low values of x, the scaling violation appears very clearly. It is due to the inclusion of the processes containing gluons.

Finally, Fig. 9.18 shows the measurement of the proton PDFs achieved at HERA. The relative importance of the sea and gluon distribution can be seen to vary significantly for Q^2 between 1.9 GeV^2 and 10 GeV^2 (note the scale reduction!). One can notice similarities with the expectation shown in Fig. 9.10.

9.6 Parton model revisited

In the following two sections we formalize the foregoing discussion and derive the expression of the QCD improved parton model for $F_2(x, Q^2)/x$ given in Eq. (9.17).

As we have seen the proton is a bound state of three quarks with strong binding. "Strong binding" says that the quark binding energy is much larger than the light quark masses: $E_{\rm bind} \gg m_q$. Compare this to the weak binding of the hydrogen atom electron: $E_{\rm bind} \ll m_e$.

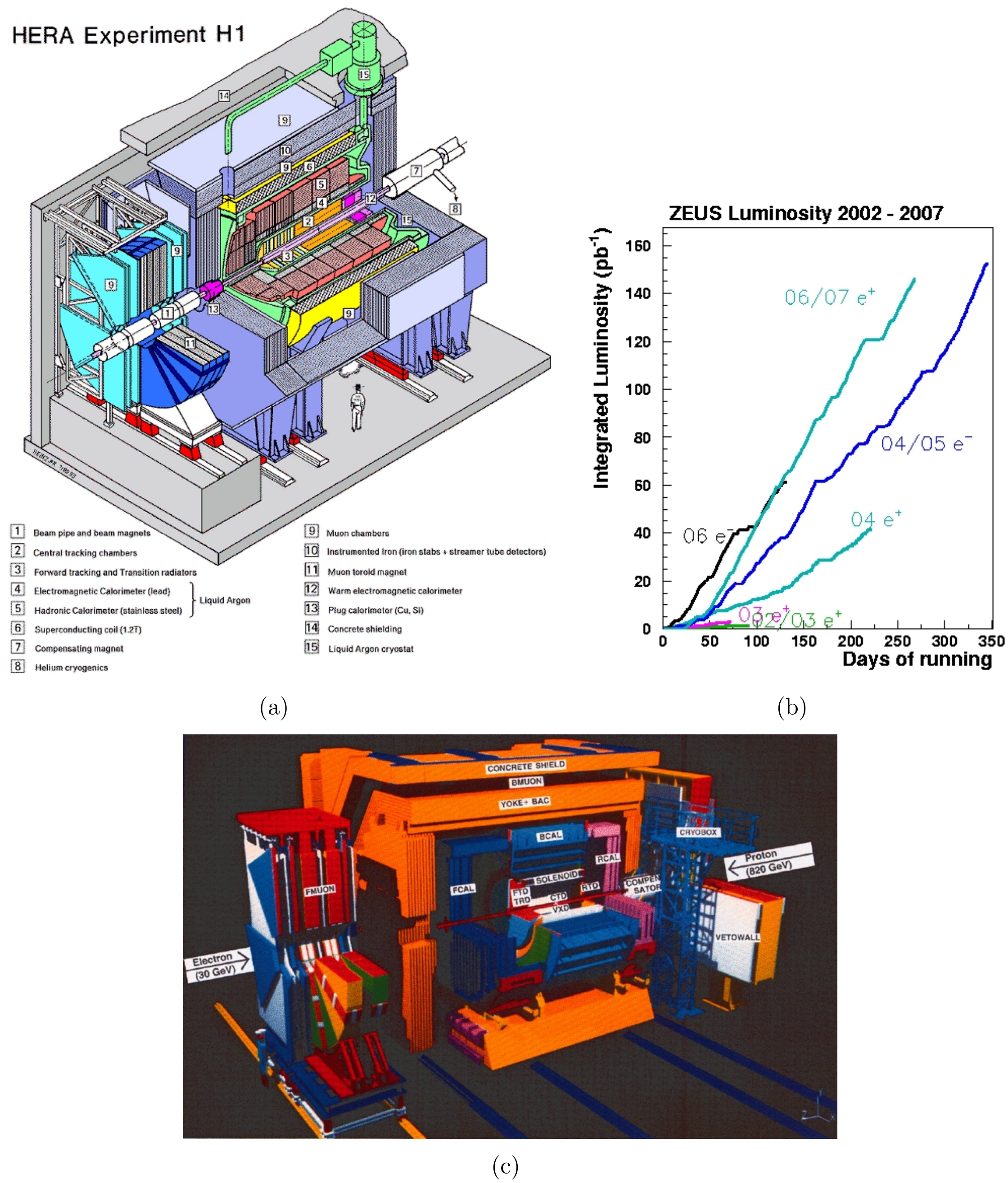


Figure 9.13: *Experiments at HERA.* (a) H1. (b) Luminosity integrated by the ZEUS during its operation. (c) ZEUS.

We consider a proton with large momentum ($|\vec{p}| \gg m_p$):

$$p^\mu = \begin{pmatrix} \sqrt{|\vec{p}|^2 + m_p^2} \\ \vec{p} \end{pmatrix} \simeq \begin{pmatrix} |\vec{p}| \\ \vec{p} \end{pmatrix}.$$

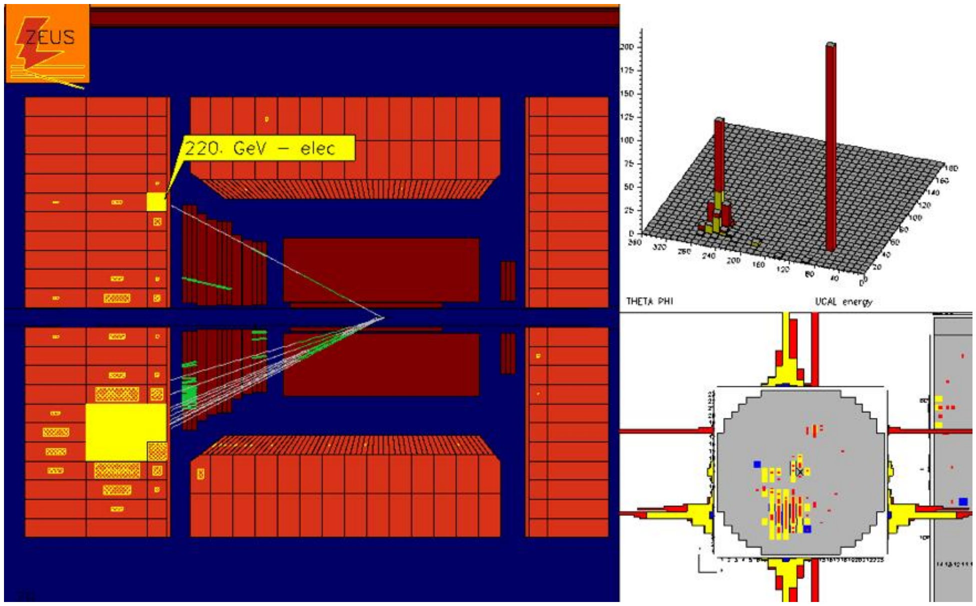
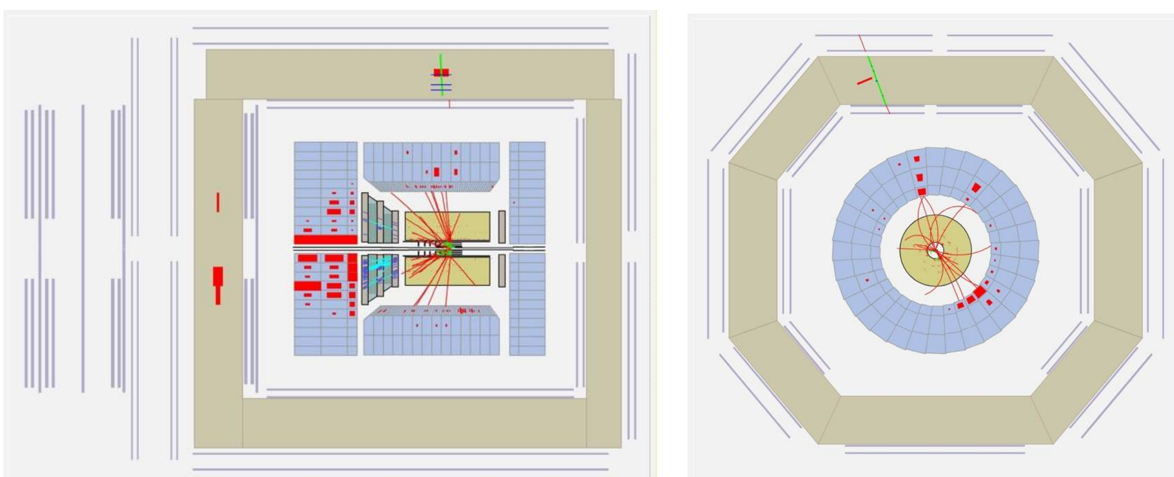


Figure 9.14: DIS event recorded by the ZEUS experiment.



(a) (b)

Figure 9.15: *Two jet event at ZEUS* (a) Side view. (b) Transverse view.

In Sect. 7.4.2 (p. 150) we discussed asymptotic freedom, namely the fact that for $Q^2 \gg \Lambda^2_{\mathrm{QCD}}$ the strong coupling constant $\bar{\alpha}_s \ll 1$. In this case the quarks of the proton are asymptotically free and therefore deep inelastic lepton-proton scattering is not an interaction with the whole proton but with just one of its constituents. This means that coherence and interference are lost (one of mutually exclusive scattering events is taking place) and deep inelastic lepton-proton scattering is an incoherent sum of lepton-quark scattering

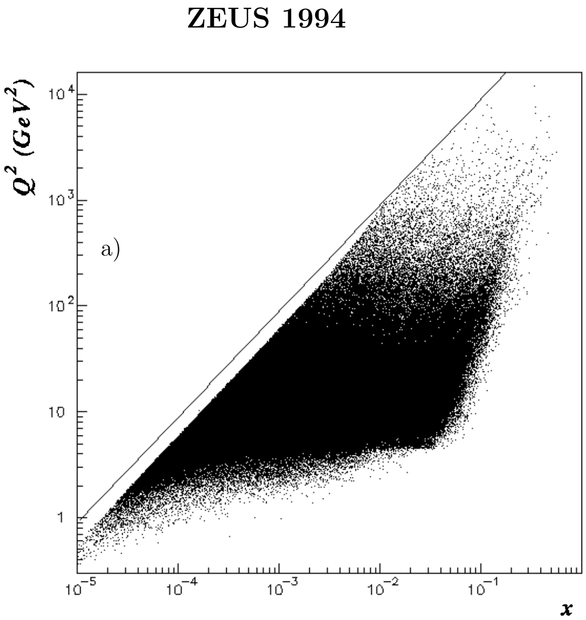


Figure 9.16: Kinematic phase-space measured by the ZEUS experiment.

processes (see Sect. 9.3.2 for diagrams) with the doubly differential cross section[2]

$$\frac{d^2\sigma}{dx dQ^2} = \sum_q \int_0^1 d\xi f_q(\xi) \frac{d^2\hat{\sigma}^{lq}}{dx dQ^2} \tag{9.18}$$

where

- $f_q(\xi)$ is a quark distribution function, i. e. the probability density of finding a quark with momentum ξp inside a proton with momentum p,
- $\xi f_q(\xi)$ is the corresponding momentum density,
- and the hat is used to denote quantities in the lepton-quark system (to distinguish them from lepton-proton system quantities).

Depending on strength and nature of the binding, one expects different behaviors of the momentum density $\xi f_q(\xi)$, as is shown in Fig. 9.19 (compare also Fig. 9.10). If the proton were pointlike the momentum density would be just a delta function, $\delta(1-\xi)$, enforcing $\xi = 1$ for the one particle involved, see Fig. 9.19(a). A proton built out of three *massive* and *weakly coupled* quarks leads to momentum densities consisting of non-ideal delta functions located at $\xi = 1/3$, $1/3\delta(1/3-\xi)$, which are insignificantly smeared out due to the ongoing exchange of binding energy between the quarks with weak, QED like coupling:

[2]Note that ξ and x are not a priori identical. Their relationship under varying assumptions is discussed below and eventually involves QCD corrections.

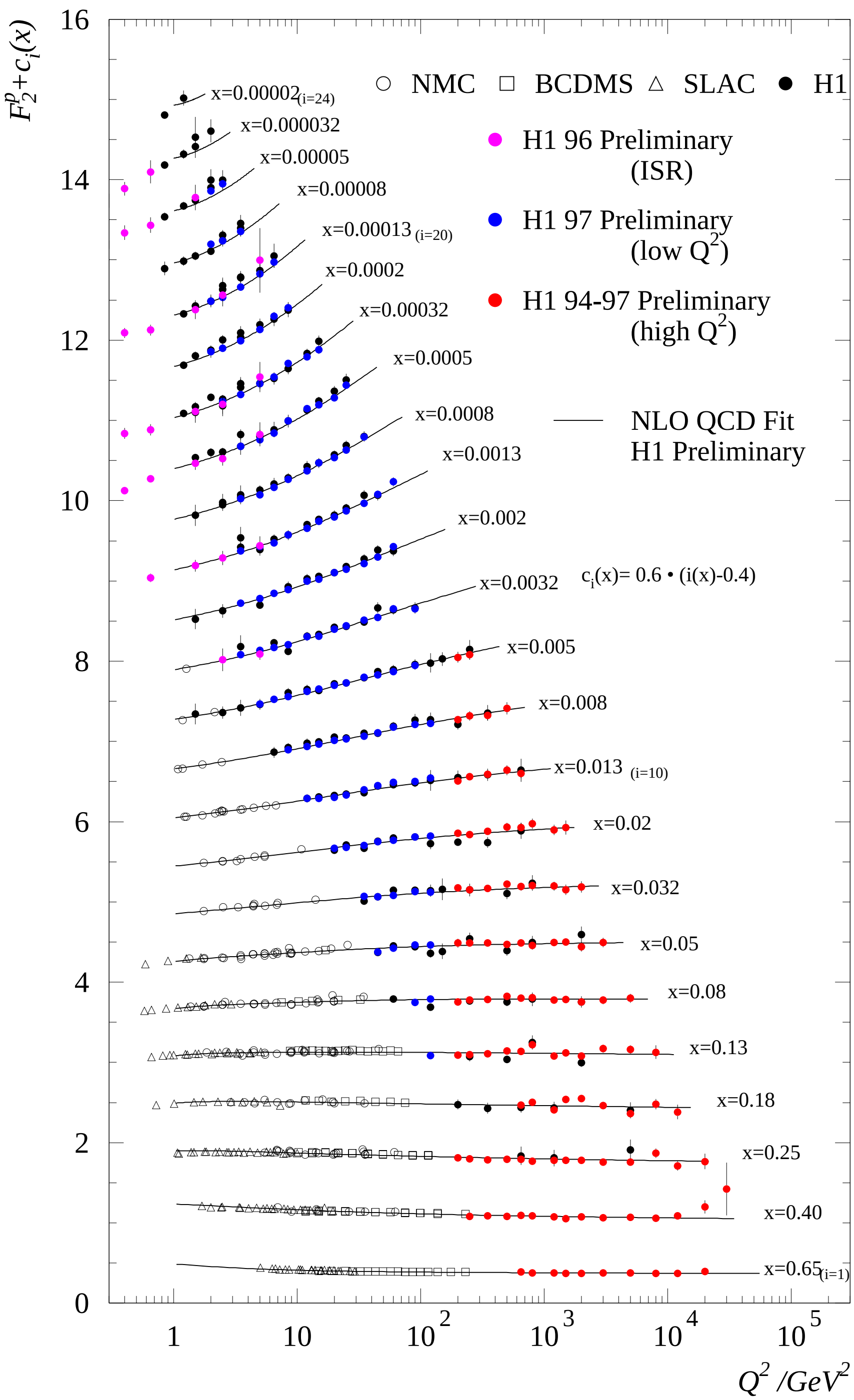
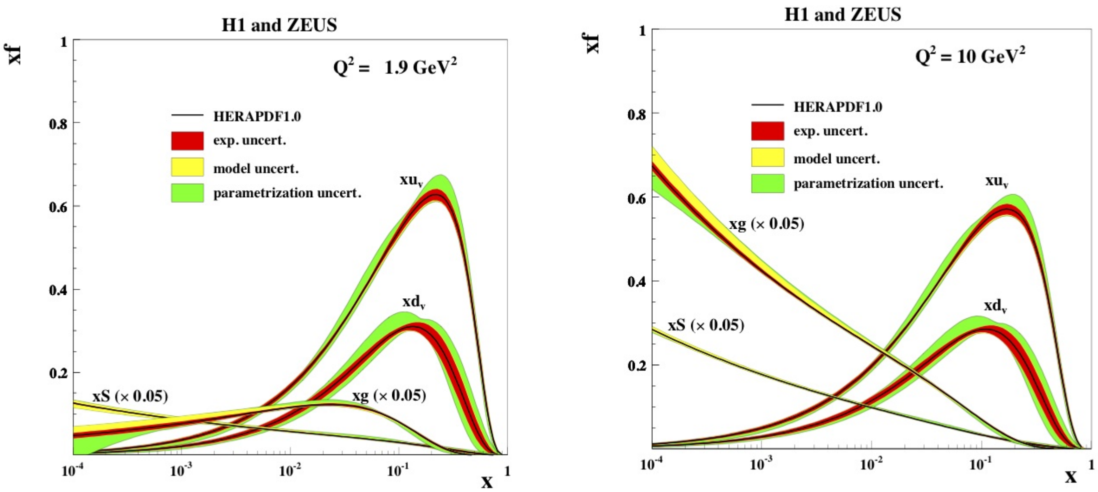


Figure 9.17: Proton structure function F_2^p measured by H1 and other experiments for various values of Q^2 and x. Scaling violations appear for $x < 10^{-2}$.

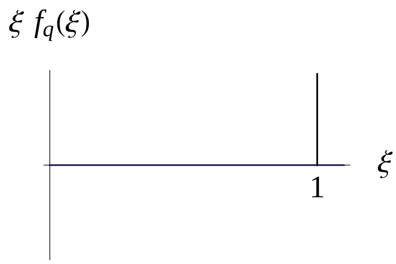
$m_p \simeq 3m_q$, see Fig. 9.19(b). If, however, the proton consisted of three *light* and *strongly coupled* quarks, $m_q \ll 1/3m_p$, the peaks of $\xi f(\xi)$ would still be located around 1/3, but, since most energy is present in the form of potential and kinetic energy, they would be



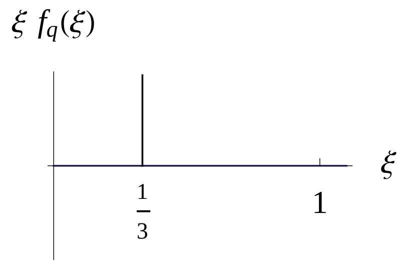
(a) (b)

Figure 9.18: *Parton distribution functions of the proton* (a) $Q^2 = 1.9\,\text{GeV}^2$. (b) $Q^2 = 10\,\text{GeV}^2$. The sea and gluon PDFs are reduced by a factor 20.

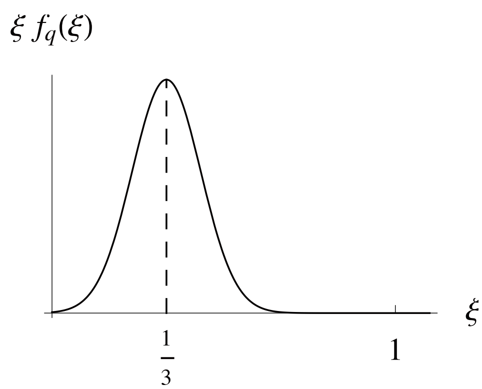
smeared out significantly at any given instant of time, as shown in Fig. 9.19(c).



(a) Pointlike proton.



(b) Three massive, weakly interacting quarks: $m_p \simeq 3m_q$.



(c) Three light, strongly coupled quarks: $m_q \ll m_p/3$.

Figure 9.19: *Quark momentum density* $\xi f_q(\xi)$.

Let us consider the kinematics of the simple parton model. The on-shell condition for the outgoing quark (see Fig. 9.20(a)) yields

$$m_q^2 = (\xi p + q)^2 \simeq 2p \cdot q\xi - Q^2 = \frac{Q^2}{x}\xi - Q^2 \Rightarrow \xi = \left(1 + \frac{m_q^2}{Q^2}\right) x \simeq x.$$

Therefore, given the assumptions made are valid, the Bjorken variable x is the momentum fraction ξ of a parton inside the proton.

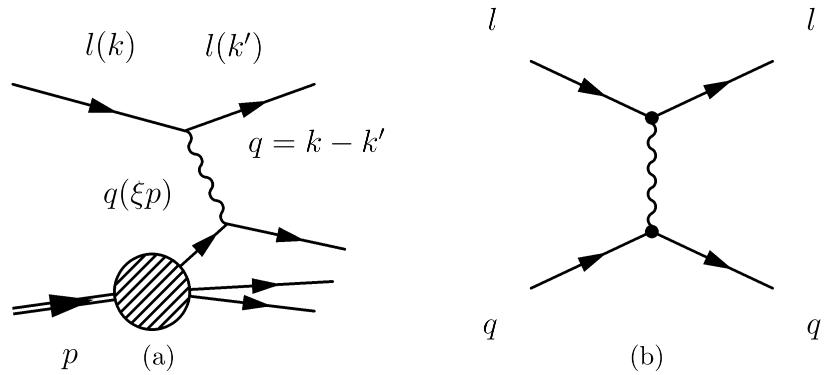


Figure 9.20: *(a) Kinematics of simple parton model and (b) Feynman diagram for lepton-quark scattering.*

To determine $d^2\hat{\sigma}^{lq}/dxdQ^2$ of lepton-quark scattering, we consider the Feynman diagram in Fig. 9.20(b) which is just a crossing of the Born level diagram for $e^+e^- \to \mu^+\mu^-$ (see Sect. 5.10, p. 94). We therefore find

$$\frac{d\hat{\sigma}^{lq}}{dt} = \frac{2\pi\alpha^2 e_q^2}{\hat{s}^2}\left(\frac{\hat{s}^2 + \hat{u}^2}{\hat{t}^2}\right)$$

where the Mandelstam variables read (the subscript ep emphasizes that s_{ep} refers to the lepton-proton system)

$$\begin{aligned}\hat{s} &= (xp + k)^2 = 2xpk = xs_{ep}\\ \hat{t} &= -Q^2 = -xys_{ep} = t\\ \hat{u} &= -\hat{s} - \hat{t} = -x(1-y)s_{ep}.\end{aligned}$$

Note that $\hat{t} = t$ depends only on the lepton kinematics. This leads to the lepton-quark differential cross section

$$\frac{d^2\hat{\sigma}^{lq}}{dxdQ^2} = \frac{2\pi\alpha^2 e_q^2}{Q^4}\left(1 + (1-y)^2\right)\delta(x - \xi).$$

Inserting this result into the parton model expression for lepton-proton scattering of Eq. (9.18) yields

$$\frac{d^2\sigma}{dxdQ^2} = \frac{4\pi\alpha^2}{xQ^4}\sum_q \int_0^1 d\xi f_q(\xi) e_q^2 \frac{x}{2}\left(1 + (1-y)^2\right)\delta(x - \xi).$$

Upon comparison with the deep inelastic scattering structure functions we find

$$F_2(x,Q^2) = \sum_q e_q^2 x f_q(x)$$
$$F_L(x,Q^2) = F_2(x,Q^2) - 2xF_1(x,Q^2) = 0$$

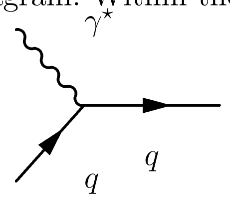
where F_L is called longitudinal structure function. We recognize that $F_2(x,Q) = F_2(x)$ ceases to be a function of two variables, but under the assumed conditions depends only on one variable, a phenomenon generally referred to as scaling. Furthermore, $F_L = 0 \Leftrightarrow 2xF_1 = F_2$ is the Callan-Gross relation, a consequence of quarks having spin 1/2 familiar from Sect. 9.3.3.

Before we go on we introduce the following notation for the distribution functions

$$f_q(x) = q(x) \quad (q = u, d, s, c, \ldots, \bar{u}, \ldots)$$
$$f_g(x) = g(x) \quad \text{(gluons)}.$$

9.7 QCD corrections to the parton model

Our discussion of the parton model involved no QCD corrections up to now; it rested on the assumption of electromagnetic interactions alone. QCD corrections will concern the quark part of our diagram. Within the parton model we just found



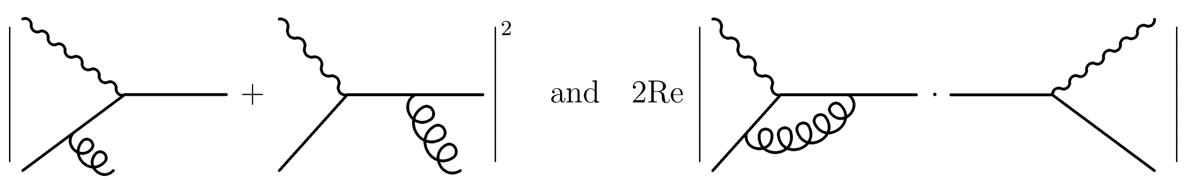
$$= \frac{4\pi\alpha e_q^2}{\hat{s}}\delta(x-\xi) =: \hat{\sigma}_0\delta(x-\xi) \tag{9.19}$$

and

$$\frac{F_2(x,Q^2)}{x} = \sum_q \int_0^1 \frac{d\xi}{\xi} q(\xi) e_q^2\, \delta\left(1 - \frac{x}{\xi}\right) \tag{9.20}$$

where $\hat{\sigma}_0$ is the QED contribution which drops out of the structure functions.

The $\mathcal{O}(\alpha_s) = \mathcal{O}(g_s^2)$ QCD corrections are given by



i. e. gluon radiation and virtual gluon exchange. The one-loop virtual gluon interference term stems from the loop corrections to the quark-photon vertex squared at $\mathcal{O}(g_s^2)$. As an example, consider the process $\gamma^\star q \to qg$ (which is a crossing of $\gamma^\star \to q\bar{q}g$):

$$|\mathcal{M}|^2 = 32\pi^2(e_q^2\alpha\alpha_s)C_F\left(-\frac{\hat{t}}{\hat{s}} - \frac{\hat{s}}{\hat{t}} + \frac{2\hat{u}Q^2}{\hat{s}\hat{t}}\right).$$

This expression for $|\mathcal{M}|^2$ is unproblematic for small $\hat{s}$, since $\hat{s}$ is fixed. However, a problem arises at small $\hat{t}$, since we have to integrate over it as it is a dynamic variable (see Sect. 3.3.2, p. 28).

For small scattering angles $-\hat{t} \ll \hat{s}$ and we have

$$p_T^2 = \frac{\hat{s}(-\hat{t})}{\hat{s} + Q^2}$$

for the transverse momentum of the outgoing gluon. Eliminating the Mandelstam variable $\hat{u}$, the differential cross section becomes

$$\frac{d\hat{\sigma}}{dp_T^2} = \frac{1}{16\pi\hat{s}^2}|\mathcal{M}|^2 \simeq \hat{\sigma}_0 \frac{\alpha_s}{2\pi} C_F \left(-\frac{1}{\hat{t}\hat{s}}\left[\hat{s} + \frac{2(\hat{s}+Q^2)Q^2}{\hat{s}}\right]\right).$$

By introducing the dimensionless variable

$$z = \frac{x}{\xi} = \frac{Q^2}{2p_q \cdot q} = \frac{Q^2}{\hat{s} + Q^2},$$

we arrive at

$$\frac{d\hat{\sigma}}{dp_T^2} = \hat{\sigma}_0 \frac{1}{p_T^2} \frac{\alpha_s}{2\pi} P_{qq}(z)$$

where

$$P_{qq}(z) = C_F \frac{1+z^2}{1-z}$$

(compare Sect. 9.4.2). Note that in the simple parton model we had $p_q = \xi p$ which is no longer the case when QCD corrections are taken into account.

To find the inclusive cross section, we have to integrate over the transverse momentum squared:

$$\frac{\hat{\sigma}^{\gamma^\star q \to qg}}{\hat{\sigma}_0} = \frac{\alpha_s}{2\pi} P_{qq}(z) \int\limits_{\mu^2}^{Q^2} \frac{dp_T^2}{p_T^2} = \frac{\alpha_s}{2\pi} P_{qq}(z) \log\frac{Q^2}{\mu^2}$$

where the infrared cutoff μ^2 has been introduced because of the collinear singularity at $p_T^2 \to 0$. The rationale is to later define observables in a way that allows to send $\mu^2 \to 0$ (compare also Sect. 8.2.1, p. 162). Having calculated the QCD corrections at $\mathcal{O}(\alpha_s)$ to the structure function in Eq. (9.20), we can state the resulting corrected expression:

$$\frac{F_2(x,Q^2)}{x} = \sum_q \int\limits_x^1 \frac{d\xi}{\xi} q(\xi) e_q^2 \left\{\delta\Big(1 - \frac{x}{\xi}\Big) + \frac{\alpha_s}{2\pi}\left[P_{qq}\Big(\frac{x}{\xi}\Big)\log\frac{Q^2}{\mu^2} + \text{finite}\right] + \mathcal{O}(\alpha_s^2)\right\} \tag{9.21}$$

which leads to some interesting consequences.[3] Observe that we found an equality of a measurable and hence finite quantity (after all, F_2 is just a specific coefficient in the parametrization of a cross section) and an expression which is divergent at the given order of perturbation theory. Since the LHS of Eq. (9.21) is fixed, the problem has to be tackled on its RHS. As a starting point, recall that we justified the form of the quark distribution functions by asymptotic freedom and neglected QCD interactions among the quarks in the first place. When QCD corrections are taken into account, the naive parton model is no longer valid. Therefore, it is necessary to redefine the parton distribution functions such that they are well-defined for the case of interacting quarks. This amounts to a redefinition of the quark distribution in the infrared region and is called mass factorization of the quark distribution:

$$q(x,\mu_F^2) = q(x) + \frac{\alpha_s}{2\pi}\int_x^1 \frac{d\xi}{\xi} q(\xi) P_{qq}\Big(\frac{x}{\xi}\Big) \log\frac{\mu_F^2}{\mu^2} \tag{9.22}$$

where $q(x,\mu_F^2)$ is a measurable, screened quark density, $q(x)$ denotes the bare (unphysical) quark density, and the integral term is the contribution from unresolvable gluon radiation with transverse momentum $\mu_F^2 \geq p_T^2 \geq \mu^2$ where μ_F^2 is the mass factorization scale at which the quark distribution is measured. Recall that the infrared cutoff μ^2 can be chosen arbitrarily small—smaller than any given detector resolution. At sufficiently small scattering angles the emitted gluon cannot be resolved by the detector as it appears to be parallel to the proton remnants. Two-jet events in deep inelastic scattering can only be excluded in the momentum range where they could be detected. Therefore, the quark distribution $q(x,\mu_F^2)$ admits gluon radiation below a predefined resolution scale μ_F.

Let us solve for $q(x)$ in Eq. (9.22) and plug it into the QCD corrected structure function in Eq. (9.21), we have

$$\begin{aligned}\frac{F_2(x,Q^2)}{x} &= \sum_q \int_x^1 \frac{d\xi}{\xi} q(\xi,\mu_F^2) e_q^2 \left\{\delta\Big(1-\frac{x}{\xi}\Big) + \frac{\alpha_s}{2\pi} P_{qq}\Big(\frac{x}{\xi}\Big)\log\frac{Q^2}{\mu^2} - \frac{\alpha_s}{2\pi} P_{qq}\Big(\frac{x}{\xi}\Big)\log\frac{\mu_F^2}{\mu^2}\right\} \\ &= \sum_q \int_x^1 \frac{d\xi}{\xi} q(\xi,\mu_F^2) e_q^2 \left\{\delta\Big(1-\frac{x}{\xi}\Big) + \frac{\alpha_s}{2\pi} P_{qq}\Big(\frac{x}{\xi}\Big)\log\frac{Q^2}{\mu_F^2}\right\}\end{aligned}$$

which is independent of the infrared cutoff μ^2 and finally, setting $\mu_F^2 = Q^2$ as in deep inelastic scattering experiments,

$$= \sum_q q(x,Q^2) e_q^2 .$$

Perturbative QCD is used to answer the question how the Q^2 dependence of the quark distribution $q(x,Q^2)$ looks like.

[3] One can observe, as was done before, that because of QCD corrections to the naive parton model scaling no longer holds, since $F_2(x,Q^2)$ ceases to be a function of the single variable x alone.

9.8 Altarelli-Parisi equations

The bare quark distribution $q(x)$ is independent of μ_F^2:

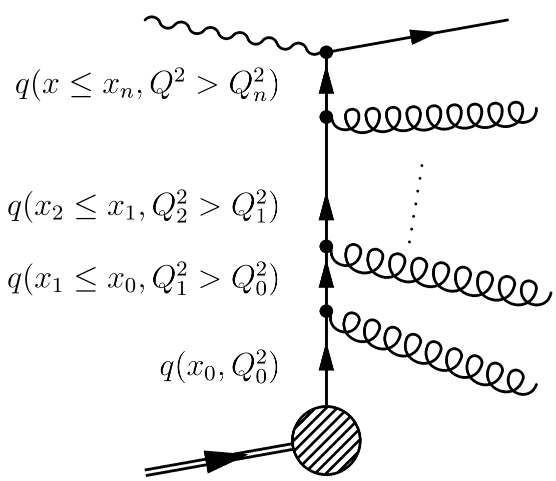
$$\mu_F^2 \frac{d}{d\mu_F^2} q(x) = 0.$$

Differentiating Eq. (9.22) with respect to $\log \mu_F^2$ we thus obtain the renormalization group equation[4] for the quark distribution:

$$\frac{\partial q(x,\mu_F^2)}{\partial \log \mu_F^2} = \frac{\alpha_s}{2\pi} \int_x^1 \frac{d\xi}{\xi} q(\xi,\mu_F^2) P_{qq}\Big(\frac{x}{\xi}\Big) \tag{9.23}$$

which means that scaling invariance is logarithmically violated.

Eq. (9.23) is known as the **Dokshitzer-Gribov-Lipatov-Altarelli-Parisi** (DGLAP) equation, or simply Altarelli-Parisi evolution equation. It is a small-p_T^2 approximation, which resums the collinear gluon radiation in the initial state at $\mathcal{O}(\alpha_s^n \log^n Q^2)$.



This diagram is a universal correction, since the emitted gluons do not know about the scattering process of the quark off the virtual photon. The DGLAP equation tells us what happens if one infinitesimally increases the resolution. It is an integro-differential equation with one "initial condition" $q(x,\mu_F^2 = \mu_0^2)$. Knowing the latter, one can compute the quark distribution at any value of μ_F^2. The procedure is analogous to the determination of the running coupling of QED (Sect. 6.1.2, p. 104) or QCD (Sect. 7.4.2, p. 150).

In using Eq. (9.23) we omitted until now, the fact that $P_{qq}(z)$ has a singularity in $z = 1$, which belongs to the integration domain. This singularity corresponds to the emitted

[4] For a concise discussion of this topic see [53, pp. 28].

gluon becoming soft. It is compensated by a singularity in the virtual corrections. As a result, $P_{qq}(z)$ is modified to become,

$$P_{qq}(z) = C_F \left(\frac{1+z^2}{(1-z)_+} + \frac{3}{2}\delta(1-z) \right),$$

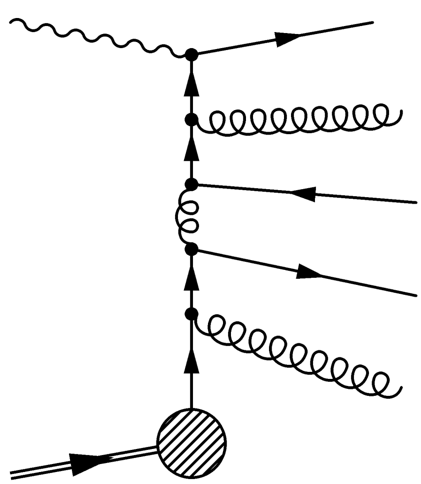
which takes into account the virtual corrections occuring at $z = 1$. We use the '+'-presciption, coming from the reguarisation procedure and defined by,

$$\int_0^1 dz \frac{f(z)}{(1-z)_+} = \int_0^1 dz \frac{f(z) - f(1)}{1-z}. \tag{9.24}$$

The factor in front of the δ-function can be inferred from the quark number conservation, which can be stated as,

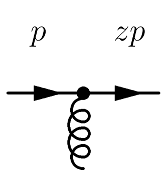
$$\int_0^1 dz P_{qq}(z) = 0. \tag{9.25}$$

Up to now, we considered only gluon radiation off a quark. However, the emission history can be made more complicated with gluons at intermediate stages of the parton cascade,



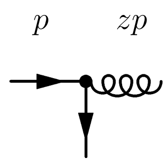
By inspection, one can find out that there are four different splitting processes at $O(\alpha_s)$:

- $q \to q$:



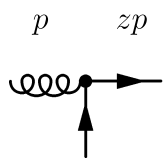
$$\rightsquigarrow P_{qq}(z) = C_F \left(\frac{1+z^2}{(1-z)_+} + \frac{3}{2}\delta(1-z) \right), \tag{9.26}$$

- $q \to g$:



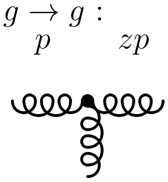
$$\leftrightsquigarrow P_{gq}(z) = C_F \left(\frac{1 + (1-z)^2}{z} \right), \tag{9.27}$$

- $g \to q$:



$$\leftrightsquigarrow P_{qg}(z) = T_F \left(z^2 + (1-z)^2 \right), \tag{9.28}$$

- $g \to g$:



$$\leftrightsquigarrow P_{gg}(z) = 2C_A \left(\frac{z}{(1-z)_+} + \frac{1-z}{z} \right) + \left(\frac{11}{6} C_A - \frac{3}{2} T_F n_f \right) \delta(1-z). \tag{9.29}$$

Those splitting functions satisfy a set of coupled DGLAP equations,

$$\frac{\partial}{\partial \log \mu_F^2} \begin{pmatrix} q(x, \mu_F^2) \\ g(x, \mu_F^2) \end{pmatrix} = \frac{\alpha_s(\mu_F^2)}{2\pi} \int_x^1 \frac{dz}{z} \begin{pmatrix} P_{qq}(z) & P_{qg}(z) \\ P_{gq}(z) & P_{gg}(z) \end{pmatrix} \begin{pmatrix} q\left(\frac{x}{z}, \mu_F^2\right) \\ g\left(\frac{x}{z}, \mu_F^2\right) \end{pmatrix}. \tag{9.30}$$

In this equation, $\frac{\alpha_s}{2\pi} P_{ji}(z)$ is the probability for $i \to j$ splitting with momentum fraction z in the transverse momentum interval $[\log \mu_F^2, \log \mu_F^2 + d \log \mu_F^2]$.

For n_f quark flavours, we get $2n_f + 1$ coupled equations (antiquarks must be taken explicitly into account). This system can be diagonalized be introducing (i labels the flavour),

- n_f valence quark distributions

$$q_i^V = q_i - \bar{q}_i, \tag{9.31}$$

- $n_f - 1$ flavour non-singlet quark distributions

$$q_i^F = \sum_{n=1}^{i-1} (q_n + \bar{q}_n - q_i - \bar{q}_i), \tag{9.32}$$

- 1 flavour singlet quark distribution

$$q^S = \sum_{n=1}^{n_f} (q_n + \bar{q}_n). \tag{9.33}$$

We also define the convolution,

$$(P \otimes q)(x, \mu_F^2) = \int_x^1 \frac{dz}{z} P(z) q\left(\frac{x}{z}, \mu_F^2\right),$$

allowing us to write,

$$\frac{\partial q_i^V}{\partial \log \mu_F^2} = \frac{\alpha_s}{2\pi} P_{qq} \otimes q_i^V \tag{9.34}$$

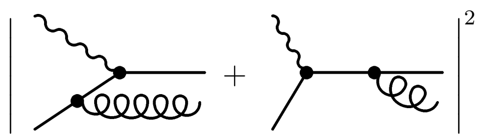
$$\frac{\partial q_i^F}{\partial \log \mu_F^2} = \frac{\alpha_s}{2\pi} P_{qq} \otimes q_i^F \tag{9.35}$$

$$\frac{\partial q^S}{\partial \log \mu_F^2} = \frac{\alpha_s}{2\pi} \left(P_{qq} \otimes q^S + 2n_f P_{qg} \otimes g\right) \tag{9.36}$$

$$\frac{\partial g}{\partial \log \mu_F^2} = \frac{\alpha_s}{2\pi} \left(P_{gq} \otimes q^S + P_{gg} \otimes g\right). \tag{9.37}$$

The factor $2n_f$ in Eq. (9.36) comes from the fact that one needs to consider quarks and antiquarks of all possible flavours. This set of equations only includes leading order corrections that are precise at 15%. The data obtained in the last years yield however results to the 5% precision, so that correction from higher orders need to be taken into account.

At NLO, $\mathcal{O}(\alpha_s^n \log^{n-1} Q^2)$, the finite term from the $\mathcal{O}(\alpha_s)$-processes is relevant,

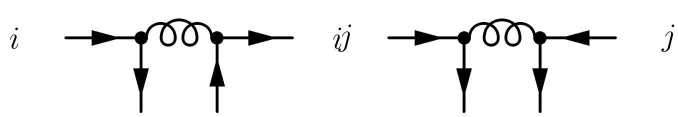


This translates in the expressions for the structure functions,

$$\frac{1}{x} F_2(x, Q^2) = \int_x^1 \frac{d\xi}{\xi} \left\{ \sum_q q\left(\xi, Q^2\right) \left[\delta\left(1 - \frac{x}{\xi}\right) + \frac{\alpha_s}{2\pi} C_{2,q}\left(\frac{x}{\xi}\right)\right] + g(\xi, Q^2) \frac{\alpha_s}{2\pi} C_{2,g}\left(\frac{x}{\xi}\right) \right\} \tag{9.38}$$

$$F_L(x, Q^2) = \mathcal{O}(\alpha_s) \neq 0 \tag{9.39}$$

We now need to compute $O(\alpha_s^2)$-corrections to the spitting functions P_{ji}. At this order, there is essentially one new spitting process with two quark-gluon vertices,



At $O(\alpha_s)$, we had implicitly $P_{qq}^V = P_{qq}^F = P_{qq}^S = P_{qq}$ in Eqs. (9.34), (9.35) and (9.36). This is no longer true at $O(\alpha_s^2)$, where all these splitting functions are different from one another. At even higher orders, no essentially new features appear, so that NLO calculations lead already quite acceptable results. These are of crucial importance for W and Z production at hadron colliders.

9.9 Solution of DGLAP equations

Looking at the set (9.30) of coupled DGLAP integro-differential equations one can expect that solving it could be a highly non-trivial task. There are basically two approaches to attack the problem :

1. Numerical solution, e.g. with the Runge-Kutta method. This approach is yielding satisfactory results for $Q_0^2 \gtrsim 2\,\mathrm{GeV}$, i.e. in the asymptotically free regime, where $\alpha_s(Q_0^2) \ll 1$,
2. Analytically, by using Mellin tranformation. This approach is especially useful to obtain a quantitative understanding and to determine the asymptotic properties.

In both cases we have to start from given initial distributions $q_i(x, Q_0^2), \bar{q}_i(x, Q_0^2), g(x, Q_0^2)$.

Mellin transformation The Mellin transform of a function $f : [0,1] \to \mathbb{R}$ is given by,

$$f(n) = M[f(x)] = \int_0^1 dx x^{n-1} f(x), \tag{9.40}$$

with inverse

$$f(x) = \frac{1}{2\pi i} \int_{a-i\infty}^{a+i\infty} dn x^{-n} f(n), \tag{9.41}$$

for $f(n)$ analytical in the half plane $\mathrm{Re}\, n > a$.

We list here some of the properties of Mellin transformations:

$$M[af(x) + bg(x)] = af(n) + bg(n) \qquad \text{(linearity)} \tag{9.42}$$

$$M\left[\frac{d^k}{dx^k} f(x)\right] = (-1)^{n-k} \frac{\Gamma(n)}{\Gamma(n-k)} f(n-k) \qquad \text{(derivative)} \tag{9.43}$$

$$M[(f \otimes g)(x)] = f(n)g(n) \qquad \text{(convolution)} \tag{9.44}$$

Armed with this new technology, we Mellin transform Eq. (9.34) with respect to the x variable to get (the following analysis is valid for the valence and flavour non-singlet quark distribution, thus, we drop the i, V/F for notational convenience),

$$\frac{\partial q(n,\mu_F^2)}{\partial \log \mu_F^2} = \frac{\alpha_s(\mu_F^2)}{2\pi} P_{qq}(n) q(n,\mu_F^2). \tag{9.45}$$

Using the evolution equation for α_s (Sect. 7.4.2, p. 153) in the leading order approximation,

$$\frac{1}{\alpha_s}\frac{\partial \alpha_s}{\partial \log \mu_F^2} = \frac{\partial \log \alpha_s}{\partial \log \mu_F^2} = -\frac{\beta_0}{4\pi}\alpha_s,$$

one gets,

$$\begin{aligned} \frac{\partial q(n,\mu_F^2)}{\partial \log \alpha_s} &= -\frac{2}{\beta_0} P_{qq}(n) q(n,\mu_F^2) \\ \frac{\partial \log q(n,\mu_F^2)}{\partial \log \alpha_s} &= -\frac{2}{\beta_0} P_{qq}(n), \end{aligned} \tag{9.46}$$

which can now be solved by integrating from $\mu_F^2 = Q_0^2$ to Q^2,

$$q(n,Q^2) = q(n,Q_0^2)\left[\frac{\alpha_s(Q_0^2)}{\alpha_s(Q^2)}\right]^{\frac{2}{\beta_0}P_{qq}(n)},$$

or, in the usually known form, using Eq. (7.44), p. 154,

$$\boxed{q(n,Q^2) = q(n,Q_0^2)\exp\left\{\frac{2}{\beta_0}P_{qq}(n)\log\frac{\log(Q^2/\Lambda^2)}{\log(Q_0^2/\Lambda^2)}\right\}}. \tag{9.47}$$

This is the solution for the quark valence and flavour non-singlet distributions.

We now turn to the two remaining distributions, namely the quark singlet and and gluon distributions. Mellin transforming Eqs. (9.36) and (9.37) yields,

$$\frac{\partial}{\partial \log \mu_F^2}\begin{pmatrix} q^S(n,\mu_F^2) \\ g(n,\mu_F^2) \end{pmatrix} = -\frac{2}{\beta_0}\begin{pmatrix} P_{qq}(n) & 2n_f P_{qg}(n) \\ P_{gq}(n) & P_{gg}(n) \end{pmatrix}\begin{pmatrix} q^S(n,\mu_F^2) \\ g(n,\mu_F^2) \end{pmatrix}. \tag{9.48}$$

The first step is the diagonalization of the matrix,

$$\begin{pmatrix} P_{qq}(n) & 2n_f P_{qg}(n) \\ P_{gq}(n) & P_{gg}(n) \end{pmatrix}.$$

Then one applies the same formalism as for the valence quark distribution discussed above. By inverse Mellin transformation, one gets the result in the variable x.

Specific values of n correspond to various physical quantities. For example, $P_{qq}(n=1) = 0$ is the Mellin transform of Eq. (9.25) and $q(n=2)$ corresponds to the fraction of the total momentum transported by the quark q. One has the momentum sum rule,

$$q^S(2,Q^2) + g(2,Q^2) = 1.$$

with the asymptotic values,

$$
\begin{aligned}
q^S(2, Q^2 \to \infty) &\to \frac{3n_f}{16+3n_f} \overset{n_f=5}{=} \frac{15}{31} \\
g(2, Q^2 \to \infty) &\to \frac{16}{16+3n_f} \overset{n_f=5}{=} \frac{16}{31}.
\end{aligned}
$$

9.10 Observables at hadron colliders

We now study processes and observables at hadron colliders and the consequences of parton evolution in this context.

The simple parton model cross section for processes at hadron-hadron colliders reads

$$\sigma_{pp} = \sum_{i,j\in\{q,g\}} \int dx_1 dx_2 f_i(x_1) f_j(x_2) \hat{\sigma}_{ij\to X}(s_{ij} = x_1 x_2 s_{pp}), \tag{9.49}$$

i. e. two partons enter into a hard collision from which a final state X emerges, as shown in Fig. 9.21(a).

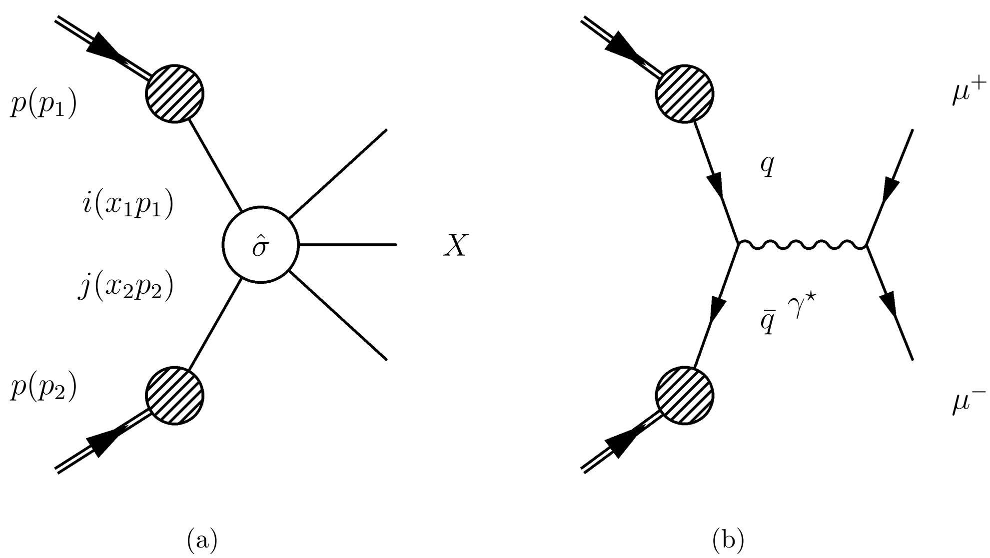


Figure 9.21: *(a) Hadron-hadron collision in naive parton model and (b) Drell-Yan process.*

As an example consider the Drell-Yan process, $pp \to \mu^+\mu^-$, shown in Fig. 9.21(b). The parton model cross section reads

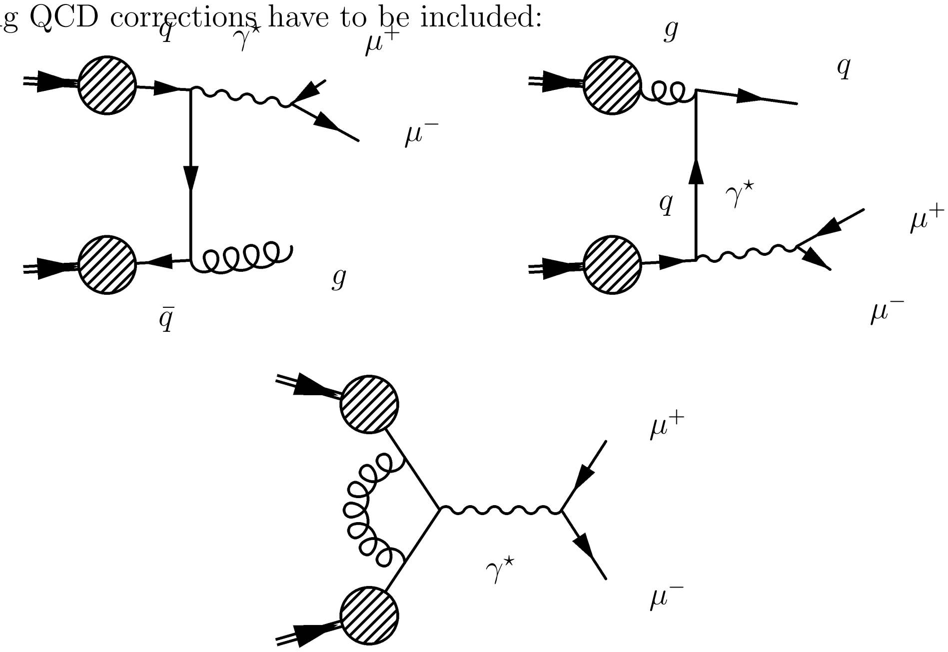
$$\sigma^{\mathrm{DY}} = \sum_q \int dx_1 dx_2 \left[q(x_1)\bar{q}(x_2) + q(x_2)\bar{q}(x_1)\right] \hat{\sigma}_{q\bar{q}\to\mu^+\mu^-} \tag{9.50}$$

where

$$\hat{\sigma}_{q\bar{q}\to\mu^+\mu^-} = \underbrace{\frac{4\pi\alpha^2}{3s_{q\bar{q}}}\frac{1}{3}}_{\hat{\sigma}_0^{\mathrm{DY}}}\, e_q^2\,\delta(1 - x_1 x_2 s_{pp}/M^2_{\mu^+\mu^-}) \tag{9.51}$$

which we basically already calculated before (Sect. 5.10, p. 94). The difference to the $e^+e^- \to \mu^+\mu^-$ result is the color factor of 1/3 and the delta function which states that the muon pair invariant mass fulfills $(p_{\mu^+} + p_{\mu^-})^2 =: M^2_{\mu^+\mu^-} = x_1 x_2 s_{pp}$.

The following QCD corrections have to be included:



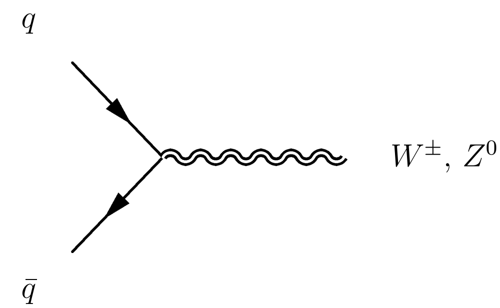
where the first two diagrams are because of parton evolution and the third diagram is a virtual correction. Setting $z = x_1 x_2 s_{pp}/M^2_{\mu^+\mu^-}$, the QCD corrected Drell-Yan cross section reads

$$\sigma^{\mathrm{DY}} = \hat{\sigma}_0^{\mathrm{DY}} \sum_q e_q^2 \int dx_1 dx_2 \Big\{ q(x_1)\bar{q}(x_2)\delta(1-z) + \frac{\alpha_s}{2\pi} C_{q\bar{q}}(z) + [q(x_1) + \bar{q}(x_1)]\, g(x_2) \frac{\alpha_s}{2\pi} C_{qg}(z) + (x_1 \leftrightarrow x_2) \Big\}$$

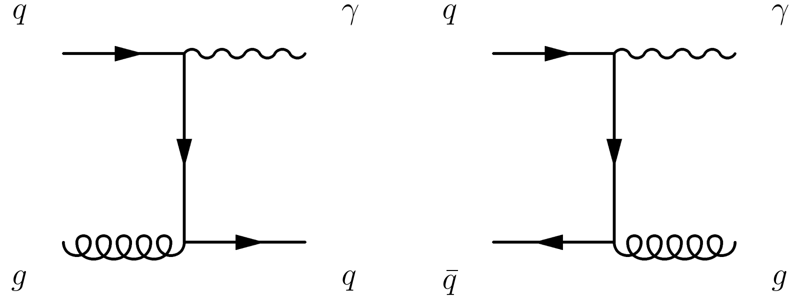
where $q(x_i)$ etc. are the QCD evolved parton distributions.

In the following some standard reactions are listed.

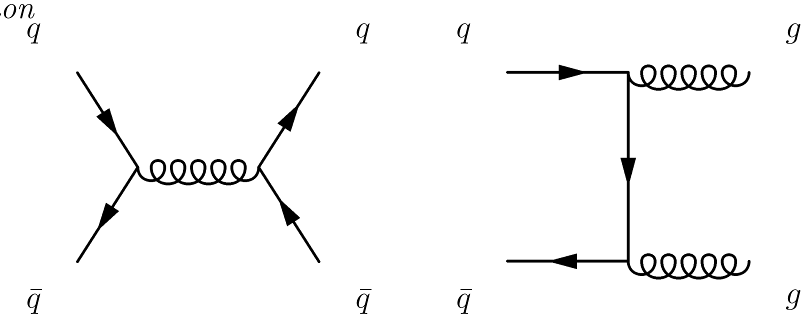
- *$W^\pm$, Z^0 production*



- $\gamma +$ jet *production*



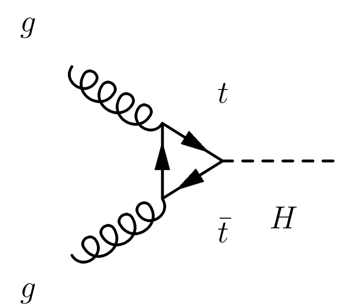
- *2-jet production*



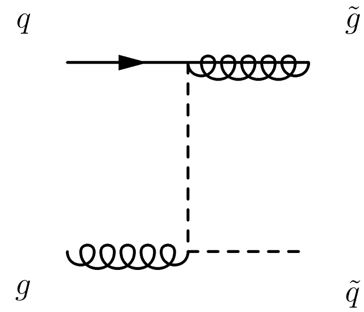
  Further processes leading to 2-jet events are $qg \to qg$, $gg \to gg$, $gg \to q\bar{q}$ and $qq \to qq$.

Examples for relevant processes in searches for new physics:

- *Higgs production*



- *SUSY particles*



A general feature of hadron-hadron colliders is that $\sqrt{s_{\text{parton-parton}}}$ is variable since the parton momentum fractions vary.[5] This allows to search for peaks in mass spectra at fixed collider energy. An example for this effect is the Z^0 peak in the $\mu^+\mu^-$ spectrum of SPS at CERN (compare also Sect. 4.4.4, p. 52).

9.11 Multiparticle production

Describing multijet final states in QCD is problematic because of two reasons.

- *Factorial growth of the number of diagrams*
  E. g. for $gg \to ng$ the number of diagrams # scales with the number of final state gluons n in the following way:

n	2	3	4	5	6	7
#	4	25	220	2485	34300	559405.

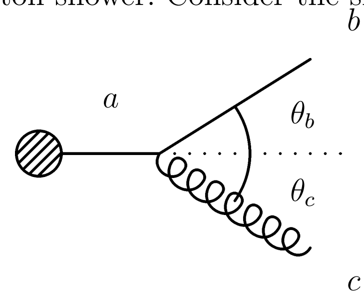
  These numbers illustrate that a computation even on the amplitude level is time-consuming.

- *Complexity of the final state phase space*
  In addition to the aforementioned problem, the final state phase space has high dimension and the integrations are constrained in various ways.

These problems can be approached by introducing approximate descriptions. One uses the fact that $|\mathcal{M}|^2$ is largest if partons are emitted into soft ($E \to 0$) or collinear ($\theta_{ij} \to 0$) regions of phase space. Therefore, the dominant contributions stem from these phase space regions.

[5] Compare this to the e^+e^- case where the center of mass energy of the actual collision is fixed by the collider energy: $s = \hat{s}$.

Let us analyze a collinear parton shower. Consider the shower subgraph



where $p_a^2 \gg p_b^2, p_c^2$ and $p_a^2 = t$. The opening angle is $\theta = \theta_b + \theta_c$ and the energy fractions are

$$z = \frac{E_b}{E_a} \qquad\qquad 1 - z = \frac{E_c}{E_a}. \tag{9.52}$$

For small angles we have

$$t = 2E_bE_c(1 - \cos\theta) = z(1-z)E_a^2\theta^2 \tag{9.53}$$

$$\frac{\theta_b}{1-z} = \frac{\theta_c}{z} = \theta. \tag{9.54}$$

For $\theta \to 0$ the matrixelement factorizes as

$$|\mathcal{M}_{n+1}|^2 = \frac{4g_s^2}{t} C_F F_{qq}(z) |\mathcal{M}_n|^2$$

where

$$F_{qq}(z) = \frac{1+z^2}{1-z} = P_{qq}(z < 1).$$

Analogous splittings involve F_{qg}, F_{gq}, and F_{gg}.

Also the phase space factorizes:

$$\begin{aligned} d\phi_n &= \ldots \frac{d^3p_a}{2E_a(2\pi)^3} \\ d\phi_{n+1} &= \ldots \frac{d^3p_b}{2E_b(2\pi)^3} \frac{d^3p_c}{2E_c(2\pi)^3}. \end{aligned}$$

Since $p_c = p_a - p_b$, we have $d^3p_c = d^3p_a$ for fixed p_b. For small θ this yields[6]

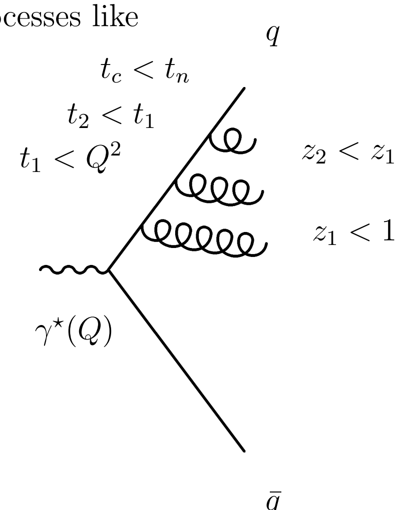
$$
\begin{aligned}
d\phi_{n+1} &= d\phi_n \frac{1}{2(2\pi)^3} \int E_b dE_b \theta_b d\theta_b d\phi \frac{dz}{1-z} \delta(z - E_b/E_a) dt \delta(t - E_a E_b \theta^2) \\
&= d\phi_n \frac{1}{4(2\pi)^3} dt dz d\phi
\end{aligned}
$$

(recall Eq. (9.52) and (9.53)).

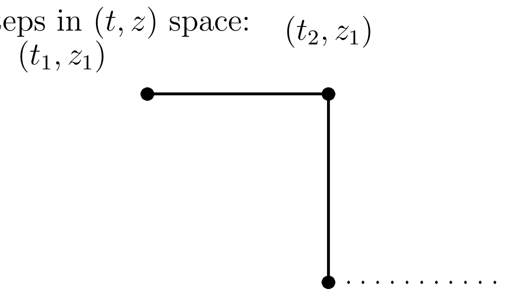
Since the matrixelement and the phase space factorize, so does the cross section:

$$
d\sigma_{n+1} = d\sigma_n \frac{dt}{t} dz \frac{d\phi}{2\pi} \frac{\alpha_s}{2\pi} C_F F(z).
$$

Therefore, multiple emission processes like



where t_c is a cutoff scale at which hadronization sets in, $t_c \gtrsim \Lambda^2_{\mathrm{QCD}}$, can be subdivided into fundamental steps in (t, z) space:



A Monte Carlo method to generate a corresponding set of final state partons proceeds as

[6] One observes that

$$
\begin{aligned}
d\phi_{n+1} &= \ldots \frac{d^3p_b}{2E_b(2\pi)^3} \frac{d^3p_c}{2E_c(2\pi)^3} = d\phi_n \frac{E_a}{E_c} \frac{d^3p_b}{(2\pi)^3 2E_b} \\
&\simeq d\phi_n \frac{E_a}{E_c} \frac{E_b dE_b}{2(2\pi)^3} \theta_b d\theta_b d\phi = d\phi_n \frac{1}{1-z} \frac{E_b dE_b}{2(2\pi)^3} \theta_b d\theta_b d\phi.
\end{aligned}
$$

And the Jacobian determinant is just $2zE_a\theta_b/(1-z)$.

follows: Starting from a simple final state (e. g. $e^+e^- \to q\bar{q}$), generate additional partons step-by-step while admitting only visible (i. e. non-soft) emission:

$$z > \varepsilon(t) \qquad\qquad (1-z) > \varepsilon(t)$$

where $\varepsilon(t)$ can be expressed in the following way:

$$\begin{aligned}
&p_a^2 = t \text{ and } p_b^2, p_c^2 > t_c \\
&p_T^2 = z(1-z)p_a^2 - (1-z)p_b^2 - zp_c^2 > 0 \\
&\Rightarrow z(1-z) > \frac{t_c}{t} \\
&\Rightarrow \varepsilon(t) = \frac{1}{2} - \frac{1}{2}\sqrt{1 - 4\frac{t_c}{t}} \simeq \frac{t_c}{t}
\end{aligned}$$

which means that the threshold $\varepsilon(t)$ gets more strict for decreasing t.

Let us define the Sudakov form factor $\Delta(t)$

$$\Delta(t) = \exp\left\{ - \int_{t_c}^{t} \frac{dt'}{t'} \int_{\varepsilon(t')}^{1-\varepsilon(t')} dz \alpha_s C_F F_{qq}(z) \right\}$$

which is the probability for a parton to evolve from t to t_c without emission of another parton. Observe that

$$\Delta(t_c) = 1$$

and the probability for a parton to evolve from $t_1 \to t_2$ without emission of another parton is given by

$$R(t_1, t_2) = \frac{\Delta(t_1)}{\Delta(t_2)}.$$

The Monte Carlo procedure is now as follows.

0. Starting point (t_1, z_1)
1. Generate a random number $R \in]0; 1[$.
2. Solve $\Delta(t_1)/\Delta(t_2) = R$ for t_2.
   - For $\Delta(t_1) > R$:
     $\Delta(t_2) > 1$: $t_2 < t_c$: no emission, parton saved for final state

- For $\Delta(t_1) < R$:
  Generate further random number $R' \in]0;1[$ and solve

$$\int\limits_{\varepsilon(t_2)}^{z_2/z_1} dz \frac{\alpha_s}{2\pi} P(z) = R' \int\limits_{\varepsilon(t_2)}^{1-\varepsilon(t_2)} dz \frac{\alpha_s}{2\pi} F(z)$$

  for z_2.

3. Use the two new partons

$$\left((t_2, z_2); \left(t_2, \frac{z_1 - z_2}{z_1} \right) \right)$$

   as starting point for another Monte Carlo step (see Fig. 9.22).

4. Repeat steps 1 to 3 until all partons fulfill $t_i < t_c$.

This procedure generates events with the same probabilities as in experiment and produces a list of final state particles which allows to perform the same analyses as on experimental data. This is how one arrives at the "theory curves" shown e. g. in some of the plots in Chap. 8.

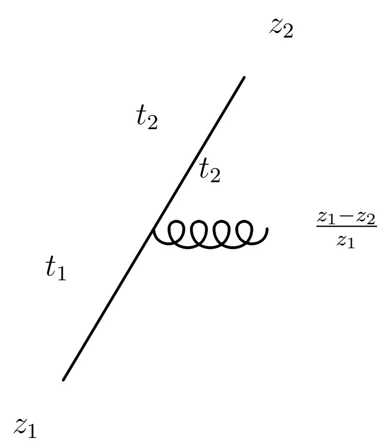


Figure 9.22: *Starting point for second Monte Carlo step.*

Chapter 10

Hadron collider physics

Literature:

- Ellis/Stirling/Webber [28]
- Dissertori/Knowles/Schmelling [27]
- Kane/Pierce [54]
- Review on QCD of the Particle Data Group [26]
- Technical Design Reports (TDRs) about the Physics Performance of ATLAS and CMS [55, 56]

With the start of the LHC at CERN on March 30, 2010, operating at the moment at a total center of mass energy of 7 TeV, a new record in particle collision energy has been achieved. Like the Tevatron at Fermilab (operating at a total center of mass energy of about 2 TeV) it is a hadron collider. The purpose of this chapter is to present the most important features of this kind of colliders and the physics studied there.

First, the purposes, advantages and weaknesses of using a hadron collider are discussed in the introduction. Then the different components of the cross-section of proton-(anti)proton interactions are presented. Next comes a digression to the topic of parton distribution functions (PDF), in particular how these are determined from the wealth of data from different experiments. An excellent knowledge of the proton structure, i. e. of the PDFs, is a necessary ingredient for obtaining precise predictions of production rates and other observables at hadron colliders. Finally specific processes are presented, such as jet production, electroweak, top and Higgs physics.

10.1 Introduction

Many measurements performed at earlier colliders have tested the standard model of particle physics to a very high accuracy. As it can be seen in Fig. 10.1, the relevant measured parameters of the model agree with their fitted values within 1 to 3 standard deviations, as obtained from a global fit of the standard model predictions to the data. Up to now there is basically no phenomenon in contradiction with the predictions of the minimal version of the standard model (with the exception of neutrino masses and oscillations). However, there are some key questions which remain unanswered so far.

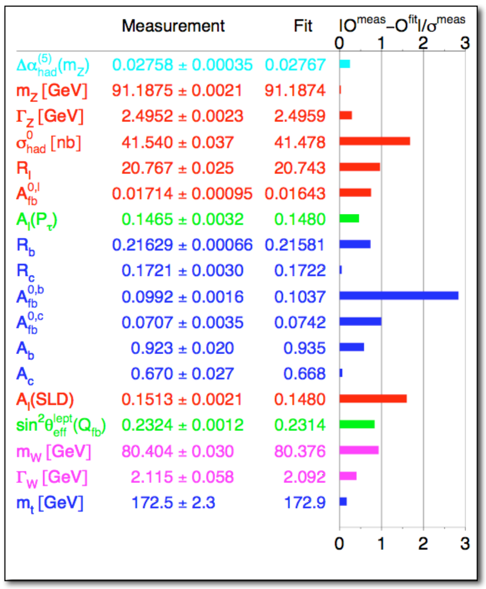


Figure 10.1: Comparison of the measured parameters of the standard model with the result of a global fit.

10.1.1 Open questions in particle physics

Mass? The question of the origin of **mass** of the fundamental constituents of matter still lies at the center of the investigations. More precisely, in its simplest form, without any spontaneous symmetry breaking, the electroweak theory [1] predicts the existence of 4 massless vector particles (gauge bosons). However, the observations show that the W^+, W^- and Z have a non-zero mass, whereas the photon γ is massless. A possible explanation is given by adding a scalar field to the model, the Higgs field (or boson). This field has a non-zero vacuum expectation value, which breaks the original symmetry (of the ground-state) and gives those particles a mass which interact with it . Since this symmetry is not broken at the Lagrangian level, one speaks of a spontaneous symmetry breaking mechanism. Predicted since the sixties, this particle has not yet been observed. Fig. 10.2 shows the most

[1] To be discussed in the next chapter.

likely Higgs mass range as obtained from global fits of the standard model to the data, with the Higgs mass as free parameter (see `http://lepewwg.web.cern.ch/LEPEWWG/`). Also shown is the mass region excluded by the LEP data ($< 114\,\mathrm{GeV}$).

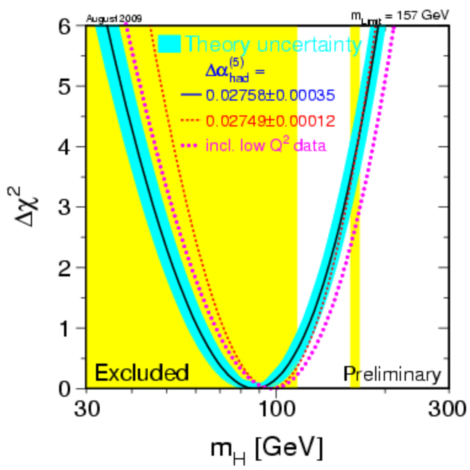


Figure 10.2: Most likely region for the Higgs mass (indicated by the minimum in the χ^2 value) as obtained from a fit of standard model predictions to LEP, SLD and Tevatron data

Unification? In the spirit of the electroweak theory of Glashow, Salam and Weinberg, which describes together the electromagnetic and weak interactions as being the low energy limit of a unified gauge theory, physicists soon thought of further unifications of the four fundamental interactions. Since the coupling of the strong interaction is decreasing with the energy, whereas the electroweak couplings are increasing, it is tempting to postulate that all three interactions would arise from a single coupling strength related to a gauge theory with extended gauge group, which "splits into three" as the energy gets below a certain (large) value. This is the basic idea behind **grand unification theories** (GUTs), which view the standard model gauge group,

$$SU(3)_{\text{color}} \times SU(2)_{\text{weak isospin}} \times U(1)_{\text{hypercharge}},$$

with three different couplings as a subgroup of a bigger "unified" gauge group G. However, a nice convergence of the electroweak and strong couplings at a single unification scale is not necessarily achieved. In case of **supersymmetry** this is achieved. Here a new fundamental symmetry is introduced, which associates to each fermion a boson and vice-versa. The supersymmetric partner of the electron e is called selectron, denoted $\tilde{e}$, which is a spin-0 particle, whereas the superpartner of the photon is called photino $\tilde{\gamma}$ and is a spin-$\frac{1}{2}$ particle. Supersymmetry is the only way to combine the internal symmetry group of a field with the Poincaré group in a non-trivial fashion. As of now, there is no quantum field theory of gravitation. Supersymmetry might provide a natural context for the inclusion of gravity (supergravity), opening the possibility for a unified theory of all interactions. The extreme weakness of gravity at the level of particle interactions has also lead physicists to conjecture that it could propagate in **extra dimensions**, whereas other interactions and matter cannot. Thus, in the usual 3+1-dimensional world we would only feel a small fraction of the total gravitational flux, which then explains the weak nature of gravity.

At this point, it should be noted that the lightest supersymmetric particle is neutral, stable and weakly interacting, thus a good candidate for dark matter. From astrophysical observations we know that dark matter represents $\sim 23\%$ of the mass of our universe. Dark matter does not interact electromagnetically, hence the name. This is based on the principle that the lightest supersymmetric particle cannot decay because of the conservation of a new quantum number, R-parity. This is an analogous explanation as the one for the stability of the electron (due to the conservation of the lepton number) or the proton (baryon number).

Flavour? From the decay width of the Z, at LEP it could be shown that there are exactly 3 types of light neutrinos, leading to the conclusion that there are three families of leptons and, by extension to the quark sector, of matter. The natural question becomes then: why 3 and not say 4? The existence of 3 families of quarks leads to **CP-violation** through the number of free parameters within the CKM matrix, related to the weak decays of quarks. A major issue is the precise measurement of its coefficients. Fig. 10.3 shows the experimental constraints on the possible values of the parameters describing the elements of the CKM matrix.

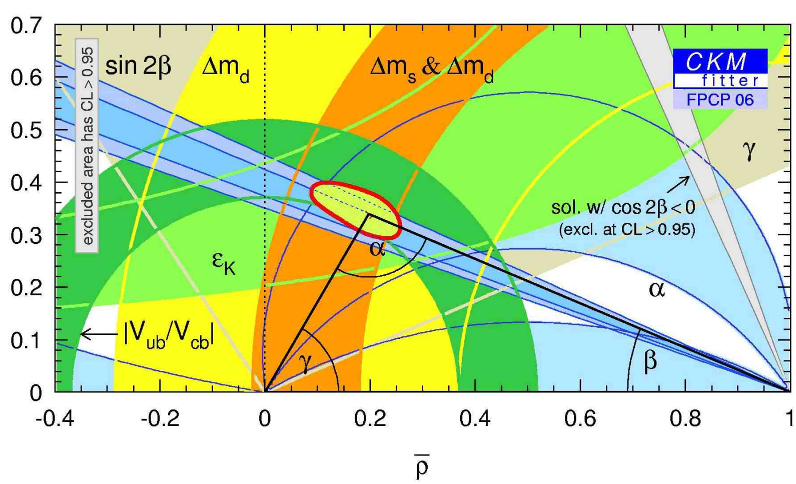


Figure 10.3: Experimental constraints on the parameters of the CKM matrix [26]

In view of these basic questions, the main goals of the experiments at the LHC are :

- Mechanism behind the electroweak symmetry breaking: search for the Higgs boson;

- Unification : test of the standard model, search for supersymmetric partners or for other physics beyond the standard model;
- Flavour : study of CP-violation in the b quark sector, by measuring properties (decays, oscillations) of B-hadrons.

10.1.2 Hadron colliders vs. e^+e^--colliders

In essence, physics at hadron colliders is much more complex than at e^+e^--colliders such as LEP or SLC, since now we are dealing with composite objects as our beam particles, whereas leptons are (as far as we know) point-like. Why then bother using hadrons?

e^+e^--colliders are precision machines : they lead to clean events, where basically all the energy of the initial state is used and the centre-of-mass system and the laboratory frame typically coincide (if both beam energies are the same). Thus the kinematics of the reaction is fixed and can be well reconstructed. Furthermore, theoretical calculations are simplified by the point-like and non-coloured initial state. On the other hand, in order to scan the energy range, the energy of the particle beam has to be changed "manually". Furthermore, the maximum energy achievable is limited (in the case of circular accelerators) by synchrotron radiation.

Hadron colliders are better suited for discoveries : the synchrotron radiation (going with the inverse fourth power of the accelerated mass) is much less relevant, and the energy range of the hard interaction is automatically scanned, since quarks and gluons can have any fraction of the proton 4-momentum. However, the complexity of the event resulting from the non-trivial proton structure and hadronization needs to be overcome and represents a challenge to and a limitation for the theoretical calculations.

10.1.3 Kinematic variables

We recapitulate here the most important kinematic quantities for hadron colliders.

Transverse (longitudinal) momentum p_T (p_L) is defined as the component of the 3-momentum perpendicular (parallel) to the beam. If θ is the angle relative to the beam and p is the modulus of the momentum, then,

$$\begin{aligned} p_T &= p\sin\theta, \\ p_L &= p\cos\theta. \end{aligned}$$

Rapidity y is defined through,

$$y = \frac{1}{2}\ln\left(\frac{E+p_L}{E-p_L}\right), \tag{10.1}$$

where $E = p^0$ is the energy of the scattered particle/jet.

Pseudorapidity η is defined through,

$$\eta = -\ln\tan\left(\frac{\theta}{2}\right), \tag{10.2}$$

with again θ being the angle between the beam and the particle/jet. For massless particles, rapidity and pseudorapidity coincide. It is customary to represent e.g. the energy deposits in the calorimeters as histograms in the η-ϕ plane, where ϕ is the angle around the detector (Fig. 10.4(a)). Tab. 10.1 gives the correspondence between angle and pseudorapidity. The barrel detectors (trackers, calorimeters) usually cover a region $|\eta| \leq 1.5$, whereas the endcap detectors go up to $|\eta| \sim 2.5 - 5$.

θ	η
$90°$	0
$10°$	2.4
$170°$	-2.4
$1°$	~ 5

Table 10.1: Correspondence between angle and pseudorapidity.

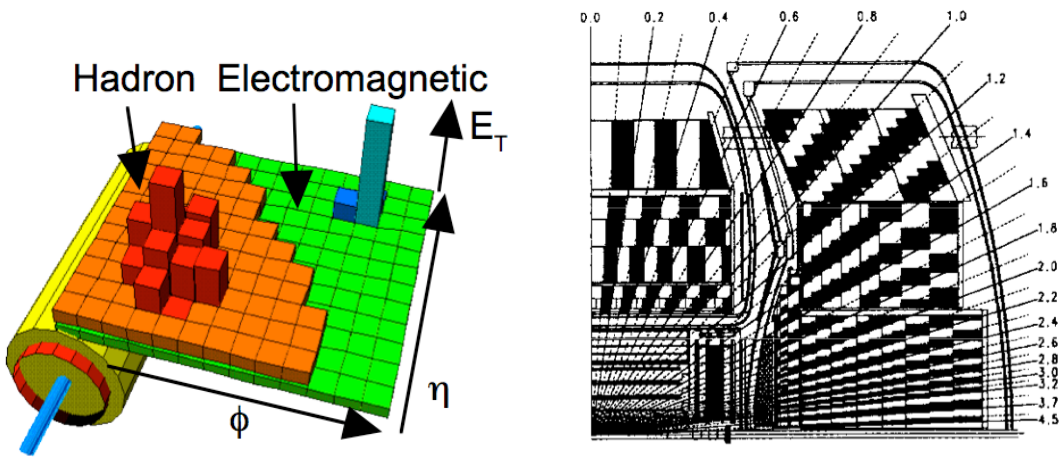
The interest of introducing rapidity and pseudorapidity lies in the fact that at hadron colliders the laboratory(detector)-frame in general does not coincide with the center of mass frame of the parton-parton collision, unless the two beams have the same energy and the parton momentum fractions fulfill $x_1 = x_2$. Typically $x_1 \neq x_2$, which leads to a longitudinal boost of the scattered system. We thus want to introduce quantities invariant under longitudinal boosts. It can be shown that the difference of two rapidities is invariant under such boosts. As a further consequence, detectors are typically built and structured in "rapidity towers" (Fig. 10.4(b)).

10.2 Components of the hadron-hadron cross section

The different components of the proton-proton cross section are shown in Fig. 10.5.

In elastic as well as in double diffractive scattering, both protons remain intact. In single diffractive scattering, one of the protons remains intact, whereas the other breaks up into several fragments. In these diffractive events, an uncolored object (a so-called "Pomeron"), which has the quantum numbers of the vacuum, is exchanged between the protons. Diffractive scattering is not well understand theoretically and it is the main field of research of the LHCf and TOTEM experiments, which focus on scattering events at very small angles.

The kind of events we will mostly focus on are called non-diffractive and correspond to the case of the complete break-up of both protons, with a cross section of $\sim 70\,\text{mb}$. The biggest fraction of this cross section is associated to soft scattering, i.e. scattering



(a) (b)

Figure 10.4: (a) η-ϕ-plane representation of a calorimeter signal. (b) Calorimeter towers of a detector, structured according to rapidity intervals.

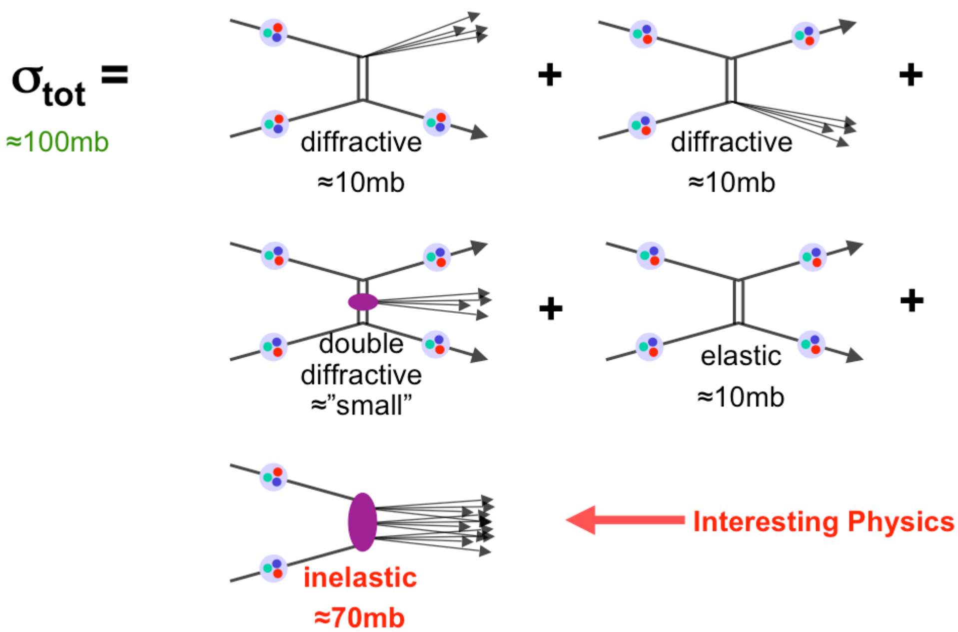


Figure 10.5: Pictorial representation of the components of the total proton-proton cross section. Here "interesting physics" refers to those processes relevant for the study of hard interactions.

where the exchanged momentum is small. The really interesting events are so-called "hard scattering events", in particular for the study of heavy objects such as energetic jets, W and Z bosons, top quarks, or the search for new heavy particles. These events are orders of magnitudes less probable than soft scattering events.

10.2.1 Soft scattering

Most of the proton-proton collisions are due to interactions with a small momentum transfer. This results in a shower of particles having a large longitudinal momentum and a small transverse momentum,

$$\langle p_T \rangle \approx 700\,[\mathrm{MeV}] \qquad\qquad \text{for } \sqrt{s} = 14\,[\mathrm{TeV}].$$

These processes cannot be reliably computed in perturbative QCD, since the coupling constant is rather big for soft processes. Thus the structure of such events is poorly known and one must rely on phenomenological models, implemented in the simulations, as well as on measurements.

Example When a proton is broken up, it produces neutral and charged pions (because of hadronization). Assuming a simple constant matrix element, from the structure of the phase-space element we realize that the produced particles should be uniformly distributed in transverse momentum squared and rapidity :

$$\frac{d^3p}{2E} = \frac{\pi}{2} dp_T^2 dy.$$

Thus, the produced particles should be distributed according to an almost flat distribution in pseudorapidity (due to the finite pion mass), as the one seen in Fig. 10.7(a) and 10.7(b). A typical soft event at 2.36 TeV measured by CMS is shown in Fig. 10.7(c). At 14 TeV one expects 4-6 charged and 2-3 neutral pions per unit of pseudorapidity, uniformly distributed in ϕ.

10.2.2 Pile-up events

Due to the very large cross section for soft scattering, the probability of having multiple proton-proton collisions during the same bunch crossing can become big, if the luminosity is large. Put in another way, interesting events – such as the production of a Higgs boson – will most probably be accompanied by other less interesting events, "polluting" the signal. The amount of additional soft proton-proton scatterings depends on the luminosity of the collider as seen in Fig. 10.8.

For example, at full LHC luminosity (10^{34} cm^{-2}sec^{-1}) there can be up to ∼25 soft collisions per bunch crossing, each generating ∼9 pions. Taking the total rapidity range typically covered by an LHC experiment to be $y_{max} = \pm 5$, we can estimate that there will be

$$25 \cdot 9 \cdot 2|y_{max}| \approx 2250$$

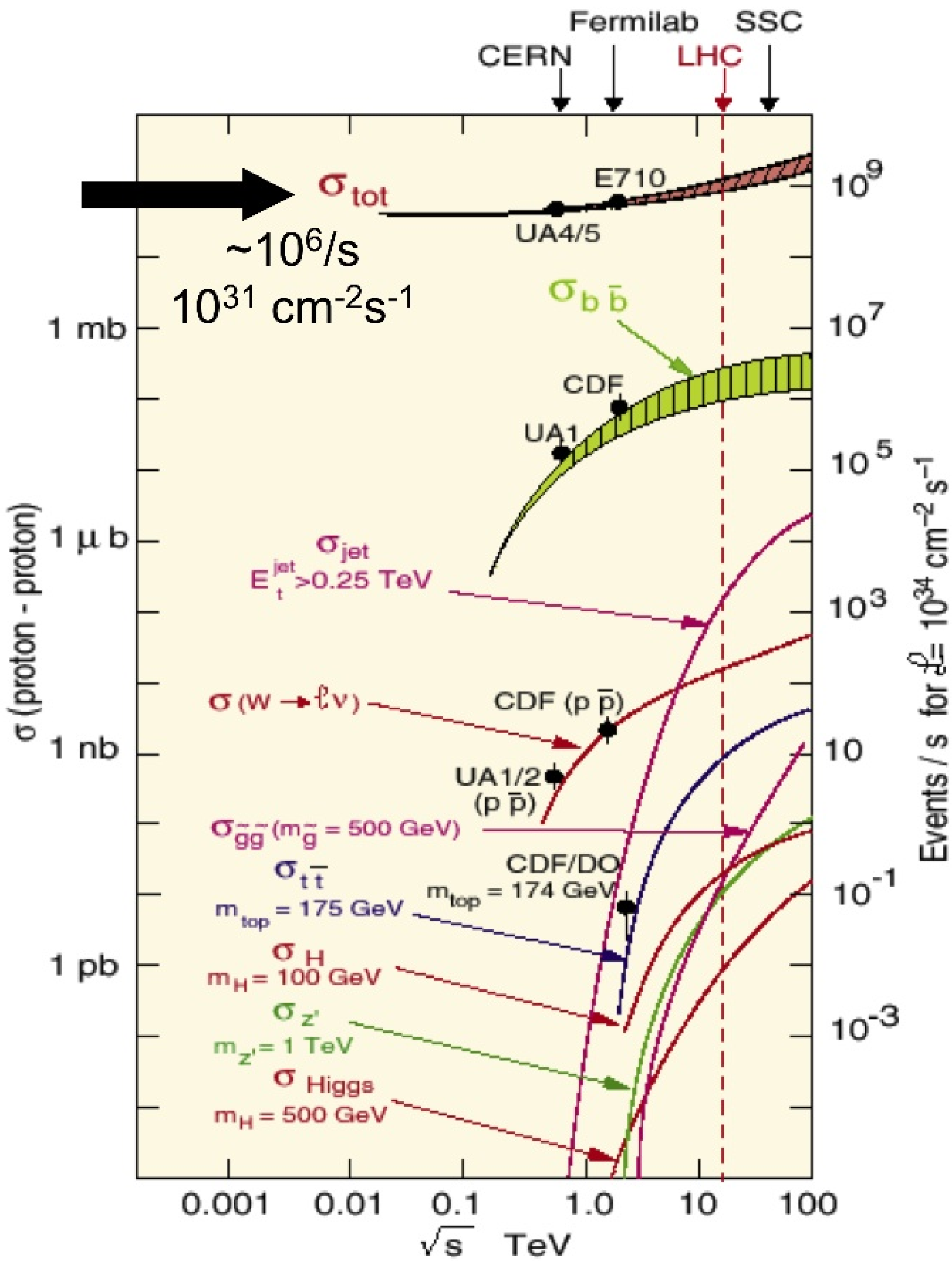
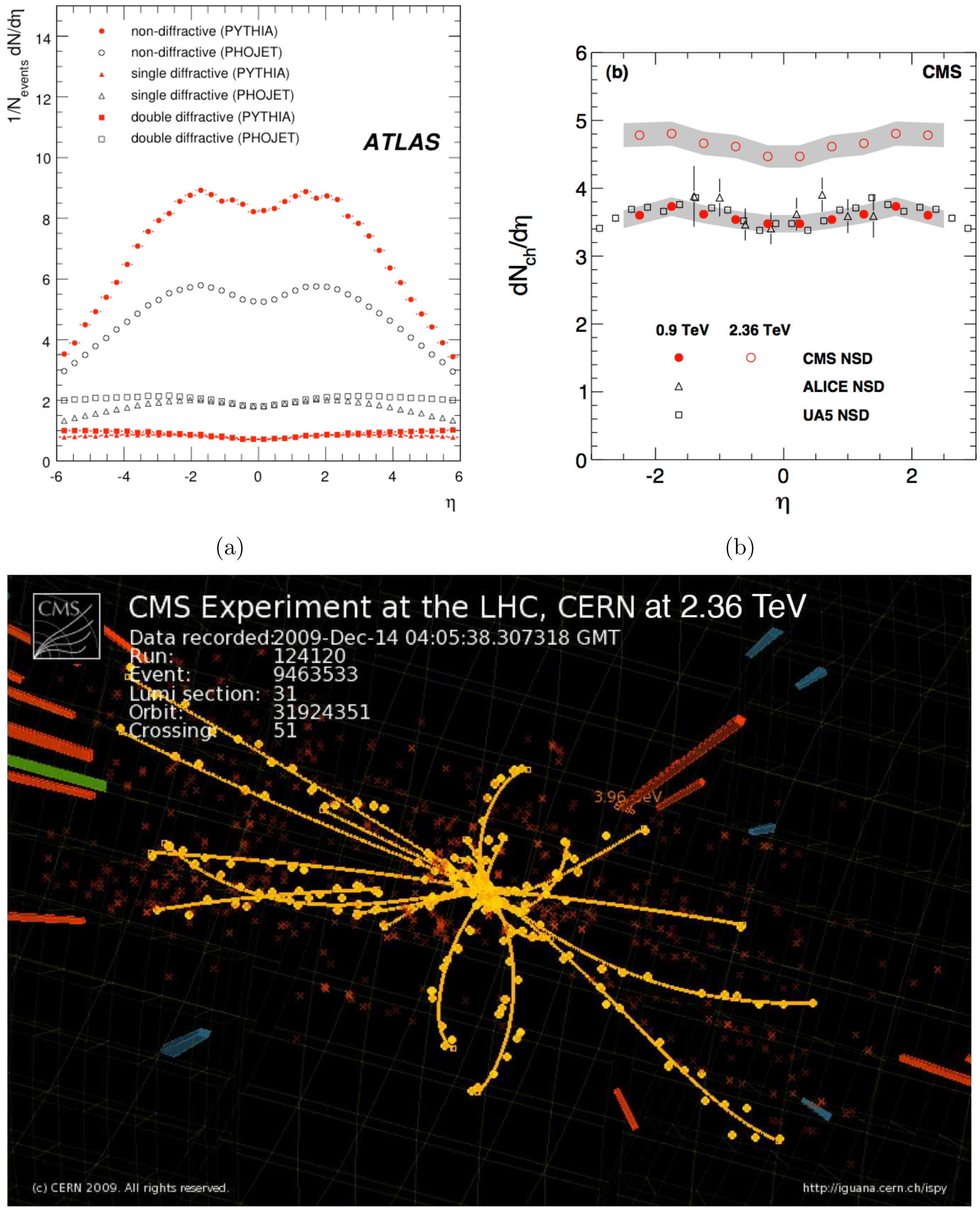


Figure 10.6: Cross sections for different processes in proton-proton scattering

pions produced that will deposit a total energy of,

$$2250 \cdot \underbrace{700\,[\text{MeV}]}_{\langle p_T \rangle} \approx 1.6\,[\text{TeV}]$$

in the calorimeters for each bunch crossing, resulting in an important background noise,



(a) (b)

(c)

Figure 10.7: (a) Pseudorapidity distribution simulated with PYTHIA and PHOJET. (b) Pseudorapidity distribution measured by CMS and ALICE and compared with UA5. (c) Soft event at 2.36 TeV recorded by CMS.

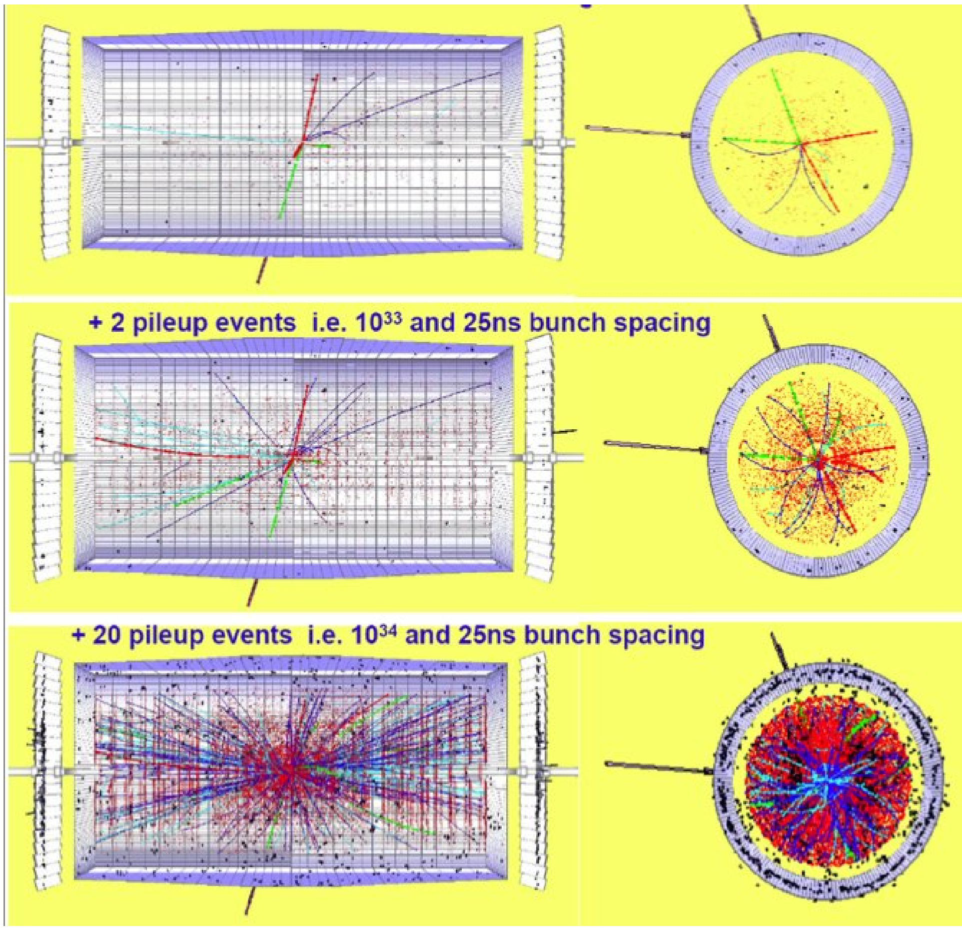


Figure 10.8: Pile-up events at different luminosities.

which has to be isolated from the interesting signal (hard scattering event).

10.3 Hard scattering

The main process of relevance for the study of energetic jets, heavy standard model particles or the discovery of new particles, is hard scattering, depicted in Fig. 10.9. Here we have a large momentum transfer (Q) involved in the scattering process. The function f_{a/h_1} denotes the PDF for the parton a in the hadron h_1, and analogously for f_{b/h_2}. Denoting by $x_{1(2)}$ the momentum fraction of $h_{1(2)}$ carried by $a(b)$, the available center of mass energy for the underlying scattering process is then (assuming massless partons)

$$\sqrt{\hat{s}} = \sqrt{x_1 x_2 s}, \tag{10.3}$$

with $s = (p_{h_1} + p_{h_2})^2$ the center of mass energy of the colliding hadrons.

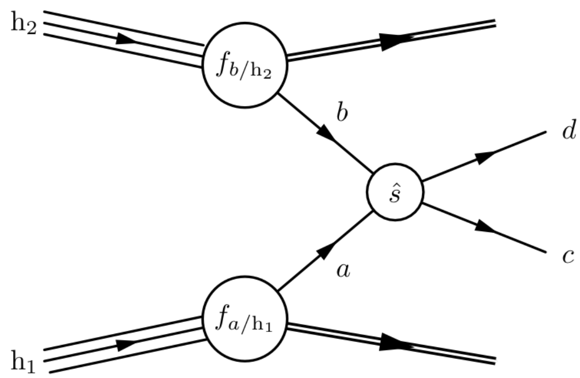


Figure 10.9: Basic Feynman graph for the description of a hard scattering process in a hadron-hadron collision.

At high energies ($\gg \Lambda_{\text{QCD}}$), we can view the resulting interaction as the incoherent sum of the interactions for any combination of the constituents [2], yielding the master formula,

$$\boxed{d\sigma^{h_1 h_2 \to cd} = \int_0^1 dx_1 \int_0^1 dx_2 \sum_{a,b} f_{a/h_1}(x_1, \mu_F^2) f_{b/h_2}(x_2, \mu_F^2) d\hat{\sigma}^{ab \to cd}(Q^2, \mu_F^2)} \tag{10.4}$$

Here μ_F^2 is the factorization scale and Q is the typical scale of the process, e.g. the momentum transfer in a t-channel or $Q = \sqrt{\hat{s}}$ in an s-channel process. Examples of parton-parton processes with a cross section $\hat{\sigma}$ can be found in Sect. 9.10, p. 230. The calculation of such cross sections can be achieved by using a given interaction theory, typically QED, QCD, electroweak theory, supersymmetry, etc.

We proceed by demonstrating that heavy particle states are produced more centrally in the detector, i.e. at low rapidity, compared to soft-particle production. For this, we consider the production of a hypothetical heavy gauge boson, Z', with mass $M \sim 1\,\text{TeV} \gg m_p$, energy E and rapidity y at a proton-proton collider. The heavy gauge boson can appear in the propagator of an s-channel quark-antiquark annihilation diagram. From the mass shell condition (which gives the largest cross section) in this propagator we have,

$$\hat{s} = x_1 x_2 s \stackrel{!}{=} M^2.$$

Since each proton has an energy $E_{\text{beam}} = \sqrt{s}/2 \gg m_p$, it is straightforward to see that (we assume w.l.o.g. that $x_1 \geq x_2$),

$$E = \frac{\sqrt{s}}{2}(x_1 + x_2)$$
$$p_L = \frac{\sqrt{s}}{2}(x_1 - x_2).$$

[2]This is nothing else than Eq. (9.49) in Sect. 9.10

Inserting those values in the definition of the rapidity, Eq. (10.1), we get,

$$\mathrm{e}^y = \sqrt{\frac{x_1}{x_2}},$$

and hence $y \to 0$ if $x_1 \to x_2$. In this case, the energy is used optimally, since the longitudinal component of the momentum of the Z' vanishes, and it becomes "easier" to produce it (Fig. 10.10). With one line of algebra, one can see that,

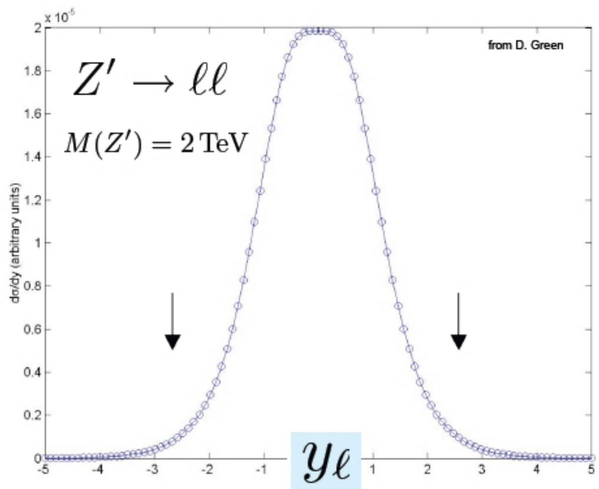


Figure 10.10: Rapidity distribution for Z' production

$$x_{1,2} = \frac{M}{\sqrt{s}}\mathrm{e}^{\pm y}, \tag{10.5}$$

i.e. to produce a Z' with rapidity y, one of the partons must have had a momentum fraction $M\mathrm{e}^{+y}/\sqrt{s}$, while the other had the momentum fraction $M\mathrm{e}^{-y}/\sqrt{s}$. This formula is often used to determine the momentum fraction carried by the partons, if the rapidity of the heavy object and its mass can be reconstructed experimentally. This is relatively easy in cases where the heavy particle decays into a lepton-antilepton pair, since the lepton momenta can be measured rather precisely.

Using Eq. (10.5), we can find the momentum fraction, which a parton should carry in order to produce a heavy particle centrally, $y \approx 0 \Leftrightarrow x = x_1 \approx x_2$ for a given collider. To produce a particle with a mass of order 100 GeV (e.g., Z and W bosons, a light Higgs), one needs $x \approx 0.05$ in the case of Tevatron ($\sqrt{s}_{\mathrm{max}} = 2\,\mathrm{TeV}$), whereas at LHC ($\sqrt{s}_{max} = 14\,\mathrm{TeV}$) one only needs $x \approx 0.007$, a momentum fraction at which the gluon PDF is already very big (compared to the PDF of valence quarks). To produce a particle with mass 5 TeV, one needs $x \approx 0.36$ at LHC, where the dominant PDF contribution comes from valence quarks.

10.4 Global PDF fits

The master formula (10.4) contains not only the underlying parton-parton cross section, calculable in the context of some theory, but also the PDFs of both hadrons. These quantities cannot be calculated from first principles. Therefore, we stop for a moment in our study of hadron colliders to review the techniques associated with the determination of the PDF of the proton.

Many measurements have been performed which probe the proton structure, such as

- F_2 measurements, in particular at HERA,
- F_3 measurements at HERA in CC[3] interactions,
- F_3 measurements in neutrino-nucleon scattering,
- measurements of the Drell-Yan process (W and Z production in hadron-hadron collisions with subsequent decays to leptons),
- Sum rules,
- Jets and direct photon production,
- Constraints on the gluon content from scaling violations, jets and heavy quark production.

Tab. 10.2 shows a typical data set which is used for a global determination (fit) of PDFs.

The general procedure goes as follows:

1. *Choose a set of experimental data* with possible restrictions in x and Q^2 in order to avoid critical phase-space regions and thus systematic uncertainties.

2. *Parametrize the PDFs at a given fixed scale*, e.g. $Q_0^2 = 4\,\mathrm{GeV}^2$, with an ansatz of the type,

$$x f_i(x, Q_0^2) = A_i \underbrace{x^{\alpha_i}}_{\text{low-}x} \underbrace{(1-x)^{\beta_i}}_{\text{large-}x},$$

   for $i = u, d, g, \bar{u}$ etc. with different coefficients.

3. *Evolve the PDFs in Q^2* using the DGLAP evolution equations (9.34)-(9.37) to bring the PDFs from the scale Q_0^2 to the scale Q^2 of the specific data set. Then fold

[3]Charged current; exchange of $W^{\pm}$ bosons.

Process/ Experiment	**Leading order subprocess**	**Parton behaviour probed**
DIS ($\mu N \to \mu X$) $F_2^{\mu p}, F_2^{\mu d}, F_2^{\mu n}/F_2^{\mu p}$ (SLAC, BCDMS, NMC, E665)*	$\gamma^* q \to q$	Four structure functions $\to$ $u + \bar{u}$ $d + \bar{d}$ $\bar{u} + \bar{d}$ s (assumed $= \bar{s}$), but only $\int xg(x, Q_0^2)dx \simeq 0.35$ and $\int(\bar{d} - \bar{u})dx \simeq 0.1$
DIS ($\nu N \to \mu X$) $F_2^{\nu N}, xF_3^{\nu N}$ (CCFR)*	$W^* q \to q'$	
DIS (small x) F_2^{ep} (H1, ZEUS)*	$\gamma^*(Z^*)q \to q$	λ $(x\bar{q} \sim x^{-\lambda_S},\ xg \sim x^{-\lambda_g})$
DIS (F_L) NMC, HERA	$\gamma^* g \to q\bar{q}$	g
$\ell N \to c\bar{c}X$ F_2^c (EMC; H1, ZEUS)*	$\gamma^* c \to c$	c $(x \gtrsim 0.01;\ x \lesssim 0.01)$
$\nu N \to \mu^+\mu^- X$ (CCFR)*	$W^* s \to c$ $\hookrightarrow \mu^+$	$s \approx \frac{1}{4}(\bar{u} + \bar{d})$
$pN \to \gamma X$ (WA70*, UA6, E706, ...)	$qg \to \gamma q$	g at $x \simeq 2p_T^\gamma/\sqrt{s} \to$ $x \approx 0.2 - 0.6$
$pN \to \mu^+\mu^- X$ (E605, E772)*	$q\bar{q} \to \gamma^*$	$\bar{q} = ...(1-x)^{\eta_S}$
$pp, pn \to \mu^+\mu^- X$ (E866, NA51)*	$u\bar{u}, d\bar{d} \to \gamma^*$ $u\bar{d}, d\bar{u} \to \gamma^*$	$\bar{u} - \bar{d}$ $(0.04 \lesssim x \lesssim 0.3)$
$ep, en \to e\pi X$ (HERMES)	$\gamma^* q \to q$ with $q = u, d, \bar{u}, \bar{d}$	$\bar{u} - \bar{d}$ $(0.04 \lesssim x \lesssim 0.2)$
$p\bar{p} \to WX(ZX)$ (UA1, UA2; CDF, D0)	$ud \to W$	u, d at $x \simeq M_W/\sqrt{s} \to$ $x \approx 0.13;\ 0.05$
$\to \ell^\pm$ **asym** (CDF)*		slope of u/d at $x \approx 0.05 - 0.1$
$p\bar{p} \to t\bar{t}X$ (CDF, D0)	$q\bar{q}, gg \to t\bar{t}$	q, g at $x \gtrsim 2m_t/\sqrt{s} \simeq 0.2$
$p\bar{p} \to$ **jet** $+ X$ (CDF, D0)	$gg, qg, qq \to 2j$	q, g at $x \simeq 2E_T/\sqrt{s} \to$ $x \approx 0.05 - 0.5$

Table 10.2: Example of data sets employed for fitting PDFs, from Stirling et al.

the PDFs with the coefficient functions/parton cross sections from NLO or NNLO perturbative QCD in order to get a structure function/cross section [4],

$$F_2(x, Q^2) = \sum_i C_i(z, Q^2/\mu_F^2) \otimes f_i(x/z, \mu_F^2)$$

$$d\sigma(Q^2) = \sum_{i,j} f_i(x, \mu_F^2) \otimes f_j(y, \mu_F^2) \otimes d\hat{\sigma}_{ij}(xyQ^2, \mu_F^2)$$

4. *Fit to the experimental data* to determine A_i, α_i, β_i for all i and use the obtained PDFs for the evolution to any other scale and the corresponding computation of cross sections.

Different groups use a different ansatz, which leads to differences in the extracted PDFs (Fig. 10.11). These agree up to a few percent, which ultimately translates into an uncertainty on the cross section given by the master formula (10.4). Therefore, a good knowledge of the structure of the proton, i.e. of the PDFs of its constituents is essential in order to be able to compute accurately cross sections at hadron colliders such as the Tevatron or the LHC.

From Fig. 10.11(a), it becomes clear that the LHC is effectively a gluon-gluon collider if one considers the production of particles around or below a scale of ~ 100 GeV. This is because of the relative importance of the gluon PDF (downscaled by a factor of 20 on the figure) in the relevant x-range (see the discussion above).

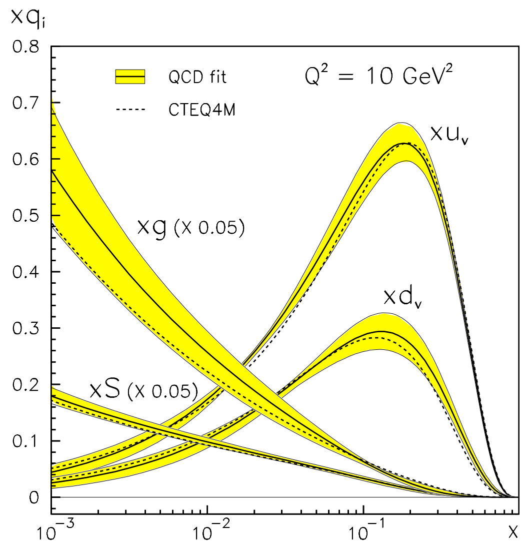
As can be seen in Fig. 10.12, the kinematic regime of the LHC is much broader than the one currently tested experimentally. Much of the relevant x range is covered by HERA, but for much smaller values of Q^2. The question arises if the DGLAP evolution is sufficient to evolve the PDFs to the full LHC kinematic range. Furthermore, one has to propagate the uncertainties on the PDFs in order to have a meaningful comparison of the predictions to data. The data acquired at the LHC will themselves serve to constrain the PDFs.

10.5 Jets

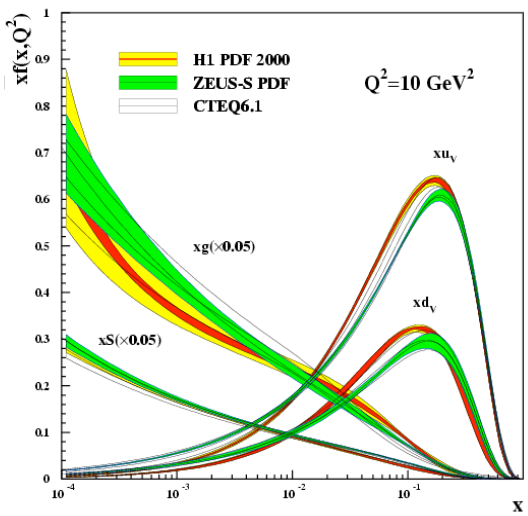
At the LHC, an important component of the inelastic cross-section after soft scattering is jet production, i.e. events where colored partons with significant transverse momentum are produced in the final state. Fig. 10.5 shows this component, labeled σ_{jet} in the case of a minimal jet energy of 250 GeV. One notes that this component is 6 orders of magnitude smaller than the total cross-section for pp-scattering.

Jet processes are important for multiple purposes. First, they are the main tool to test precisely perturbative QCD. Second, they allow to test if the quarks are composite objects.

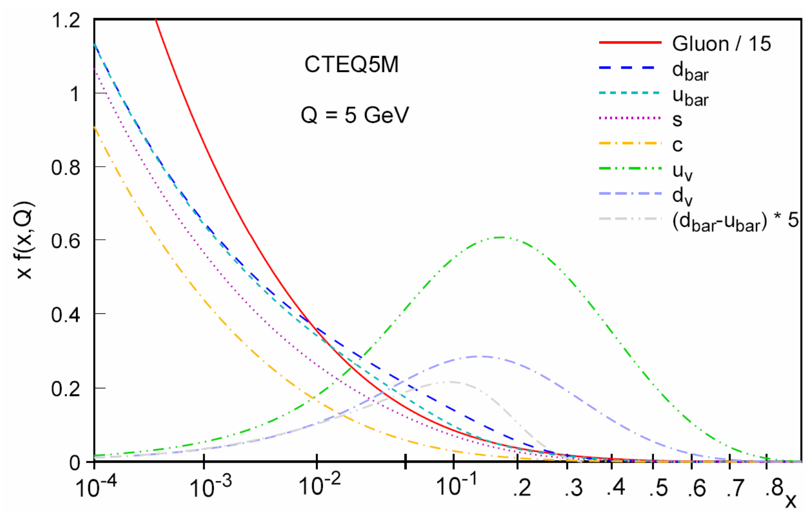
[4] It is implicitly understood that one integrates over the z variable or the x and y variables respectively, see Sect. 9.8.



(a)



(b)



(c)

Figure 10.11: (a) PDFs for $Q^2 = 10\,\mathrm{GeV}^2$ from Botje. (b) PDFs for $Q^2 = 10\,\mathrm{GeV}^2$ from HERA collaborations. (c) PDFs for $Q^2 = 5\,\mathrm{GeV}^2$ from CTEQ.

Finally, they represent a part of the background for other more rare processes and must thus be extensively understood in order to be able to filter out the signal.

Fig. 10.13 shows the differential production cross section at zero rapidity (center of the detector) as a function of the transverse momentum of the jet for the Tevatron and the LHC (note the logarithmic scale). We see that the Tevatron almost cannot produce jets with transverse energy bigger than 800 GeV, whereas the LHC can access for the same rate about 4.5 TeV.

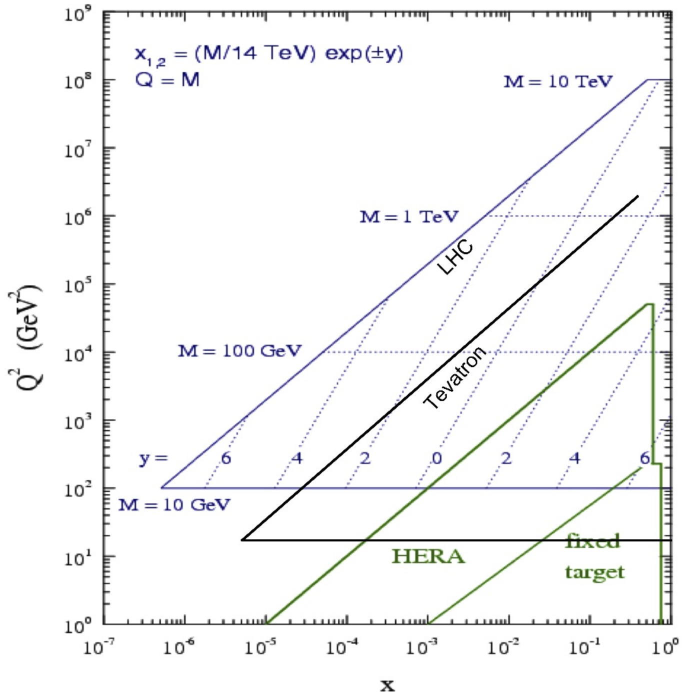


Figure 10.12: Q^2-x range of LHC, Tevatron and HERA.

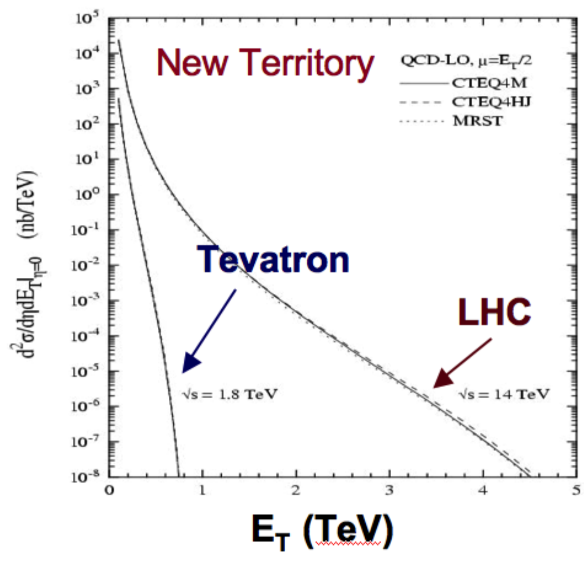


Figure 10.13: Differential cross section for jet production at zero rapidity as a function of the transverse momentum for the Tevatron and the LHC.

The relevant elementary processes (represented by $\hat{s}$ in Fig. 10.9) for jet production are shown in Fig. 10.14. These processes can all be achieved at both Tevatron and LHC since sea partons are dominant at low x. Since the color factor for a three-gluon vertex (3) is almost twice the one for a quark-gluon vertex ($\frac{4}{3}$), jets are more likely to be produced through gg-collisions. Also, the gluon PDF dominates at low x.

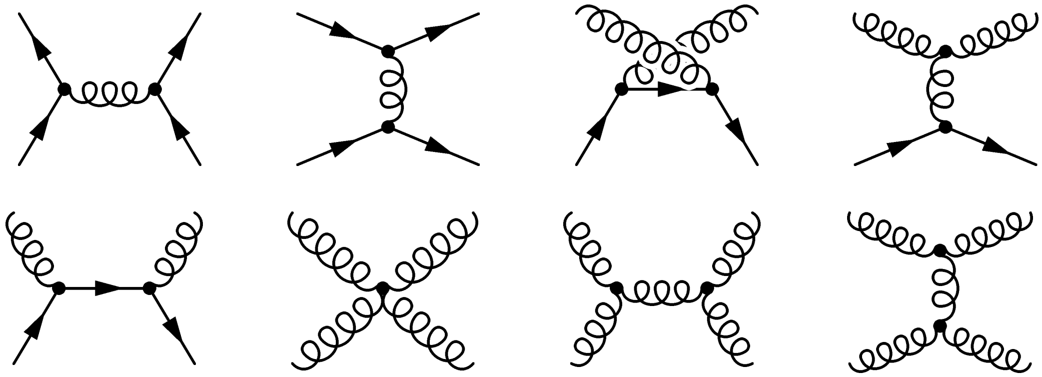
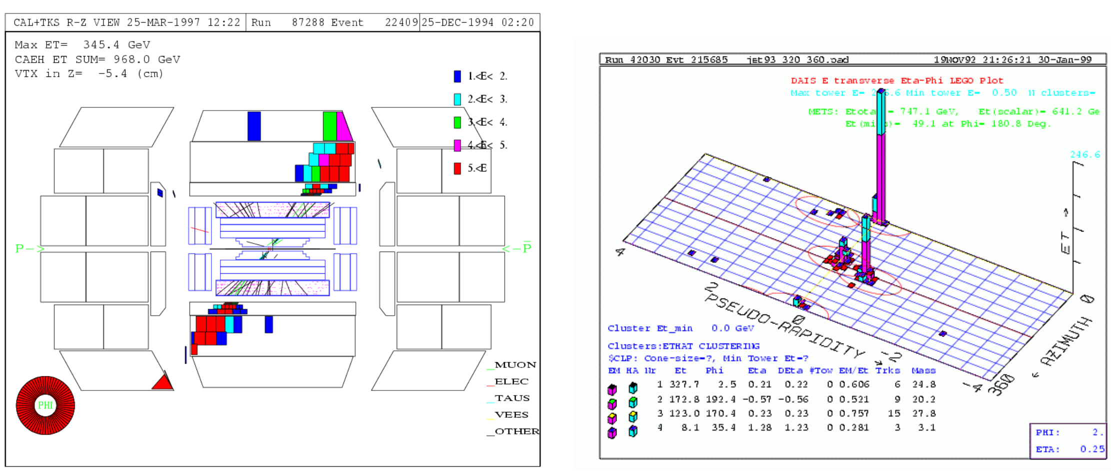


Figure 10.14: Elementary processes at hadron colliders.

10.5.1 Jet algorithms

Section 8.2, p. 161, contains a discussion of jet algorithms used at e$^+$*e*$^-$*-colliders.*

CONE algorithms Fig. 10.15 shows some typical jet events at the DØ and CDF experiments at Tevatron, a $p\bar{p}$-collider.



(a) (b)

Figure 10.15: *Jet events.* (a) at DØ, (b) at CDF.

From this type of events it seems sensible to define jets via a cone with opening,

$$R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2},$$

where η is the pseudorapidity and ϕ the angle around the beam axis. This is the basis of the CONE class of jet algorithms.

The CONE algorithm can be represented by the following algorithmic flow, starting from a list of seeds and a given R:

1. Is the list of seeds exhausted?
   - Yes : send list of protojets to recombination/splitting algorithm
   - No : continue to 2.
2. Compute centroid using R. Is the new axis the same as the old one?
   - Yes : continue to 3.
   - No : return to 2.
3. Was the cone already found?
   - Yes : remove it from the list of seeds.
   - No : add it to the list of protojets.
4. Return to 1.

The computation of the centroid is achieved by doing an energy weighting of the (η, ϕ)-coordinates of the energy deposits inside a cone of a given R. An energy deposit i is part of the cone C if,

$$i \in C : \sqrt{(\eta^i - \eta^C)^2 + (\phi^i - \phi^C)^2} \leq R,$$

where,

$$\eta^C := \frac{1}{E_T^C} \sum_{i \in C} E_T^i \eta^i, \qquad \phi^C := \frac{1}{E_T^C} \sum_{i \in C} E_T^i \phi^i, \qquad E_T^C := \sum_{i \in C} E_T^i$$

One of the major drawbacks of the CONE algorithm is that it is neither infrared nor collinear safe. A new algorithm called SISCone has been developed recently that solves this issue.

Recombination algorithms (k_T-type) We are now going to present a class of algorithms called k_T-recombination algorithms [57], having the following properties:

- Infrared and collinear safe,
- No overlapped jets,
- Every particle/detector tower is unambiguously assigned to a single jet,

- No biases from seed towers
- Sensitive to soft particles, area could depend on pile-up.

We start with a set of 4-momenta $\{p_i\}_{i=1,\dots,n}$ with coordinates (η_i, ϕ_i). One then defines the metric,

$$
\begin{aligned}
d_{ij} &= \min(p_{T,i}^2, p_{T,j}^2)\frac{\Delta R_{ij}}{D^2} \qquad\qquad i > j \\
d_{ii} &= p_{T,i}^2,
\end{aligned} \tag{10.6}
$$

with,

$$
\begin{aligned}
\Delta R_{ij}^2 &= (\eta_i - \eta_j)^2 + (\phi_i - \phi_j)^2 \\
D &\sim 0.4 - 1.
\end{aligned}
$$

Next, we determine the minimum d_{min} of the set $\{d_{ij} | i \geq j\}$. If $d_{min} = d_{kl}, k \neq l$, we combine the 4-momenta k and l: $p_k + p_l \to p_{kl}$. If instead $d_{min} = d_{kk}$, we identify it as a jet of its own and take p_k out of the list of 4-momenta. One then restarts with the new set until there are no 4-momenta left. This algorithm ends up with a list of jets having $\Delta R \geq D$.

A deviation from this algorithm is obtained by modifying the metric, for a given $p \in \mathbb{Z}$,

$$
\begin{aligned}
d_{ij} &= \min(p_{T,i}^{2p}, p_{T,j}^{2p})\frac{\Delta R_{ij}}{D^2} \qquad\qquad i > j \\
d_{ii} &= p_{T,i}^{2p}.
\end{aligned} \tag{10.7}
$$

One then speaks of,

- $p = 1$: regular k_T jet algorithm,
- $p = 0$: Cambridge/Aachen jet algorithm,
- $p = -1$: anti-k_T jet algorithm.

Interestingly enough, the anti-k_T jet algorithm yields jets with a cone structure. The soft particles are first clustered with hard particles instead of being combined with other soft particles. Fig. 10.16 shows the jet shapes for different recombination algorithms.

Further difficulties In the context of jet physics, pile-up events (typically containing a hard scattering and some soft proton-proton interactions) generate a homogeneous background that needs to be substracted before applying the jet algorithms since this energy is not carried by the original jets themselves.

Another challenge is the identification of the underlying event. As an example, consider the case in which 4 jets are observed. They could come from a true 4-jet event (2 quarks + 2 initial or final state gluon radiation) or from two independent 2-jet events (see Sect. 10.7 for a discussion of this phenomenon).

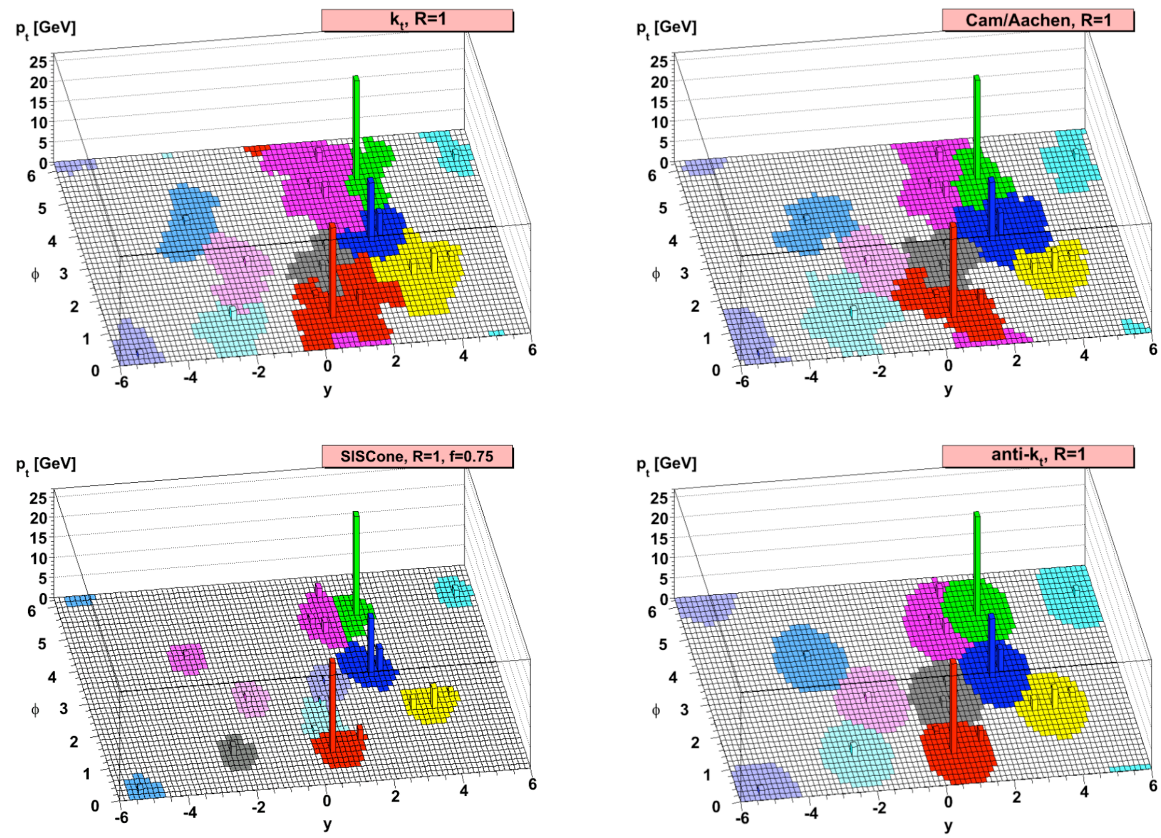


Figure 10.16: Jet areas for different values of the p-parameter in the modified k_T jet algorithm and for the SISCone jet algorithm for the same input data.

10.5.2 Measurements

The goal of a jet algorithm is to make it possible to measure cross-sections in an inclusive manner as a function of the transverse energy of the jet E_T. Fig. 10.17 shows a comparison of the Monte Carlo simulation JETRAD with the data from DØfor small rapidities.

If there is some new physics, e.g. if quarks have a substructure, the high-energy tail would be shifted from the simulation values.

One can compare the predicted value in perturbative QCD and the experimental data through the relation,

$$\left\langle \frac{d^2\sigma}{dE_T d\eta} \right\rangle = \frac{N}{\Delta E_T \Delta\eta \epsilon \mathcal{L}_{int}},$$

where N is the number of events in the bin $(E_j, \eta_j) \in [E_T, E_T+\Delta E_T] \times [\eta, \eta+\Delta\eta]$, ϵ denotes the efficiency in reconstructing jets (typically obtained using Monte Carlo simulations) and $\mathcal{L}_{int}$ is the integrated luminosity.

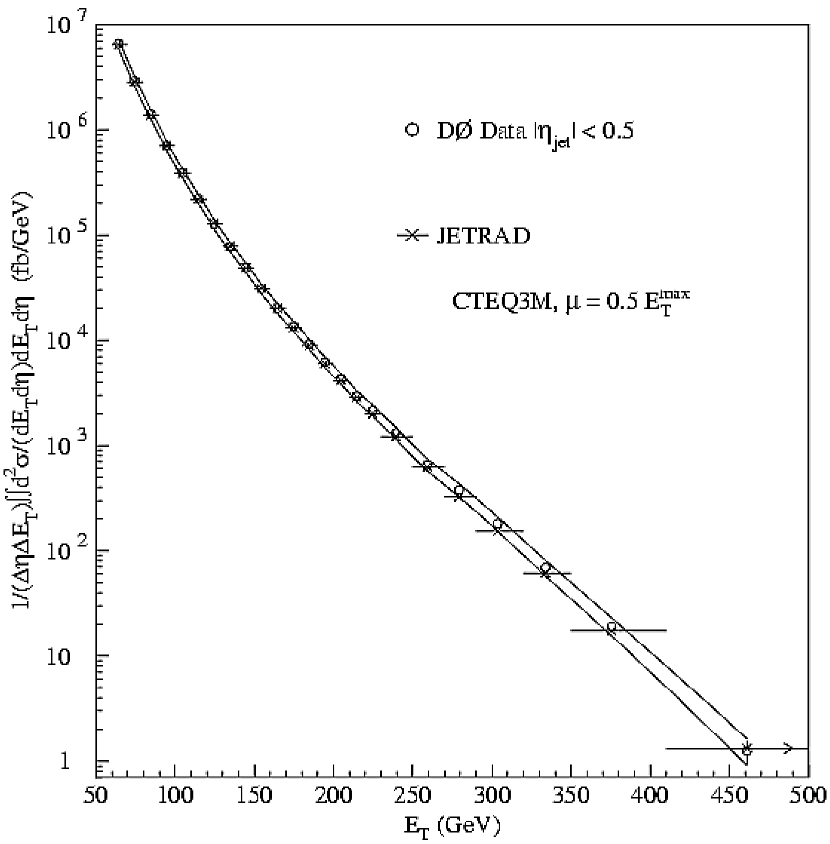


Figure 10.17: Differential jet cross-section as a function of E_T. Monte Carlo simulation (JETRAD) and DØdata.

A decisive requirement for a precise measurement, a test of QCD, and to be able to "see" new physics, is to have a very good energy calibration. Indeed, the double-differential cross section is very steeply falling:

$$\frac{d^2\sigma}{dE_T d\eta} \propto E_T^{-6},$$

and the propagation of the error becomes important,

$$\frac{\delta N}{N} \approx 6\frac{\delta E_T}{E_T}.$$

In fact, the slope is so steep that the energy resolution can distort the spectrum. The number of measured events with a given E_T can be expressed as the convolution,

$$N(E_T^{meas}) = \int_0^\infty N(E_T^{true}) \cdot Resol(E_T^{meas}, E_T^{true}) dE_T^{true}. \tag{10.8}$$

It is usual to assume a Gaussian resolution function,

$$Resol(E_T^{meas}, E_T^{true}) \propto \exp\left[-\frac{(E_T^{meas} - E_T^{true})^2}{2\sigma_{E_T}^2}\right],$$

where σ_{E_T} is the typical energy resolution of the detector. Although the resolution is symmetric around E_T^{true} the steepness causes it to have more influence on one side than on the other, leading to the distortion of the spectrum. This problem can be minimized by choosing the bin-width $\Delta E_T \approx \sigma_{E_T}$.

Beside these measurement problems, one must account also for the errors/uncertainties on the theory's side (non-perturbative effects) or of the proton PDFs (see Sect. 10.4) when comparing measurement with theory. Fig. 10.18 shows the typical relative uncertainty range on the energy of the jet from the experimental and theoretical point of view for $\sqrt{s} = 10\,\mathrm{TeV}$ at CMS. We see that the jet energy scale and the PDFs induce the largest uncertainties.

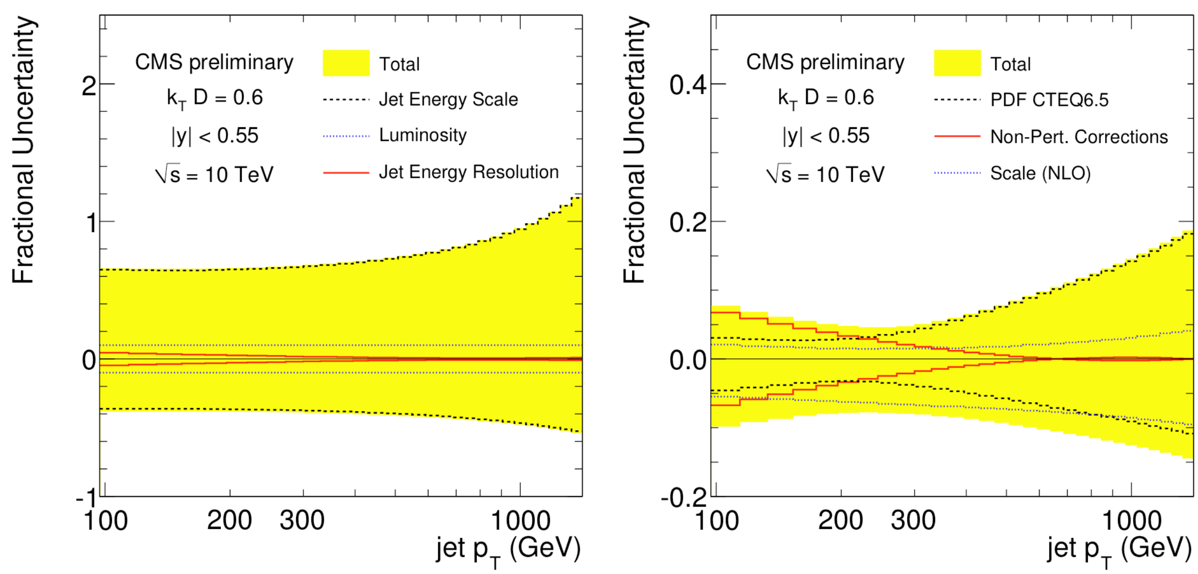
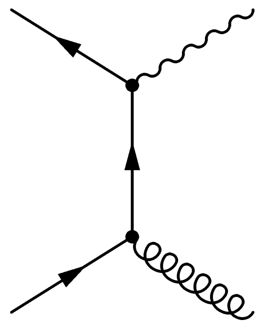


Figure 10.18: Experimental and theoretical part of the fractional uncertainty as a function of the jet transverse energy E_T.

10.5.3 Jet energy scale

In order to calibrate the (transverse) energy of a jet, the most useful process is $q\bar{q} \to g\gamma$,



.

Indeed, the energy of the photon can be measured to a high precision (1-2%) with the electromagnetic calorimeter. From conservation of momentum, the (transverse component of the) 3-momenta of the jet and the photon must add up to zero and their energies are then the same. The selection of this type of event is achieved by requiring that the photon is well isolated, that there is no secondary jet and that the photon and the jet must be well separated in the transverse plane. Fig. 10.19 shows a typical event of this type as observed at DØ. However, a bias cannot be fully avoided since soft charged particles might not make it to the calorimeter due to the strong magnetic fields. Also, an additional second soft jet can spoil the momentum balance.

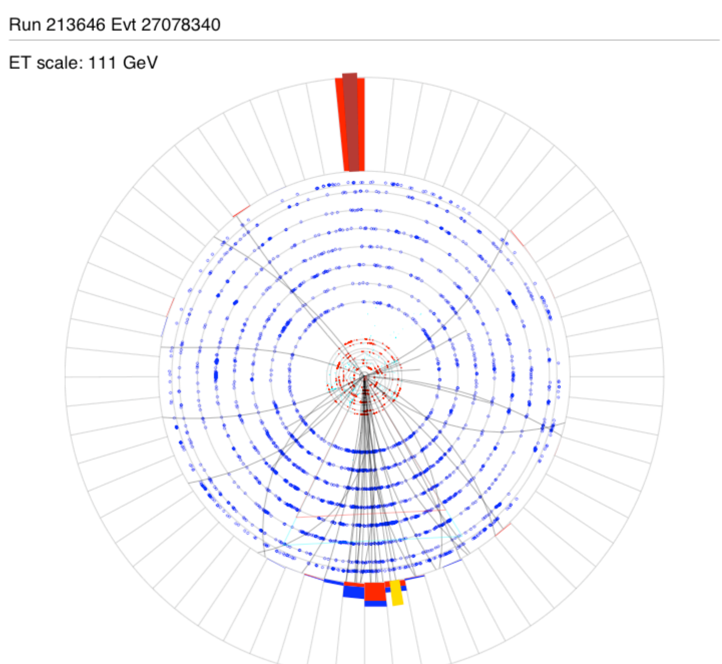
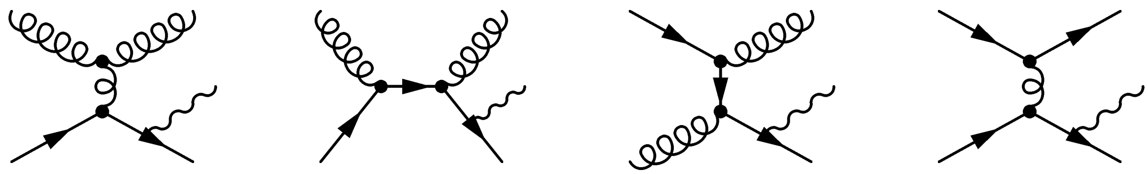


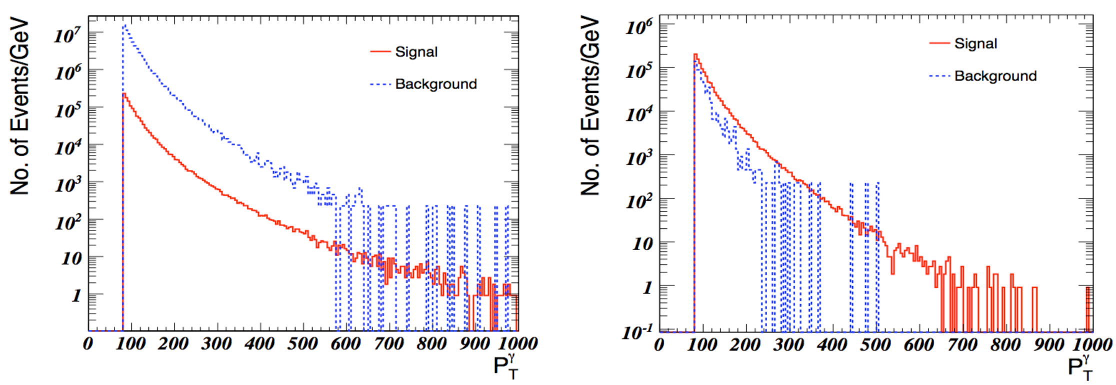
Figure 10.19: Event display of DØwith a jet and a photon used to calibrate the jet energy scale.

10.5.4 Isolation

It is often the case that the signal we would like to observe is smaller than the background (e.g. Higgs). The way out is the introduction of some filtering procedure to reject background events. This can be for example a veto on events presenting an energy deposit in a given cone about a photon. For an observed jet + photon event, one background consists in a photon radiation off the final state quark, yielding a 2-jets + photon event,



In this specific case, a sufficient requirement is that the photon should be isolated, i.e. there are no energy deposits nor charged tracks in a cone around the photon. This does not exclude possible "fake" photons from a boosted pion decay, $\pi^0 \to \gamma\gamma$. Fig. 10.20 shows the background and signal before and after the isolation cut for jet + photon events.



(a) (b)

Figure 10.20: *Isolation.* Signal and background (a) before and (b) after isolation cut.

10.5.5 Di-jet events

To look for new physics (e.g. a heavy gauge boson Z' of some grand unified theory) a simple procedure consists in making a histogram of the invariant mass of all di-jet events. A new resonance would then manifest itself as a peak in this histogram. In the eventuality that there is no special feature, we can test higher order QCD corrections (see Fig. 10.21).

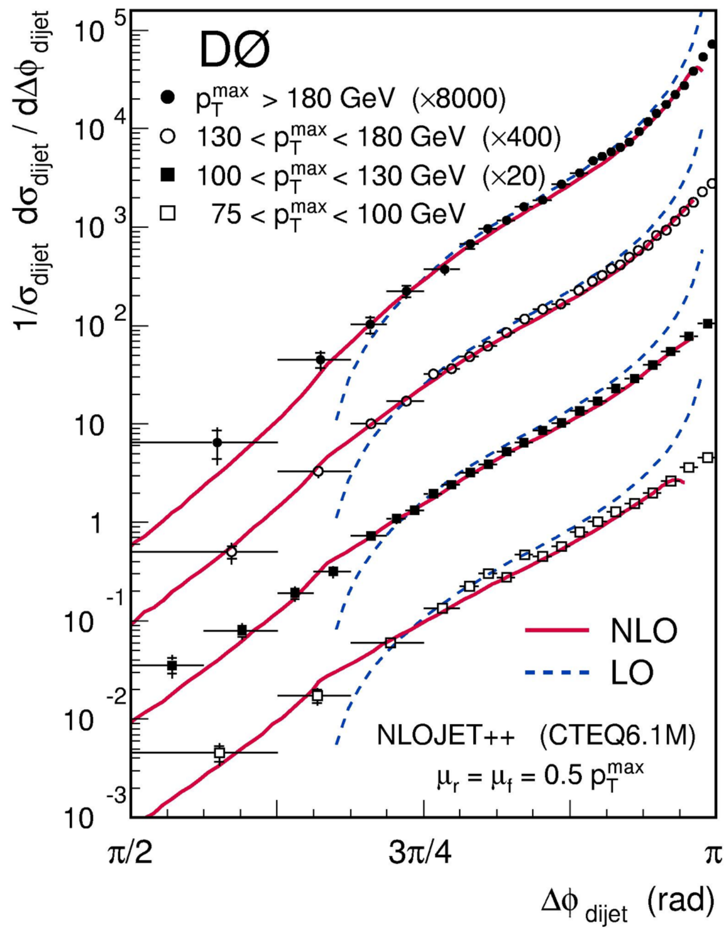


Figure 10.21: Measured dijet angular distributions at DØcompared to LO and NLO QCD calculations.

10.6 W and Z production

After jets the second most abundantly produced type of events at LHC are the ones containing the massive gauge bosons of the weak interaction : $W^{\pm}$ and Z^0 (Fig. 10.5). For a luminosity of $10^{34}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$ (design luminosity of LHC) at $\sqrt{s} = 14\,\mathrm{TeV}$ there will be about 100 W bosons produced per second. These bosons can decay into leptons and are thus easy to separate from the hadronic signal (jets) : one filters events with high-p_T and isolated leptons.

Since the weak gauge bosons do not couple to gluons and a high center of mass energy is needed, the valence quarks are determinant for their production. At LHC, W bosons can be produced via $u\bar{d} \to W^+$ and $d\bar{u} \to W^-$, and since there are more u-valence quarks than d-valence quarks in the proton, there will be more W^+ produced than W^-.

10.6.1 Predictions

The production cross sections for weak gauge bosons in pp-collisions are known to NNLO in perturbative QCD. Fig. 10.22 shows some of the diagrams contributing to the production of Z bosons with two leptons in the final state.

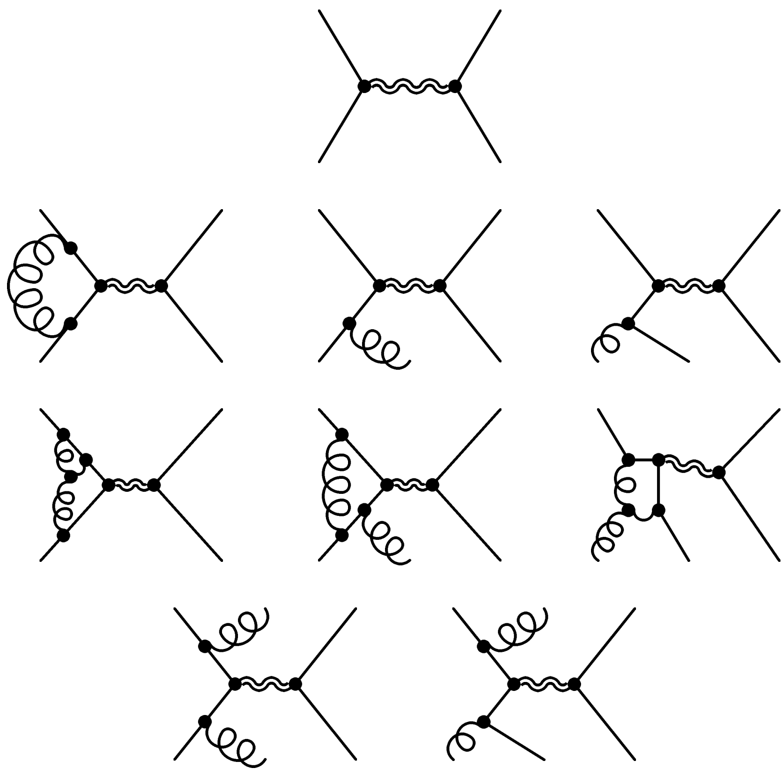
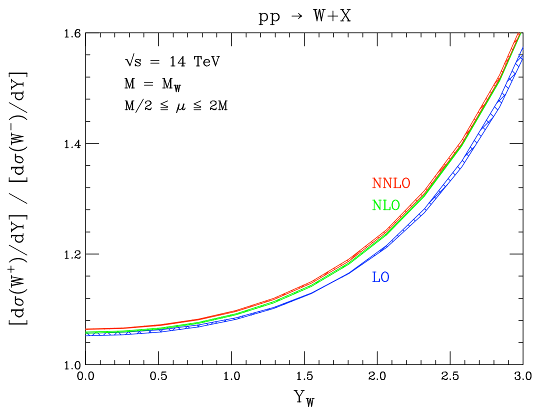


Figure 10.22: LO, NLO and NNLO Feynman diagrams for Z-production

Fig. 10.23 shows the double differential cross-section for W- and Z-production at LHC at LO, NLO and NNLO. One observes the stabilization of the shape and the small uncertainty at NNLO (at zero rapidity : 0.5-0.7% for the W and 0.1% for the Z). However, for the total production cross-sections significant uncertainties from the PDFs interfere and cause 4-5% of relative error (Remember the master formula, Eq. (10.4)).

10.6.2 Experimental signature

Events involving weak gauge bosons are relatively easy to spot, due to their clean signature. We focus here on a decay involving at least one lepton and disregard decays involving hadrons.



(a) (b)

Figure 10.23: *Differential production cross-section at LO, NLO and NNLO perturbative QCD as a function of the rapidity.* (a) W boson (b) Z boson.

Z **: pair of charged leptons** A Z boson decays (in its visible mode!) into a pair of charged leptons. These carry a large transverse momentum p_T, are well isolated, have opposite charge (bending direction) and have an invariant mass (Sect. 4.4.4, p. 52) in a typical range of $70 - 110\,\mathrm{GeV}$. Fig. 10.24 shows the topology of a typical Z event.

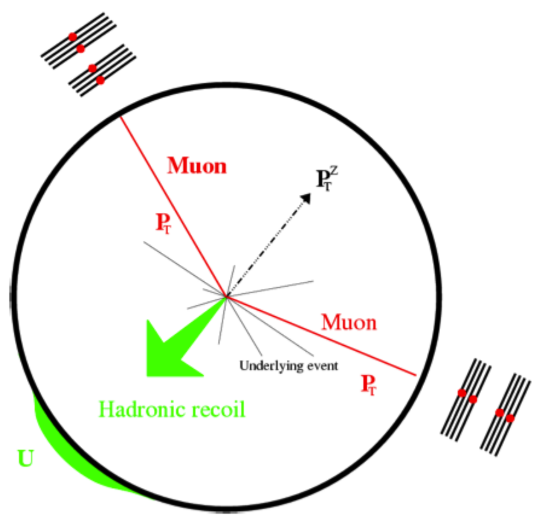


Figure 10.24: Typical dileptonic signature for a Z event.

W **: single charged lepton** A W boson decays into a charged lepton and its corresponding neutrino. The charged lepton has a large transverse momentum p_T and is well isolated. By summing the energies and momenta, one can deduce the missing p_T of the neutrino that escapes undetected. Fig. 10.25 shows the topology of a typical W event.

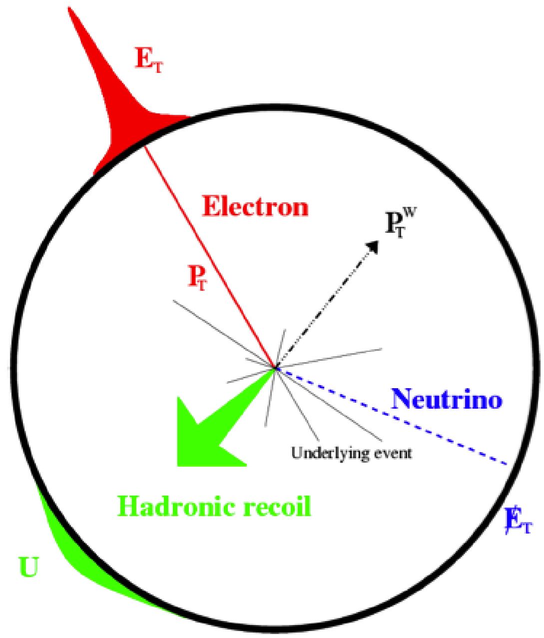


Figure 10.25: Typical signature for a W event.

Fig. 10.26 shows the experimental data (37584 candidates for a W production, $\mathcal{L}_{int} = 72\,\mathrm{pb}^{-1}$ of data) from CDF at Tevatron and the Monte Carlo simulation for different channels before and after a missing E_T cutoff : $\not{E}_T > 25\,\mathrm{GeV}$. The low-$E_T$ events (denoted QCD) correspond to collimated jets erroneously interpreted as electrons.

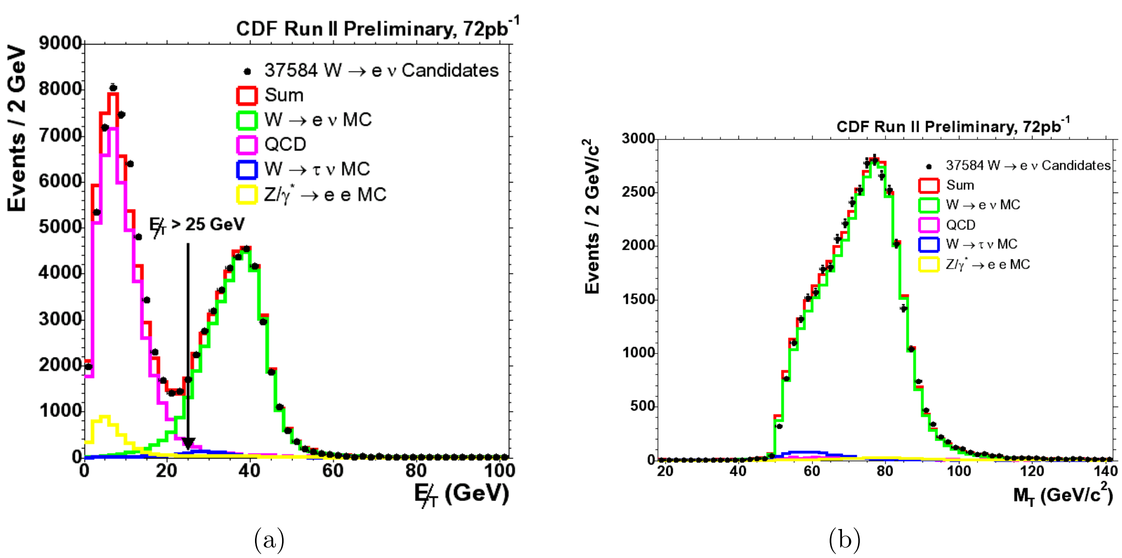
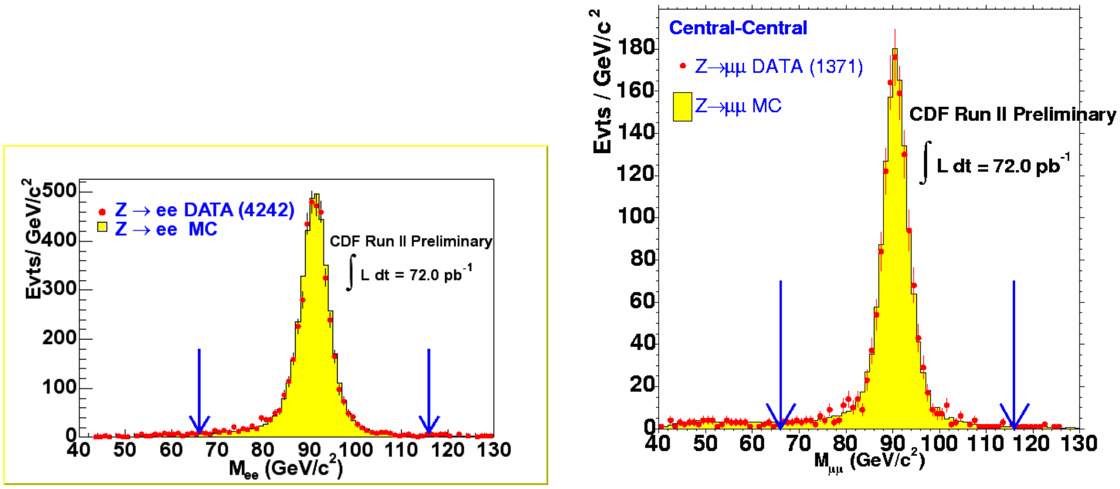


Figure 10.26: *Histogram of missing E_T associated with W production.* (a) Raw data. (b) Data after $\not{E}_T > 25\,\mathrm{GeV}$ cut.

Fig. 10.27 shows the invariant mass of e^+e^- (4242 candidates) and $\mu^+\mu^-$-pairs (1371 candidates) around the Z pole and the Monte Carlo simulation with $\mathcal{L}_{int} = 72\,\text{pb}^{-1}$ of data from CDF at Tevatron. The larger number of charged leptons (making the identification of the process easier) makes the background very small.



(a) (b)

Figure 10.27: *Histogram of the invariant mass of lepton pairs associated with Z production.* (a) Electron decay channel. (b) Muon decay channel.

The total production cross section for W and Z bosons are respectively (CDF, electron and muon channels):

$$\sigma_W = 2775 \pm 10(stat) \pm 53(sys) \pm 167(lum)[\text{pb}]$$
$$\sigma_Z = 254.9 \pm 3.3(stat) \pm 4.6(sys) \pm 15.2(lum)[\text{pb}].$$

Fig. 10.28 shows the evolution of the measured production cross-sections for weak gauge bosons at UA1, UA2, CDF and DØcompared to the theoretical prediction.

Fig. 10.29 shows the expected experimental missing E_T and invariant mass distribution at LHC after collection of $10\,\text{pb}^{-1}$ of data at 10 TeV for W and Z production respectively. Selection will be achieved by requiring isolated leptons and a transverse energy of 30 resp. 20 GeV.

W/Z + jets An important background for many searches are the W/Z + jets events. Indeed the topology of a SUSY event will be typically jet + n leptopns + $\not{E}_T$, $n \geq 0$, which is for example identical if the Z of a Z + jet event decays in 2 neutrinos, resulting also in a large $\not{E}_T$.

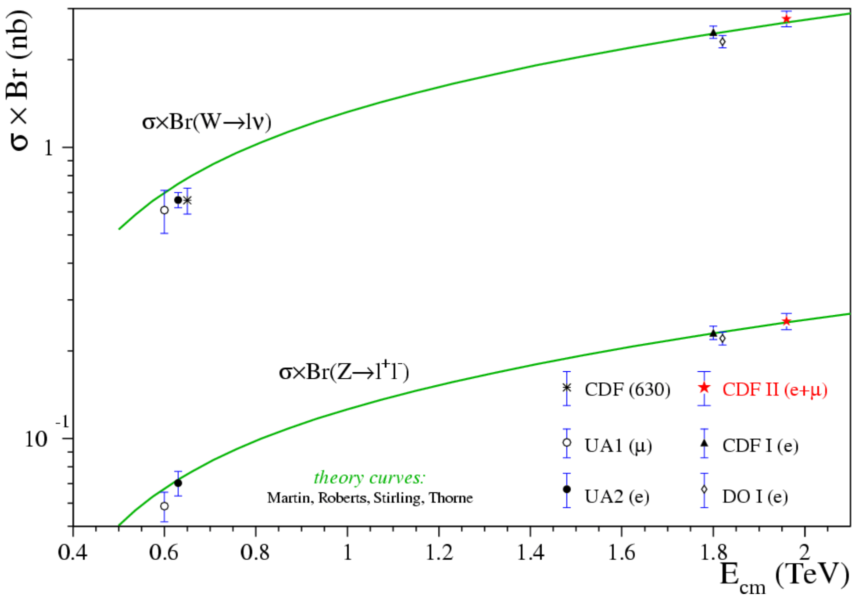
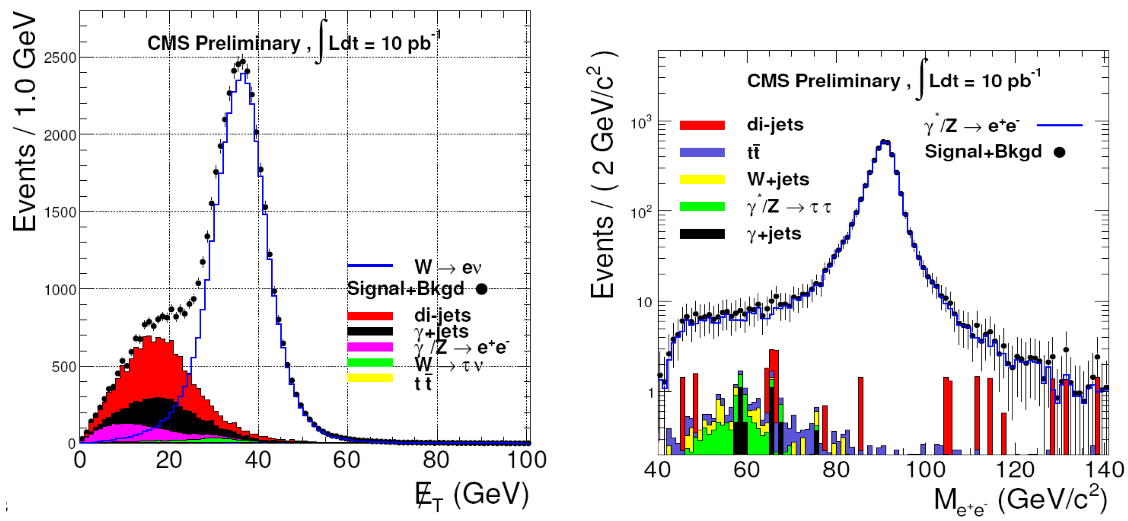


Figure 10.28: Evolution of the production cross-section for W and Z bosons.

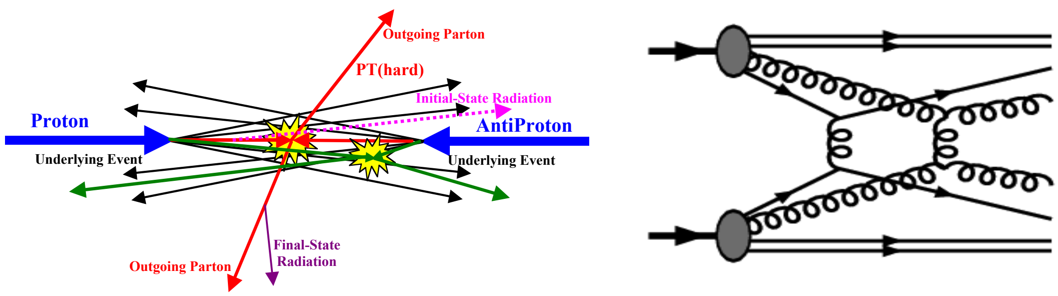


(a) (b)

Figure 10.29: *Simulation of the different signal and backgrounds for CMS.* (a) $W^- \to e^- \bar{\nu}_e$. (b) $Z \to e^+e^-$.

10.7 Underlying event and multi-parton interactions

So far we have neither discussed the role of the remnants left over e. g. from a hard scattering process like the one depicted in Fig. 10.9 nor the possibility of multiple-parton scattering. Since the scattering partons carry color, so do the remnants. Therefore, soft particle production out of the color field between parton and remnant is to be expected.



(a) (b)

Figure 10.30: *Underlying event.* The underlying event is everything except for the hard scattering component of the collision. This includes initial and final state radiation of soft gluons, spectators, and remnants (a), as well as multi-parton interactions (b).

Furthermore, as was discussed in connection with parton evolution, gluons may be radiated off before the partons engage in the actual scattering. In summary this means that many soft particles, not directly related to the hard scattering process, are around in the detector constituting the so-called underlying event (see Fig. 10.30(a)).

The underlying event is defined to be everything except for the hard scattering component of the collision, i. e. initial and final state soft gluon radiation, spectators, remnants, and multiple-parton interactions.

The momentum scale of the interaction is set by the parton hard scattering. There is the possibility of further partons engaging in scattering; one then speaks of multiple-parton scattering (see Fig. 10.30(b)). Since high p_T values are improbable, any further parton scatterings will, if they happen, do so at a lower p_T scale than the initial hard scattering. It is in this way that multi-parton scattering contributes to the background of soft hadrons potentially obscuring interesting results of hard parton scattering processes. Calculations concerning multi-parton interactions are hard and thus only phenomenological models with some parameters to be tuned exist. In tuning these parameters for LHC, the issue is their energy dependence.

Now that the problem is stated, let us examine the possibilities to study the underlying event by taking a look at corresponding observables. One possibility is to work with charged jets, using minimum bias and and jet triggers. Looking for the highest p_T (leading jet), the direction $\phi = 0$ is defined (see Fig. 10.31). Since the underlying event should be uniformly distributed in ϕ, the transverse region, where neither the leading jet nor the back-to-back jet are relevant, is particularly sensitive to the underlying event. The underlying event depends on the leading jet p_T and one wants to measure how many particles are in the transverse region per rapidity and angle and their transverse momentum, i. e. the charged density $dN/d\eta d\phi$ and the transverse momentum density $dp_{T,\text{sum}}/d\eta d\phi$. Another possibility is to work with Drell-Yan muon pair production (see Fig. 9.21(b)), using muon triggers. In this case, after removing the muon pair, everything else is by definition the underlying event.

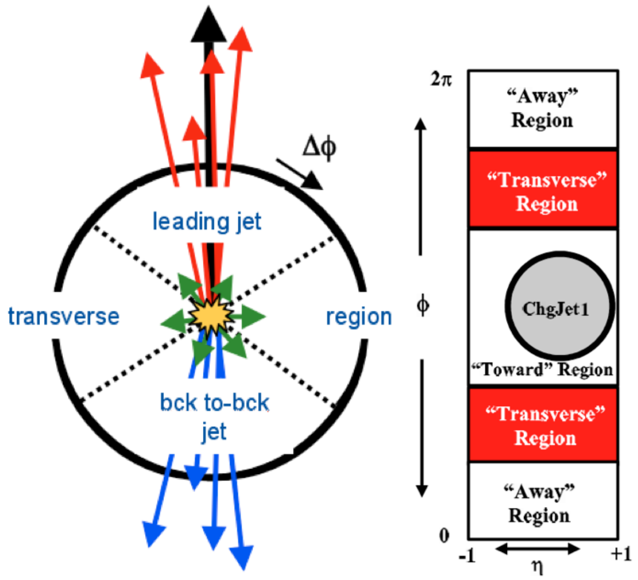


Figure 10.31: *Leading jet and transverse region.*

Examples of results for the charged density, $dN/d\eta d\phi$, as function of ϕ and leading jet p_T are shown in Fig. 10.32. Besides the expected peaks at the position of the leading jet and in the opposite direction, one can also observe that the underlying event depends on the leading jet transverse momentum. This behavior is also shown in the RHS plot, which in addition illustrates the dependence of underlying event studies on phenomenological models and the values of their parameters.

To conclude this section on the underlying event, let us briefly mention handles to estimate multiple partonic interaction rates: One can count pairs of mini-jets (two additional jets balanced on their own) in minimum bias interactions, reconstructing them using charged tracks. Another possibility is to look for the production of same-sign W pairs.[5]

10.8 Top production

Since the top quark which was discovered in the nineties at Tevatron is much heavier than the other quarks and leptons (see Fig. 10.33(a)), one might suspect a special link to the Higgs which, after all, should be responsible for nonvanishing masses. Because of its large mass, the top decays immediately into bW^+, such that no top-mesons can be produced.

Why is it important to measure the top mass (besides in its own right)? First of all, m_t, combined with m_W, yields an indirect constraint on the Higgs mass (see Fig. 10.33(b)). Furthermore, the measurement of m_t serves to test the overall consistency of the standard model (or of something beyond that), if the Higgs is found. The Higgs contributions to

[5] See [58].

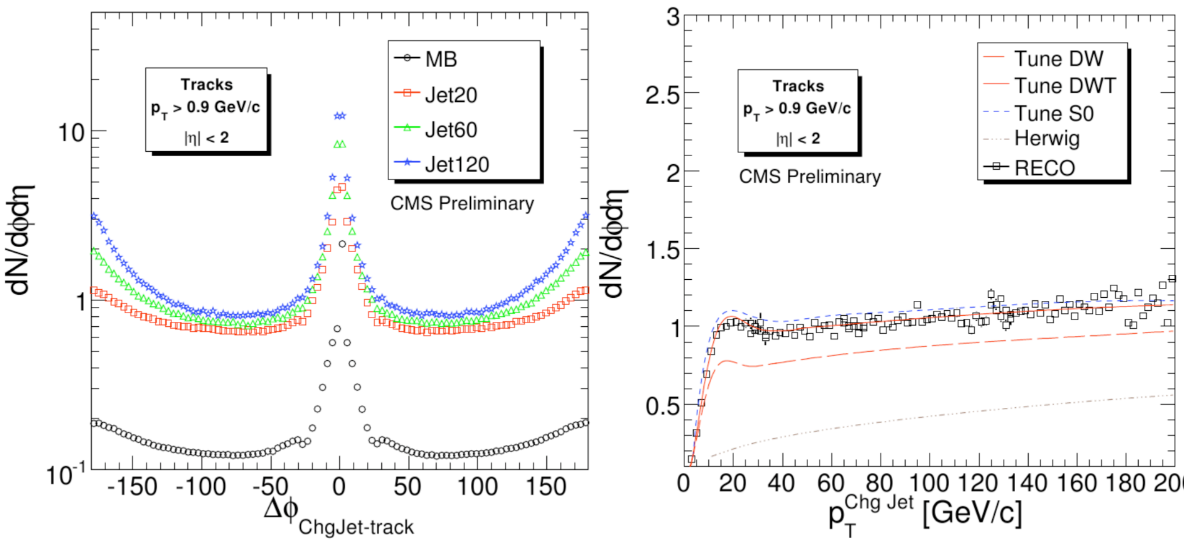


Figure 10.32: *Underlying event studies at CMS.*

cross sections depend on m_t. One can therefore check how well corresponding predictions agree with the data as a function of m_t. The ellipses in Fig. 10.33(b) state restrictions on the Higgs mass from measurements of m_t and m_W and can accordingly be shrunk by more precise measurements of these masses.

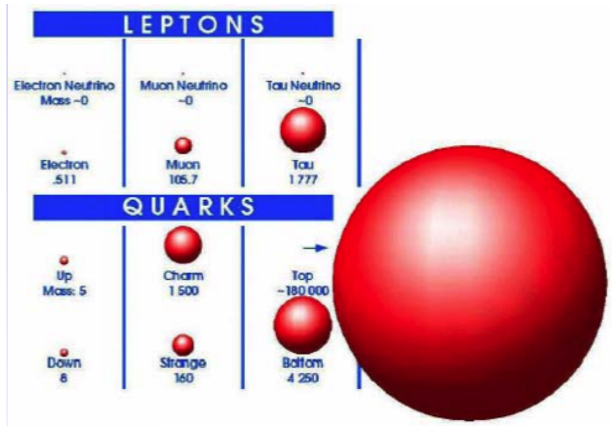
Possible measurements related to the top quark include the production cross section, the production via a heavy intermediate state Z' (resonance production), along with mass, spin and charge. A summary of top quark physics is given in Fig. 10.34.

The examples of decay modes given here indicate the type of events originating from top production: They involve many jets and possibly missing energy. But what exactly does the shaded blob (in Fig. 10.34) hide? Two possible diagrams for top production are shown in Fig. 10.35. Initial state gluon radiation may produce additional hadrons X or the $t\bar{t}$ pair may be produced in pair creation by two gluons.

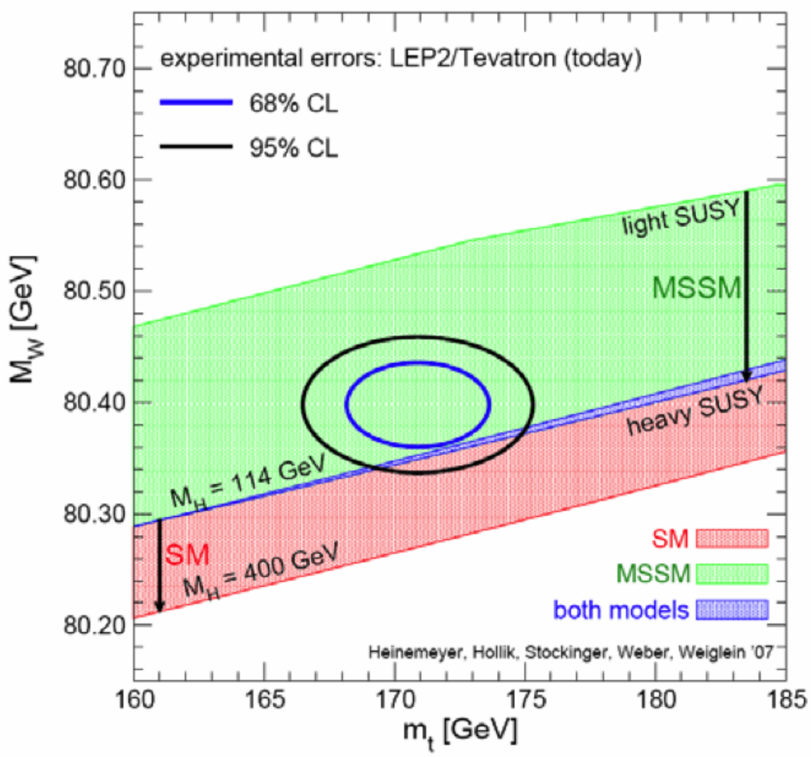
As mentioned before, the top almost immediately and exclusively decays into W^+b: $\mathrm{BR}(t \to W^+b) \sim 100\%$. According to the subsequent decays of the thus produced Ws one classifies the top decay channels as follows:

- *Dilepton channel.* Both Ws decay via $W \to l\nu$ ($l = e$ or μ; 5%);
- *Lepton + jet channel.* One W decays via $W \to l\nu$ ($l = e$ or μ; 30%);
- *All-hadronic channel.* Both Ws decay via $W \to q\bar{q}$ (44%).

Therefore, important experimental signatures are leptons or lepton pairs, missing transverse momentum (ν), and b jets. In terms of detection, the all-hadronic channel causes some difficulties, since the QCD background has a comparable magnitude. Figure 10.36(a) shows some features of an event that can be used to search for jets originating from b



(a)



(b)

Figure 10.33: *Top quark mass (a) and constraints on Higgs mass by m_t and m_W (b).*

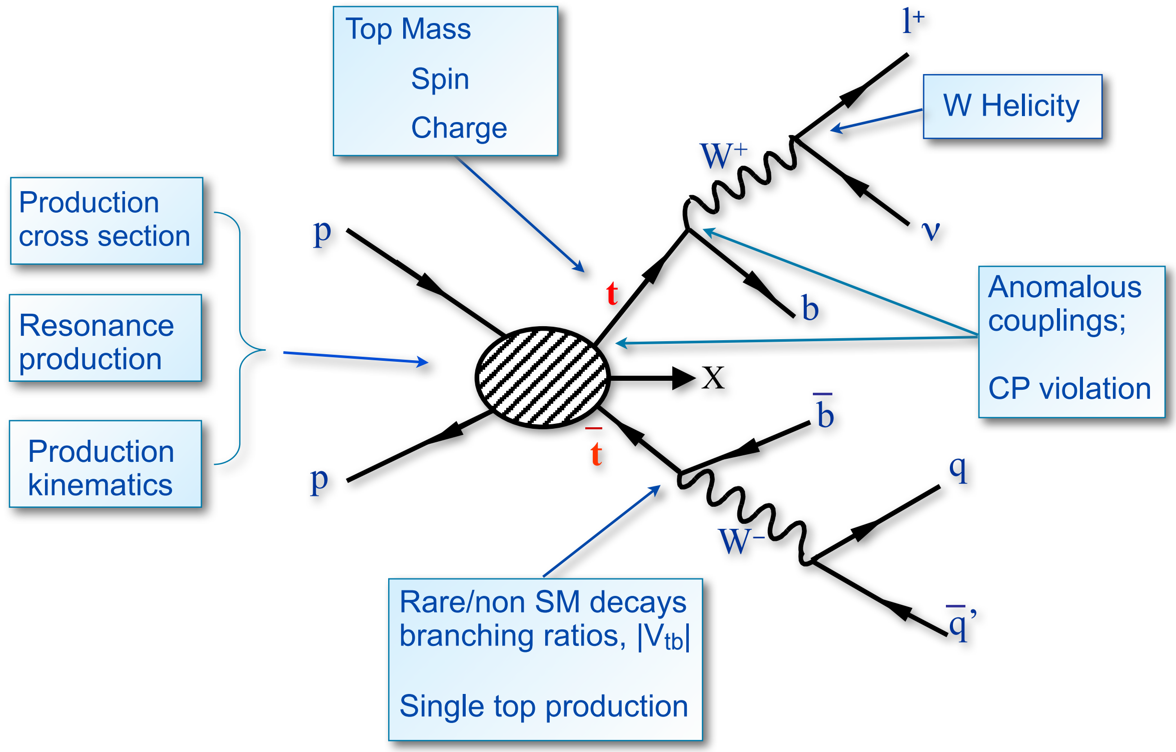


Figure 10.34: *Top physics summary.*

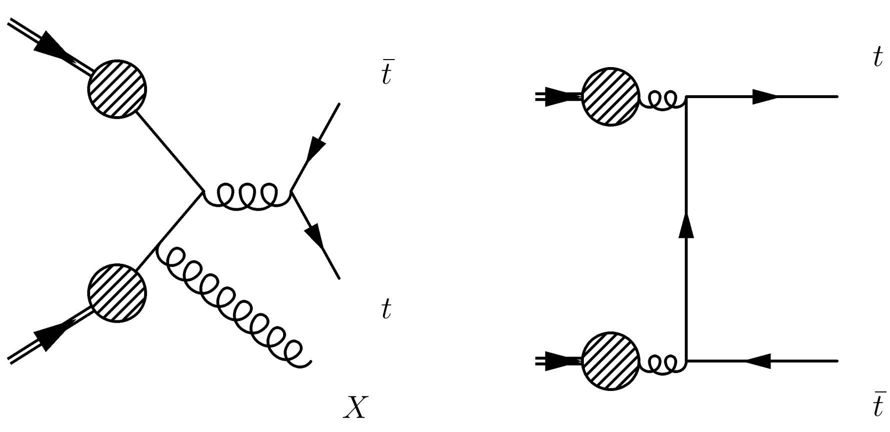
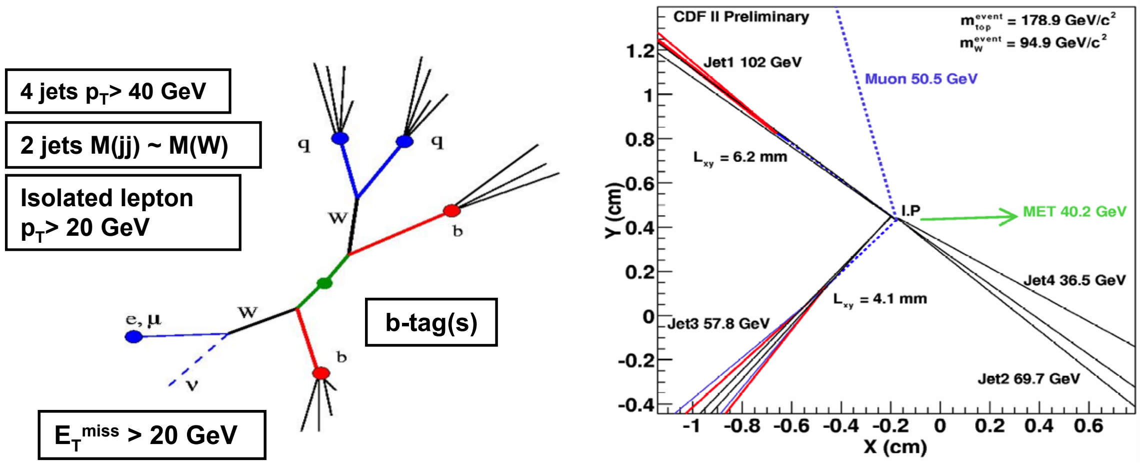


Figure 10.35: *Two possibilities for top production.* In (a) initial state radiation produces additional hadrons X while in (b) the top pair is produced by pair creation.

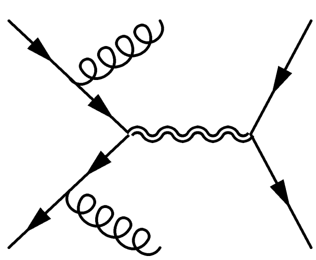
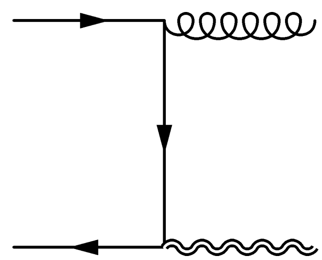


(a) (b)

Figure 10.36: b *tagging.* Event features used to identify b jets (a) and vertex close-up of a top decay. b jets can be identified by looking for displaced vertices. They arise because B mesons can travel some millimeters before decaying (b).

quarks (b tags). From the invariant mass of the jets the top mass can be reconstructed; however, it can be difficult to correctly combine the observed jets. Since b tagging is important for top identification, excellent silicon vertex and pixel detectors are needed to measure displaced tracks originating from secondary vertices. These secondary vertices arise because the B meson lifetime allows it to travel some millimeters before decay. Therefore, displaced vertices can be used to find b jets, see Fig. 10.36(b).

Results of measurements which employ the discussed criteria for b tagging are shown in Fig. 10.37.[6] On the LHS semi-leptonic events (one b tag) are counted, while the RHS lists events with two b tags (which excludes one-jet events). The background of the measured top signal stems from the production of W + jets by diagrams like the following:

One can observe that the background signal relies on gluon radiation for jet production and is therefore rather limited in jet multiplicity. An example for the top mass reconstruction from lepton + jets events is shown in Fig. 10.38.

[6]For a collection of Tevatron results on the top mass and production cross section see e. g. `http://www-cdf.fnal.gov/physics/new/top/top.html` or `http://www-d0.fnal.gov/d0_publications/d0_pubs_list_runII_bytopic.html#top`.

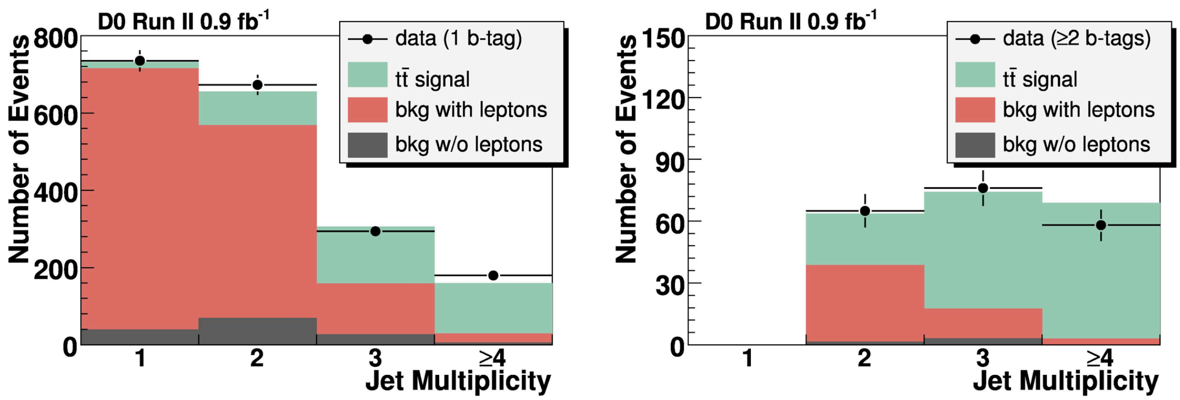


Figure 10.37: *Jet multiplicity and b tagging.*

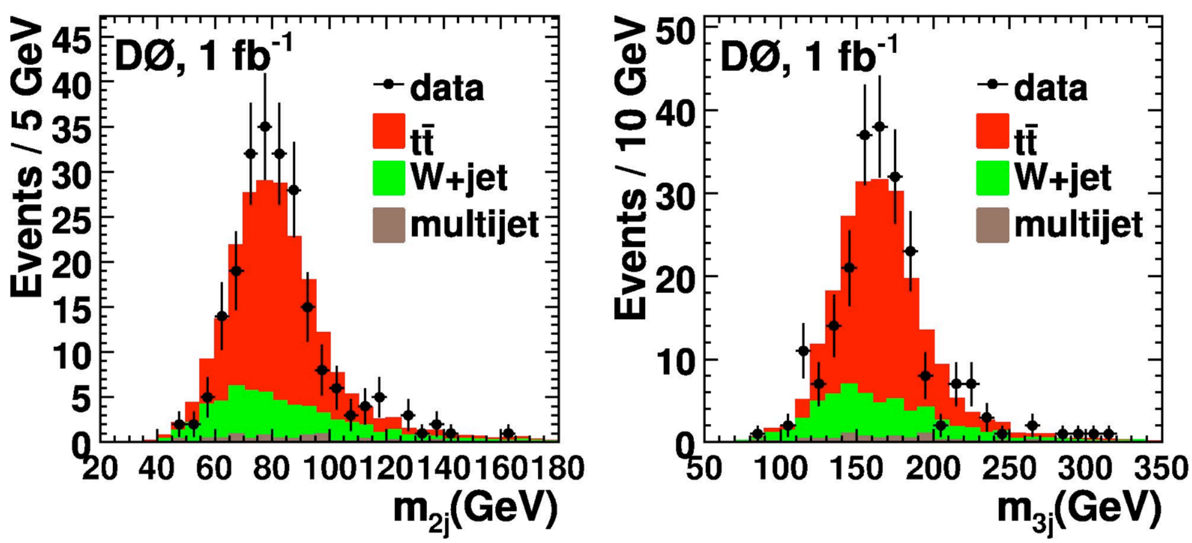
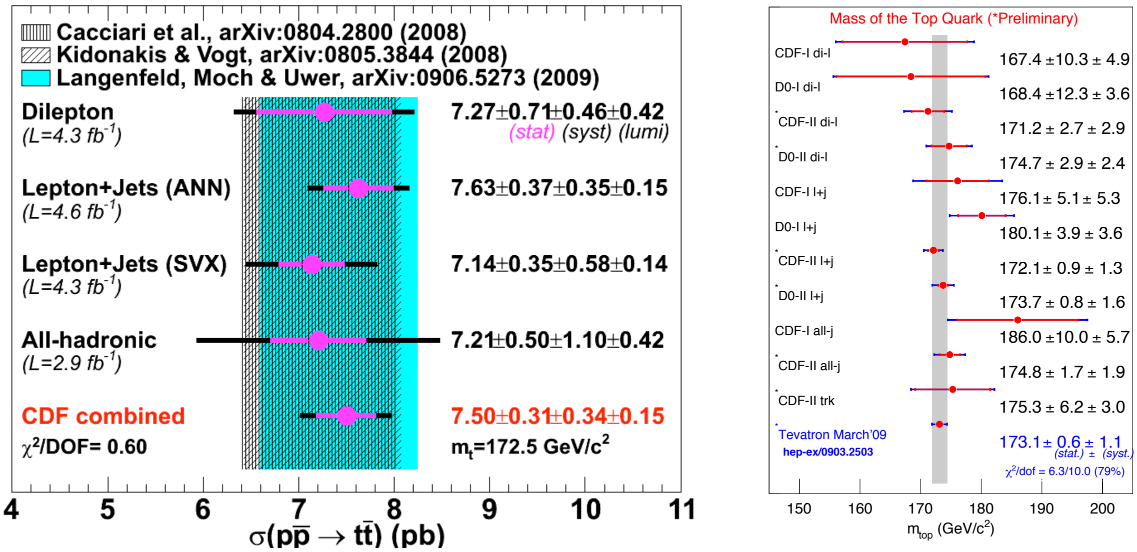


Figure 10.38: *Mass reconstruction.* Comparison between data and Monte Carlo two-jet (m_{2j}) and three-jet (m_{3j}) invariant mass distributions.



(a) (b)

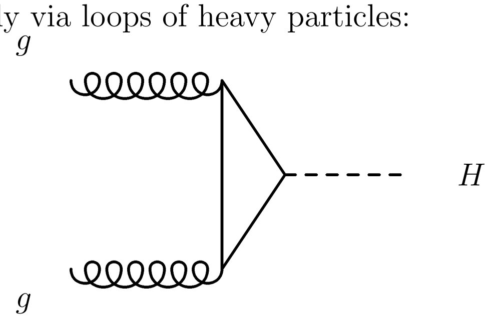
Figure 10.39: *Tevatron results for top production cross section and mass.*

By combining measurements for different decay channels (see Fig. 10.39(a)) the CDF experiment determined the top production cross section to be $\sigma^{p\bar{p}\to t\bar{t}}/\text{pb} = 7.50 \pm 0.31 \pm 0.34 \pm 0.15$ (statistical, systematic, and integrated luminosity errors). This is compared to theoretical predictions for a top mass of $m_t = 172.5\,\text{GeV}$ and $\sqrt{s} = 1.96\,\text{TeV}$. Top mass results obtained by considering various channels are given in Fig. 10.39(b).

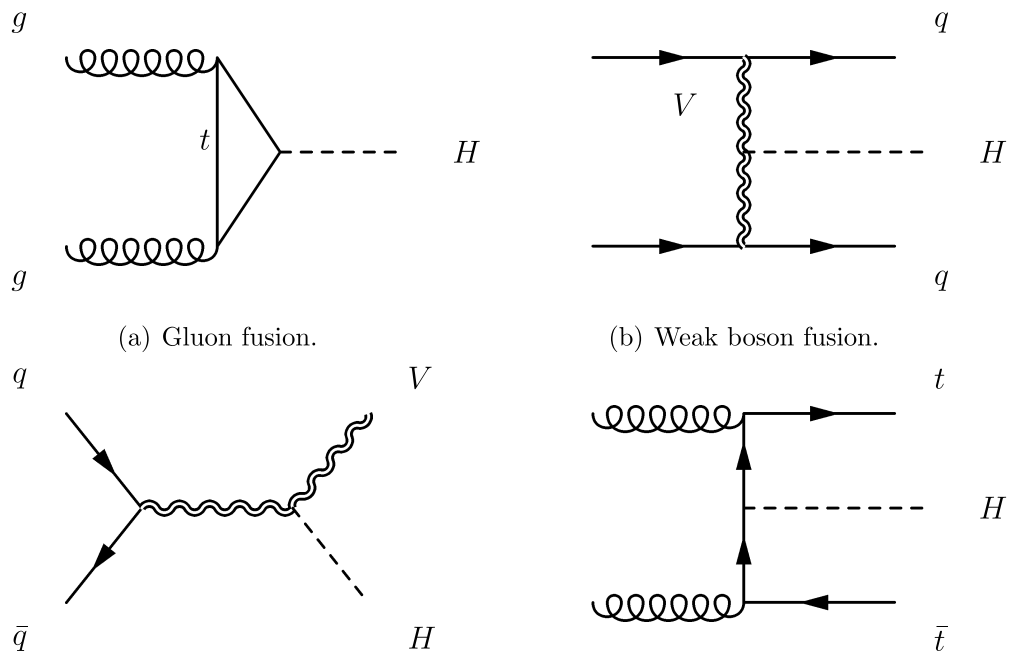
10.9 Searches for a SM Higgs and SUSY

We conclude this chapter on collider physics by discussing ways to produce and detect a standard model Higgs and SUSY particles.

Let us first examine Higgs production. The Higgs couples to particles with mass, while it couples to g and γ indirectly via loops of heavy particles:



This motivates the first of the four production diagrams shown in Fig. 10.40. Gluon fusion is the most likely one of these processes at hadron colliders (if $m_H \sim 100-200\,\text{GeV}$), since for small parton momentum fractions x gluons are dominating in the proton PDFs (see



(a) Gluon fusion. (b) Weak boson fusion.

(c) Higgs strahlung. (d) Associated production.

Figure 10.40: *Higgs production at hadron colliders.*

Fig. 9.18(b)). Figure 10.41 shows as functions of m_H the corresponding standard model cross sections $\sigma^{pp\to H+X}$ at $\sqrt{s} = 14\,\text{TeV}$. Again, one observes that the gluon fusion cross section is dominant; the subdominant mechanisms are important for measuring the Higgs couplings.

To appreciate the challenges in Higgs detection, we now discuss Higgs decay. Branching ratios and width predictions are shown in Fig. 10.42(a) and 10.42(b), respectively. The Higgs couplings to fermions grow with their masses and the coupling of H to W and Z grows as m_H^2. Therefore, the branching ratios strongly depend on the Higgs mass. If m_H is around 120 GeV the dominant channel is decay to b quarks. This basically leads to two-jet events which compete with a large QCD background. Although the 2γ channel only has a branching ratio of ~ 0.002 it is still useful since in this case detection is easier as in the b quark case. Also, together with jets, the tau channel seems feasible. In the case of $m_H = 120 - 200\,\text{GeV}$ the W and Z channels are dominant. Figure 10.42(b) shows the total Higgs width as a function of the Higgs mass: Only for m_H less than 200 GeV a narrow resonance is to be expected. In the most likely mass region there is a considerable spread in possible values for the the total Higgs width.

Combining Higgs production cross sections and branching ratios, we can (in parts recapitulatory) discuss some experimental signatures:

- *Two-photon final states.*
  Excellent detector resolution, isolation and rejection of QCD background jets is required.

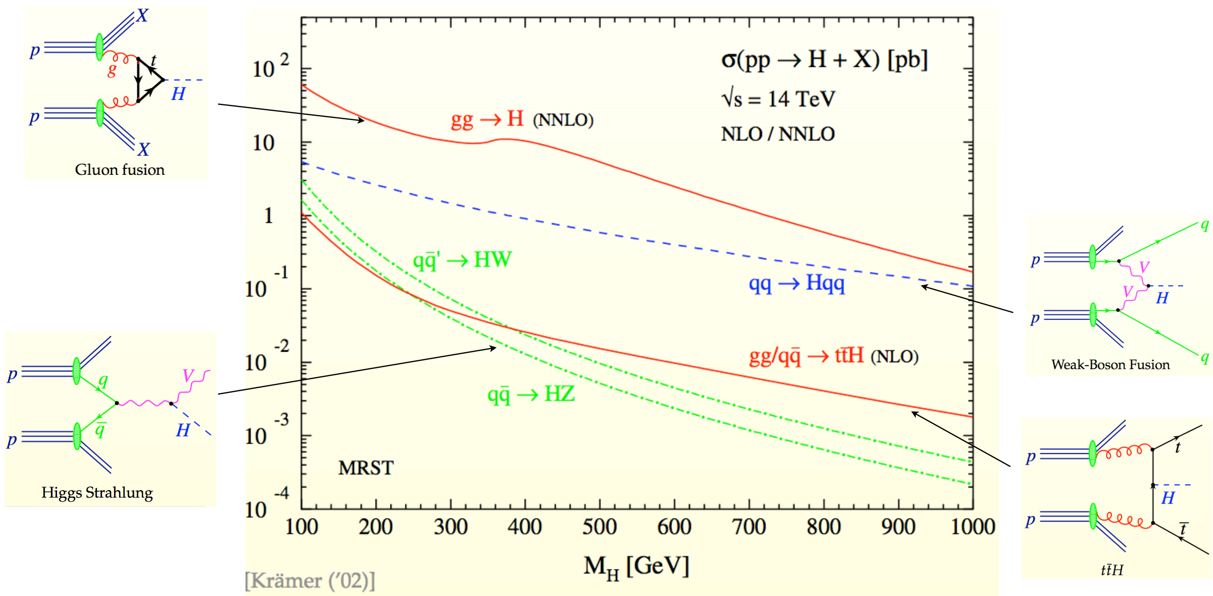
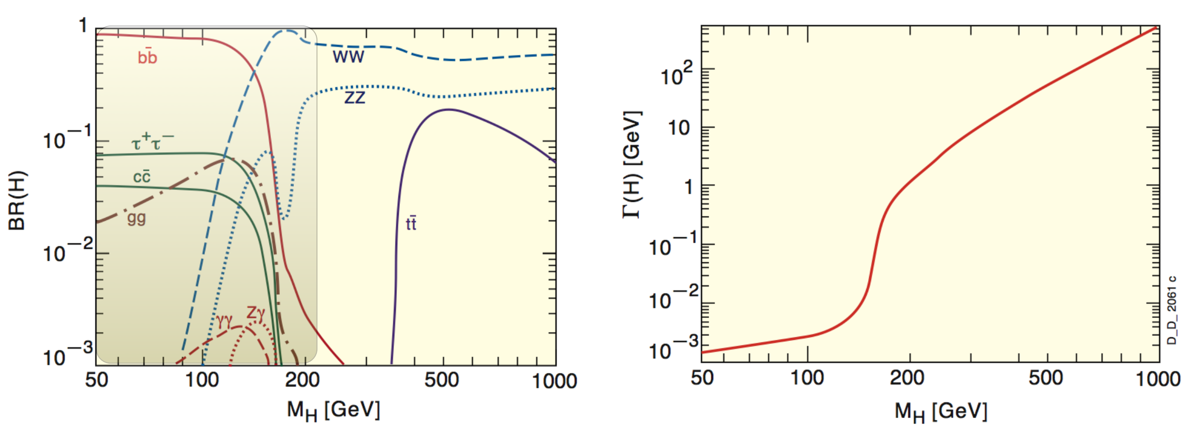


Figure 10.41: *Higgs production cross section as function of Higgs mass.*



(a) (b)

Figure 10.42: *Higgs branching ratios (a) and total width (b).*

- *Lepton final states (μ, e or τ).*
  As in the 2γ case the final state has to be isolated. This measurement relies on the lepton momentum resolution and, if necessary, τ identification.

- *Lepton + neutrino final states.*
  Here lepton identification and missing energy resolution are important. In addition, the background from W and t pairs has to be rejected.

- *Associated Higgs production ($b\bar{b}H$, $t\bar{t}H$).*
  b tagging as well as jet spatial and energy resolution are important. Background signal from hadronic top decays.

- *Higgs production via vector boson fusion.*
  The two jets in forward direction have to be identified: "very forward jet tagging". This signature (a rapidity gap appears if the Higgs is produced by vector boson fusion) will help distinguish the signal from the hadronic top decay and underlying event background.

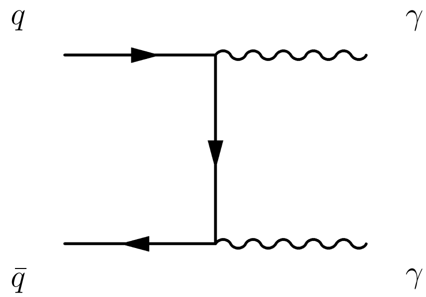
The final states can be classified according to whether mass reconstruction is possible: For the final states $\gamma\gamma$, $4l$, and $b\bar{b}$ the mass can be fully reconstructed. In these cases the background is obtained from the "sidebands" surrounding the signal box. For hadronic final states an excellent jet E_T resolution is needed. Final states containing neutrinos form a second class for which no exact mass reconstruction is possible. Such decays are for example $H \to W^+W^- \to l^+\nu l^-\bar{\nu}$ or decays into tau pairs. In these cases one will look for Jacobian peaks in the transverse mass spectrum, while the background will be determined from sideband measurements if possible.

As we have discussed, there are three important Higgs discovery channels:

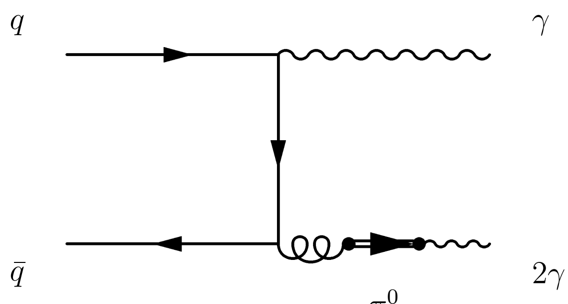
- $m_H \simeq 114 - 140\,\text{GeV}$: $\gamma\gamma(H \to \gamma\gamma)$;
- $m_H \simeq 140 - 175\,\text{GeV}$: $2l + \not{E}_T(H \to WW^{(\star)})$ and $4l(H \to ZZ^{(\star)})$;
- $m_H \simeq 175 - 600\,\text{GeV}$: $4l(H \to ZZ^{(\star)})$.

Note that there are further possibilities under detailed study which appear more difficult for now. These are vector boson fusion with decay into taus and associated Higgs production with Higgs decays into b quark pairs which may turn out to be extremely difficult.

As an example for event selection and background treatment in measuring the important Higgs discovery channels consider the decay $H \to \gamma\gamma$. In this case the event selection would proceed as follows: Search for two isolated photons such that $p_{T,1} > 25\,\text{GeV}$, $p_{T,2} > 40\,\text{GeV}$, and $|\eta| < 2.5$ and identify the primary vertex. This procedure will yield about 30% selection efficiency. Estimating the background from the sidebands will yield an uncertainty smaller than 1% for an integrated luminosity of $20\,\text{fb}^{-1}$. The problem is that the reducible background will be large. Figure 10.43 shows a plot of expected background and signal. This QCD background arises for example from diagrams analogous to electron-positron pair annihilation:



The spectrum of these background photon pairs will just decrease with invariant mass without peaks, as is also shown in Fig. 10.43. Note that the simulated Higgs peaks shown there are amplified by a factor of 10. Therefore, integrated luminosities of much more than $1\,\mathrm{fb}^{-1}$ are needed to see a signal significantly above the background. There is also the possibility that one photon is produced immediately and instead of a second photon a gluon is radiated off which forms a π^0 that subsequently decays into two almost parallel photons which are finally detected as one. Overall, the event will therefore look like pair annihilation,



and it will contribute to the background since the large probability of radiating off the initial gluon outweighs the small probability of it forming one π^0 carrying almost all its momentum.

As a second example consider the channel $H \to ZZ^{(\star)} \to 4l$. In this case the selection goes as follows: Look for four isolated and well reconstructed leptons; because they originate from Z decays, they can be either two e^+e^- pairs (see Fig. 10.44(a)) or two $\mu^+\mu^-$ pairs or an e^+e^- and a $\mu^+\mu^-$ pair (see Fig. 10.44(b)). The transverse momentum should be above $5 - 10\,\mathrm{GeV}$. For $m_H \sim 140 - 150\,\mathrm{GeV}$ the expected signal should be larger than the background produced by top decays: $t\bar{t} \to WbWb \to l\nu\, cl\nu\, l\nu\, cl\nu$. This background contribution is considerable, since $\sigma \times \mathrm{BR} \sim 1300\,\mathrm{fb}$. In oder to reduce it, criteria based on isolation of the detected leptons and secondary vertexing can be used.

10.9.1 The road to discovery

There are three scenarios for an early discovery which vary in their experimental difficulty.

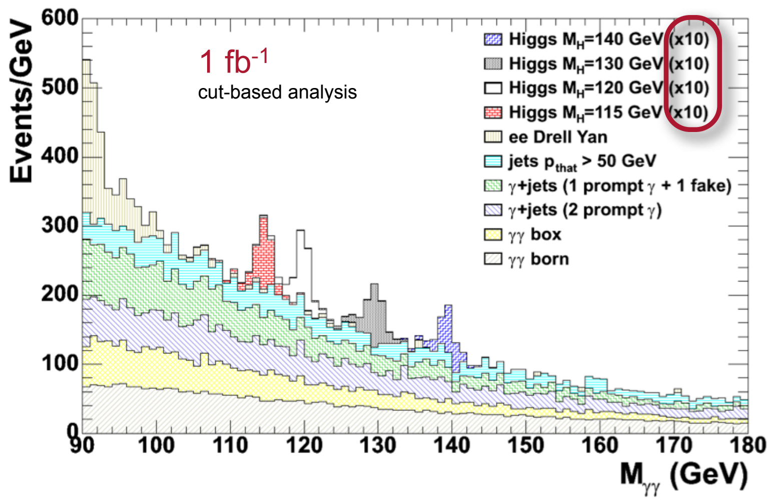
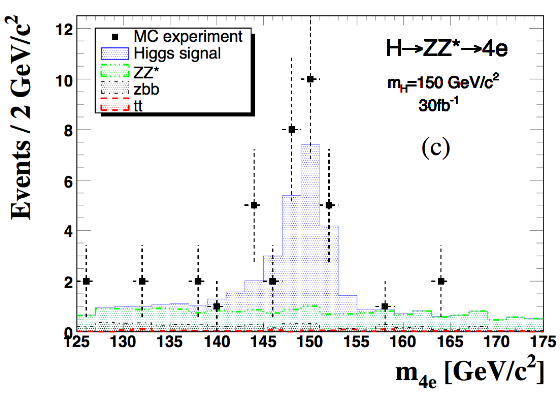
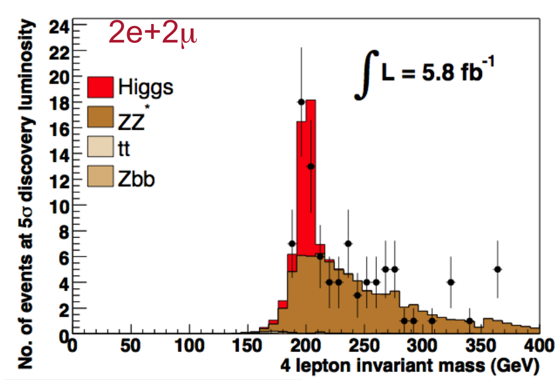


Figure 10.43: *Invariant mass a of photon pair for the Higgs decay $H \to \gamma\gamma$.* Note that the Higgs peaks are increased by a factor of 10.

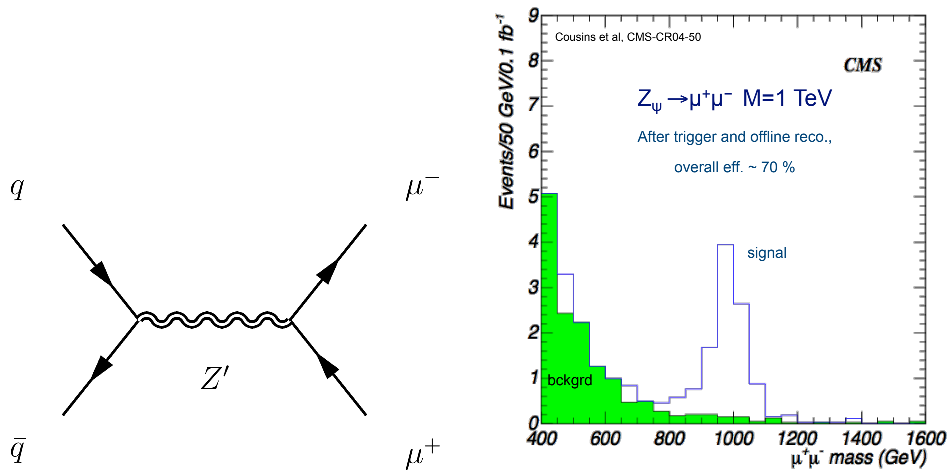


(a)



(b)

Figure 10.44: *Mass reconstruction in $H \to 4l$ decays.* Note that (a) shows the $4e$ final state case while (b) is the $2e2\mu$ case.



(a) (b)

Figure 10.45: *Diagram for production and decay of a new heavy resonance Z' (a) and expected signal for decay into a lepton pair (b).*

1. *An easy case.*
   A *new resonance* decaying into e^+e^- or $\mu^+\mu^-$, e. g. $Z' \to e^+e^-$ of mass $1 - 2\,\text{TeV}$ would be easily detectable.

2. *An intermediate case.*
   SUSY (See below.)

3. *A difficult case.*
   As we have seen, a *light Higgs* with $m_H \sim 115 - 120\,\text{GeV}$ would be difficult to detect since, with many other interactions happening at the same momentum scale as the Higgs mass scale, the background would be large.

The easy case is the production of new heavy gauge bosons, as predicted by GUT, dynamical EWSB, etc. which are generically called Z'. The diagram would look like in Fig. 10.45(a) and the background would be low and mainly stem from the Drell-Yan process (see Fig. 9.21(b)). The clear two-lepton signature combined with the low background should yield a clear signal as shown in Fig. 10.45(b).

Let us now turn to the intermediate case, the search for SUSY at the LHC. If SUSY exists at the EW scale, a discovery at the LHC should be easy. What helps is that squarks and gluinos are colored and are therefore produced via the strong interaction, which means large production cross sections. These then decay via cascades into the lightest SUSY particles (LSP) and other SM particles (leptons and jets) (see Fig. 10.46(a)). Thus the final states contain leptons, jets and missing energy. The general procedure will be as follows:

1. Look for deviations from the SM predictions, e. g. in the multi-jet + E_T^{miss} signature.

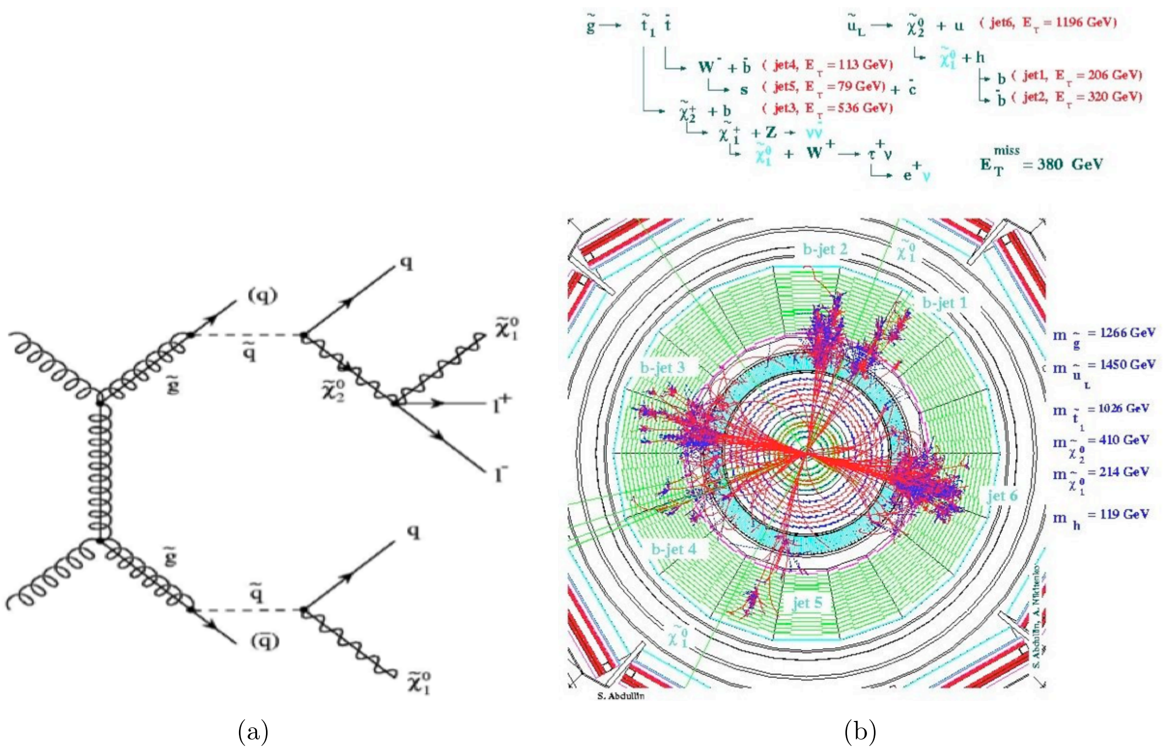


Figure 10.46: *Diagram for strong production and subsequent decay of SUSY particles (a) and SUSY event display simulation (b).*

2. Establish the SUSY mass scale by using inclusive variables such as the effective mass

$$M_{\text{eff}} = \not{E}_T + \sum_{\text{jets}} p_T(\text{jet}).$$

3. Determine the model parameters (difficult). The strategy is to select particular decay chains and to use kinematics to determine the mass combinations.

Because of the mentioned features SUSY events promise to be very spectacular: There will be many hard jets, large missing energy (from two LSPs and many neutrinos), and many leptons. A corresponding event display simulation is shown in Fig. 10.46(b). As one can see from the following numbers, for low SUSY mass scales the LHC should become a real SUSY factory (numbers for $\sqrt{s} = 14\,\text{TeV}$):

M/GeV	σ/pb	$\#_{\text{events per year}}$
500	100	$10^6 - 10^7$
1000	1	$10^4 - 10^5$
2000	0.01	$10^2 - 10^3$

Having said that, SUSY detection is still not easy, for it relies on good reconstruction and understanding of multi-jet backgrounds and missing transverse energy. A typical

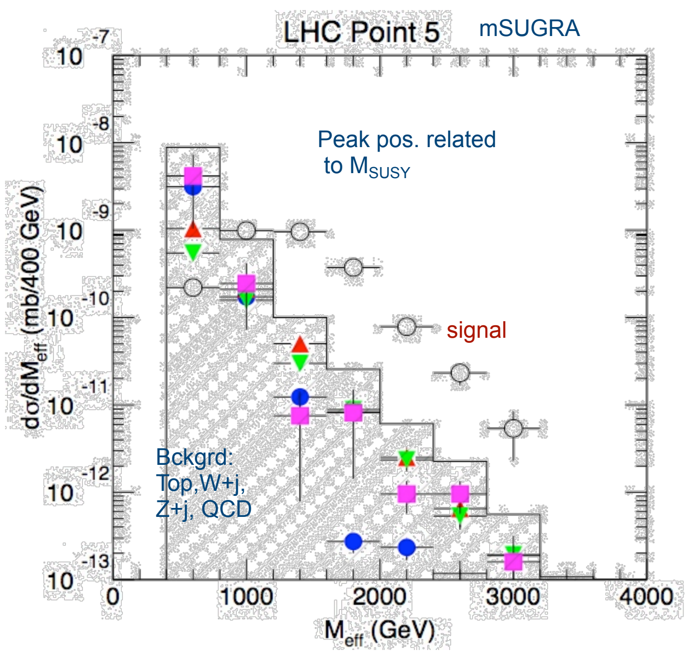


Figure 10.47: *Typical SUSY signal and backgrounds.*

selection would be based on the following criteria: $N_{\rm jet} > 4$, $E_T > 100, 50, 50, 50\,\text{GeV}$, and $\not{E}_T > 100\,\text{GeV}$. One would then hope to find a signal as shown in Fig. 10.47, where the effective mass variable $M_{\rm eff}$ is used.

Chapter 11

Electroweak interactions

Literature:

- Böhm/Denner/Joos [59]

In this chapter a unified theory of electromagnetic and weak interactions is discussed. The energy scale of this unification corresponds to the mass of the vector bosons: $E_{\mathrm{EW}} \sim M_W$, $M_Z \sim 100\,\mathrm{GeV}$. At low energies, in contrast, there are two distinct interactions, the electromagnetic interaction described by QED, and the weak interaction described by Fermi's theory. Some signals are also present in low energy atomic physics, e. g. electroweak interference and parity violation.

11.1 Introduction – the weak force

A comparison of strong, electromagnetic and weak interactions is given in the following table:

Interaction	Involved	$\sim \tau/\mathrm{s}$
Strong	quarks	10^{-23}
Electromagnetic	charged leptons and quarks	10^{-16}
Weak	all leptons and quarks	$10^{-6} - 10^{-8}$

One can observe that the timescales involved in weak decays are much larger than the ones of strong or electromagnetic decays. Thus, since $\tau \sim 1/\mathrm{coupling}^2$, the weak coupling is supposed to be some orders of magnitude smaller than the strong coupling (see also Sect. 7.3.3).

Weak processes are classified according to the leptonic content of their final state:

- *Leptonic.*
  E. g. $\mu^+ \to e^+ + \bar{\nu}_\mu + \nu_e$; $\quad \nu_e + e^- \to \nu_e + e^-$.

- *Semi-leptonic.*
  E. g. $\tau^+ \to \rho^+ + \bar{\nu}_\tau$.
- *Hadronic (non-leptonic).*
  E. g. $K^0 \to \pi^+ + \pi^-$; $\Lambda^0 \to n + \pi^0$.

The weak interaction violates parity (P) and charge conjugation (C) symmetry. It also violates CP and T, much more weakly, though. Also flavor is not conserved in weak interactions (see Sect. 7.3.2). If $m_\nu \neq 0$, neutrino oscillations occur and lepton family number is not conserved either.

Let us review some of the experimental results for the weak interaction.

Existence of neutrinos. Consider nuclear β^- decay, assuming a two-particle final state: $n \to p + e^-$. Since $m_e \ll m_n$, m_p, the recoil can be neglected and so

$$\begin{aligned} m_n &= E_p + E_e \\ m_n &\simeq m_p + p_e \\ p_e &\simeq m_n - m_p. \end{aligned}$$

This result means that for a two-body decay monoenergetic electrons are to be expected. However, the measured electron spectrum is continuous (see Fig. 11.1(a)). To solve this problem, Fermi and Pauli introduced an invisible neutrino carrying part of the decay energy: $n \to p + e^- + \bar{\nu}_e$ (see Fig. 11.1(b)). The Fermi theory amplitude for this process reads

$$\mathcal{M} = \frac{G_F}{\sqrt{2}} (\bar{\psi}_p \gamma^\mu \psi_n)(\bar{\psi}_e \gamma_\mu \psi_{\bar{\nu}}), \tag{11.1}$$

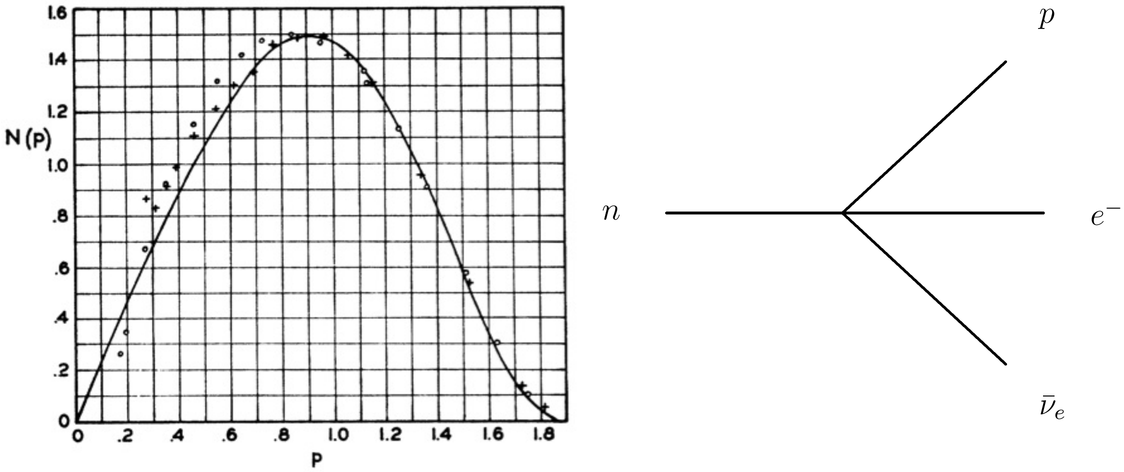
where $G_F \sim 1/(300\,\mathrm{GeV})^2$ is the Fermi constant. Note that the expression in Eq. (11.1) has vector structure and therefore does not violate parity. This point will be revisited later on.

Leptonic decays of $\pi^\pm$. Since $\pi^\pm$ is the lightest hadron, it cannot decay into other hadrons. Furthermore, electromagnetic decay (like in the case of $\pi^0 \to \gamma\gamma$) is forbidden by charge conservation. Thus no other channels are obscuring the study of the leptonic decay $\pi^+ \to \mu^+ + \nu_\mu$.

Non-observation of $\mu \to e + \gamma$. Although energetically possible, the decay $\mu^- \to e^- + \gamma$ is not observed in experiment. This leads to the introduction of a new quantum number called lepton number L, where

$$\begin{aligned} L_l &= 1 \qquad && e^-,\ \mu^-,\ \nu_e,\ \ldots \\ L_{\bar{l}} &= -1 \qquad && e^+,\ \mu^+,\ \bar{\nu}_e,\ \ldots. \end{aligned}$$

The leptonic muon decay conserving lepton number per family reads $\mu^- \to e^- + \nu_\mu + \bar{\nu}_e$.



(a) (b)

Figure 11.1: *β^- decay spectrum (a) and diagram (b).* (a) shows an electron momentum spectrum for the β^- decay of ^{64}Cu, source: [60, p. 14].

Parity violation. One famous instance of parity violation is the so-called τ-θ puzzle (1956). It consists in the finding that the Kaon K^+ decays into two final states with opposite parity:

$$
\begin{aligned}
&K^+ \begin{cases} \theta \to \pi^+\pi^0 \\ \tau \to \pi^+\pi^+\pi^- \end{cases} \\
&P\,|\pi\pi\rangle = (-1)(-1)(-1)^l = +1 \\
&P\,|\pi\pi\pi\rangle = (-1)^3(-1)^{l_{\pi_1\pi_2}}(-1)^{l_{\pi_3}} = -1,
\end{aligned}
$$

where l denotes angular momentum eigenvalues. The above is true for $J_{K^+} = 0$, since then, by conservation of angular momentum, $l = 0$ and $l_{\pi_1\pi_2} \oplus l_{\pi_3} = 0$ such that $l_{\pi_1\pi_2} = l_{\pi_3}$. Lee and Young introduced the idea that θ and τ are the same particle K^+ (fitting into its multiplet, see Fig. 7.6) which undergoes a flavor changing decay.

Another famous example for the demonstration of parity violation in weak interactions is the Wu experiment (1957). The idea is to consider β decay of nuclei polarized by an external magnetic field:

$$
\vec{B} \qquad \underbrace{{}^{60}\mathrm{Co}}_{J=5} \to \underbrace{{}^{60}\mathrm{Ni}^\star}_{J=4} + \underbrace{e^- + \bar{\nu}_e}_{J_z=1}
$$

The Cobalt nuclei are aligned to the external magnetic field and are in a state with $J = 5$. By conservation of angular momentum, the electron and neutrino spins have to be parallel (the decay product ^{60}Ni* is fixed). Since, to fulfill momentum conservation, they are emitted in opposite directions, the electron and its neutrino must have opposite chirality. It is observed that electrons are emitted preferentially opposite to the $\vec{B}$ field direction:

$$\begin{aligned}&\Gamma\left({}^{60}\mathrm{Co}\to{}^{60}\mathrm{Ni}^\star+e_L^-+\bar{\nu}_{e,R}\right)\\>&\Gamma\left({}^{60}\mathrm{Co}\to{}^{60}\mathrm{Ni}^\star+e_R^-+\nu_{e,L}\right)=P\left\{\Gamma\left({}^{60}\mathrm{Co}\to{}^{60}\mathrm{Ni}^\star+e_L^-+\bar{\nu}_{e,R}\right)\right\}.\end{aligned}$$

Thus left-handed leptons and right-handed antileptons $(e_L^-, \bar{\nu}_{e,R})$ are preferred over right-handed leptons and left-handed antileptons $(e_R^-, \bar{\nu}_{e,L})$. Recall (Sect. 5.2.4) that one uses the projectors $P_{\substack{R\\L}} = \frac{1}{2}(\mathbb{1} \pm \gamma_5)$ to indicate the chirality basis: $u_{L,R} = P_{L,R}u$.

These observations gave rise to the $V - A$ theory of weak interactions, described in Sect. 11.3 below.

11.2 γ_5 and $\varepsilon_{\mu\nu\rho\sigma}$

Recall that the amplitude in Eq. (11.1) does not violate parity. Therefore it has to be modified such that parity violation is included. To achieve this aim, the matrix γ^μ which forms the vector $\bar{\psi}\gamma^\mu\psi$ has to be replaced by a linear combination of elements of the set

$$\{\mathbb{1},\ \gamma^\mu,\ \sigma^{\mu\nu},\ \gamma_5\gamma^\mu,\ \gamma_5\}$$

where $\sigma^{\mu\nu} = \frac{i}{2}[\gamma^\mu, \gamma^\nu]$ and $\gamma_5 = i\gamma^0\gamma^1\gamma^2\gamma^3$. Using these matrices we can form the following field bilinears whose names are inspired by their transformation behavior under proper and improper Lorentz transformations[1]

$\bar{\psi}\psi$	scalar
$\bar{\psi}\gamma^\mu\psi$	vector
$\bar{\psi}\sigma^{\mu\nu}\psi$	tensor
$\bar{\psi}\gamma_5\psi$	pseudoscalar
$\bar{\psi}\gamma^\mu\gamma_5\psi$	pseudovector.

In Sect. 5.2.4 we discussed operators on spinor spaces, including helicity,

$$h = \frac{1}{2}\vec{\sigma}\cdot\frac{\vec{p}}{|\vec{p}|}\otimes\mathbb{1} \qquad\qquad P_\pm = \frac{1}{2}(\mathbb{1}\pm h),$$

and chirality,

$$\gamma_5 \qquad\qquad P_{\substack{R\\L}} = \frac{1}{2}(\mathbb{1}\pm\gamma_5).$$

Recall that in the high energy limit chirality and helicity have the same eigenstates.

The chirality matrix γ_5 has the following useful properties (see also Sect. 5.9)

[1] See e. g. [14, p. 64].

- $\gamma_5^2 = \mathbb{1}$;
- $\{\gamma_5, \gamma_\mu\} = 0$;
- $\gamma_5^\dagger = i\gamma^3\gamma^2\gamma^1\gamma^0 = \gamma_5$;
- $\text{Tr}\gamma_5 = 0$;
- Dirac-Pauli representation: $\gamma_5 = \begin{pmatrix} 0 & \mathbb{1} \\ \mathbb{1} & 0 \end{pmatrix}$.

Now let us define the totally antisymmetric ε tensor in four dimensions:

$$\varepsilon^{\mu\nu\rho\sigma} = \begin{cases} +1, & \{\mu, \nu, \rho, \sigma\} \text{ even permutation of } \{0, 1, 2, 3\} \\ -1, & \{\mu, \nu, \rho, \sigma\} \text{ odd permutation of } \{0, 1, 2, 3\} \\ 0 & \text{else} \end{cases}, \tag{11.2}$$

such that

$$\begin{aligned} \varepsilon^{0123} &= +1 \\ \varepsilon^{\mu\nu\rho\sigma} &= -\varepsilon_{\mu\nu\rho\sigma}. \end{aligned}$$

The product of two such ε tensors is then given by

$$\varepsilon^{\mu\nu\rho\sigma}\varepsilon^{\mu'\nu'\rho'\sigma'} = -\det \begin{pmatrix} g^{\mu\mu'} & g^{\mu\nu'} & g^{\mu\rho'} & g^{\mu\sigma'} \\ g^{\nu\mu'} & g^{\nu\nu'} & g^{\nu\rho'} & g^{\nu\sigma'} \\ g^{\rho\mu'} & g^{\rho\nu'} & g^{\rho\rho'} & g^{\rho\sigma'} \\ g^{\sigma\mu'} & g^{\sigma\nu'} & g^{\sigma\rho'} & g^{\sigma\sigma'} \end{pmatrix}$$

resulting in

$$\begin{aligned} \varepsilon^{\mu\nu\rho\sigma}\varepsilon_{\mu\nu}{}^{\rho'\sigma'} &= -2(g^{\rho\rho'}g^{\sigma\sigma'} - g^{\rho\sigma'}g^{\sigma\rho'}) \\ \varepsilon^{\mu\nu\rho\sigma}\varepsilon_{\mu\nu\rho}{}^{\sigma'} &= -6g^{\sigma\sigma'} \\ \varepsilon^{\mu\nu\rho\sigma}\varepsilon_{\mu\nu\rho\sigma} &= -24 = -4!. \end{aligned}$$

Using the definition in Eq. (11.2), one can express γ_5 as

$$\gamma_5 = i\gamma^0\gamma^1\gamma^2\gamma^3 = -\frac{i}{4!}\varepsilon_{\mu\nu\rho\sigma}\gamma^\mu\gamma^\nu\gamma^\rho\gamma^\sigma.$$

Here are some traces involving γ_5:

- $\text{Tr}\gamma_5 = 0$;
- $\text{Tr}(\gamma_5\gamma^\mu\gamma^\nu) = 0$;

- $\mathrm{Tr}(\gamma_5\gamma^\alpha\gamma^\beta\gamma^\gamma\gamma^\delta) = -4i\varepsilon^{\alpha\beta\gamma\delta}$
  Observe that interchanging two matrices in the trace above yields a minus sign, furthermore the trace vanishes if two indices are identical. Hence the trace is proportional to the ε-tensor:

$$a\varepsilon^{\alpha\beta\gamma\delta} = \mathrm{Tr}(\gamma_5\gamma^\alpha\gamma^\beta\gamma^\gamma\gamma^\delta).$$

  Multiplying both sides by $\varepsilon_{\alpha\beta\gamma\delta}$ yields

$$\begin{aligned} -24a &= \mathrm{Tr}(\gamma_5\gamma^\alpha\gamma^\beta\gamma^\gamma\gamma^\delta)\varepsilon_{\alpha\beta\gamma\delta} \\ &= 24i\mathrm{Tr}(\gamma_5\gamma_5 = \mathbb{1}) \\ \Rightarrow a &= -4i. \end{aligned}$$

11.3 The $V - A$ amplitude

The correct linear combination of bilinears replacing the vector $\bar{\psi}\gamma^\mu\psi$ in Eq. (11.1) in order to achieve parity violation turns out to be the "vector minus axialvector", or $V - A$, combination $\bar{\psi}\gamma^\mu\psi - \bar{\psi}\gamma^\mu\gamma_5\psi$.[2]

Adjusting the amplitude in Eq. (11.1) accordingly yields for the β^- decay amplitude

$$\mathcal{M}(n \to pe^-\bar{\nu}_e) = \frac{G_F}{\sqrt{2}}[\bar{u}_p\gamma^\mu(\mathbb{1} - \gamma_5)u_n][\bar{u}_e\gamma_\mu(\mathbb{1} - \gamma_5)u_{\nu_e}] \tag{11.3}$$

and analogously for the muon decay

$$\mathcal{M}(\mu^- \to \nu_\mu e^-\bar{\nu}_e) = \frac{G_F}{\sqrt{2}}[\bar{u}_{\nu_\mu}\gamma^\mu(\mathbb{1} - \gamma_5)u_\mu][\bar{u}_e\gamma_\mu(\mathbb{1} - \gamma_5)u_{\nu_e}]. \tag{11.4}$$

Let us analyze the general form and properties of $V - A$ amplitudes. Their structure is that of a current-current interaction:

$$\mathcal{M} = \frac{4}{\sqrt{2}}G_F J_i^\mu J_{j,\mu}^\dagger \tag{11.5}$$

where

$$J_i^\mu = \bar{u}_{i^0}\gamma^\mu\frac{1}{2}(\mathbb{1} - \gamma_5)u_{i^-} \tag{11.6}$$

$$J_{j,\mu}^\dagger = \bar{u}_{j^-}\gamma_\mu\frac{1}{2}(\mathbb{1} - \gamma_5)u_{j^0}. \tag{11.7}$$

Note the following properties of this kind of amplitudes:

[2] An axialvector is a pseudovector, since the prefix "pseudo" is used for cases where an extra minus sign arises under the parity transformation (in contrast to the non-pseudo case).

1. $\gamma^\mu(\mathbb{1}-\gamma_5)$ selects left-handed fermions,

$$\gamma_5 u_L = \gamma_5 P_L u = \gamma_5 \frac{1}{2}(\mathbb{1}-\gamma_5) = -\frac{1}{2}(\mathbb{1}-\gamma_5)u = -u_L,$$

   and right-handed antifermions, as desired.

2. G_F is universal.

3. Parity and charge conjugation alter the outcome of experiments, but here CP is conserved:

$$\begin{array}{lll}
\Gamma\left(\pi^+ \to \mu_R^+ + \nu_L\right) \neq \Gamma\left(\pi^+ \to \mu_L^+ + \nu_R\right) & \not{P} & ✗ \\
\Gamma\left(\pi^+ \to \mu_R^+ + \nu_L\right) \neq \Gamma\left(\pi^- \to \mu_R^- + \bar{\nu}_L\right) & \not{C} & ✗ \\
\Gamma\left(\pi^+ \to \mu_R^+ + \nu_L\right) = \Gamma\left(\pi^- \to \mu_L^- + \bar{\nu}_R\right) & CP & ✓.
\end{array}$$

11.4 Muon decay – determination of G_F

Consider the decay

$$\mu^-(p) \to e^-(p') + \bar{\nu}_e(k') + \nu_\mu(k),$$

see Fig. 11.2. The amplitude is given by

$$\mathcal{M} = \frac{G_F}{\sqrt{2}}[\bar{u}(k)\gamma^\mu(\mathbb{1}-\gamma_5)u(p)][\bar{u}(p')\gamma_\mu(\mathbb{1}-\gamma_5)v(k')].$$

Recall that the differential decay rate reads

$$d\Gamma = \frac{1}{2E_\mu}|\mathcal{M}|^2(2\pi)^4 dR_3(p',k,k')$$

where

$$dR_3(p',k,k') = \frac{d^3p'}{(2\pi)^3 2E_{p'}}\frac{d^3k}{(2\pi)^3 2E_k}\frac{d^3k'}{(2\pi)^3 2E_{k'}}\delta^{(4)}(p-p'-k-k').$$

For $m_\nu = m_e = 0$ this yields

$$\frac{d\Gamma}{dE_{p'}} = \frac{m_\mu G_F^2}{2\pi^3} m_\mu^2 E_{p'}^2 \left(3 - \frac{4E_{p'}}{m_\mu}\right)$$

and

$$\Gamma = \int_0^{m_\mu/2} dE_{p'} \frac{d\Gamma}{dE_{p'}} = \frac{G_F^2 m_\mu^5}{192\pi^3} = \frac{1}{\tau}.$$

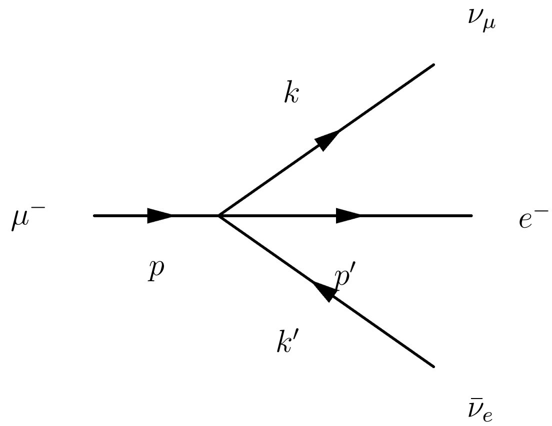


Figure 11.2: *Leptonic muon decay.*

The measured muon lifetime is

$$\tau = 2.1970 \cdot 10^{-6}\,\mathrm{s} = 2.9960 \cdot 10^{-10}\,\mathrm{eV};$$

assuming a muon mass of

$$m_\mu = 105.658 \cdot 10^6\,\mathrm{eV},$$

this yields

$$G_F = 1.166 \cdot 10^{-5}\,\mathrm{GeV}^{-2} \simeq \frac{1}{(300\,\mathrm{GeV})^2}$$

which is a dimensionful ($[G_F] = m^{-2}$) quantity. This hints to the fact that there are some problems with Fermi's theory:

1. It deals with massless fermions only.

2. It is not renormalizable. This problem, along with the dimensionful coupling, is typical for an effective theory, a low energy approximation of a more general theory, in this case the GWS theory.

3. It violates unitarity at high energies. E. g. one finds that the cross section for electron-neutrino scattering is divergent for $E_{\mathrm{CM}} \to \infty$:

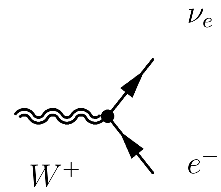
$$\sigma^{e^- + \nu_e \to e^- \nu_e} = \frac{4G_F^2}{\pi} E_{\mathrm{CM}}^2.$$

   One can show that the optical theorem yields the following unitarity constraint for the S-wave: $G_F^2 s^2 \lesssim 1$. Thus Fermi's theory is a good approximation only for $\sqrt{s} \lesssim 1/\sqrt{G_F}$ and it breaks down for higher energies.

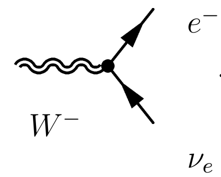
11.5 Weak isospin and hypercharge

From the earlier analysis, we consider the currents of the weak interaction as charged currents [3],

$$j_\mu = j_\mu^+ = \bar{u}_\nu \gamma_\mu \frac{1}{2}(1-\gamma_5)u_e = \bar{\nu}\gamma_\mu \frac{1}{2}(1-\gamma_5)e = \bar{\nu}_L \gamma_\mu e_L =$$



$$j_\mu^\dagger = j_\mu^- = \bar{u}_e \gamma_\mu \frac{1}{2}(1-\gamma_5)u_\nu = \bar{e}\gamma_\mu \frac{1}{2}(1-\gamma_5)\nu = \bar{e}_L \gamma_\mu \nu_L =$$



.

These currents correspond to transitions between pairs of fermions whose charge differs by one unit. For this reason, one speaks of **charged currents** (CC). These two currents are the ones associated with (weak) decays of muons and neutrons.

In analogy to the case of isospin, where the proton and neutron are considered as the two isospin eigenstates of the nucleon, we postulate a **weak isopin** doublet structure ($T = \frac{1}{2}$),

$$\chi_L = \begin{pmatrix} \nu \\ e \end{pmatrix}_L \qquad \begin{matrix} T_3 = +\frac{1}{2} \\ T_3 = -\frac{1}{2} \end{matrix} \;, \tag{11.8}$$

with raising and lowering operators,

$$\tau_\pm = \frac{1}{2}(\tau_1 \pm i\tau_2),$$

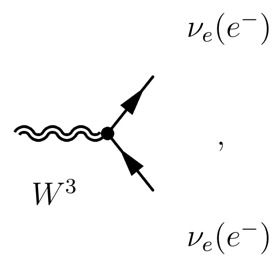
where the τ_i are the usual Pauli matrices. With this formalism, one can write the charged currents as,

$$j_\mu^+ = \bar{\chi}_L \gamma_\mu \tau_+ \chi_L \tag{11.9}$$

$$j_\mu^- = \bar{\chi}_L \gamma_\mu \tau_- \chi_L \tag{11.10}$$

The next step consists in *postulating* an $SU(2)$ symmetry of these currents. In the case of isospin, this leads to the prediction of three currents mediated by the pions $\pi^\pm, \pi^0$. We thus expect a third current to exist, which does not change the charge and is thus called **neutral current** (NC),

$$j_\mu^3 = \bar{\chi}_L \gamma_\mu \frac{1}{2}\tau_3 \chi_L = \bar{\nu}_L \gamma_\mu \frac{1}{2}\nu_L - \bar{e}_L \gamma_\mu \frac{1}{2} e_L =$$



, (11.11)

[3] The 'plus' + and 'dagger' † shall not be confused.

yielding a weak isospin triplet of weak currents,

$$j^i_\mu = \bar{\chi}_L \gamma_\mu \frac{1}{2} \tau_i \chi_L \qquad\qquad i = 1, 2, 3,$$

with algebra,

$$[\tau_i, \tau_j] = i\varepsilon_{ijk}\tau_k.$$

Now, we compare these to the electromagnetic current,

$$j^{em}_\mu = \bar{e}\gamma_\mu Q e = \bar{e}_R \gamma_\mu Q e_R + \bar{e}_L \gamma_\mu Q e_L, \tag{11.12}$$

where Q is the electromagnetic charge operator. This current is invariant under $U(1)_Q$, the gauge group of QED associated to the electromagnetic charge. It is however not invariant under the $SU(2)_L$ which we postulated for the weak currents : it contains e_L instead of χ_L.

To solve this issue, we construct an $SU(2)_L$-invariant $U(1)$-current,

$$j^Y_\mu = \bar{e}_R \gamma_\mu Y_R e_R + \bar{\chi}_L \gamma_\mu Y_L \chi_L, \tag{11.13}$$

where the **hypercharges** Y_R and Y_L are the conserved charge operators associated to the $U(1)_Y$ symmetry. It is different for left and right handed leptons.

We now want to write j^{em}_μ as a linear combination of j^3_μ and $\frac{1}{2} j^Y_\mu$ (the factor $\frac{1}{2}$ is a matter of convention). One gets,

$$\bar{e}_R \gamma_\mu Q e_R + \bar{e}_L \gamma_\mu Q e_L = \bar{\nu}_L \gamma_\mu \frac{1}{2} \nu_L - \bar{e}_L \gamma_\mu \frac{1}{2} e_L + \frac{1}{2} \bar{e}_R \gamma_\mu Y_R e_R + \frac{1}{2} \bar{\chi}_L \gamma_\mu Y_L \chi_L,$$

from which we read out,

$$Y_R = 2Q \qquad Y_L = 2Q + 1. \tag{11.14}$$

with the weak isospin third components,

$$\begin{array}{rl} T_3(e_R) = 0 & \text{singlet, blind to the weak interaction} \\ \left.\begin{array}{r} T_3(\nu_L) = +\frac{1}{2} \\ T_3(e_L) = -\frac{1}{2} \end{array}\right\} & \text{doublet,} \end{array}$$

one can then write the relation,

$$\boxed{Y = 2Q - 2T_3}. \tag{11.15}$$

In Tab. 11.1 and 11.2, we summarise the quantum numbers of leptons and quarks. It should be noted that the right handed neutrino ν_R does not carry $SU(2)_L$ or $U(1)_Y$ charges, and thus decouples from the electroweak interaction.

	T	T_3	Q	Y
ν_L	$1/2$	$1/2$	0	-1
e^-_L	$1/2$	$-1/2$	-1	-1
ν_R	0	0	0	0
e^-_R	0	0	-1	-2

Table 11.1: Weak quantum numbers of leptons

	T	T_3	Q	Y
u_L	$1/2$	$1/2$	$2/3$	$1/3$
d_L	$1/2$	$-1/2$	$-1/3$	$1/3$
u_R	0	0	$2/3$	$4/3$
d_R	0	0	$-1/3$	$-2/3$

Table 11.2: Weak quantum numbers of quarks

11.6 Construction of the electroweak interaction

As in the case of QED (Sec. 5.12, p.100) and QCD (Sec. 7.4, p. 140), we expect the electroweak interaction to be mediated by gauge fields. In the case of QED, we had,

$$\mathcal{L}_{int}^{\text{QED}} = -iej_\mu^{em} A^\mu,$$

where e is the $(U(1)_Q$-)coupling, j_μ^{em} the $(U(1)_Q$-)current, and A^μ the $(U(1)_Q$-)gauge field (photon). We copy this for the current triplets and singlet :

$$\mathcal{L}_{int}^{\text{EW}} = -igj_\mu^i W^{i\mu} - i\frac{g'}{2} j_\mu^Y B^\mu, \tag{11.16}$$

where we introduced the $SU(2)_L$-gauge field triplet $W^{i\mu}$ and singlet B^μ associated to the weak isospin and weak hypercharge respectively.

From those we can construct the massive charged vector bosons,

$$W^{\pm\mu} = \frac{1}{\sqrt{2}}(W^{1\mu} \mp iW^{2\mu}),$$

as well as the neutral vector bosons (mass eigenstates) as a linear combination of $W^{3\mu}$ and B^μ,

$$\begin{aligned} A^\mu &= B^\mu \cos\theta_w + W^{3\mu}\sin\theta_w && \text{massless} \ \to \gamma, \\ Z^\mu &= -B^\mu \sin\theta_w + W^{3\mu}\cos\theta_w && \text{massive} \ \to Z^0, \end{aligned}$$

where θ_w is called the **weak mixing angle** (or sometimes Weinberg angle).

Substituting these quantities in the interaction Lagrangian of the neutral electroweak current, we obtain,

$$\begin{aligned}-igj_\mu^3 W^{3\mu} - i\frac{g'}{2}j_\mu^Y B^\mu = &- i\left(g\sin\theta_w j_\mu^3 + g'\cos\theta_w \frac{j_\mu^Y}{2}\right)A^\mu \\ &- i\left(g\cos\theta_w j_\mu^3 - g'\sin\theta_w \frac{j_\mu^Y}{2}\right)Z^\mu.\end{aligned}$$

The first term corresponds to the electromagnetic current, for which we had $j_\mu^{em} = j_\mu^3 + \frac{1}{2}j_\mu^Y$, implying,

$$\boxed{g\sin\theta_w = g'\cos\theta_w = e}, \tag{11.17}$$

and thus linking the three couplings together. One often uses e and $\sin\theta_w$ as parameters for the standard model to be measured experimentally.

The second term corresponds to the weak neutral current. From $j_\mu^Y = 2(j_\mu^{em} - j_\mu^3)$, we get,

$$j_\mu^{\mathrm{NC}} = \frac{g}{\cos\theta_w}(j_\mu^3 - \sin^2\theta_w j_\mu^{em}). \tag{11.18}$$

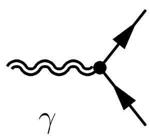
11.7 Electroweak Feynman rules

Vertices The Feynman rules for vertices stemming from,

$$\mathcal{L}_{int}^{\mathrm{EW}} = \mathcal{L}_{int}^{\mathrm{QED}} + \mathcal{L}_{int}^{\mathrm{CC}} + \mathcal{L}_{int}^{\mathrm{NC}},$$

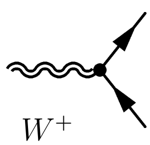
can be computed as follows,

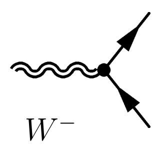
$$i\mathcal{L}_{int}^{\mathrm{QED}} = -ie\bar{\psi}_f\gamma_\mu Q\psi_f A^\mu$$

$\Rightarrow$ 

$$= -ieQ_f\gamma_\mu$$

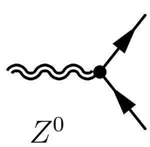
$$\begin{aligned}
i\mathcal{L}_{int}^{\mathrm{CC}} &= -i\frac{g}{\sqrt{2}}(\bar{\chi}_L\gamma_\mu\tau_+\chi_L)W^{+\mu} - i\frac{g}{\sqrt{2}}(\bar{\chi}_L\gamma_\mu\tau_-\chi_L)W^{-\mu} \\
&= -i\frac{g}{\sqrt{2}}\bar{\nu}\gamma_\mu\left(\frac{1-\gamma_5}{2}\right)eW^{+\mu} - i\frac{g}{\sqrt{2}}\bar{e}\gamma_\mu\left(\frac{1-\gamma_5}{2}\right)\nu W^{-\mu}
\end{aligned}$$

$\Rightarrow$ 

$=$ 

$$= -i\frac{g}{\sqrt{2}}\gamma_\mu\left(\frac{1-\gamma_5}{2}\right)$$

$$\begin{aligned}
i\mathcal{L}_{int}^{\mathrm{NC}} &= -i\frac{g}{\cos\theta_w}\bar{\psi}_f\gamma_\mu\left[\left(\frac{1-\gamma_5}{2}\right)T_3 - \sin^2\theta_w Q\right]\psi_f Z^\mu \\
&= -i\frac{g}{\cos\theta_w}\bar{\psi}_f\gamma_\mu\frac{1}{2}(c_V^f - c_A^f\gamma_5)\psi_f Z^\mu
\end{aligned}$$

$\Rightarrow$ 

$$= -i\frac{g}{\cos\theta_w}\gamma_\mu\frac{1}{2}(c_V^f - c_A^f\gamma_5)$$

where c_V^f and c_A^f are the vector and axial vector couplings of the fermion type f. A simple calculation yields,

$$c_V^f = T_3^f - 2\sin^2\theta_w Q^f \tag{11.19}$$

$$c_A^f = T_3^f. \tag{11.20}$$

Tab. 11.3 lists the couplings for the various types of fermions.

	Q^f	c_V^f	c_A^f
ν	0	$1/2$	$1/2$
e	-1	$-1/2 + 2\sin^2\theta_w$	$-1/2$
u	$2/3$	$1/2 - 4/3\sin^2\theta_w$	$1/2$
d	$-1/3$	$-1/2 + 2/3\sin^2\theta_w$	$-1/2$

Table 11.3: Vector and axial vector couplings of fermions.

Propagator of a massive vector boson Form Eq. (11.17), we see that e and g should be of the same order of magnitude (since we know experimentally that $\sin^2\theta_w \approx 0.23$). This leads to the question : why is the weak interaction so much weaker than the electromagnetic one? This can be made evident by looking at the typical lifetime of weakly decaying particles (as the neutron or the muon) compared with electromagnetic decays. The answer lies in the large mass of the weak gauge bosons $W^\pm$ and Z^0.

The components $X^\mu = W^{+\mu}, W^{-\mu}, Z^\mu$ fulfill the Klein-Gordon equation,

$$(\Box + M^2)X^\mu = 0, \qquad \partial_\mu X^\mu = 0 \text{ (gauge fixing)},$$

which results in the propagator,

$$i\frac{\sum_\lambda (\varepsilon^\mu_\lambda)^* \varepsilon^\nu_\lambda}{p^2 - M^2}.$$

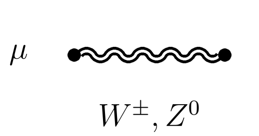
The polarisation sum $\Pi^{\mu\nu}$ must take the form,

$$\Pi^{\mu\nu} = \sum_\lambda (\varepsilon^\mu_\lambda)^* \varepsilon^\nu_\lambda = A g^{\mu\nu} + B p^\mu p^\nu.$$

Using the identities,

$$p_\mu p^\mu = M^2, \qquad p_\mu \Pi^{\mu\nu} = p_\nu \Pi^{\mu\nu} = 0, \qquad g_{\mu\nu}\Pi^{\mu\nu} = 3,$$

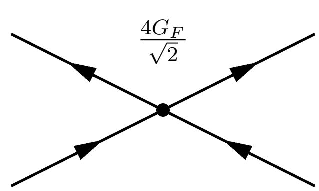
coming from the on-shell condition, the conservation of current and the count of polarization states (for a massive particle) respectively, we get $A = -1$ and $B = M^{-2}$, making us able to write,



$$= i\frac{-g^{\mu\nu} + p^\mu p^\nu / M^2}{p^2 - M^2}.$$

So unless momentum transfer is not of the order of $M \gtrsim 100\,\text{GeV}$, the propagator gets suppressed drastically by the mass.

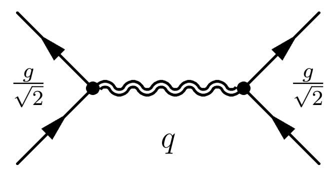
Relation of the Fermi $V-A$-interaction In $V-A$-theory, we have a 4 point vertex,



yielding the matrix element,

$$\mathcal{M}^{V-A} = \frac{4G_F}{\sqrt{2}} j^\mu j^\dagger_\mu.$$

The same process, viewed as the exchange of a low momentum ($q^2 \ll M_W^2$) vector boson,



corresponds to the matrix element,

$$\mathcal{M}^{\mathrm{EW}} \approx \left(\frac{g}{\sqrt{2}} j^\mu\right) \frac{1}{M_W^2} \left(\frac{g}{\sqrt{2}} j^\dagger_\mu\right),$$

yielding the relation,

$$\boxed{G_F = \frac{\sqrt{2} g^2}{8 M_W^2}}. \tag{11.21}$$

From this relation, the first estimates of the mass of the $W^\pm$ bosons were $50 - 100\,\mathrm{GeV}$.

11.8 Spontaneous symmetry breaking: Higgs mechanism

The ad hoc introduction of non-vanishing vector boson masses runs into a serious problem: One would have to include into the Lagrangian the usual mass term

$$\mathcal{L}_M = -\frac{m^2}{2} A_\mu A^\mu \tag{11.22}$$

which violates gauge invariance (the boson field transforms as $A_\mu \to A_\mu - \partial_\mu \alpha(x)$). If the "massive vector bosons" are indeed to be massive, gauge symmetry needs to be broken in some way, since the inclusion of a mass term requires breaking of gauge symmetry. To avoid problems at the theory level caused by broken gauge symmetry, the idea is to retain gauge symmetry in this respect, while physical states are less symmetric than the Lagrangian. This situation can e. g. also be found in solid state physics: Consider a ferromagnet modeled as a collection of spins. As long as no magnetization is imposed, this system is rotationally invariant. A non-vanishing magnetization breaks this symmetry, in that it singles out a specific direction. Symmetry breaking occurs due to the influence of changing a continuos parameter (magnetization) caused by the environment. This does not affect the rotational invariance of the theory describing the ferromagnet and two physical states with different imposed directions are related by a transformation corresponding to the symmetry that is broken by imposing directions.

Let us start out with an example: Consider a real scalar field with a four-point interaction (which is to the complex scalar field what is the Ising model to the isotropic ferromagnet mentioned above):

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \phi)^2 - \left(\frac{1}{2}\mu^2\phi^2 + \frac{1}{4}\lambda\phi^4\right) \tag{11.23}$$

$$= T - V \tag{11.24}$$

where $-1/2\mu^2\phi^2$ is a mass term and $-1/4\lambda\phi^4$ is an interaction term corresponding to the four-point vertex. Because the potential needs to be bounded from below, $\lambda > 0$. Observe that $\mathcal{L}$ is even in ϕ and therefore invariant under the transformation $\phi \to -\phi$.

The vacuum state of this theory corresponds to a minimum of the potential:

$$\frac{\partial V}{\partial \phi} = \phi(\mu^2 + \lambda\phi^2) \stackrel{!}{=} 0. \tag{11.25}$$

Depending on the sign of μ^2, one can distinguish two cases.

a) $\mu^2 > 0,\ \lambda > 0$.
   In this case the vacuum state is reached for $\phi = 0$, see Fig. 11.3(a).

b) $\mu^2 < 0,\ \lambda > 0$.
   Here, $\phi = 0$ is still an extremum, but has turned into a local maximum. In addition there are two minima at

$$\phi = \pm\sqrt{\frac{-\mu^2}{\lambda}} = \pm v$$

   which correspond to two vacua, degenerate in energy, see Fig. 11.3(b). In this case, the symmetry transformation $\phi \to -\phi$, which leaves the Lagrangian in Eq. (11.23) invariant, changes two distinct physical states into each other.

A perturbative calculation is an expansion around the vacuum sate. If we consider case b), this means $\phi = v$ or $\phi = -v$. Therefore, the symmetry $\phi \to -\phi$ is broken, although the Lagrangian has this symmetry irrespective of the signs of μ^2 and λ. Let us choose the positive sign vacuum state and expand:

$$\phi(x) = v + \eta(x) \tag{11.26}$$

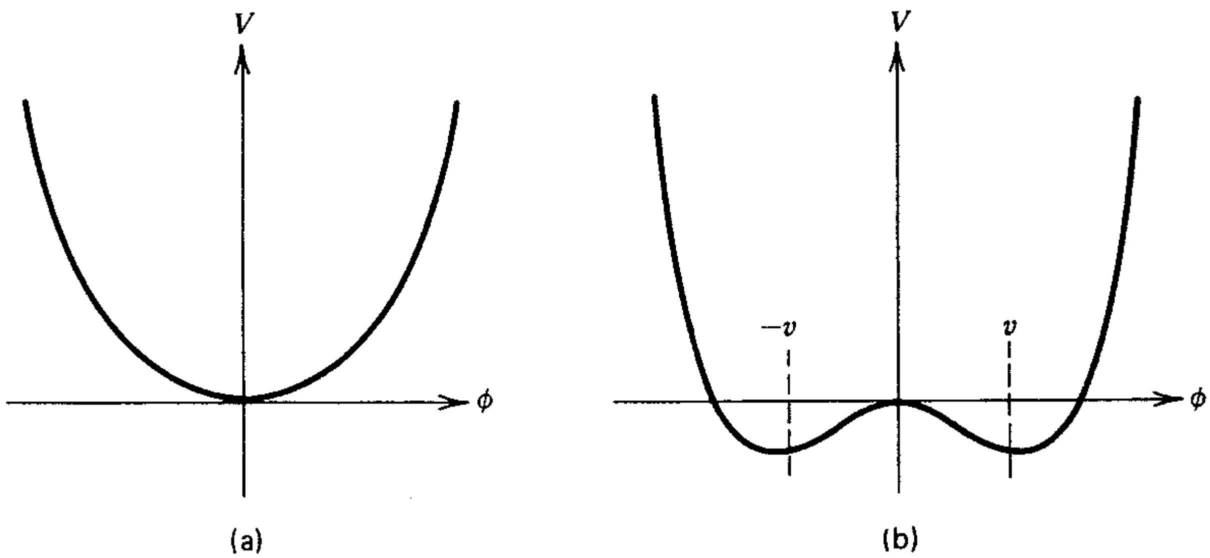


Figure 11.3: *The Potential $V(\phi)$ for (a) $\mu^2 > 0$ and (b) $\mu^2 < 0$ and $\lambda > 0$.* Source: [1, p. 322].

where $\eta(x)$ is some perturbation around v. Inserting this expansion into the Lagrangian yields

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \eta)^2 - \lambda v^2 \eta^2 - \lambda v \eta^3 - \frac{1}{4}\lambda \eta^4 + \text{const.} \tag{11.27}$$

Here, the first term is a kinetic term for η with mass $m_\eta = \sqrt{2\lambda v^2} = \sqrt{-2\mu^2}$ and the second an third terms are the thee-pint and four-point interaction terms, respectively.

Two other examples for spontaneous symmetry breaking are

- The alignment of spins in a ferromagnet which violates rotational invariance and
- The bending of an elastic bar under a force aligned with its symmetry axis, see Fig. 11.4.

These examples share the following feature: Variation of some continuous parameter is associated with a transition between two phases with differing degree of symmetry.

Above we considered a discrete symmetry of the Lagrangian; we now turn to the spontaneous breaking of a continuos symmetry, namely of *global* gauge symmetry. Consider now a complex scalar field:

$$\phi = \frac{1}{\sqrt{2}}(\phi_1 + i\phi_2) \tag{11.28}$$

$$\mathcal{L} = (\partial_\mu \phi)^*(\partial^\mu \phi) - \mu^2 \phi^* \phi - \lambda(\phi^* \phi)^2. \tag{11.29}$$

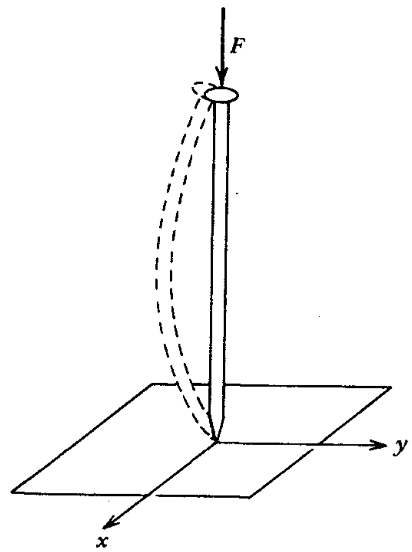


Figure 11.4: *Bending of an elastic bar.* Source: [1, p. 324].

The Lagrangian is invariant under global $U(1)$ transformations $\phi \to e^{i\alpha}\phi$. In the case $\lambda > 0,\ \mu^2 < 0$ the minimum of the potential $V(\phi)$ is a circle in the $\phi_1,\ \phi_2$ plane with

$$\phi_1^2 + \phi_2^2 = v^2 = -\frac{\mu^2}{\lambda}, \tag{11.30}$$

see Fig. 11.5. Out of the infinitely many distinct vacua, degenerate in energy, we choose $\phi_1 = v,\ \phi_2 = 0$. Again, we can expand around the ground state, this time in two orthogonal directions: $\eta(x)$ denotes the perturbation in the steepest ascent direction and $\xi(x)$ is the perturbation in the orthogonal direction (potential valley, see Fig. 11.5):

$$\phi(x) = \frac{1}{\sqrt{2}}\left[v + \eta(x) + i\xi(x)\right]. \tag{11.31}$$

Inserting this expansion into the Lagrangian in Eq. (11.29) yields

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \xi)^2 + \frac{1}{2}(\partial_\mu \eta)^2 + \mu^2\xi^2 + \text{const} + \mathcal{O}\left((\eta,\xi)^3\right) \tag{11.32}$$

where we identify a mass term $-1/2 m_\eta^2 \eta^2$ with $m_\eta = -2\mu^2$ while for the ξ field there is only a kinetic and no mass term.[4] This is because η is an excitation along the potential direction while ξ corresponds to a rotation along the circle of vacua. Here, the process of spontaneous symmetry breaking leads from a more symmetric phase with two massive fields to a less symmetric phase with a massive and a massless field.

[4]This massless scalar is a Goldstone boson. The Goldstone theorem says that for every broken continuous symmetry there is a massless boson.

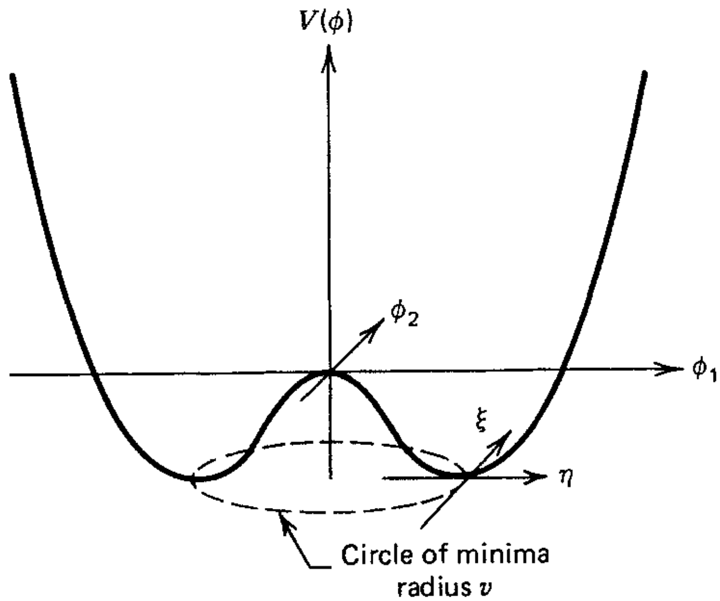


Figure 11.5: *The potential $V(\phi)$ for a complex scalar field for the case $\mu^2 < 0$ and $\lambda > 0$.* Source: [1, p. 325].

Let us now turn to the spontaneous breaking of *local* gauge symmetry. Consider a complex scalar field and local $U(1)$ gauge transformations:

$$\phi \to \phi' = \phi e^{ie\alpha(x)}. \tag{11.33}$$

Gauge invariance of the Lagrangian requires the covariant derivative

$$D_\mu = \partial_\mu + ieA_\mu \tag{11.34}$$

with the massless $U(1)$ gauge field A_μ transforming as

$$A_\mu \to A'_\mu = A_\mu - \partial_\mu \alpha(x). \tag{11.35}$$

A gauge invariant Lagrangian reads

$$\mathcal{L} = (\partial^\mu - ieA^\mu)\phi^*(\partial_\mu + ieA_\mu)\phi - \mu^2\phi^*\phi - \lambda(\phi^*\phi)^2 - \frac{1}{4}F^{\mu\nu}F_{\mu\nu}. \tag{11.36}$$

As before, we consider the case $\mu^2 < 0,\ \lambda > 0$; v and the expansion are

$$v^2 = -\frac{\mu^2}{\lambda} \qquad\qquad \phi(x) = \frac{1}{\sqrt{2}}\left[v + h(x)\right] e^{i\frac{\xi(x)}{v}} \tag{11.37}$$

where in this case wee keep the finite rotation due to ξ to preserve gauge freedom. This allows to absorb $\xi(x)$ into a redefinition of the gauge field:

$$A_\mu \to \hat{A}_\mu = A_\mu - \frac{1}{v}\partial_\mu \xi(x). \tag{11.38}$$

Combining expansion and redefinition with the Lagrangian in Eq. (11.36) yields

$$\mathcal{L} = \frac{1}{2}(\partial_\mu h)^2 - \lambda v^2 h^2 + \frac{1}{2}e^2 v^2 \hat{A}_\mu^2 - \lambda v h^3 - \frac{1}{4}\lambda h^4 + \frac{1}{2}e^2 \hat{A}_\mu^2 h^2 + v e^2 \hat{A}_\mu^2 - \frac{1}{4}\hat{F}^{\mu\nu}\hat{F}_{\mu\nu}. \tag{11.39}$$

The particle spectrum of this theory is as follows.

- There is a massive scalar field h (Higgs) of mass $m_h = \sqrt{2\lambda v^2}$.
- The Goldstone field has been absorbed into $\hat{A}_\mu$ and is no longer present in the Lagrangian.
- There is a massive $U(1)$ vector field $\hat{A}_\mu$ of mass $m_A = ev$.

It is important to notice that the vacuum state $\phi = v/\sqrt{2}$ is charged under the gauge interaction.

Finally, let us consider the degrees of freedom for the Lagrangian given in terms of ϕ and A and in terms of h and $\hat{A}$:

$\mathcal{L}$	Fields	d. o. f.
$\mathcal{L}$ in $\phi,\ A$	ϕ complex, scalar	2
	A^μ massless, spin-1 vector	2
$\mathcal{L}$ in $h,\ \hat{A}$	h real, scalar	1
	$\hat{A}^\mu$ massive, spin-1 vector	3

This acquiring of a mass by a spin-1 vector boson is also what happens to the photons belonging external fields in superconductors: Since the propagation of the massive photons is exponentially suppressed, the field is correspondingly excluded (Meißner-Ochsenfeld effect).

11.9 Gauge boson masses in $SU(2)_L \times U(1)_Y$

For constructing a gauge invariant Lagrangian, we define the covariant derivative in $SU(2)_L \times U(1)_Y$:

$$D_\mu = \partial_\mu - ig\frac{1}{2}\vec{\tau} \cdot \overrightarrow{W}_\mu - ig'\frac{1}{2}YB_\mu. \tag{11.40}$$

The corresponding Lagrangian for a complex scalar field reads

$$\mathcal{L} = [iD^\mu\phi]^\dagger[iD_\mu\phi] - \mu^2\phi^\dagger\phi - \lambda[\phi^\dagger\phi]^2 \tag{11.41}$$

where ϕ is an $SU(2)$ doublet (choose to arrange fields such that $Y = 1$):

$$\phi = \frac{1}{\sqrt{2}}\begin{pmatrix}\phi_1 + i\phi_2 \\ \phi_3 + i\phi_4\end{pmatrix} = \begin{pmatrix}\phi^+ \\ \phi^0\end{pmatrix}. \tag{11.42}$$

This is also called a Higgs doublet.

Again let us consider the case $\mu^2 < 0$ and $\lambda > 0$. We may choose the following vacuum state: $\phi_1 = \phi_2 = \phi_4 = 0$ and $\phi_3 = v$ and expand, which, up to a phase, yields

$$v^2 = -\frac{\mu^2}{\lambda} \qquad\qquad \phi = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ v + h(x) \end{pmatrix}. \tag{11.43}$$

This choice of vacuum breaks the $SU(2)_L$ and $U(1)_Y$ gauge symmetries, since it is hypercharged. The $U(1)_Q$ symmetry of electromagnetism, though, is conserved, because $Q\phi = (T_3 + Y/2)\phi = 0$ and the photon remains massless. What is the particle spectrum for this theory, given the vacuum expectation value chosen above? Inserting $\phi_0 = 1/\sqrt{2}(0, v)^T$ into the relevant term of the Lagrangian in Eq. (11.41), $[D^\mu\phi]^\dagger[D_\mu\phi]$, gives the answer:

$$\begin{aligned}
\left|\left(-i\frac{g}{2}\vec{\tau}\cdot\vec{W}_\mu - i\frac{g'}{2}B_\mu\right)\phi\right|^2 &= \frac{1}{8}\left|\begin{pmatrix} gW^3_\mu + g'B_\mu & g(W^1_\mu - iW^2_\mu) \\ g(W^1_\mu + iW^2_\mu) & -gW^3_\mu + g'B_\mu \end{pmatrix}\begin{pmatrix} 0 \\ v \end{pmatrix}\right|^2 \\
&= \frac{1}{8}v^2 g^2\left|(W^1)^2 + (W^2_\mu)^2\right| + \frac{1}{8}v^2(g'B_\mu - gW^3_\mu)(g'B^\mu - gW^{3\mu}) \\
&= \left(\frac{1}{2}vg\right)^2 W^+_\mu W^{-\mu} + \frac{1}{8}v^2(g'B_\mu - gW^3_\mu)^2
\end{aligned}$$

which, using $Z_\mu = (gW^3_\mu - g'B_\mu)/\sqrt{g^2 + g'^2}$,

$$= M_W^2 W^+_\mu W^{-\mu} + \frac{1}{2}M_Z^2 Z_\mu Z^\mu$$

where

$$\boxed{M_W = \frac{1}{2}vg} \qquad\qquad \boxed{M_Z = \frac{1}{2}v\sqrt{g^2 + g'^2}}. \tag{11.44}$$

Using $g'/g = \tan\theta_w$ yields the following relation between the W and the Z mass:

$$\boxed{\frac{M_W}{M_Z} = \cos\theta_w}. \tag{11.45}$$

Finally, knowing the W mass, we can use Fermi's constant to obtain an estimate for the vacuum expectation value v:

$$G_F = \frac{\sqrt{2}g^2}{8M_W^2} = \frac{1}{\sqrt{2}v^2} \rightarrow \boxed{v = 246\,\mathrm{GeV}}.$$

11.10 Fermion masses

The usual mass term for quarks and leptons (we focus on the $T_3 = -\frac{1}{2}$ fermions, i.e. down quarks and electrons) takes the form,

$$\mathcal{L}_{m-} = -m\bar{\psi}\psi = -m\left(\bar{\psi}_R\psi_L + \bar{\psi}_L\psi_R\right),$$

where ψ_L is a component of the $SU(2)_L$-doublet χ_L, and ψ_R is an $SU(2)_L$-singlet. Because of its form, this mass term cannot be invariant under the action of the gauge group $SU(2)_L$ (ψ_R transforms trivially, whereas ψ_L necessarily changes).

The solution consists in pairing ψ_L with an adjoint doublet, the **Higgs doublet**, that we have already introduced earlier to give masses to the vector bosons by means of spontaneous symmetry breaking. A gauge invariant mass term is obtained by coupling to the Higgs doublet, e.g. for the electron (also valid for all $T_3 = -\frac{1}{2}$ fermions):

$$\begin{aligned}\mathcal{L}_{m-} &= -G^e\left[\begin{pmatrix}\bar{\nu}_e & \bar{e}\end{pmatrix}_L\begin{pmatrix}\phi^+\\ \phi^0\end{pmatrix}e_R + \bar{e}_R\begin{pmatrix}\bar{\phi}^+ & \bar{\phi}^0\end{pmatrix}\begin{pmatrix}\nu_e\\ e\end{pmatrix}_L\right]\\ &= -\frac{G^e v}{\sqrt{2}}\left(\bar{e}_L e_R + \bar{e}_R e_L\right) - \frac{G^e}{\sqrt{2}}h\left(\bar{e}_L e_R + \bar{e}_R e_L\right),\end{aligned} \tag{11.46}$$

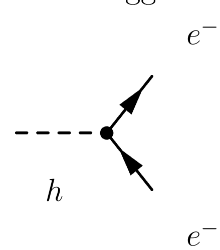
where G^e denotes the **Yukawa coupling** of the electron, and we used,

$$\begin{pmatrix}\phi^+\\ \phi^0\end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix}0\\ v + h(x)\end{pmatrix}.$$

We can now read out of Eq. (11.46),

$$\boxed{m_e = \frac{G^e v}{\sqrt{2}}}, \tag{11.47}$$

and the coupling of the electron to the Higgs field,



$$= -i\frac{m_e}{v}.$$

Since $m_e = 511\,\text{keV}$ and $v = 246\,\text{GeV}$, this vertex factor is very small for the electron. In the case of the top, $m_t = 172\,\text{GeV}$ and the vertex factor is much bigger. In the event the Higgs mass is big enough ($m_h > 2m_t$), thus kinematically allowing this decay mode, the branching ratio,

$$BR(h \to t\bar{t}) = \frac{\Gamma(h \to t\bar{t})}{\Gamma(h \to \text{anything})},$$

would be significant.

The vacuum is charged under both $SU(2)_L$ and $U(1)_Y$ but not electrically. Because of this, the photon stays massless, even after $SU(2)_L \times U(1)_Y$ has been broken. Therefore the vacuum expectation value (VEV) of the Higgs fields concentrates on the *neutral* component of the doublet, i.e. the second component having $T_3 = -\frac{1}{2}$ (otherwise the vacuum would also be charged electrically, giving a mass to the photon). Up to now, we have been able to give a gauge invariant mass term to the charged leptons and d-type quarks (d, s, b) all having $T_3 = -\frac{1}{2}$. It appears that we are not able to give a mass term to the neutrinos (neutral leptons) and u-type quarks (u, c, t) having $T_3 = +\frac{1}{2}$ without introducing another Higgs doublet [5].

In the case of $SU(2)$ (but not in general), we are allowed to use at this end the charge conjugate of the Higgs doublet,

$$\phi^c = i\tau_2\phi^\dagger = \begin{pmatrix} \bar{\phi}^0 \\ -\bar{\phi}^+ \end{pmatrix} \to \frac{1}{2}\begin{pmatrix} v + h(x) \\ 0 \end{pmatrix}, \tag{11.48}$$

which has $Y = -1$, because ϕ and ϕ^c are equivalent, i.e. can be connected by a unitary transformation.

Example For quarks we get,

$$\begin{aligned}
\mathcal{L}_{m-} + \mathcal{L}_{m+} = &- G^d \left[\begin{pmatrix} \bar{u} & \bar{d} \end{pmatrix}_L \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} d_R + \bar{d}_R \begin{pmatrix} \bar{\phi}^+ & \bar{\phi}^0 \end{pmatrix} \begin{pmatrix} u \\ d \end{pmatrix}_L \right] \\
&- G^u \left[\begin{pmatrix} \bar{u} & \bar{d} \end{pmatrix}_L \begin{pmatrix} \phi^0 \\ -\phi^+ \end{pmatrix} u_R + \bar{u}_R \begin{pmatrix} \bar{\phi}^0 & -\bar{\phi}^+ \end{pmatrix} \begin{pmatrix} u \\ d \end{pmatrix}_L \right] \\
= &- m_d\bar{d}d - \frac{m_d}{v}h\bar{d}d - m_u\bar{u}u - \frac{m_u}{v}h\bar{u}u.
\end{aligned} \tag{11.49}$$

We conculde by emphasising that all fermion masses are generated in a gauge invariant way through coupling of the field to the Higgs VEV v. The coupling of each fermion to the Higgs boson h is proportional to the mass of the particle. The origin of mass is reduced to a Yukawa coupling of the different fermions to the Higgs field.

11.11 Lagrangian of the electroweak standard model

The theory of the electroweak interaction was formulated between 1961 and 1967 by Sheldon Lee Glashow, Abdus Salam and Steven Weinberg. All three received the Physics Nobel Prize in 1979 although the $W^\pm$ and Z^0 had not yet been observed directely. Deep inelastic scattering of spin-polarized electrons off nuclei gave evidence for a minute parity

[5] This is the case in extensions of the standard model, e.g. for the minimal supersymmetric standard model (MSSM), where we have a Higgs doublet for each value of T_3.

violating interaction (all interactions except the weak interaction conserve parity). The first evidence for neutral currents (mediated by the Z^0 boson) were found in 1973 in the bubble chamber Gargamelle at CERN. Direct observation of both $W^\pm$ and Z^0 was achieved in 1983 by the experiments UA1 also at CERN, leading to the Physics Nobel Prize of 1984 for Carlo Rubbia and Simon van der Meer.

The Lagrangian of the electroweak theory can be decomposed as,

$$\mathcal{L}^{\mathrm{EW}} = \mathcal{L}_{gauge} + \mathcal{L}_{matter} + \mathcal{L}_{Higgs} + \mathcal{L}_{Yukawa},$$

with,

$$\begin{aligned}
\mathcal{L}_{gauge} &= -\frac{1}{4}\overrightarrow{W}_{\mu\nu}\cdot\overrightarrow{W}^{\mu\nu} - \frac{1}{4}B_{\mu\nu}B^{\mu\nu} && (11.50)\\
W^i_{\mu\nu} &= \partial_\mu W^i_\nu - \partial_\nu W^i_\mu - ig\varepsilon^{ijk}W^j_\mu W^k_\nu \\
B_{\mu\nu} &= \partial_\mu B_\nu - \partial_\nu B_\mu, \\
\mathcal{L}_{matter} &= \sum_L \bar{L}\gamma^\mu\left(i\partial_\mu + g\frac{1}{2}\vec{\tau}\cdot\overrightarrow{W}_\mu + g'\frac{Y}{2}B_\mu\right)L + \sum_R \bar{R}\gamma^\mu\left(i\partial_\mu + g'\frac{Y}{2}B_\mu\right)R, && (11.51)\\
\mathcal{L}_{Higgs} &= \left|\left(i\partial_\mu + g\frac{1}{2}\vec{\tau}\cdot\overrightarrow{W}_\mu + g'\frac{Y}{2}B_\mu\right)\phi\right|^2 - V(\phi), && (11.52)\\
V(\phi) &= -\mu^2\phi^\dagger\phi + \lambda(\phi^\dagger\phi)^2 \\
\mathcal{L}_{Yukawa} &= -\sum_{f-} G^f_-(\bar{L}\phi R + \bar{R}\bar{\phi}L) - \sum_{f+} G^f_+(\bar{L}\phi^c R + \bar{R}\bar{\phi}^c L), && (11.53)
\end{aligned}$$

where L denotes a left-handed fermion doublet, R a right-handed fermion singlet, $G^f_\pm$ the fermion Yukawa coupling for $T_3 = \pm\frac{1}{2}$. All terms in $\mathcal{L}^{\mathrm{EW}}$ are invariant under $SU(2)_L$ and $U(1)_Y$ gauge transformations.

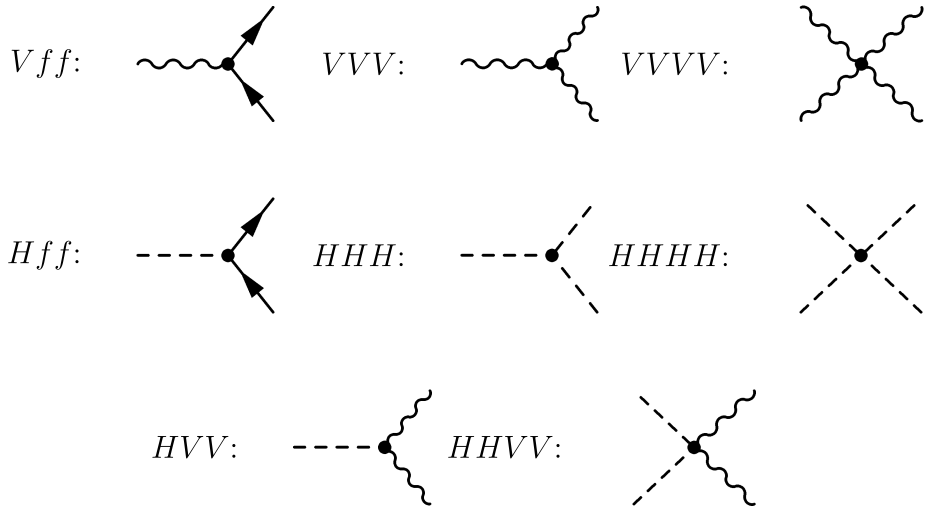
After the spontaneous symmetry breaking, we have,

$$\phi(x) = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ v + h(x) \end{pmatrix},$$

yielding the masses through the Higgs mechanism:

$$\begin{aligned}
M_W &= 2gv = 80.4\,[\mathrm{GeV}] && (11.54)\\
M_Z &= \frac{M_W}{\cos\theta_w} = 91.19\,[\mathrm{GeV}] && (11.55)\\
M_f &= \frac{G^f v}{\sqrt{2}} \qquad m_e = 511\,[\mathrm{keV}], \ldots, m_t = 172\,[\mathrm{GeV}] && (11.56)\\
M_h &= v\sqrt{2\lambda} > 114\,[\mathrm{GeV}] \qquad (\mathrm{LEP}) && (11.57)
\end{aligned}$$

We now classify the vertices of the electroweak Lagrangian (V : vector boson, f : fermion, H : Higgs boson):



Care must be taken in choosing the fields as for example photon can interact with W bosons because they carry an electric charge, but not with the Z boson. All diagrams not involving a Higgs bosons have been observed experimentally so far.

11.12 Properties of the Higgs boson

The decay width of the Higgs boson $\Gamma = \frac{1}{\tau}$ for a two particle final state is (see Eq. (3.15), p. 24),

$$\Gamma_H = \frac{1}{2M_H}\frac{1}{(2\pi)^2}\sum_f \int \frac{d^3p_1}{2E_1}\frac{d^3p_2}{2E_2}\delta^{(4)}(p_f - p_H)|\mathcal{M}_{fH}|^2,$$

where f denotes the final state : $b\bar{b}, t\bar{t}, W^+W^-, Z^0Z^0, \tau^+\tau^-, \ldots$ and $m_1 = m_2 = m_f$.

$|\mathcal{M}_{fH}|^2$ cannot depend on individual components of p_1 or p_2, and we can hence factorize the phase space,

$$R_2 = \int \frac{d^3p_1}{2E_1}\frac{d^3p_2}{2E_2}\delta^{(4)}(p_f - p_H) = \frac{\pi}{2M_H^2}\sqrt{\lambda\left(M_H^2, m_f^2, m_f^2\right)} = \frac{\pi}{2}\sqrt{1 - \frac{4m_f^2}{M_H^2}},$$

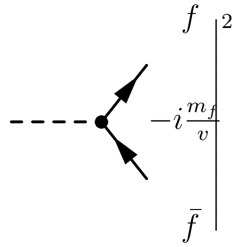
and hence,

$$\Gamma_H = \frac{1}{16\pi M_H}\sum_f \sqrt{1 - \frac{4m_f^2}{M_H^2}}|\mathcal{M}_{fH}|^2 = \sum_f \Gamma_{H\to f}. \tag{11.58}$$

We now look at the different final states separately :

Decay into fermions Leptons :

$$|\mathcal{M}_{lH}|^2 = \sum_{s,f} \left| \quad \right|^2$$



$$= \frac{m_f^2}{v^2} \mathrm{Tr}\left((\not{p}_f + m_f)(\not{p}_{\bar{f}} - m_f) \right) = \frac{4m_f^2}{v^2}\left(\frac{M_H^2}{2} - 2m_f^2\right),$$

where s denotes the spin and f the flavour.

Quarks :

$$|\mathcal{M}_{qH}|^2 = \sum_c |\mathcal{M}_{lH}|^2 = 3|\mathcal{M}_{lH}|^2,$$

where c denotes the color.

Plugging these into Eq. (11.58), we get the partial widths,

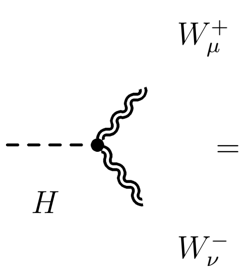
$$\Gamma_{H\to l^+l^-} = \frac{1}{8\pi^2 v^2} m_f^2 M_H \left(1 - \frac{4m_f^2}{M_H^2}\right)^{\frac{3}{2}} \tag{11.59}$$

$$\Gamma_{H\to q\bar{q}} = \frac{3}{8\pi^2 v^2} m_q^2 M_H \left(1 - \frac{4m_q^2}{M_H^2}\right)^{\frac{3}{2}}. \tag{11.60}$$

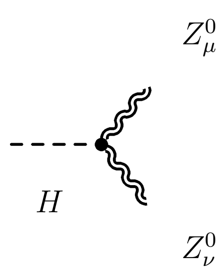
We remark at this point that the dominant decay mode (corresponding to the largest partial width) is always into the heaviest kinematically allowed fermion. In the case of a light Higgs boson ($M_H < 2M_{W,Z}$), the dominant channels would be into $b\bar{b}$ and $\tau^+\tau^-$.

The partial width for a decay into fermions is proportional to the mass of the Higgs boson, so there is no upper limit to M_H.

Decay into gauge bosons The relevant vertices are,

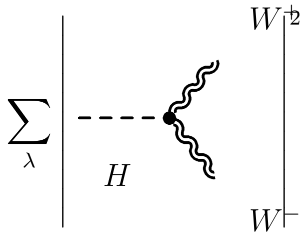


$$= igM_W g_{\mu\nu},$$



$$= \frac{igM_Z}{\cos\theta_w} g_{\mu\nu},$$

Summing the moduli squared over the polarizations, we get,



$$= g^2 M_W^2 \left(-g^{\mu\rho} + \frac{p_1^\mu p_2^\rho}{M_W^2}\right)\left(-g_{\mu\rho} + \frac{p_{1\mu} p_{2\rho}}{M_W^2}\right)$$

$$= \frac{g^2 M_H^4}{4M_W^2}\left(1 - 4\frac{M_W^2}{M_H^2} + 12\frac{M_W^4}{M_H^4}\right),$$

and an analogous result for the decay $H \to Z^0 Z^0$. The partial widths are then, respectively,

$$\Gamma_{H\to W^+W^-} = \frac{1}{16\pi v^2} M_H^3 \left(1 - \frac{4M_W^2}{M_H^2}\right)^{\frac{1}{2}} \left(1 - 4\frac{M_W^2}{M_H^2} + 12\frac{M_W^2}{M_H^4}\right) \tag{11.61}$$

$$\Gamma_{H\to Z^0Z^0} = \frac{1}{32\pi v^2} M_H^3 \left(1 - \frac{4M_Z^2}{M_H^2}\right)^{\frac{1}{2}} \left(1 - 4\frac{M_Z^2}{M_H^2} + 12\frac{M_Z^2}{M_H^4}\right), \tag{11.62}$$

where the factor $\frac{1}{2}$ in the second line is a symmetry factor for identical bosons.

In the case of a decay into gauge bosons, the partial width is proportional to the third power of the Higgs mass. This implies that for a heavy Higgs boson ($M_H > 2M_{W,Z}$), the decay into gauge bosons will be dominant over the decay into fermions, the only competing fermionic decay being $H \to t\bar{t}$ (for $M_H \approx 2m_t$). Fig. 10.42(a) and (b), show the different branching ratios and total width as a function of M_H.

Due to this power dependence, one remarks by plugging the known values of M_W, M_Z and v that if $M_H \approx 1\,\mathrm{TeV}$, $\Gamma_H \approx M_H$ and the interpretation of the Higgs particle as a resonance of the $\mathcal{S}$-matrix is no longer possible. This yields an upper bound for the Higgs mass in the framework of the standard model. A mass of the order of 1 TeV would imply a coupling $\lambda \approx 1$ requiring some non-perturbative approach (as in QCD for $Q \approx \Lambda^{\mathrm{QCD}}$).

11.13 Tests of electroweak theory

In the previous sections the theory of electroweak interactions was discussed, in particular it was shown how *massive* gauge bosons emerge; in this section we discuss experimental tests of the theory, including the consistency of the standard model parameters, the W and Z boson discovery and measurements of the width. We discuss the forward-backward asymmetries, as well as examples of Higgs boson searches. An introduction to the latter topic is given in Sect. 10.9, here we focus on a specific case study, namely searches for heavy Higgs decaying into W boson pairs.

11.13.1 Parameters of the standard model and historical background

A summary[6] of the experimental values of the standard model parameters is shown in Fig. 10.1. The stated deviations are a measure for the consistency of the standard model. As can be seen from the bars, which visualize the deviation of the measured from the best fitting values, assuming the standard model to be correct, in units of measurement standard deviations, the majority of the measured parameters is compatible within 1σ. A notable exception is the variable $A_{\mathrm{fb}}^{0,b}$, an asymmetry measured in the b sector.

Electroweak unification was accomplished theoretically in the sixties by Glashow, Salam and Weinberg. The predictions derived from this theory were consistent with the observed charged current interactions (flavor-changing exchange of $W^{\pm}$ bosons, see e. g. Fig. 1.1(b)). However, as we have seen in Sect. 11.5, the theory also predicts neutral current interactions (via Z^0 exchange and γ/Z^0 interference) which had never been observed up to that time. In fact, until 1973 all observed weak interactions were consistent with the existence of only charged bosons $W^{\pm}$. The first neutral current interaction was observed at CERN in 1973 with the "Gargamelle" experiment in the following reaction:

$$\nu_\mu + \text{nucleus} \to \nu_\mu + p + \pi^- + \pi^0$$

which can be explained by a flavor conserving weak interaction, i. e. a weak neutral current. This discovery made urgent the question of how to observe W and Z bosons directly to test electroweak predictions.

11.13.2 W and Z boson discovery, mass and width measurements

Electroweak theory predicted bosons with masses $M_W \sim 83\,\mathrm{GeV}$ and $M_Z \sim 93\,\mathrm{GeV}$. Therefore, to produce W and Z bosons, a particle collider was needed capable of producing particles with mass $\sim 100\,\mathrm{GeV}$. A the time, two candidates were available at CERN. The ISR with $\sqrt{s} = 61\,\mathrm{GeV}$ was too weak and also the SPS, which consisted of a 400 GeV proton beam against a fixed target, did not provide sufficient center of mass energy (recall that for fixed target experiments $\sqrt{s} = \sqrt{2mE}$, see Sect. 4.1.1).

This problem was solved by the S$p\bar{p}$S machine, designed by Rubbia and van der Meer, a proton-antiproton collider at $\sqrt{s} = 540\,\mathrm{GeV}$. It had a luminosity of $5 \cdot 10^{27}\,\mathrm{cm}^{-2}\mathrm{s}^{-1}$, achieved with three against three bunches with $\sim 10^{11}$ particles per bunch. The first collisions took place in 1981.

LEP, which later on delivered part of the precision data discussed in this chapter was an electron-positron collider, while S$p\bar{p}$S was a hadron collider.[7] Figure 11.6 shows the

[6] `http://lepewwg.web.cern.ch/LEPEWWG/`

[7] A general comparison of these types of colliders can be found in Sect. 10.1.2.

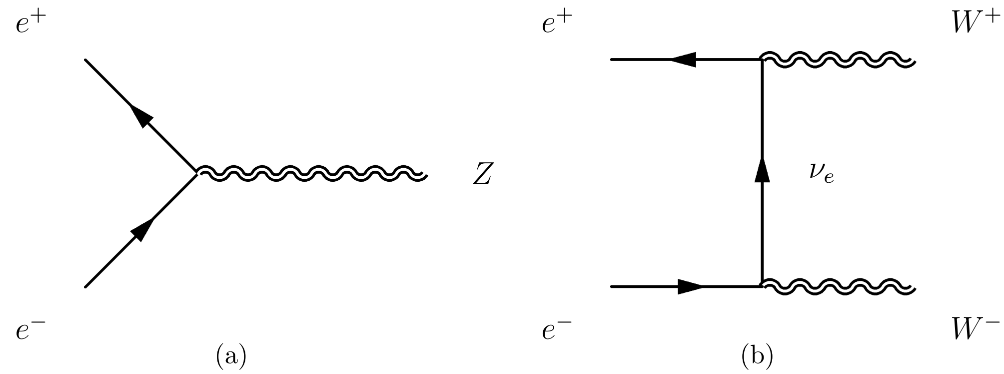


Figure 11.6: *Z (a) and W (b) boson production at electron-positron colliders.*

relevant production diagrams for e^+e^- colliders while Fig. 11.7 shows a hadron collider production diagram along with the dominant background diagram (see also Fig. 10.14). In the electron-positron case, beam energies of about $M_Z/2$ are sufficient to produce Z bosons (see Fig. 11.6(a)), while $W^\pm$ bosons can only be produced in pairs, requiring a higher center of mass energy (see Fig. 11.6(b)). Now compare this to the hadron collider case shown in Fig. 11.7(a): To produce a Z boson, flavor conservation is required such that processes like $u\bar{u} \to Z^0$ and $d\bar{d} \to Z^0$ contribute. The production of $W^\pm$ bosons involves quarks of different flavors, such as $u\bar{d} \to W^+$ and $d\bar{u} \to W^-$. What has been said so far concerns *production* of W and Z bosons, what about their *detection*? Consider first the decay into quark-antiquark pairs: The cross section of "usual" two-jet production, e. g. via gluon exchange (see Fig. 11.7(b)) is much larger than the one of hadronic vector boson decays. In other words, the cross section for W production is small compared to the total cross section:

$$\frac{\sigma(\bar{p}p \to WX \to e\nu X)}{\sigma_T(p\bar{p})} \simeq 10^{-8}.$$

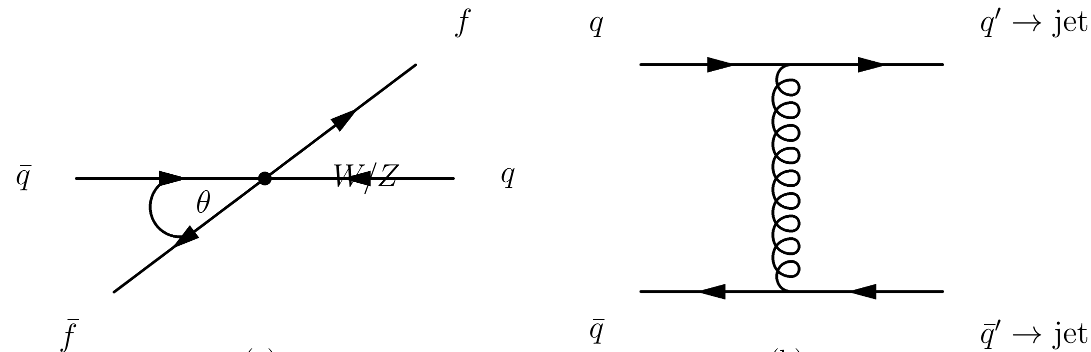
Therefore, it is preferred to look for W and Z decays into leptons, where the background is smaller:[8]

$$\begin{aligned} W^\pm &\to e^\pm \overset{(-)}{\nu}_e,\ \mu^\pm \overset{(-)}{\nu}_\mu,\ \tau^\pm \overset{(-)}{\nu}_\tau \\ Z^0 &\to e^+e^-,\ \mu^+\mu^-, \tau^+\tau^-. \end{aligned}$$

11.13.2.1 *W* discovery and mass measurement

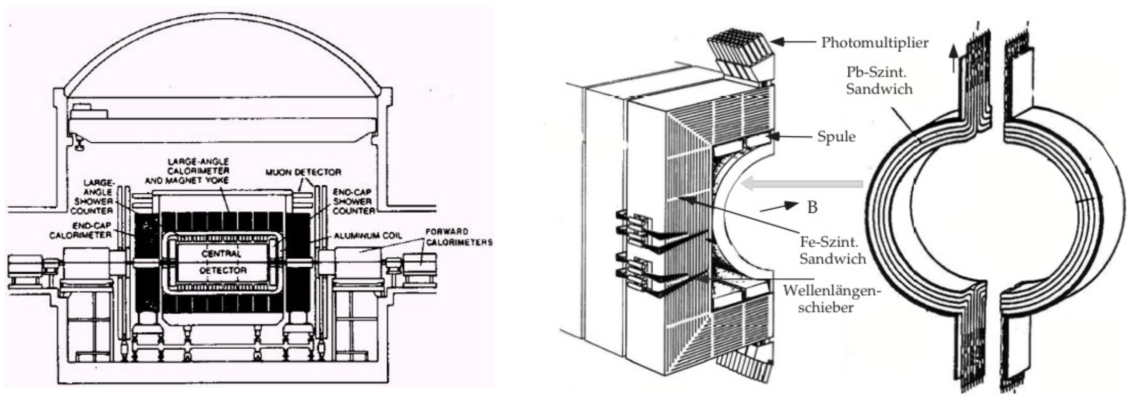
The UA1 experiment at the S$p\bar{p}$S collider was an hermetic particle detector optimized for the $W^\pm \to e^\pm\nu_e/\bar{\nu}_e$ measurement. It featured for the first time the general design principles of collider detectors (see also Sect. 4.3.3): tracking devices inside a magnetic

[8]The Z^0 boson may also decay into neutrino-antineutrino pairs, which makes it possible to determine the number of neutrino families with $m_\nu < M_Z/2$, see below.



(a) (b)

Figure 11.7: *(a) Sketch of the kinematics of W and Z boson production at hadron colliders and diagram of a process leading to two jets (b).*



(a) (b)

Figure 11.8: *UA1 experiment.* A cross section along the beam line, featuring the important components of collider experiment detectors is shown in (a), while (b) shows the electromagnetic and hadronic calorimeters. Source: [8, p. 305].

field, followed by electromagnetic calorimeters, hadron calorimeters and muon chambers (see Fig. 11.8(a)). Since $M_W \sim 80\,\mathrm{GeV}$, the electromagnetic calorimeter resolution is optimized for $40\,\mathrm{GeV}$ electrons to $\pm 500\,\mathrm{MeV}(1\%)$. Because the photomultipliers had to be placed outside the magnetic field of the coil, the hadron calorimeter is sandwiched in the return yoke (see Fig. 11.8(b)): Showering in the lead layers, the particles then produce light in the szintillator layers which is transferred to the photomultipliers via light-guides.

To understand how to search for the W decay in the data, we look at the final-state kinematics. Since the neutrino cannot be detected, there is no direct information on its momentum. However, due to momentum conservation one can write

$$\vec{p}_\perp(\nu) = -\vec{p}_\perp(H) - \vec{p}_\perp(e)$$

where $\vec{p}_\perp(\nu)$ is the neutrino transverse momentum while $\vec{p}_\perp(H)$ and $\vec{p}_\perp(e)$ denote the total hadron transverse momentum and the electron transverse momentum, respectively.

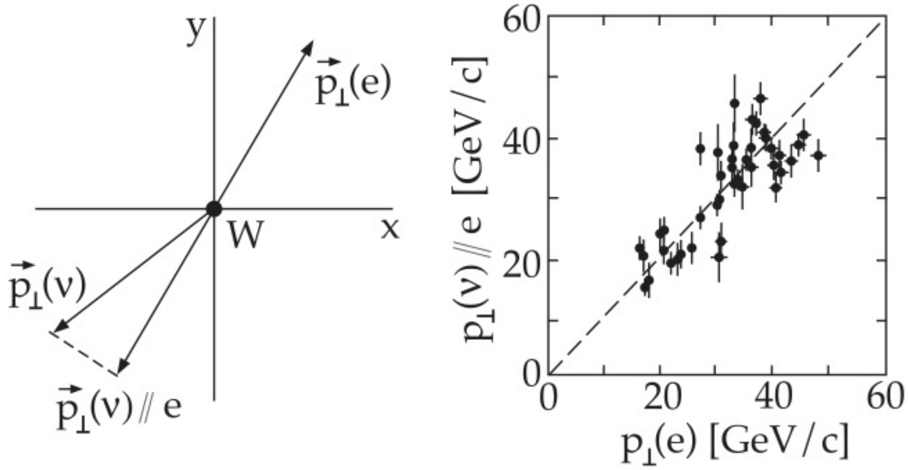


Figure 11.9: *Transverse momenta in a leptonic W decay.* On the LHS one sees a sketch of the electron and neutrino transverse momenta. $\vec{p}_\perp \| e$ is the component of the neutrino transverse momentum parallel to $\vec{p}_\perp(e)$. The correlation between these momenta is shown in the RHS Subfig. Source: [8, p. 305].

Momenta are considered in the transverse plane to avoid leakage along the beam lines. Since the W boson is not always produced at rest and the detector resolution is finite, the neutrino transverse momentum $\vec{p}_\perp(\nu)$ is not exactly anti-parallel to the electron transverse momentum (see Fig. 11.9). Nevertheless, there is still a strong correlation between $\vec{p}_\perp(e)$ and the neutrino transverse momentum projected along the electron transverse momentum $\vec{p}_\perp(\nu) \| e$ (see Fig. 11.9).

We discuss now how to measure the W boson mass using the electron transverse momentum spectrum (see also exercises). Electron emission is assumed to be isotropic ($dN/d\cos\theta = \text{const}$) and detector effects are emulated with Monte Carlo simulation. One can rewrite the spectrum as

$$\frac{dN}{dp_\perp} = \frac{dN}{d\cos\theta}\frac{d\cos\theta}{dp_\perp} = \text{const}\,\frac{d\cos\theta}{dp_\perp}$$

where θ is the electron polar angle. Using the kinematics of Sect. 2.1 and $|\vec{p}_\perp| = |\vec{p}|\sin\theta$, we have

$$p_\perp = \frac{M_W}{2}\sin\theta = \frac{M_W}{2}\sqrt{1-\cos^2\theta},$$

which yields

$$\frac{dp_\perp}{d\cos\theta} = \frac{M_W}{2}\frac{\cos\theta}{\sin\theta} = \frac{M_W}{2}\frac{\sqrt{1-\sin^2\theta}}{\sin\theta} = \left(\frac{M_W}{2}\right)^2\frac{\sqrt{1-\frac{4p_\perp^2}{M_W^2}}}{p_\perp}.$$

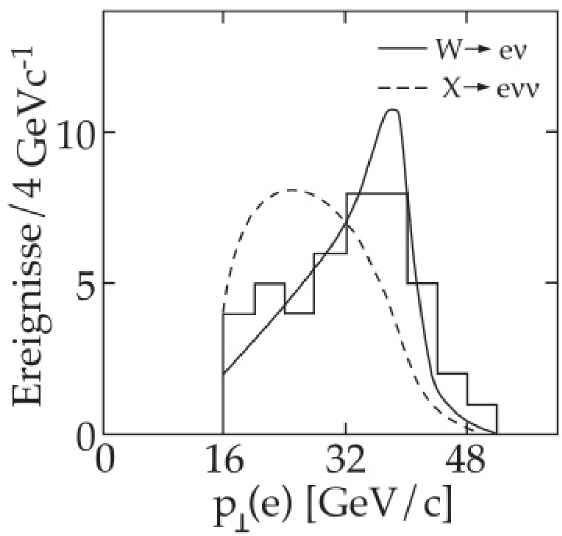


Figure 11.10: *Momentum distribution of the electron perpendicular to the beam (43 events).* The histogram shows the data while the continuous and dashed lines show the Monte Carlo expectation for a two-body decay and three-body decay scenarios, respectively. Source: [8, p. 306].

We thus find

$$\frac{dN}{dp_\perp} \propto \frac{p_\perp}{\sqrt{M_W^2 - 4p_\perp^2}}. \tag{11.63}$$

The denominator vanishes at $M_W = 2p_\perp$, which allows to determine the W boson mass from a measurement of the electron transverse momentum spectrum (see Fig. 11.10).

A summary of experimental results for the W boson mass is shown in Fig. 11.11.

11.13.2.2 W and Z width

Using the kinematics discussed Chap. 3, one can calculate the partial width of the W boson. From Eq. (3.15) we have

$$\Gamma = \frac{1}{2M_W}\frac{1}{(2\pi)^2}\int dR_2|\mathcal{M}_{fi}|^2$$

and Eq. (3.29) reads

$$dR_2 = \frac{1}{8s}\sqrt{\lambda(s, m_e^2, m_\nu^2)}d\Omega.$$

Combining these results yields

$$\frac{d\Gamma}{d\Omega} = \frac{1}{64\pi^2 M_W}|\mathcal{M}_{fi}|^2.$$

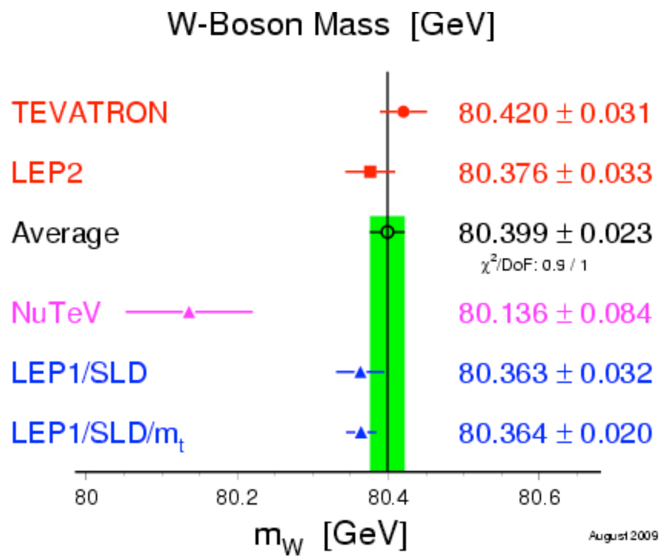


Figure 11.11: *Summary of the current W boson mass measurements.* Source: [63].

Using the following result for the matrix element:

$$|\mathcal{M}_{fi}|^2 = \frac{g^2 M_W^2}{4}(1 - \cos\theta),$$

where θ is the electron polar angle in the center of mass frame, and integrating over θ, one finds for $M_W = 80\,\mathrm{GeV}$

$$\Gamma(W \to e\nu) = \frac{g^2 M_W}{48\pi} = \frac{G_F}{\sqrt{2}}\frac{M_W^3}{6\pi} = 224\,\mathrm{MeV}. \tag{11.64}$$

To obtain the total width (for the W^- case) from the partial widths, we consider the following points:

1. All leptonic decays (e, μ, τ) have the same width.
2. $\bar{u}d$ and $\bar{c}s$ are similar to the leptonic channels ($\cos\theta_c \sim 1$).
3. The other hadronic decays ($\bar{u}s$, $\bar{c}d$, $\bar{u}b$, $\bar{c}b$) with quarks of different families are Cabibbo-suppressed.

Keeping these facts in mind, we have to sum over three lepton currents and two quark currents to find the total width Γ_T. Each quark current can be realized in three colors, therefore:

$$\Gamma_T(W) = 3\,\text{lepton currents} + (3\,\text{colors} \times 2\,\text{quark currents}) \tag{11.65}$$

$$= 9\Gamma(W \to e\nu) = 2.02\,\mathrm{GeV}. \tag{11.66}$$

We now consider the Z boson decay. The Z resonance in the hadronic cross section for e^+e^- annihilation can be used to count the number of neutrino families with $m_\nu < M_Z/2$. One way to accomplish this is to derive a standard model prediction for the Z decay widths as a function of the number of neutrino families N_ν which can be compared to the experimental data.

First we calculate the partial width of the Z boson decaying into neutrino pairs (see also exercises for the explicit calculation). It can be obtained from the W boson case with some substitutions: Using the Feynman rules given in Sect. 11.7, one finds, since $c_V^\nu = c_A^\nu = 1/2$, that substituting

$$g \to \frac{g}{\sqrt{2}\cos\theta_w}, \;\; M_W \to M_Z$$

in the partial W width in Eq. (11.64) does the trick:

$$\Gamma(Z \to \nu\bar{\nu}) = \frac{g^2 M_Z}{96\pi \cos^2\theta_w} = \frac{G_F}{\sqrt{2}}\frac{M_Z^3}{12\pi} = 165\,\mathrm{MeV}, \tag{11.67}$$

assuming $M_Z = 91\,\mathrm{GeV}$. To obtain the total width of the Z boson, one has to sum over all partial widths, originating from all the allowed decays into quarks and leptons. Solving exercise sheet 9[9] we showed that for the general fermionic case the Z partial width is

$$\Gamma(Z \to f\bar{f}) = \frac{g^2}{48\pi\cos^2\theta_w}\sqrt{M_Z^2 - 4m_f^2}\left\{[c_V^f]^2\left(1 + \frac{2m_f^2}{M_Z^2}\right) + [c_A^f]^2\left(1 - \frac{4m_f^2}{M_Z^2}\right)\right\}.$$

Neglecting m_f, one finds that the total Z width is proportional to the sum

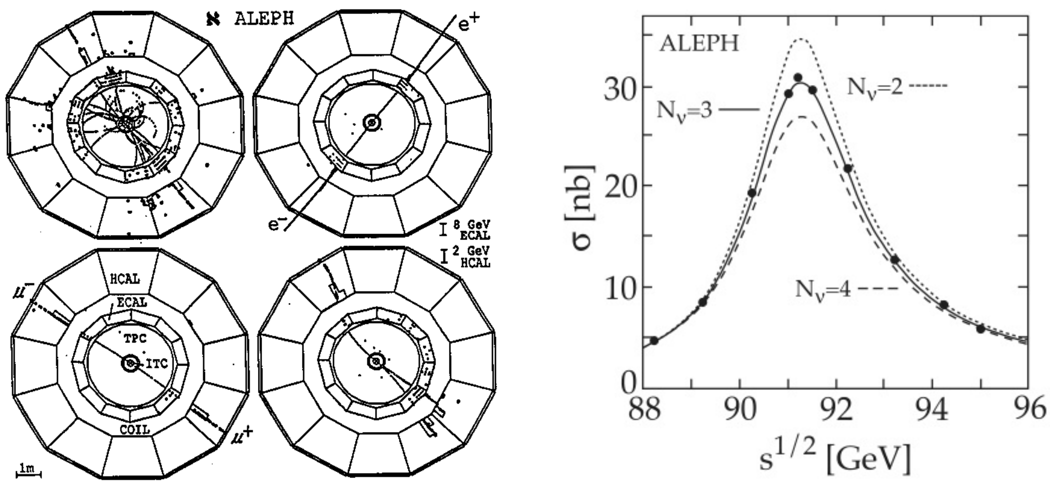
$$\sum_{m_f < M_Z/2}^{\text{fermions}} \left([c_V^f]^2 + [c_A^f]^2\right)$$

which can be calculated using Tab. 11.3. Note that only the following fermionic final states contribute:

- three neutrino pairs: $\nu_e\bar{\nu}_e,\ \nu_\mu\bar{\nu}_\mu,\ \nu_\tau\bar{\nu}_\tau$;
- three other halves of the doublets: $e^+e^-,\ \mu^+\mu^-,\ \tau^+\tau^-$;
- two quark pairs with $T_3 = +1/2$: $u\bar{u},\ c\bar{c}$ and finally
- three quark pairs with $T_3 = -1/2$: $d\bar{d},\ s\bar{s},\ b\bar{b}$.

Assuming $\sin^2\theta_w = 0.23$, the total Z width is

$$\Gamma_T(Z) = \frac{g^2 M_Z}{48\pi\cos^2\theta_w}\sum_{m_f < M_Z/2}^{\text{fermions}}\left([c_V^f]^2 + [c_A^f]^2\right) = 2.41\,\mathrm{GeV}.$$



(a) (b)

Figure 11.12: *ALEPH event displays of Z decays and Z jets cross-section as function of* $\sqrt{s}$*.* Subfigure (a) shows typical events in the ALEPH detector. Starting in the top left corner and proceeding in clockwise order, one has $e^+e^- \to$ hadrons, $e^+e^- \to e^+e^-$, $e^+e^- \to \mu^+\mu^-$ and $e^+e^- \to \tau^+\tau^-$. Source: [61, p. 15]. The Z cross section fit is shown in (b). The dots show the measurement while the expectation from scenarios with different number of neutrino families are shown by the continuous and dashed lines. Source: [8, p. 312].

One can measure the hadronic cross section for e^+e^- annihilation around the Z peak as a function of $\sqrt{s}$ to constrain the number of neutrino families. This is done by a fit to a modified Breit-Wigner distribution,

$$\sigma(s) = \frac{12\pi\Gamma(e^+e^-)\Gamma(f\bar{f})}{M_Z^2} \frac{s}{(s-M_Z^2)^2 + M_Z^2\Gamma_T^2(Z)}, \tag{11.68}$$

for the Z resonance. One also has to take into account γ/Z interference, the $1/s$ dependent QED contribution, and quite substantial corrections due to initial and final state radiation. To measure the relevant cross sections, one selects (e. g. hadronic) events, which is done using on their basic properties, such as number of tracks (see Fig. 11.12(a)). Since the cross section is given by $\sigma = N/(\varepsilon\mathcal{L}_{\text{int}})$, the precision of the result depends on the precision of the integrated luminosity measurement, as well as the trigger and its efficiency. A best fit to the hadronic cross section yields for the number of light neutrino families

$$N_\nu = 2.994 \pm 0.012$$

(see Fig. 11.12(b)). Note that because of the kinematics of $1 \to 2$ decay, this does not exclude heavy ($m_\nu > M_Z/2$) quark and neutrino families.

As we have seen, since the cross section is inversely proportional to the integrated luminosity, the luminosity error propagates into the cross section error. Therefore, it is essential

[9] `http://www.itp.uzh.ch/~pfmonni/PPPII_FS10/sheet9.pdf`

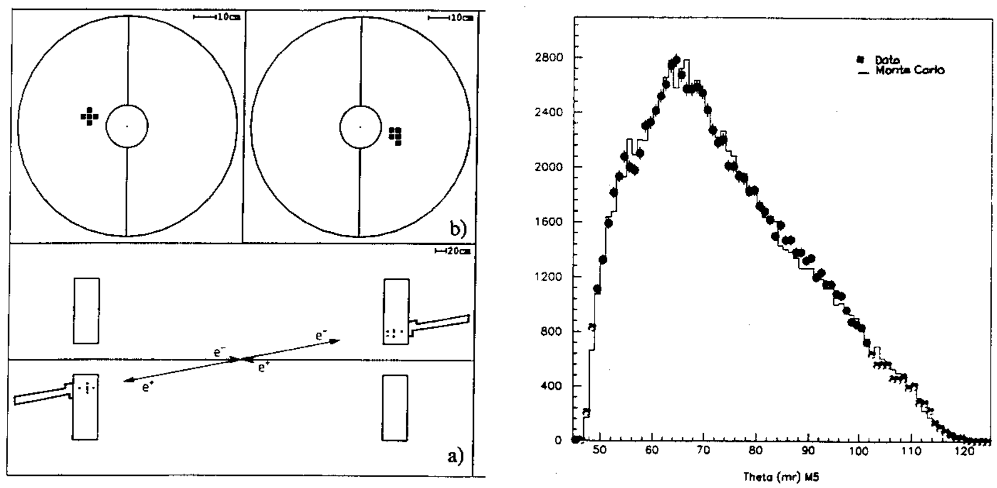


Figure 11.13: *Luminosity measurement in ALEPH using the Bhabha scattering.* On the left a small angle electron-positron scattering event is shown. (a) shows a cut including the beam direction and (b) is a view along the beam of the two luminosity calorimeters. A comparison of measured and simulated polar angle of the scattered electron is shown on the right. Source: [61, p. 20].

to determine the luminosity with high accuracy. This is done by measuring the rate of Bhabha scattering, which can be precisely calculated. As we have seen in Sect. 6.2.4, the corresponding cross section is divergent as the electron polar angle goes to zero (see also Fig. 11.13). This procedure yields a final precision of about 3% for the luminosity measurement.

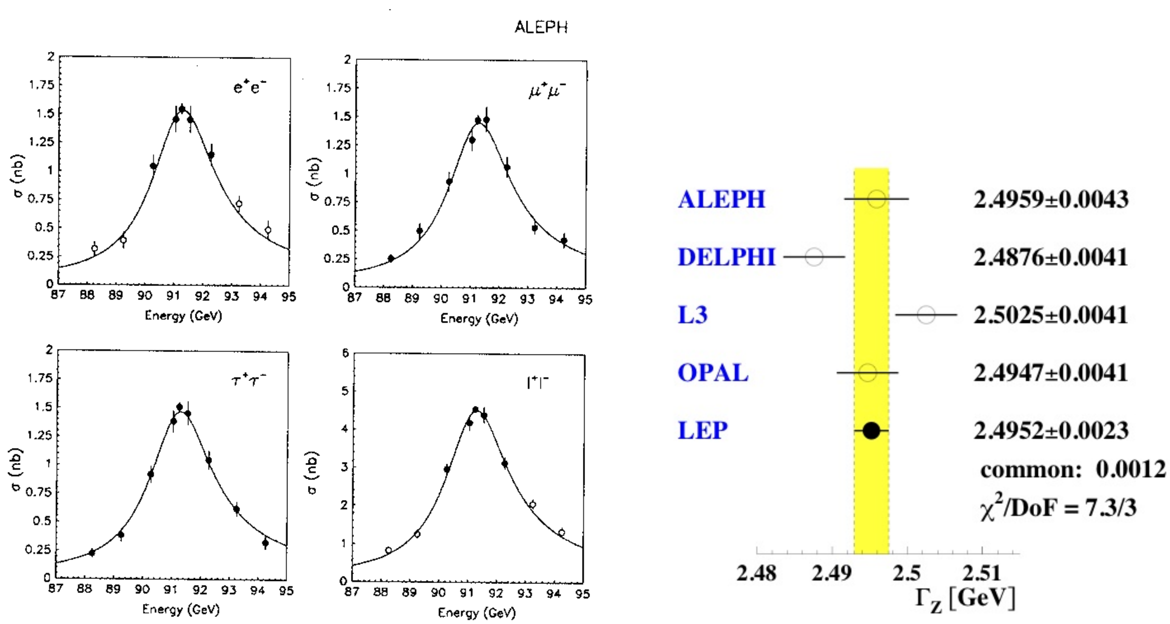
Selecting leptonic events, one can perform the same measurement as the one shown for the hadronic case (see Fig. 11.14(a); note that the cross sections are considerably smaller). This delivers the partial widths $\Gamma(l\bar{l})$ and thus allows for a test of lepton universality. Remembering our discussion of the total Z width, one finds for the leptonic widths (e. g. for muons) the following prediction:

$$\frac{\Gamma(\mu^+\mu^-)}{\Gamma_T} = \frac{[c_V^\mu]^2 + [c_A^\mu]^2}{\sum_{m_f<M_Z/2}^{\text{fermions}} \left([c_V^f]^2 + [c_A^f]^2\right)} = 3.4\%.$$

The corresponding experimental result is

$$\frac{\Gamma(\mu^+\mu^-)}{\Gamma_T} = (3.366 \pm 0.007)\%.$$

A summary of the LEP results for the Z boson width is shown in Fig. 11.14(b). To conclude this section, let us put our discussion into an historic and energetic context: Figure 11.15 shows the cross section for $e^+e^- \to$ hadrons as measured by various experiments at center of mass energies up to 200 GeV. For center of mass energies smaller than about 50 GeV, the



(a) (b)

Figure 11.14: *Cross sections for electron-positron annihilation into leptons around the Z pole measured by ALEPH (a) and LEP summary of the Z width measurements (b).* Source: [61, p. 24].

cross section agrees with the $1/s$ prediction obtained by QED alone (quark mass effects included, see Sect. 8.1). Around 90 GeV the Z resonance is the dominant contribution. The figure shows also the cross section for W production from $e^+e^- \to W^+W^-$.

11.13.3 Forward-backward asymmetries

As we have begun to discuss in Sect. 6.2.5, the weak contributions to electron-positron annihilation cross sections result in forward-backward asymmetries (in the angle between the outgoing fermion and the incident positron), which are not predicted by QED alone (see e. g. Fig. 6.17). Solving exercise sheet 8[10], we showed that the differential cross section for $e^+e^- \to f\bar{f}$, obtained by squaring the sum of the γ and the Z exchange diagram, can be written as

$$\frac{d\sigma_f}{d\Omega} = \frac{\alpha^2 N_c^f}{4s} \left[F_1(s)(1+\cos^2\theta) + 2F_2(s)\cos\theta \right] \tag{11.69}$$

where

$$\begin{aligned} F_1(s) &= Q_f^2 - 2v_e v_f Q_f \mathrm{Re}\chi + (v_e^2 + a_e^2)(v_f^2 + a_f^2)|\chi|^2 \\ F_2(s) &= -2a_e a_f Q_f \mathrm{Re}\chi + 4v_e a_e v_f a_f |\chi|^2 \end{aligned}$$

[10] `http://www-theorie.physik.unizh.ch/~pfmonni/PPPII_FS10/sheet8.pdf`

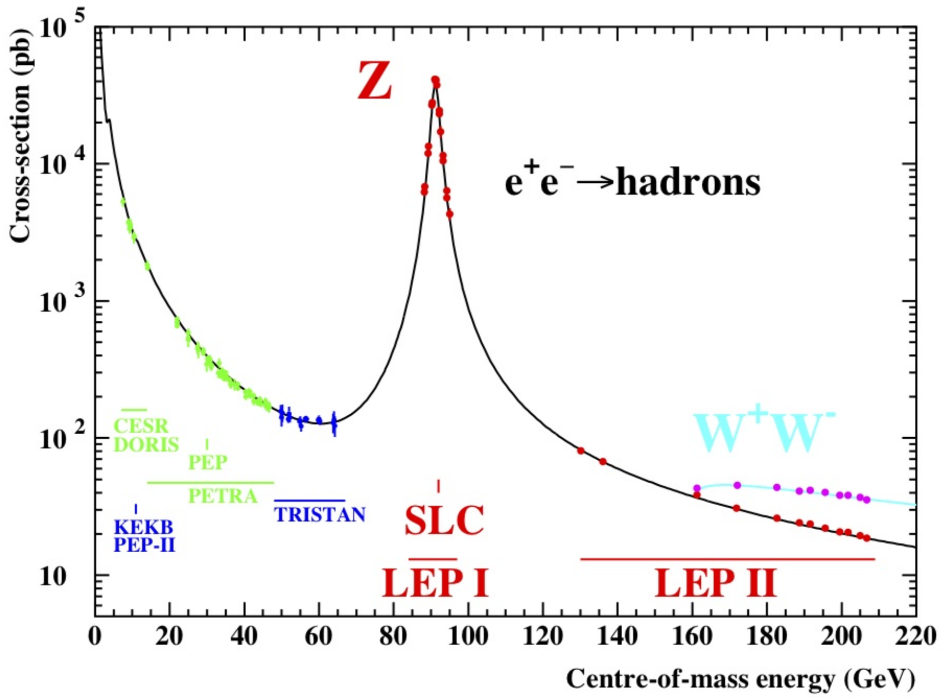


Figure 11.15: *Summary of the* $e^+e^- \to$ *hadrons cross section measurements as a function of the center of mass energy* $\sqrt{s}$.

with

$$\chi = \frac{s}{s - M_z^2 + iM_Z\Gamma_T(Z)}$$

the Breit-Wigner term (compare Eq. (11.68)) and

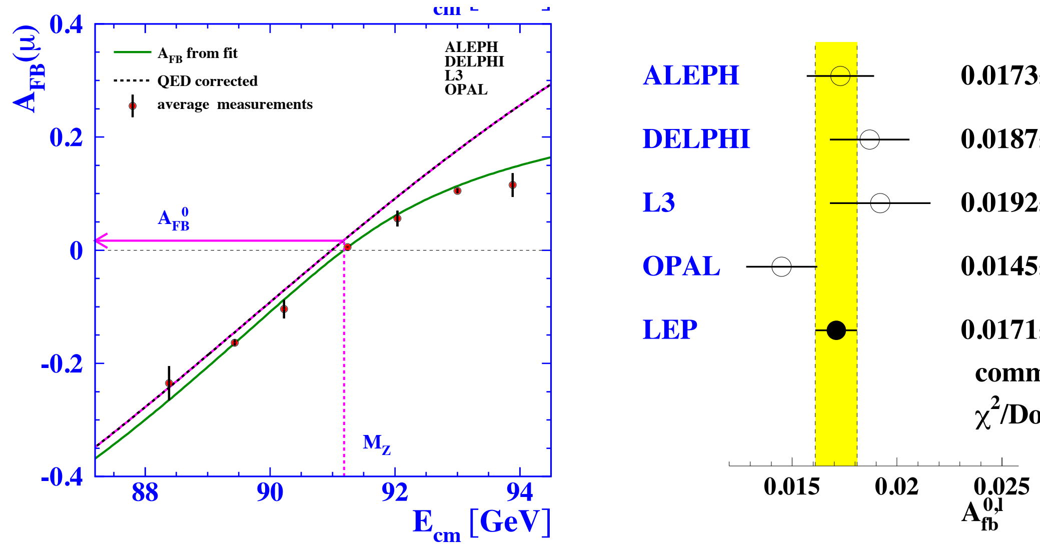
$$v_f \equiv \frac{c_V^f}{2\sin\theta_w\cos\theta_w}$$
$$a_f \equiv \frac{c_A^f}{2\sin\theta_w\cos\theta_w}.$$

To get a quantitative estimate of the forward-backward asymmetry, we define the following quantity

$$A_{\text{FB}} = \frac{\mathcal{I}(0,1) - \mathcal{I}(-1,0)}{\mathcal{I}(0,1) + \mathcal{I}(-1,0)} \tag{11.70}$$

where we have defined the integral $\mathcal{I}(a,b)$ as

$$\mathcal{I}(a,b) \equiv \int_a^b d\cos\theta \frac{d\sigma}{d\cos\theta}. \tag{11.71}$$



(a) (b)

Figure 11.16: *LEP results for forward-backward asymmetry A_{FB}.* (a) shows a plot of the LEP data for $A_{\rm FB}$ as a function of $\sqrt{s}$ and (b) shows a summary of the numerical values at $\sqrt{s} = M_Z$ and the combined result.

Thus forward-backward asymmetry means $A_{\rm FB} \neq 0$. In terms of $F_1,\ F_2$ defined above, we have

$$A_{\rm FB} = \frac{3}{4}\frac{F_2}{F_1} = \frac{3v_e a_e v_f a_f}{(v_e^2 + a_e^2)(v_f^2 + a_f^2)} = 3\frac{(v/a)_e (v/a)_f}{[1 + (v/a)_e^2][1 + (v/a)_f^2]}. \tag{11.72}$$

Therefore, at the Z peak the asymmetry $A_{\rm FB}$ is sensitive to the ratio of vector to axial vector couplings $v/a = c_V^f/c_A^f$. Recalling the definition of c_V^f and c_A^f (see Sect. 11.7), we see that in the electroweak theory the c_V/c_A ratio depends on $\sin^2\theta_w$:

$$c_V/c_A = 1 - 4|Q|\sin^2\theta_w. \tag{11.73}$$

Furthermore, rewriting Eq. (11.69) using Eq. (11.72) yields

$$\frac{d\sigma}{d\cos\theta} \propto 1 + \cos^2\theta + \frac{8}{3}A_{\rm FB}\cos\theta \tag{11.74}$$

(see Fig. 6.17). Figure 11.16(a) shows results for $A_{\rm FB}$ by the four LEP experiments. The corresponding numerical values are shown in Fig. 11.16(b). Combining these results gives

$$A_{\rm FB} = 0.0171 \pm 0.0010$$

for the forward-backward asymmetry at $\sqrt{s} = M_Z$.

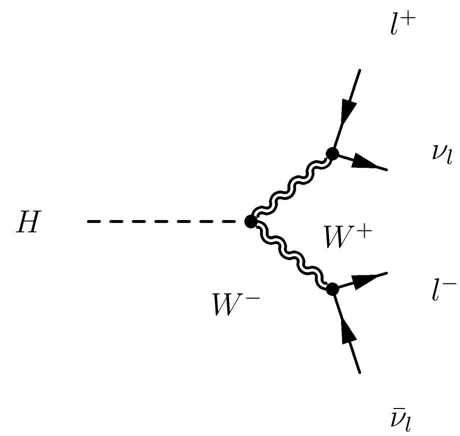


Figure 11.17: *Feynman diagram for the decay of a heavy Higgs into a W^+W^- pair.*

11.13.4 Searches for heavy Higgs decays into W pairs

Having studied extensively the observable consequences of non-vanishing gauge boson masses, we now turn to the source of this phenomenon. In Sect. 11.12 we discussed properties of the Higgs boson, including its partial widths for decay into W and Z boson pairs. Sect. 10.9 introduces the principles of Higgs production and searches; here we focus on searches of heavy Higgs in the the $H \to W^+W^-$ channel.

Recall from Sect. 10.9 that for $m_H \simeq 140-175$ GeV the important Higgs discovery channel is $H \to W^+W^-$, which yields two leptons and missing transverse energy in the final state (see Fig. 11.17).

Figure 11.18 shows the orders of magnitude of various production cross sections at Tevatron. Note the difference of about ten orders of magnitude between the production cross sections for heavy flavors and Higgs bosons. In addition, also the production cross sections for $Z/\gamma^\star$ and standard model W^+W^- pair production not involving Higgs boson exchange are orders of magnitude larger than the Higgs production cross section.

How does one select events in the desired final states? To reduce the background as much as possible, the following cuts are applied:

- *Total missing energy larger than* 20 GeV.
  This requirement reduces the $Z/\gamma^\star \to$ leptons background.

- *Invariant mass of two leptons larger than* 15 GeV.
  This requirement reduces the background from semi-leptonic decays of heavy quarks.

The remaining background is due to standard model W pair production not involving Higgs bosons (see Fig. 11.19). Therefore, the remaining task is to reject this kind of electroweak background obscuring the $H \to W^+W^-$ signal. To achieve this aim, one can exploit the fact that the standard model Higgs is a scalar (i. e. it has spin 0). W bosons, on the other hand, have spin 1. To conserve angular momentum, the two decay leptons

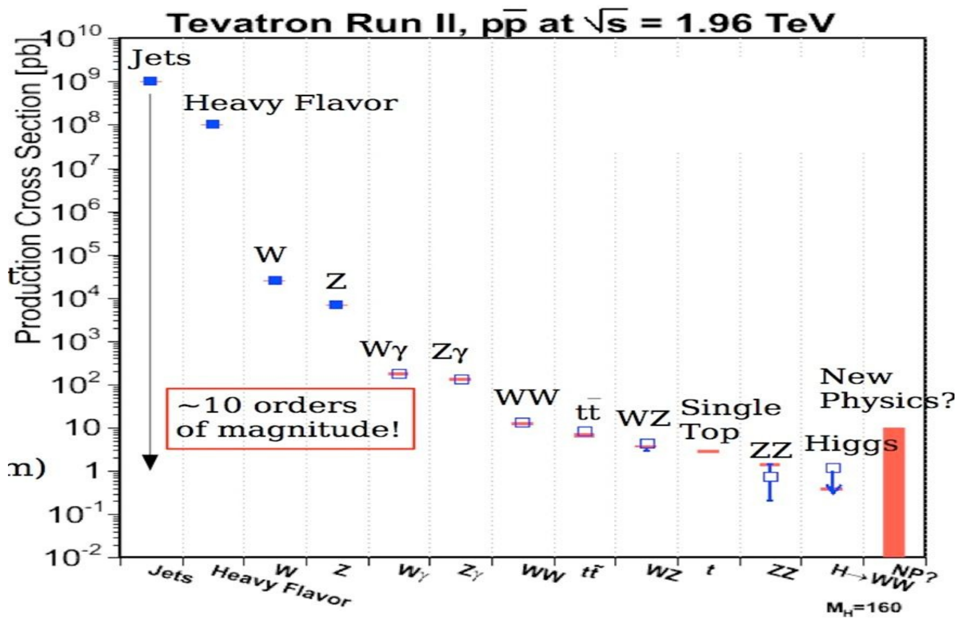


Figure 11.18: *Various production cross sections at Tevatron.* Note that the scale is logarithmic.

are almost collinear. Therefore, it is convenient to measure the opening angle between the lepton pair in the transverse plane, $\Delta\phi_{l^+l^-}$. This allows to select only events with small opening angle: $\Delta\phi < 2\,\mathrm{rad}$. Figure 11.20 shows plots for the ee, $\mu\mu$ and $e\mu$ case: The left column shows the signal plus a considerable amount of background by various processes unrelated to Higgs production. The right column shows $\Delta\phi$ after all cuts but the $\Delta\phi < 2$ cut are applied (the $\Delta\phi$ cut is indicated by arrows). If no event survives all cuts, it is possible to set an exclusion limit on the Higgs mass. A combined Tevatron (DØ and CDF) result using an amount of data corresponding to $\mathcal{L}_{\mathrm{int}} \sim 5\,\mathrm{fb}^{-1}$ excluding the mass range from 162 to 166 GeV at 95% CL is shown in Fig. 11.21. The current combined Tevatron and LEP standard model Higgs mass fit and excluded regions[11] are shown in Fig. 10.2.

[11] http://lepewwg.web.cern.ch/LEPEWWG/

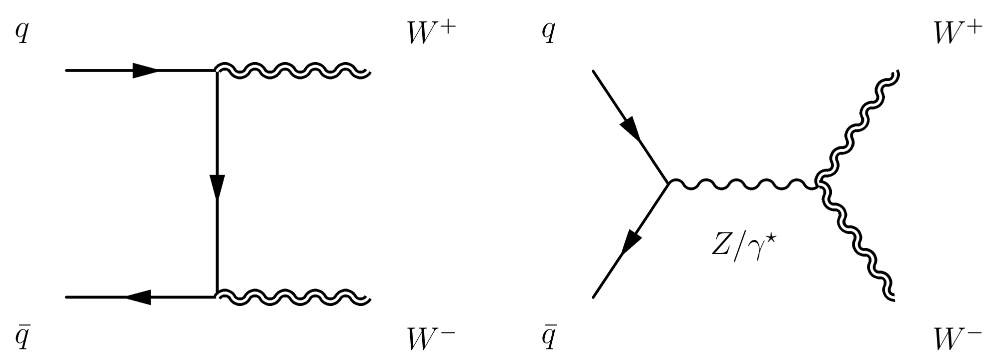


Figure 11.19: *Examples of W^+W^- production diagrams at hadron colliders not involving Higgs boson exchange.*

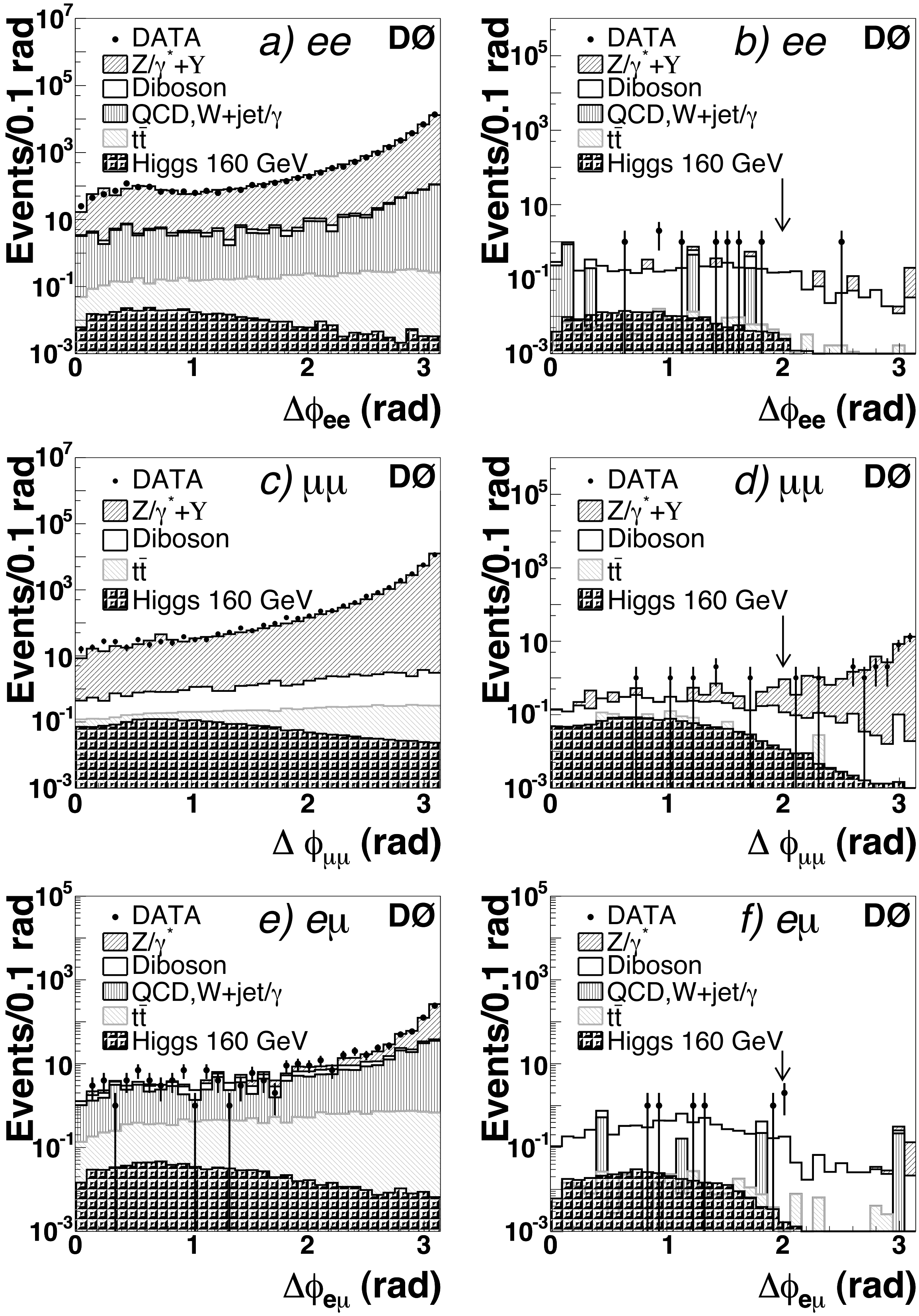


Figure 11.20: *Distribution of the opening angle $\Delta\phi_{ll'}$ after applying the initial transverse momentum cuts (a), (c), (e) and after all cuts, except for the $\Delta\phi$ cut (b), (d), (f).* Source: [64].

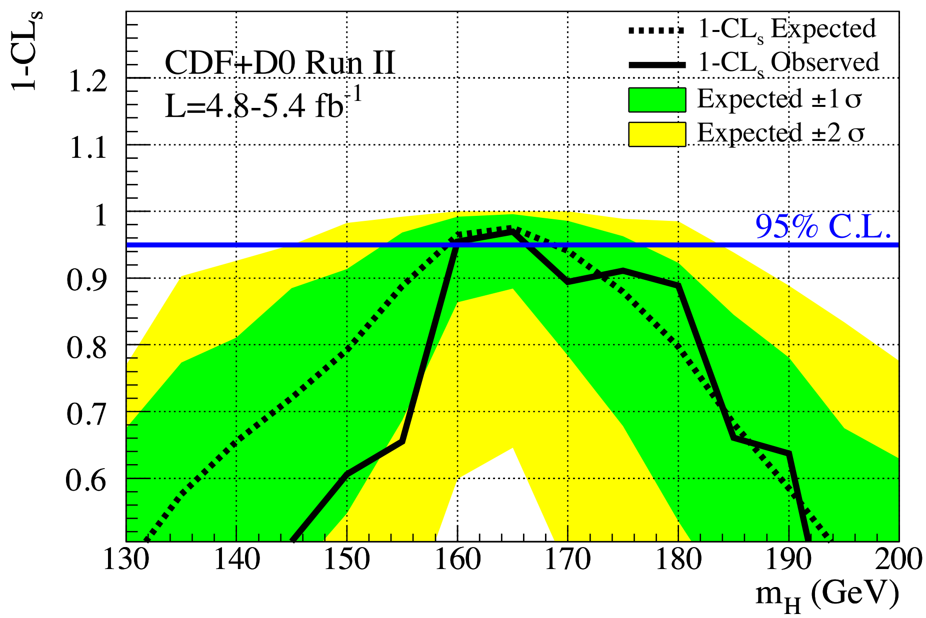


Figure 11.21: *Higgs mass range exclusion with combined Tevatron results.* Source: [65].

Chapter 12

Flavor physics

Quarks and leptons can be ordered in flavour doublets, each column being called a family,

$$\begin{array}{lccccc} \text{Quarks:} & \begin{pmatrix} u \\ d \end{pmatrix} & \begin{pmatrix} c \\ s \end{pmatrix} & \begin{pmatrix} t \\ b \end{pmatrix} & \begin{array}{c} Q = \frac{2}{3} \\ Q = -\frac{1}{3} \end{array} & , \\ \text{Leptons:} & \begin{pmatrix} \nu_e \\ e \end{pmatrix} & \begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix} & \begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix} & \begin{array}{c} Q = 0 \\ Q = -1 \end{array} & . \end{array}$$

These arrangements correspond to an approximate flavor $SU(6)$ symmetry. The isospin $SU(2)$ of p, n (Sect. 7, p. 127) or the flavor $SU(3)$ symmetry of u, d, s (Sect. 7.3, p. 133) are much better fulfilled since the mass differences between the different particles are much smaller than the masses themselves.

12.1 Cabibbo angle

The structure of the charged currents,

$$j_\mu^\pm = \bar{\chi}_L \gamma_\mu \tau_\pm \chi_L,$$

allows transitions within a single doublet, e.g. $d \to u$, $c \to s$, $t \to b$, but not between different doublets. This would imply that the lightest particle of each doublet should be stable (the electromagnetic and strong interactions do not allow flavor changing processes, since photons and gluons do not carry any flavor quantum numbers), a fact which is in contradiction with the observation that our universe is composed almost exclusively of particles of the first family, consisting of the lightest particles.

Assuming that the weak eigenstates of the d-type quarks [1] are linear combinations of the mass eigenstates one can reproduce the observed phenomenology. Let us first consider the

[1] Some authors prefer to rotate the u-type quarks. We follow here the most common version.

case of two quark families for simplicity. We have the weak eigenstate doublets,

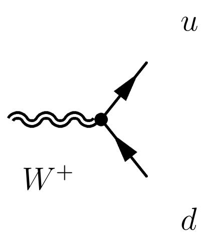
$$\begin{pmatrix} u \\ d' \end{pmatrix} \quad \begin{pmatrix} c \\ s' \end{pmatrix},$$

and we assume that the weak eigenstates $|d'\rangle$ and $|s'\rangle$ are linear combinations of the mass eigenstates $|d\rangle$ and $|s\rangle$,

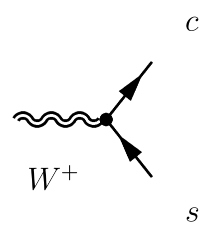
$$\begin{aligned} |d'\rangle &= \cos\theta_c\,|d\rangle + \sin\theta_c\,|s\rangle \\ |s'\rangle &= -\sin\theta_c\,|d\rangle + \cos\theta_c\,|s\rangle\,, \end{aligned} \tag{12.1}$$

where θ_c is called the **Cabibbo angle**.

Since decaying particles and decay products are mass eigenstates, this trick allows transitions between different families. Using Eq. (12.1), we can write vertex factors between mass eigenstates,

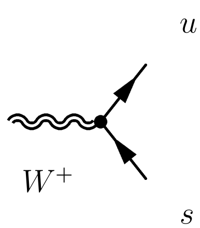


$\propto \cos\theta_c$

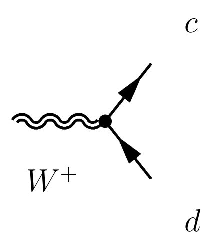


$\propto \cos\theta_c$,

called **Cabibbo preferred decays**, and,



$\propto \sin\theta_c$



$\propto -\sin\theta_c$,

called **Cabibbo suppressed decays**. If the weak and mass eigenstates would be the same, $\theta_c = 0$ and the second series of decay could not occur. The kaons are unstable but have a relatively long lifetime, since the decay of the s quark is Cabibbo supressed.

The introduction of the Cabibbo angle also destroys the universality of the Fermi constant,

$$G_F^{n\to pe^-\bar{\nu}_e} = \cos\theta_c G_F^{\mu^-\to e^-\nu_\mu\bar{\nu}_e}, \tag{12.2}$$

with the experimentally measured value,

$$\cos\theta_c \approx 0.974. \tag{12.3}$$

We can now rewrite the interaction Lagrangian for the charged current coupling to quarks,

$$\begin{aligned} i\mathcal{L}_{int}^{W^\pm,q} =& - i\frac{g}{\sqrt{2}} \begin{pmatrix} \bar{u} & \bar{c} \end{pmatrix} \gamma_\mu \frac{1-\gamma_5}{2} U \begin{pmatrix} d \\ s \end{pmatrix} W^{+\mu} \\ & - i\frac{g}{\sqrt{2}} \begin{pmatrix} \bar{d} & \bar{s} \end{pmatrix} U^T \gamma_\mu \frac{1-\gamma_5}{2} \begin{pmatrix} u \\ c \end{pmatrix} W^{-\mu}, \end{aligned} \tag{12.4}$$

with,

$$U = \begin{pmatrix} \cos\theta_c & \sin\theta_c \\ -\sin\theta_c & \cos\theta_c \end{pmatrix} \in U(2). \tag{12.5}$$

We remark at this point, that $U = U^*$ or in other words $U \in O(2)$ implying that $U^\dagger = U^T$.

12.2 Cabibbo-Kobayashi-Maskawa matrix

In 1973, before the observation of c, b and t quarks, the existence of three families and its implications were already hypothesised.

Analogously to Eq. (12.4), we write for three families,

$$\begin{aligned} i\mathcal{L}_{int}^{W^\pm,q} =& - i\frac{g}{\sqrt{2}} \begin{pmatrix} \bar{u} & \bar{c} & \bar{t} \end{pmatrix} \gamma_\mu \frac{1-\gamma_5}{2} V \begin{pmatrix} d \\ s \\ b \end{pmatrix} W^{+\mu} \\ & - i\frac{g}{\sqrt{2}} \begin{pmatrix} \bar{d} & \bar{s} & \bar{b} \end{pmatrix} V^\dagger \gamma_\mu \frac{1-\gamma_5}{2} \begin{pmatrix} u \\ c \\ t \end{pmatrix} W^{-\mu}, \end{aligned} \tag{12.6}$$

where $V \in U(3)$.

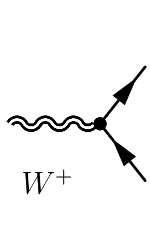
Recall that for a matrix $V \in U(N)$:

- V contains N^2 real parameters ($2N^2$ entries minus N^2 from the unitarity condition $V^\dagger V = \mathbb{1}$),
- $2N-1$ relative phases can be factorized by a phase redefinition of the quantum fields.

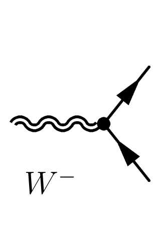
Thus V contains $N^2 - (2N-1) = (N-1)^2$ independent real parameters. On the other hand, a matrix $O \in O(N)$ is determined by $\frac{1}{2}N(N-1)$ independent real parameters (Euler angles).

Comparing V and O, we have, $N_a = \frac{1}{2}N(N-1)$ real angles and $N_p = (N-1)^2 - N_a = \frac{1}{2}(N-1)(N-2)$ complex phases. It then easy to see that we always have complex phases for $N \geq 3$, implying $V^* \neq V$.

Looking at the vertex factors connected through a CP-transformation,



$\propto V_{ji} \quad \neq$



$\propto V_{ij}^*,$

we conclude that the weak interaction violates CP invariance for $N \geq 3$ through complex phases in the CKM matrix V.

12.3 Neutrino mixing

Literature:

- Fukugita/Yanagida [66]

As in the case of d-type quarks, one can consider the phenomenology implied by neutrinos whose mass eigenstates (ν_1, ν_2 and ν_3) are not the same as the weak eigenstates (ν_e, ν_μ and ν_τ). The interaction Lagrangian becomes,

$$\begin{aligned} i\mathcal{L}_{int}^{W^\pm,l} = &-i\frac{g}{\sqrt{2}} \begin{pmatrix} \bar{\nu}_1 & \bar{\nu}_2 & \bar{\nu}_3 \end{pmatrix} U^\dagger \gamma_\mu \frac{1-\gamma_5}{2} \begin{pmatrix} e \\ \mu \\ \tau \end{pmatrix} W^{+\mu} \\ &- i\frac{g}{\sqrt{2}} \begin{pmatrix} \bar{e} & \bar{\mu} & \bar{\tau} \end{pmatrix} \gamma_\mu \frac{1-\gamma_5}{2} U \begin{pmatrix} \nu_1 \\ \nu_2 \\ \nu_3 \end{pmatrix} W^{-\mu}, \end{aligned} \tag{12.7}$$

with U the unitary neutrino mixing matrix, also called **Pontecorvo-Maki-Nakagawa-Sakata (PMNS) matrix**. As in the case of quarks, the existence of three neutrino families would let room for a CP violation in the neutrino sector . Up to now, it has not been possible to observe it experimentally.

In order to treat neutrino oscillations, it is important to remember the following facts about neutrinos:

- They are always produced as eigenstates of the weak interaction, e.g. $\pi^- \to \mu^- \bar{\nu}_\mu$,
- They are always detected as eigenstates of the weak interaction, e.g. $\nu_\mu p \to \mu^- X$,
- But they propagate in the vacuum as mass eigenstates.

Assuming two lepton families (e, μ), we write the weak eigenstates as,

$$\begin{aligned} |\nu_e\rangle &= \cos\theta\,|\nu_1\rangle + \sin\theta\,|\nu_2\rangle \\ |\nu_\mu\rangle &= -\sin\theta\,|\nu_1\rangle + \cos\theta\,|\nu_2\rangle\,. \end{aligned} \tag{12.8}$$

The time evolution of the mass eigenstates is given by,

$$|\nu_i, t\rangle = \mathrm{e}^{-iE_i t}\,|\nu_i, 0\rangle\,, \tag{12.9}$$

such that the evolution of the weak eigenstates is given by,

$$|\nu_\alpha, t\rangle = \sum_i U_{\alpha i}\mathrm{e}^{-iE_i t}\,|\nu_i, 0\rangle\,. \tag{12.10}$$

Since we know experimentally that $m_{\nu_i} < \mathrm{eV}, \mathrm{keV} \ll E \approx \mathrm{MeV}$, we can safely assume that they are ultrarelativistic and make the approximation,

$$E_i = \sqrt{\vec{p}^{\,2} + m_i^2} \approx |\vec{p}| + \frac{m_i^2}{2|\vec{p}|} = |\vec{p}| + \frac{m_i^2}{2E} \qquad (|\vec{p}| \gg m_i) \tag{12.11}$$

Inserting this in Eq. (12.9) we get,

$$\begin{aligned} |\nu_\alpha, t\rangle &= \mathrm{e}^{-i|\vec{p}|t}\left(U\begin{bmatrix} \mathrm{e}^{-i\frac{m_1^2 t}{2E}} & 0 \\ 0 & \mathrm{e}^{-i\frac{m_2^2 t}{2E}} \end{bmatrix}U^\dagger\right)_{\alpha\beta}|\nu_\beta, t\rangle \\ &\approx \mathrm{e}^{-i|\vec{p}|t}\left(U\begin{bmatrix} 1 - \frac{im_1^2 t}{2E} & 0 \\ 0 & 1 - \frac{im_2^2 t}{2E} \end{bmatrix}U^\dagger\right)_{\alpha\beta}|\nu_\beta, t\rangle\,, \end{aligned}$$

and, using,

$$U^\dagger m^\dagger m U = m^2_{Diag} = \begin{pmatrix} m_1^2 & 0 \\ 0 & m_2^2 \end{pmatrix},$$

we obtain (reexpressing $1 + iX = \mathrm{e}^{iX}$),

$$|\nu_\alpha, t\rangle = \mathrm{e}^{-i|\vec{p}|t}\left(\mathrm{e}^{-i\frac{m^\dagger m}{2E}t}\right)_{\alpha\beta}|\nu_\beta, 0\rangle\,. \tag{12.12}$$

We can interpret Eq. (12.12) as the solution of the Schrödinger equation,

$$i\frac{d}{dt}|\nu_\alpha, t\rangle = \left(|\vec{p}|\delta_{\alpha\beta} + \frac{(m^\dagger m)_{\alpha\beta}}{2E}\right)|\nu_\beta, t\rangle\,. \tag{12.13}$$

We now compute the $m^\dagger m$ matrix,

$$\begin{aligned} m^\dagger m = U m^2_{Diag} U^\dagger &= \begin{pmatrix} m_1^2 \cos^2\theta + m_2 \sin^2\theta & \frac{1}{2}(m_2^2 - m_1^2)\sin 2\theta \\ \frac{1}{2}(m_2^2 - m_1^2)\sin 2\theta & m_1^2 \sin^2\theta + m_2^2 \cos^2\theta \end{pmatrix} \\ &= \frac{m_1^2 + m_2^2}{2}\mathbb{1} + \frac{\Delta m^2}{2}\begin{pmatrix} -\cos 2\theta & \sin 2\theta \\ \sin 2\theta & \cos 2\theta \end{pmatrix}, \end{aligned}$$

with $\Delta m^2 = m_2^2 - m_1^2$. The term proportional to the identity does not induce a mixing and corresponds to a trivial phase factor. Inserting this result in Eq. (12.13) and dropping the diagonal term, we get,

$$\begin{aligned} i\frac{d}{dt}\begin{pmatrix} |\nu_e, t\rangle \\ |\nu_\mu, t\rangle \end{pmatrix} &= \frac{\Delta m^2}{4E}\begin{pmatrix} -\cos 2\theta & \sin 2\theta \\ \sin 2\theta & \cos 2\theta \end{pmatrix}\begin{pmatrix} |\nu_e, t\rangle \\ |\nu_\mu, t\rangle \end{pmatrix} \\ &= H_{vac}\begin{pmatrix} |\nu_e, t\rangle \\ |\nu_\mu, t\rangle \end{pmatrix}, \end{aligned}$$

with solution,

$$\begin{pmatrix} |\nu_e, t\rangle \\ |\nu_\mu, t\rangle \end{pmatrix} = \mathrm{e}^{-iH_{vac}t}\begin{pmatrix} |\nu_e, 0\rangle \\ |\nu_\mu, 0\rangle \end{pmatrix}.$$

Writing,

$$\mathrm{e}^{-iH_{vac}t} = \begin{pmatrix} A_{ee}(t) & A_{e\mu}(t) \\ A_{\mu e}(t) & A_{\mu\mu}(t) \end{pmatrix},$$

and using,

$$H_{vac} = \frac{\Delta m^2}{2E}\left(\sin(2\theta)\sigma_1 - \cos(2\theta)\sigma_3\right),$$

we get,

$$\mathrm{e}^{-iH_{vac}t} = \cos\left(\frac{\Delta m^2}{2E}t\right)\mathbb{1} - i\sin\left(\frac{\Delta m^2}{2E}t\right)\left(\sin(2\theta)\sigma_1 + \cos(2\theta)\sigma_3\right). \tag{12.14}$$

We finally get the transition amplitude from the projection of $|\nu_e, t\rangle$ onto $\langle\nu_e|$:

$$\langle\nu_e|\nu_e, t\rangle = A_{ee}(t) = \cos\left(\frac{\Delta m^2}{2E}t\right) - i\sin\left(\frac{\Delta m^2}{2E}t\right)\cos 2\theta,$$

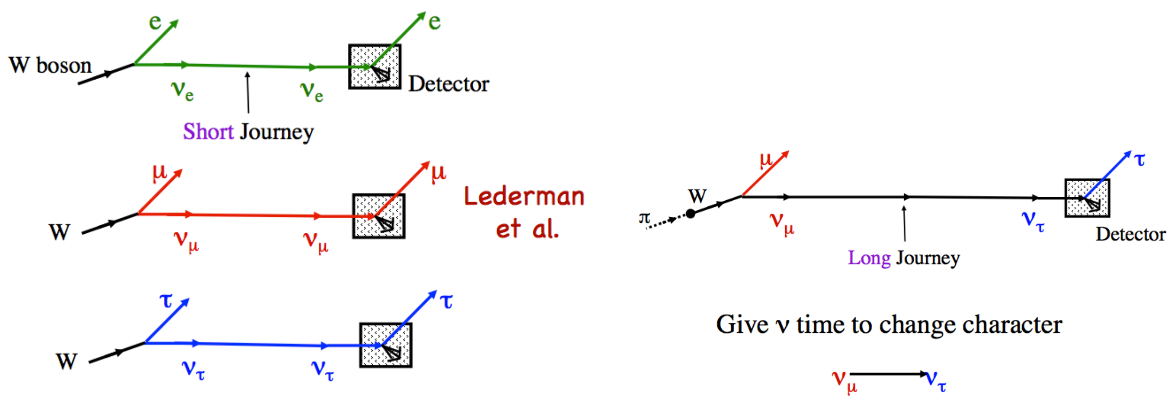
and the transition probability,

$$P_{\nu_e\to\nu_e}(t) = |\langle\nu_e|\nu_e, t\rangle|^2 = 1 - \sin^2(2\theta)\sin^2\left(\frac{\Delta m^2}{2E}t\right) \tag{12.15}$$

$$P_{\nu_e\to\nu_\mu}(t) = |\langle\nu_\mu|\nu_e, t\rangle|^2 = \sin^2(2\theta)\sin^2\left(\frac{\Delta m^2}{2E}t\right) \tag{12.16}$$

A useful formula to estimate the distance over which full oscillations take place is (since the neutrino is ultrarelativistic $L = t$),

$$\frac{\Delta m^2 L}{4E} \approx 1.27\frac{\Delta m^2[\mathrm{eV}^2]L[\mathrm{m}]}{E[\mathrm{MeV}]}. \tag{12.17}$$



(a) (b)

Figure 12.1: *Neutrino production and detection.* During a sufficiently long journey, the neutrinos may change character (b). Source: B. Kayser.

12.4 Neutrino physics

In the previous sections we have seen that neutrino oscillation can be accounted for by assuming that neutrino flavor eigenstates are not identical to the mass eigenstates. Here we will again take a look at the two-neutrino case, discuss what can be measured in experiment and extend the theoretical treatment of oscillation to the three-neutrino case. Based on these results, we will proceed to the discussion of phenomenological aspects. It will become clear that to measure absolute neutrino masses, different experiments than the ones documenting neutrino oscillations are necessary. Their discussion will conclude this section.[2]

12.4.1 Neutrino oscillation theory revisited

Consider the charged-current interaction or W boson decay $W \to e\nu_e$ (see Fig. 12.1(a)). Since the electron (positron) produced together with its anti-neutrino (neutrino) can be detected and identified, the neutrino flavor at the time of production is fixed and in principle known (see also [70]). Detection of the neutrino proceeds via the inverse process, by lepton number conservation producing again an electron (positron), if the flavor is conserved while the neutrino travels from its place of production to the detector. The analogue holds for μ and τ.

However, if neutrinos have mass, it is possible for them to change their flavor, given the journey to the detector is long enough (see Fig. 12.1(b)). As we have seen, a difference in the mass eigenvalues $\delta m \neq 0$ is a necessary condition for oscillation to occur. Recently, a first candidate for a direct observation of the flavor change $\nu_\mu \to \nu_\tau$ was reported.[3]

[2]This section is heavily based on lectures by E. Lisi at the CHIPP PhD school, Jan. 2010 [67, 68, 69].
[3]`http://operaweb.lngs.infn.it/IMG/pdf/OPERA_press_release_May_2010_english-5.pdf`

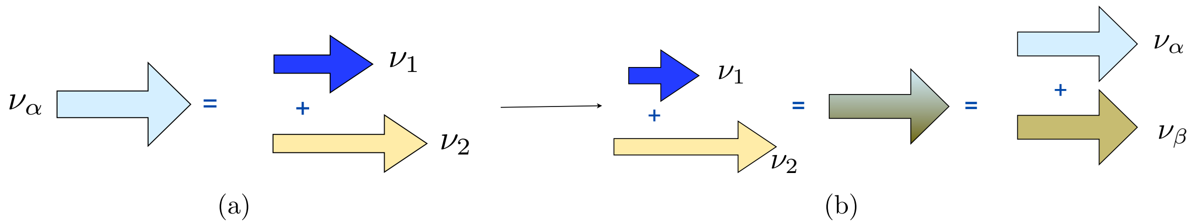


Figure 12.2: *Neutrino mixing in the two-neutrino case.*

In Sect. 12.4.2 we will discuss further experimental evidence that such flavor oscillations actually occur. This means that neutrino flavor is not a constant of motion. From electroweak theory we know that left-handed neutrinos ν_l are produced together with the corresponding lepton l in charged-current interactions (see Sect. 11.5). Recall that the right-handed neutrino carries neither $SU(2)_L$ nor $U(1)_Y$ charge and thus decouples from the electroweak interactions. Recent experiments, probing probabilities $P(\nu_\alpha \to \nu_\beta)$, have found that flavor is not conserved over macroscopic distances, especially in the so-called disappearance mode:

$$P(\nu_e \to \nu_e) < 1$$
$$P(\nu_\mu \to \nu_\mu) < 1$$

means that one finds less events than expected from the production rate, i. e. individual lepton number is not conserved.

These phenomena can be explained by neutrino mixing: For neutrinos, flavor eigenstates $\{\nu_\alpha\}$ are not identical to mass eigenstates $\{\nu_i\}$ and thus they can be expressed as linear combinations of each other. For the left-handed fields this reads, in analogy to the CKM matrix,

$$\nu_{\alpha L} = \sum_{i=1}^{3} U_{\alpha i} \nu_{iL} \tag{12.18}$$

for $\alpha = e, \mu, \tau$. Here $U = U^\dagger$ is called PMNS (Pontecorvo-Maki-Nakagawa-Sakata) matrix with $U \to U^*$ for $\nu \to \bar{\nu}$.

So, how does this setup bring about neutrino mixing? At production we start out with a pure flavor eigenstate ν_α which is according to Eq. (12.18) a certain superposition of mass eigenstates, say ν_1 and ν_2 (see Fig. 12.2(a)). If the eigenvalues of the mass eigenstates are different, so are their energies: $E_1 \neq E_2$. Thus the free time evolution operator introduces different phases and the superposition changes while traveling the distance $L \simeq ct$. Now, neutrino detection is a projection to one *flavor* eigenstate, such that, depending on the mixing angle θ and the mass difference δm^2, the number of produced neutrinos of flavor α may differ from the number of detected neutrinos of this flavor (see Fig. 12.2(b)). Recall that for the two-neutrino case the superpositions can be written as

$$\begin{pmatrix} \nu_\alpha \\ \nu_\beta \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \nu_1 \\ \nu_2 \end{pmatrix}$$

where θ is the mixing angle. This ansatz predicts the phenomena of "disappearance",

$$P(\nu_\alpha \to \nu_\alpha) = P(\nu_\beta \to \nu_\beta) = P(\bar{\nu}_\alpha \to \bar{\nu}_\alpha) = P(\bar{\nu}_\beta \to \bar{\nu}_\beta) = 1 - \sin^2 2\theta \sin^2 \frac{\Delta_{12}}{2},$$

and "appearance",

$$P(\nu_\alpha \to \nu_\beta) = P(\nu_\beta \to \nu_\alpha) = P(\bar{\nu}_\alpha \to \bar{\nu}_\beta) = P(\bar{\nu}_\beta \to \bar{\nu}_\alpha) = \sin^2 2\theta \sin^2 \frac{\Delta_{12}}{2}$$

where $\Delta_{12} \equiv \Delta m^2 t/(2E) \simeq \Delta m^2 L/(2E)$. Stating the above in another way, we can say that in the two-neutrino case the transition probability is

$$P(\nu_\alpha \to \nu_\beta) = \sin^2 2\theta \sin^2 \frac{\Delta m^2 L}{4E}$$

where

$$\frac{\Delta m^2 L}{4E} = 1.27 \left(\frac{\Delta m^2}{\mathrm{eV}^2}\right) \left(\frac{L}{\mathrm{km}}\right) \left(\frac{\mathrm{GeV}}{E}\right).$$

Let us define the oscillation wavelength

$$\lambda_{\mathrm{osc}} = \frac{4\pi E}{\Delta m^2}$$

and rewrite the transition probability accordingly:

$$P(\nu_\alpha \to \nu_\beta) = \underbrace{\sin^2(2\theta)}_{\text{mixing term}} \underbrace{\sin^2 \left(\pi \frac{L}{\lambda_{\mathrm{osc}}}\right)}_{\text{oscillation term}}. \tag{12.19}$$

The LHS of Eq. (12.19) is determined in experiment by counting events and normalizing. Since the mixing angle θ is fixed, so is the mixing term on the RHS. However the oscillation term can be influenced by the experimental design: Although Δm^2 is fixed, the experimenter is free to choose the source-detector distance L and can, by selecting the production process, influence the neutrino energy E and thus λ_{osc}. We now discuss the behavior of Eq. (12.19) for different sizes of L/λ_{osc}.

A) $L/\lambda_{\mathrm{osc}} \ll 1$. E. g. this is realized for $\Delta m^2 \sim 10^{-5}\,\mathrm{eV}^2$ and $E \sim 1\,\mathrm{MeV}$ which is the energy scale of nuclear reactions; at the same time L needs to be small, e. g. $L \sim 1\,\mathrm{km}$. Since the argument of the oscillation term is small, it can be approximated by the first term of the Taylor series:

$$\sin^2 \left(\pi \frac{L}{\lambda_{\mathrm{osc}}}\right) \simeq \left(\pi \frac{L}{\lambda_{\mathrm{osc}}}\right)^2.$$

Therefore the transition probability is small and the effect might be very difficult to measure, depending on the experimental resolution.

B) $\pi L/\lambda_{\text{osc}} \simeq 1$. E. g. consider the case that L and E are such that $\pi L/\lambda_{\text{osc}} \simeq \pi/2$, i. e. the oscillation term is at its first maximum. Possible numbers are: $\Delta m^2 \simeq 10^{-3}\,\text{eV}^2$, $E = 1\,\text{GeV}$ (energy scale of accelerators and cosmic rays) and $L \simeq 1000\,\text{km}$. In this case

$$1.27\Delta m^2 \frac{L}{E} \simeq 1.3 \simeq \frac{\pi}{2}$$

such that the sensitivity to the mixing term is maximized.

C) $L/\lambda_{\text{osc}} \gg 1$: For instance, this is the case if $\Delta m^2 \simeq 10^{-5}\,\text{eV}^2$, $L =$ distance earth-sun $\sim 150 \cdot 10^6\,\text{km}$, $E \sim 1\,\text{MeV}$. Therefore, fast oscillation is taking place which leads to a measurement of the average due to uncertainties in E and L:

$$\left\langle \sin^2\left(\pi \frac{L}{\lambda_{\text{osc}}}\right)\right\rangle = \frac{1}{2} \Rightarrow P(\nu_\alpha \to \nu_\beta) = \frac{1}{2}\sin^2(2\theta).$$

To conclude this comment on orders of magnitude, let us take a look at the detector sizes needed in neutrino experiments. The number of events is given by the product of cross section and integrated luminosity:

$$N_{\text{events}} = \Phi \sigma_{\nu p} T N_p \tag{12.20}$$

where $\Phi \sim 10^{10-12}\,\text{m}^{-2}\text{s}^{-1}$ is the flux of incoming neutrinos, $\sigma_{\nu p} \sim 10^{-45}\,\text{m}^{-2}$ is the cross section[4] of neutrino-proton scattering, $T \sim 1y \simeq 10^7\,\text{s}$ is the observation time and N_p is the number of protons in the target. One can see that, although one can try to increase the flux or measure longer, the main problem is the small cross section $\sigma_{\nu p}$. The only parameter left to tune is the number of protons N_p: To find a reasonable number of events, one has to choose e. g. $N_p > 10^{30}$ which corresponds to about 10^7 mol, i. e. we are talking about detector sizes of tons and kilotons.

Having discussed the behavior of the oscillation term, we can think about what an experiment may be sensitive to. As we have seen, for fast oscillations (large Δm^2) the $\sin^2$ is averaged over and there is, due to uncertainty in E and L no sensitivity on the mass difference (see Fig. 12.3). If the experiment does not find an oscillation signal, one can exclude the RHS region of the curve. To constrain the parameter space, various experiments with different sensibilities are needed.

To attack the case of three light neutrinos, we have to consider a 3×3 mixing matrix. One possible parametrization is ($\Gamma_\delta = \text{diag}(1, 1, e^{i\delta})$)

$$\begin{aligned} U &= O_{23}\Gamma_\delta O_{13}\Gamma_\delta^\dagger O_{12} \\ &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta_{23} & \sin\theta_{23} \\ 0 & -\sin\theta_{23} & \cos\theta_{23} \end{pmatrix} \begin{pmatrix} \cos\theta_{13} & 0 & \sin\theta_{13}e^{-i\delta} \\ 0 & 1 & 0 \\ -\sin\theta_{13}e^{i\delta} & 0 & \cos\theta_{13} \end{pmatrix} \begin{pmatrix} \cos\theta_{12} & \sin\theta_{12} & 0 \\ -\sin\theta_{12} & \cos\theta_{12} & 0 \\ 0 & 0 & 1 \end{pmatrix}. \end{aligned}$$

[4] This is only a rough estimate.

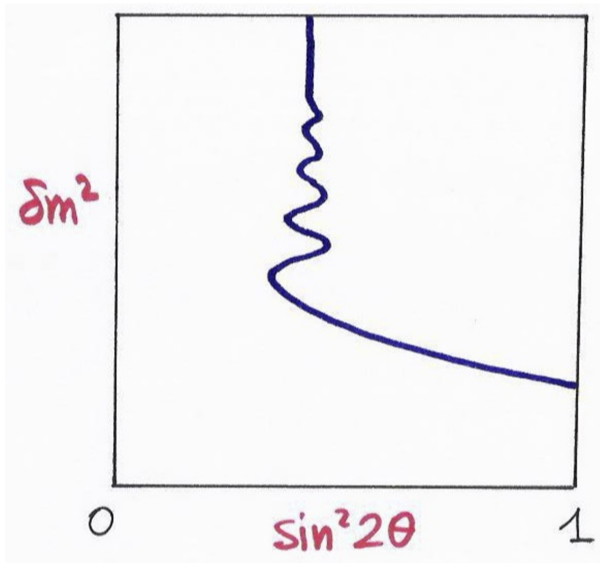


Figure 12.3: *Oscillation experiment sensitivity.* Source: [71].

Experiment shows that $\sin^2\theta_{23} \sim 0.5$ which means almost maximal mixing, $\sin^2\theta_{13} \lesssim$ few %, $\delta = ?$ (small) and $\sin^2\theta_{12} \sim 0.3$. This structure is very different from the CKM case, where the diagonal elements are dominant. What about mass differences in the three-neutrino case? We do not know the absolute ν masses, but they roughly fulfill $m_i \lesssim 1\,\text{eV}$. For ultrarelativistic neutrinos in vacuum we may expand the energy as

$$E = \sqrt{\vec{p}^{\,2} + m_i^2} \simeq |\vec{p}| + \frac{m_i^2}{2E}.$$

Since the oscillation phase is caused by $\Delta E \propto \Delta m_{ij}^2$, this is what oscillation experiments probe. For three neutrinos there are two independent mass differences. For historical reasons the small splitting δm^2 is called "solar" mass^2 splitting:

$$\delta m^2 \simeq 7.7 \cdot 10^{-5}\,\text{eV}^2,$$

for the same reason the large splitting is called "atmospheric" mass^2 splitting:

$$\Delta m^2 \simeq 2.4 \cdot 10^{-3}\,\text{eV}^2.$$

Note that, because $\delta m^2/\Delta m^2 \simeq 1/30$, it is very difficult to be sensitive to both mass splittings in the same experiment (L/E is fixed). The absolute masses m_i are unknown, and thus it is possible to arrange the mass eigenstates in two ways, corresponding to the labeling convention

$$\begin{aligned} \delta m^2 &= m_2^2 - m_1^2 > 0 \\ |\Delta m^2| &= |m_3^2 - m_{1,2}^2| \end{aligned}$$

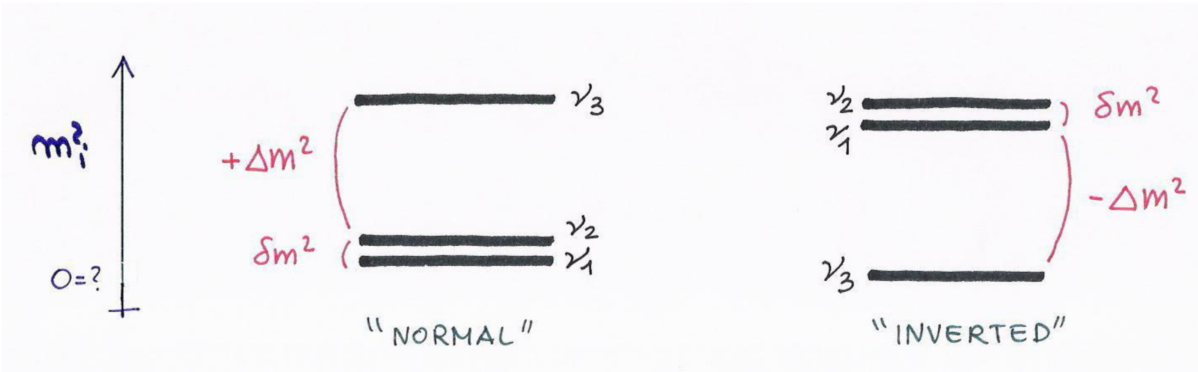


Figure 12.4: *Normal and inverted mass hierarchies for the three-neutrino case.* Source: [71].

(see Fig. 12.4).

To find simple expressions for the oscillation probabilities in the three-neutrino case, we apply two approximations: We neglect the complex phase ($\delta = 0$) and we assume that only one mass scale is relevant:

$$|\delta m^2| \ll |\Delta m^2| \text{ and } |\delta m^2| \ll \frac{E}{L}.$$

This simplified three-neutrino oscillation is described by three parameters only: the mass difference Δm^2, and the mixing angles θ_{13} and θ_{23}. This allows to write the oscillation probabilities as follows [72]:

$$P(\nu_e \to \nu_e) = 1 - \sin^2 2\theta_{13} \sin^2 \frac{\Delta m^2 L}{4E} \tag{12.21}$$

$$P(\nu_e \to \nu_\mu) = \sin^2 2\theta_{13} \sin^2 \theta_{23} \sin^2 \frac{\Delta m^2 L}{4E} \tag{12.22}$$

$$P(\nu_\mu \to \nu_\tau) = \sin^2 2\theta_{13} \sin^2 \theta_{23} \sin^2 \frac{\Delta m^2 L}{4E} \tag{12.23}$$

$$P(\nu_e \to \nu_\tau) = \sin^2 2\theta_{13} \cos^2 \theta_{23} \sin^2 \frac{\Delta m^2 L}{4E} \tag{12.24}$$

$$P(\nu_\mu \to \nu_\tau) = \cos^4 \theta_{13} \sin^2 2\theta_{23} \sin^2 \frac{\Delta m^2 L}{4E}. \tag{12.25}$$

Note that the last equation gives the oscillation probability measured at the OPERA experiment (mentioned above). Not neglecting the CP violating phase δ, one has

$$P(\nu_\alpha \to \nu_\beta) = \delta_{\alpha\beta} - 4 \sum_{i<j} \mathrm{Re} J^{ij}_{\alpha\beta} \sin^2 \left(\frac{\Delta m^2_{ij} L}{4E} \right) - 2 \sum_{i<j} \mathrm{Im} J^{ij}_{\alpha\beta} \sin \left(\frac{\Delta m^2_{ij} L}{2E} \right) \tag{12.26}$$

where $\Delta m_{ij} = m_i^2 - m_j^2$ and $J^{ij}_{\alpha\beta} = U_{\alpha i} U^*_{\beta i} U^*_{\alpha j} U_{\beta j}$. CP violation would be caused by the imaginary part in Eq. (12.26); if it indeed existed, there would be CP violation not only in the quark sector, but also in the lepton sector.

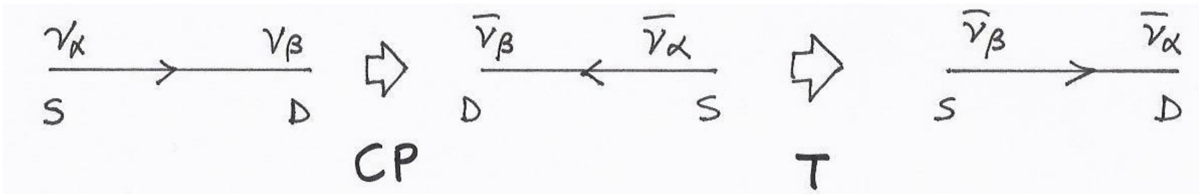


Figure 12.5: *Action of CP and T transformations on the $\nu_\alpha \to \nu_\beta$ process from source (S) to detector (D).* Source: [71].

Let us now take a closer look at the justification of the oscillation probabilities in Eq. (12.21) to (12.25). First consider the influence of symmetries. Figure 12.5 shows the action of CP and T transformations on the $\nu_\alpha \to \nu_\beta$ process from source (S) to detector (D). CP mirrors the setup and trades particles for antiparticles while T reverses the flow of time. This can be summarized as follows:

$$\begin{array}{rll} CP \text{ invariance} & P(\nu_\alpha \to \nu_\beta) = P(\bar\nu_\alpha \to \bar\nu_\beta) & (\nu \leftrightarrow \bar\nu) \\ T \text{ invariance} & P(\nu_\alpha \to \nu_\beta) = P(\nu_\beta \to \nu_\alpha) & (\alpha \leftrightarrow \beta) \\ & P(\bar\nu_\alpha \to \bar\nu_\beta) = P(\bar\nu_\beta \to \bar\nu_\alpha) & \\ CPT \text{ invariance} & P(\nu_\alpha \to \nu_\beta) = P(\bar\nu_\beta \to \bar\nu_\alpha) & (\nu \leftrightarrow \bar\nu)\&(\alpha \leftrightarrow \beta) \end{array}$$

Looking at Eq. (12.26), one sees that $(\alpha \leftrightarrow \beta)$ or $(\nu \leftrightarrow \bar\nu)$ amount to $(U \leftrightarrow U^*)$. Therefore, CP invariance requires $U = U^*$, while CPT invariance holds in any case. If the experiments are such that the two approximations used to obtain Eq. (12.21) to (12.25) are valid, the corresponding expressions read

$$\begin{aligned} P(\nu_\alpha \to \nu_\alpha) &= 1 - 4|U_{\alpha 3}|^2(1 - |U_{\alpha 3}|^2)\sin^2\left(\frac{\Delta m^2 L}{4E}\right) \\ P(\nu_\alpha \to \nu_\beta) &= 4|U_{\alpha 3}|^2|U_{\beta 3}|^2\sin^2\left(\frac{\Delta m^2 L}{4E}\right) \qquad \alpha \neq \beta. \end{aligned}$$

Using $|U_{e3}|^2 = \sin^2\theta_{13},\ |U_{\mu 3}|^2 = \cos^2\theta_{13}\sin^2\theta_{23},\ |U_{\tau 3}|^2 = \cos^2\theta_{13}\cos^2\theta_{23}$, one recovers Eq. (12.21) to (12.25). Measurements based on these results are neither sensitive to the type of mass hierarchy nor to CP violation. Also there is no sensitivity to δm^2 and θ_{12}. Finally, there is no difference between the expressions for ν and $\bar\nu$. Table 12.1 shows a summary of the experiments for which the said approximation, $\Delta m^2 L/(4E) \simeq 1$, holds. These include atmospheric neutrino experiments (ATM), long-baseline accelerator experiments (LBL) and short-baseline reactor experiments (SBR). Note that the first two oscillation probabilities reduce to the two-neutrino form for $\theta_{13} \to 0$ and the second two are constant for $\theta_{13} \to 0$.

At the other side of the mass spectrum, there are experiments mainly sensitive to δm^2

Experiment	Measurement
OPERA (LBL)	$P(\nu_\mu \to \nu_\tau) \simeq \mathrm{c}_{13}^4 \sin^2 2\theta_{23} \sin^2(\Delta m^2 L/(4E))$
K2K, MINOS (LBL), atmospheric	$P(\nu_\mu \to \nu_\mu) \simeq 1 - 4\mathrm{c}_{13}^2 \mathrm{s}_{23}^2 (1 - \mathrm{c}_{13}^2 \mathrm{s}_{23}^2) \sin^2(\Delta m^2 L/(4E))$
ATM, LBL	$P(\nu_\mu \to \nu_e) \simeq \mathrm{s}_{23}^2 \sin^2 2\theta_{13} \sin^2(\Delta m^2 L/(4E))$
CHOOZ (SRB)	$P(\nu_e \to \nu_e) \simeq 1 - \sin^2 2\theta_{13} \sin^2(\Delta m^2 L/(4E))$

Table 12.1: *Summary of neutrino experiments with* $\Delta m^2 L/(4E) \simeq \infty$*.* $\mathrm{s}_{ij}^2 = \sin^2\theta_{ij}$ and $\mathrm{c}_{ij}^2 = \cos^2\theta_{ij}$.

where

$$\frac{\delta m^2 L}{4E} \simeq \mathcal{O}(1) \tag{12.27}$$

$$\frac{\Delta m^2 L}{4E} \gg 1. \tag{12.28}$$

In this case

$$P(\nu_e \to \nu_e) \simeq \cos^4\theta_{13} \left[1 - \sin^2 2\theta_{12} \sin^2\left(\frac{\delta m^2 L}{4E}\right)\right] + \sin^4\theta_{13} \tag{12.29}$$

which holds e. g. for the KamLAND long-baseline reactor experiments. Note that also in this case there is no dependence on hierarchy, neutrino-antineutrino interchange and CP violation.

To conclude the theory part, let us summarize the above discussion. We have worked out approximate oscillation probabilities as a function of dominant mass mixing parameters for different classes of experiments (see Fig. 12.6). Furthermore, we have seen that the smallness of θ_{13} and of $\delta m^2/\Delta m^2$ make it difficult to probe CP violation and the hierarchy via oscillations in current experiments. Finally [73, p. 215], matter effects can occur if the neutrinos under consideration experience different interactions by passing through matter. In the Sun and the Earth ν_e can have neutral-current and charged-current interactions with leptons because of the existence of electrons, while for ν_μ and ν_τ only neutral-current interactions are possible. This is not being discussed any further here, see e. g. [73].

12.4.2 Phenomenology – experiments and current knowledge

Figure 12.7 shows combined results of neutrino experiments. In the excluded regions, no oscillations are observed; note that the (more or less) symmetric shape in the upper part of the plot is because for the three-neutrino case (and because of matter effects) the dependence is not only on $\sin^2 2\theta$, such that octant symmetry, $P(\theta) = P(\pi/2 - \theta)$, (see also Fig. 12.3) does not hold in general and the second octant has to be unfolded (see Fig. 12.8). In any case, one realizes that there are many experimental results available.

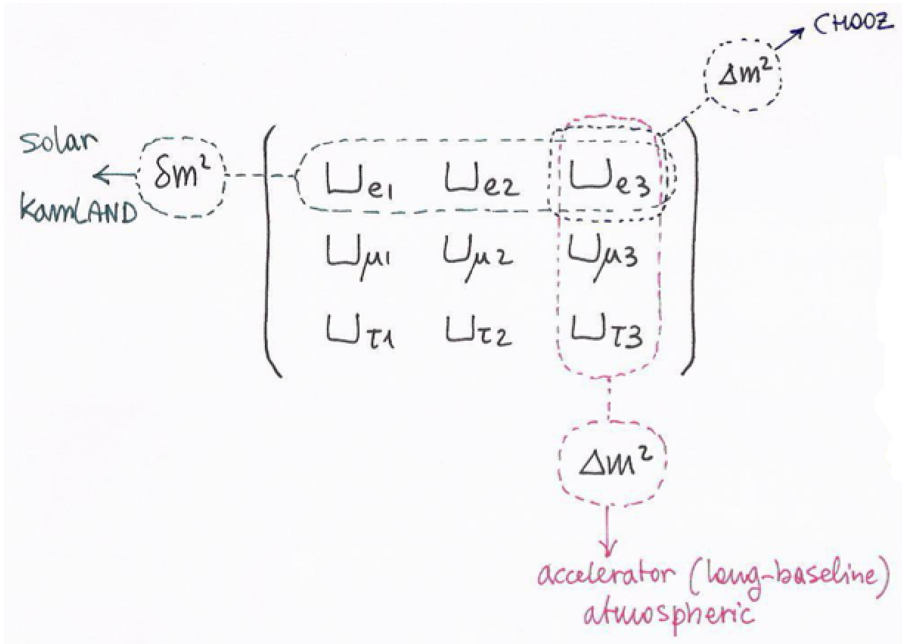


Figure 12.6: *Summary of experimental sensitivities to the neutrino mixing matrix.* Source: [71].

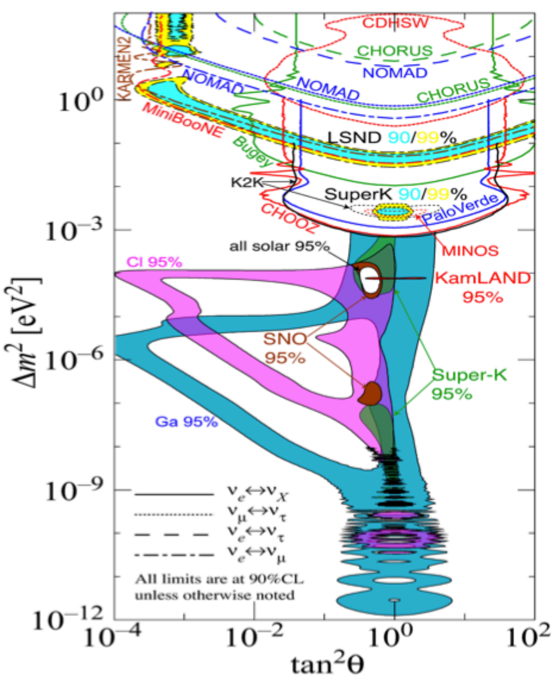


Figure 12.7: *Summary of neutrino oscillation experiments.* Source: Particle Data Group 2009.

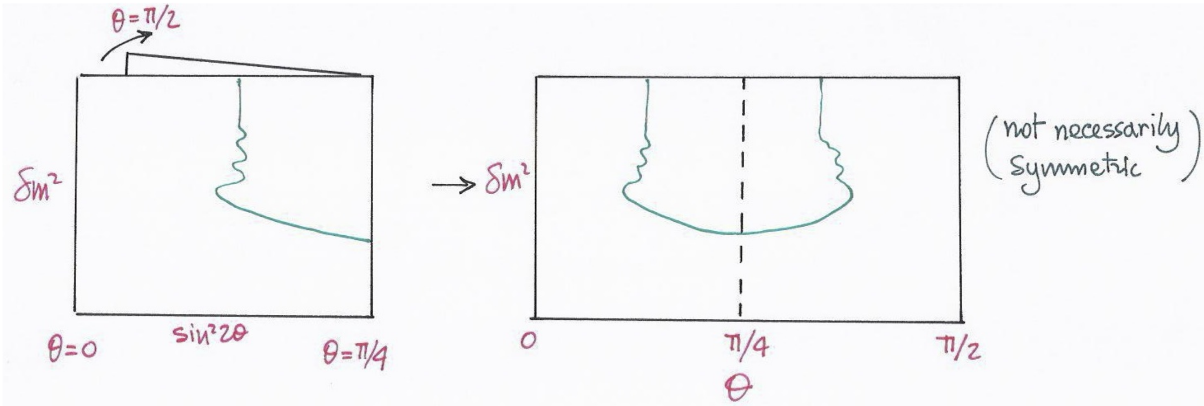


Figure 12.8: *Oscillation experiment sensitivity as a function of θ, rather than* $\sin^2 2\theta$. Source: [71].

Their three-neutrino interpretation is summarized in Fig. 12.9; the numerical values (with one digit accuracy) read:

$$\begin{aligned}
\delta m^2 &\sim 8 \cdot 10^{-5}\,\text{eV}^2\\
\Delta m^2 &\sim 3 \cdot 10^{-3}\,\text{eV}^2\\
m_\nu &< \mathcal{O}(1)\,\text{eV}\\
\text{sign}(\Delta m^2) &= ?\\
\sin^2\theta_{12} &\sim 0.3\\
\sin^2\theta_{23} &\sim 0.5\\
\sin^2\theta_{13} &\sim \text{few}\,\%\\
\delta(CP) &= ?.
\end{aligned}$$

Figure 12.10 gives an overview of which type of experiment contributed to the individual parts of the present knowledge on neutrino mass properties. In the following we discuss how such information is constrained by the following types of experiments:

- Short-baseline reactor;
- Atmospheric;
- Long-baseline accelerator and
- Solar.

The short-baseline reactor experiment CHOOZ. Figure 12.11 shows the general setup of the CHOOZ experiment. Nuclear fission in a reactor produces antineutrinos via neutron decay: $n \to p + e^- + \bar{\nu}_e$, leading to production rates as high as $\sim 6 \cdot 10^{20}\,\text{s}^{-1}$, the

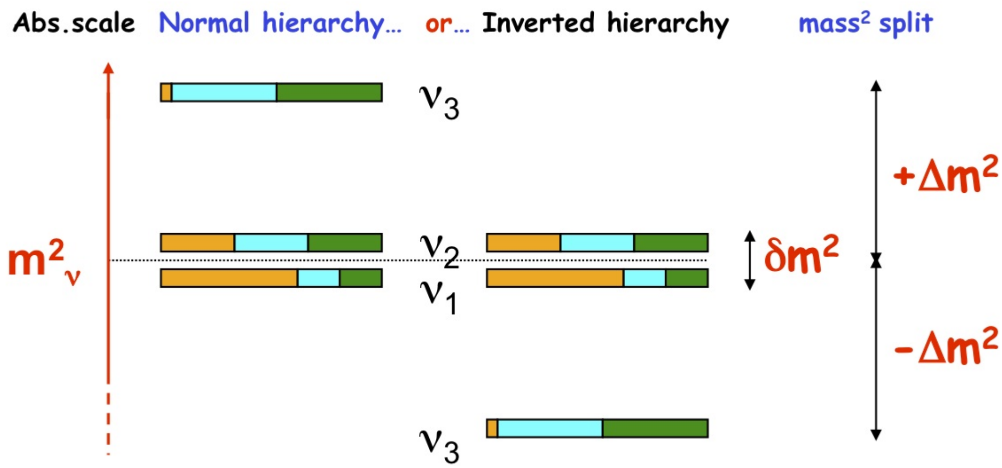


Figure 12.9: *Summary of the current knowledge on neutrino oscillations.* Source: [71].

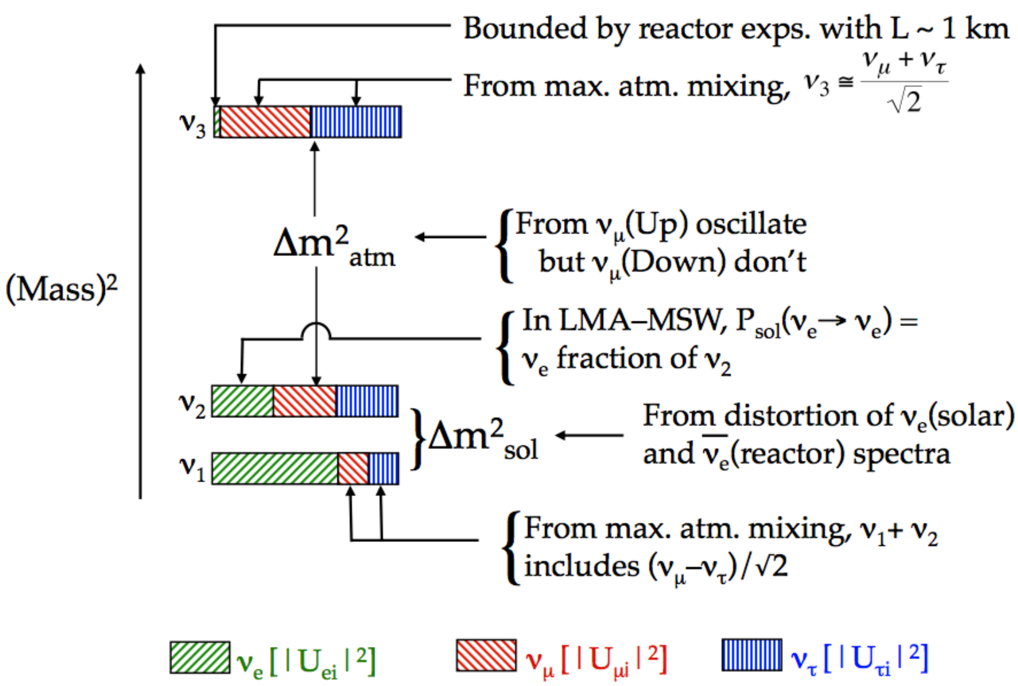


Figure 12.10: *Origin of the current knowledge on neutrino oscillations.* Source: B. Kayser.

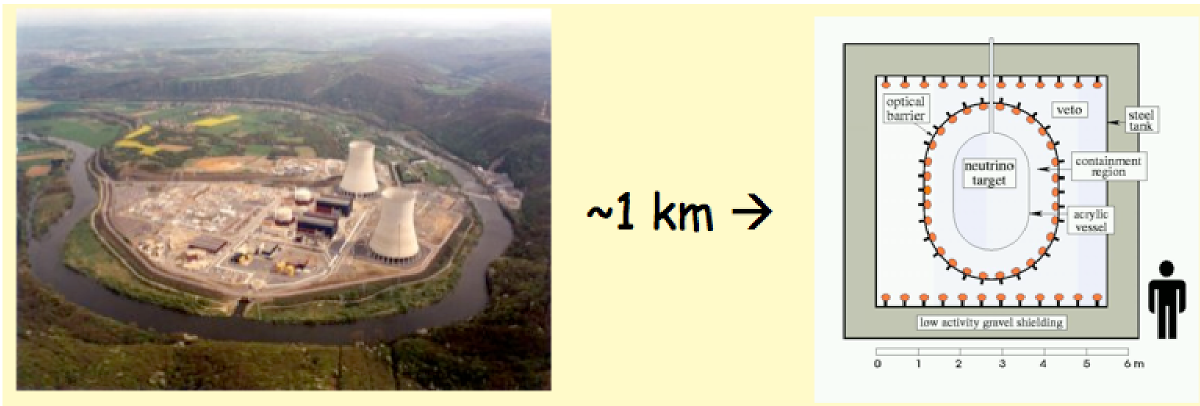


Figure 12.11: *Setup of short-baseline reactor experiments.* Source: [71].

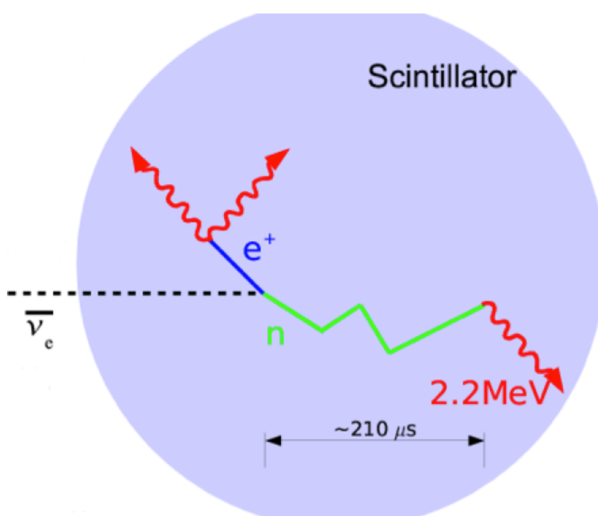
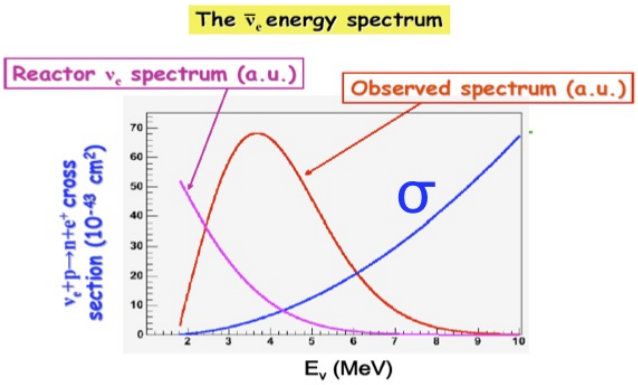


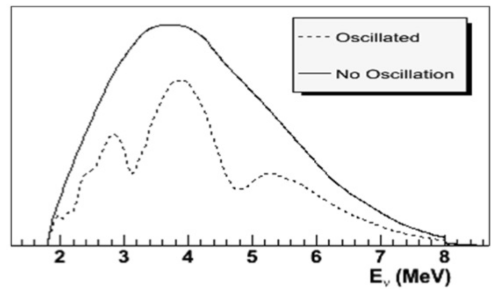
Figure 12.12: *Neutrino detection via inverse beta decay.*

energy being of the order of MeV. Detection is accomplished by inverse β-decay: $\bar{\nu}_e + p \to e^+ + n$; $n + p \to d + \gamma$, i. e. an incoming antineutrino hits a proton in the scintillator which acts both as target and detector, producing a positron and a neutron (see Fig. 12.12). In the scintillator, the positron annihilates with an electron to produce two photons, both at 511 keV. Some 210 μs later the neutron is captured, producing an excited state, which decays emitting a photon of about 2.2 MeV. Taken together, due to their energy and temporal pattern, the three photons produced in total constitute a clear signature. In particular, the fact that the third γ is delayed allows for good background rejection. What does one expect assuming that there are no oscillations visible with this setup? The reactor antineutrino spectrum is shown in Fig. 12.13(a) together with the cross section for inverse β-decay. Convoluting both distributions yields the observed spectrum. However, if there are oscillations the picture changes (see Fig. 12.13(b)). As one can see in Fig. 12.13(c), the CHOOZ results are in agreement (within a few % error) with the assumption that there are no oscillations happening. Based on the one-mass scale dominance interpretation discussed above, one uses the disappearance formula in Tab. 12.1 to produce the exclusion plot shown in Fig. 12.13(d). To reduce systematics (by using a second close detector), there is worldwide activity to build a new reactor experiment with higher θ_{13} resolution.

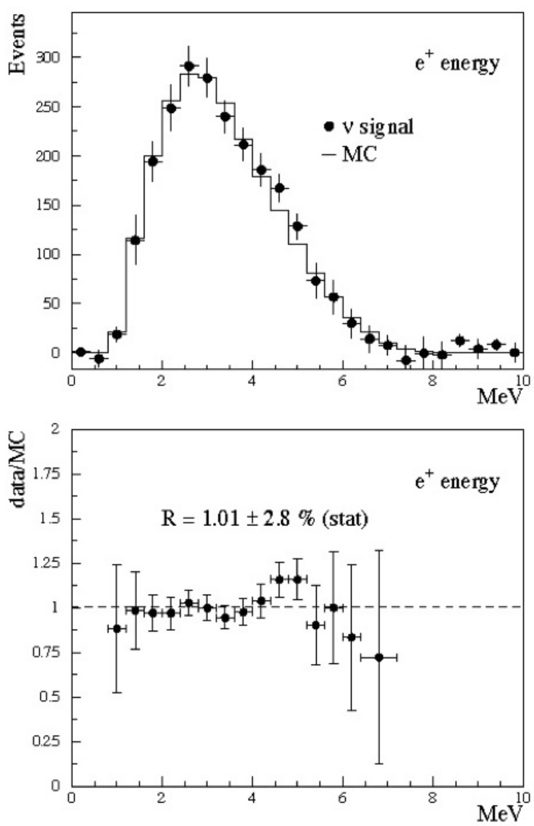
Atmospheric neutrinos: the Super-Kamiokande breakthrough. Figure 12.14(a) shows the zenith angle dependence of the number of events in the 50 kt Super-Kamiokande detector. One observes that there is a deficit in μ-like events in the up-going direction, whereas the electron-like events follow more or less the expectations. Atmospheric neutrinos with electron or muon flavor are produced as secondary (anti)particles in decays of mesons produced by cosmic rays hitting the atmosphere (see Fig. 12.15(b)). Although the primary flux is affected by large normalization uncertainties, the neutrino flavor ratio (about twice as much μ neutrinos than electron-neutrinos) is robust within a few per-cent. As we have seen, the idea is to look up and down, since the neutrino flux from opposite directions is the same, because for the opposite side the increased flux dilution ($\sim 1/r^2$) is compensated by the larger production surface ($\sim r^2$) (see Fig. 12.14(b)). The actual detection employs again charged-current interactions in the target. It is possible to distinguish the muonic from the electronic final state by means of the Cherenkov ring sharpness: Producing showers in the target, the electron/positron smears out its Cherenkov ring (see Fig. 12.16). This method does not allow for charge discrimination and τ events are not reconstructed. A summary of the zenith distributions at Super-Kamiokande is shown in Fig. 12.17. One can observe that the distribution of electronic events is more or less in agreement with the expectation for no mixing, while there is a deficit in muonic events from below, compared to the expectation for no oscillation. Observations over several decades of L/E show the same results. How to interpret them? In terms of oscillations this means that the channel $\nu_\mu \to \nu_e$ is non-existing or subdominant (in agreement with CHOOZ) and that the channel $\nu_\mu \to \nu_\tau$ is dominant. Recall that the one-mass scale



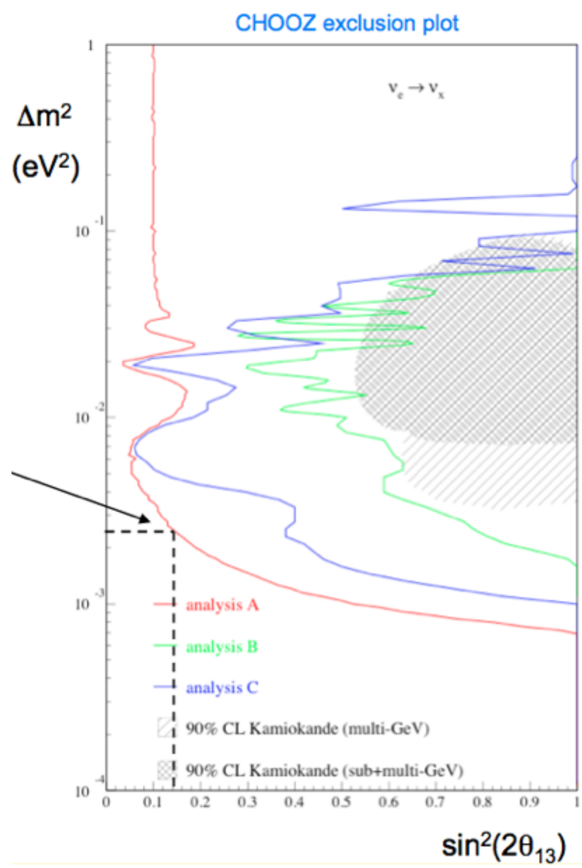
(a)



(b)

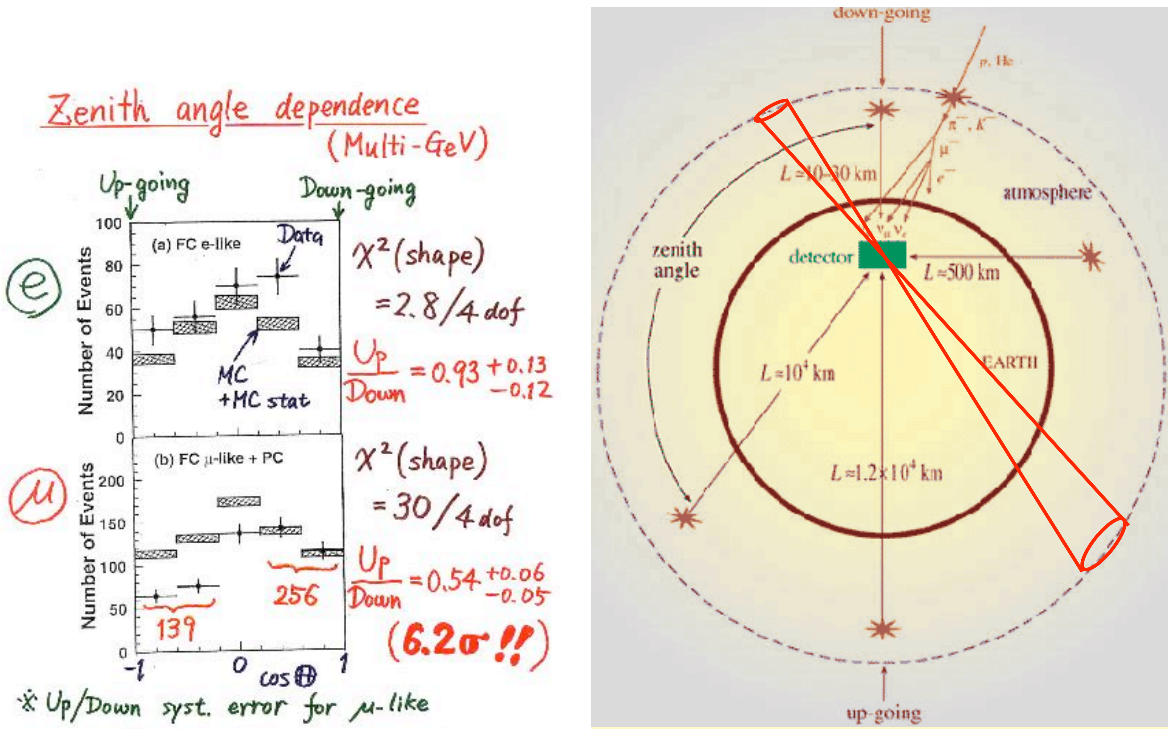


(c)



(d)

Figure 12.13: *Results of the short-baseline reactor experiment CHOOZ.*



(a) (b)

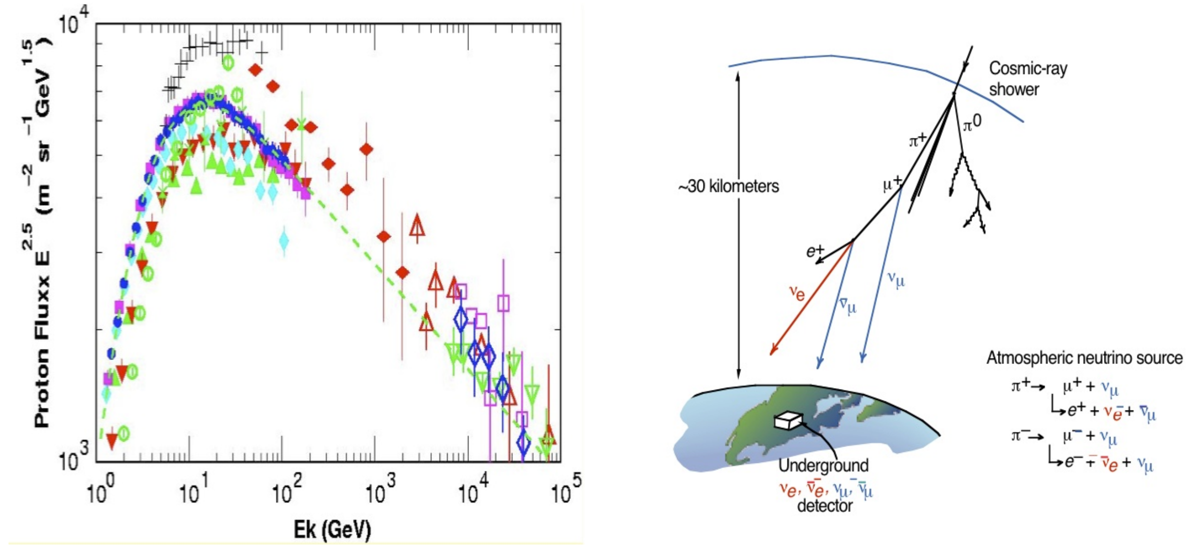
Figure 12.14: *Zenith angle dependence of μ-like events in the Super-Kamiokande experiment.* Source: T. Kajita at Neutrino '98, Takayama.

approximation for $\theta_{13} = 0$ reads

$$P(\nu_\mu \to \nu_\tau) = \sin^2 2\theta_{23} \sin^2 \left(\frac{\Delta m^2 L}{4E} \right). \tag{12.30}$$

The results are consistent with other atmospheric experiments using different techniques (MACRO, Soudan2) but with lower statistics. Performing a dedicated L/E analysis in Super-Kamiokande, it is even possible to "see" one half-period of the oscillation (distorted by convolution with resolution, see Fig. 12.18(a)). Overall, the Super-Kamiokande measurement yields strong constraints on the parameters Δm^2 and θ_{23} (see Fig. 12.18(b)).

Long-baseline neutrino experiments. With long-baseline experiments it is possible to reproduce atmospheric μ-neutrino physics under controlled conditions (known flux etc.). Sketches of such experiments in the US, Japan and Europe are shown in Fig. 12.19. An example of neutrino beam production is shown in Fig. 12.20. Protons hitting a fixed target produce pions which in turn decay into muons and muon neutrinos. To obtain a focussed beam, the pions have to be focussed in the first place. This is achieved with magnetic lenses, so called "horns". Due to the production mode via pion decay, there is a small contamination by electron neutrinos. Far detection of the neutrinos is achieved by



(a) (b)

Figure 12.15: *Production of atmospheric neutrinos.* The absolute value of the primary flux is not known precisely (a), but the flavor ratio is robust within a few percent (b).

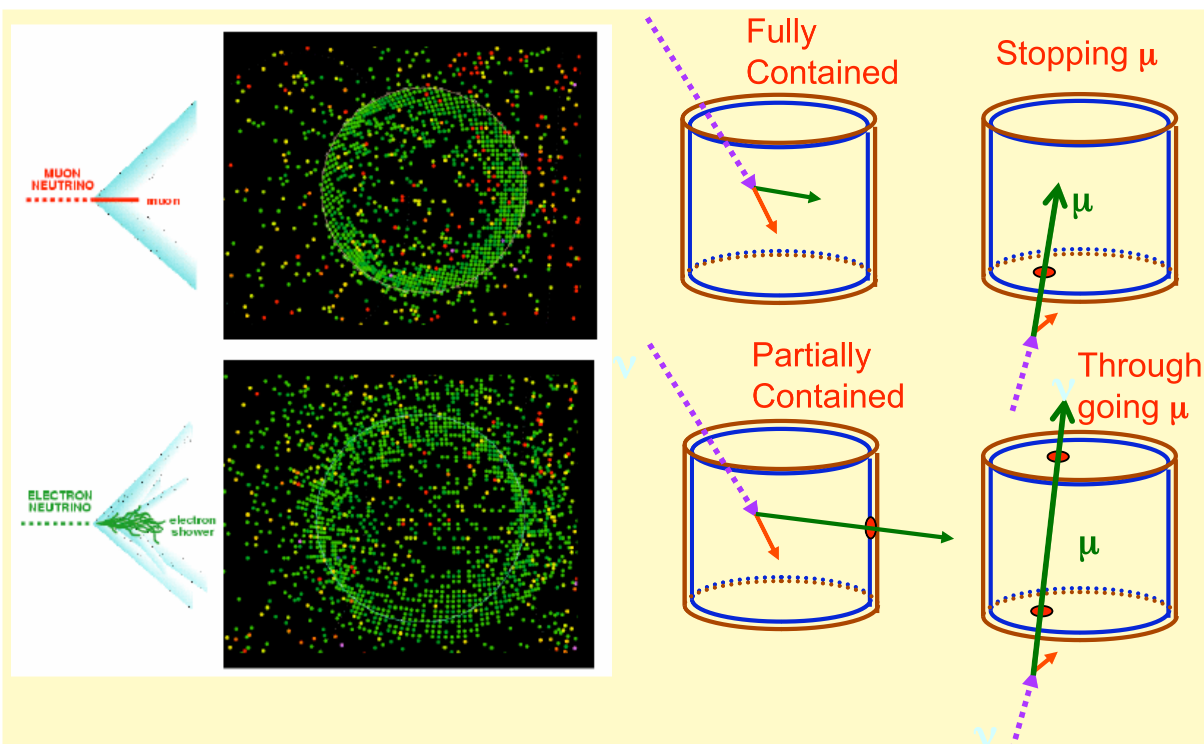


Figure 12.16: *Detection in Super-Kamiokande.* Parent neutrinos are detected via charged-current interactions in the water target.

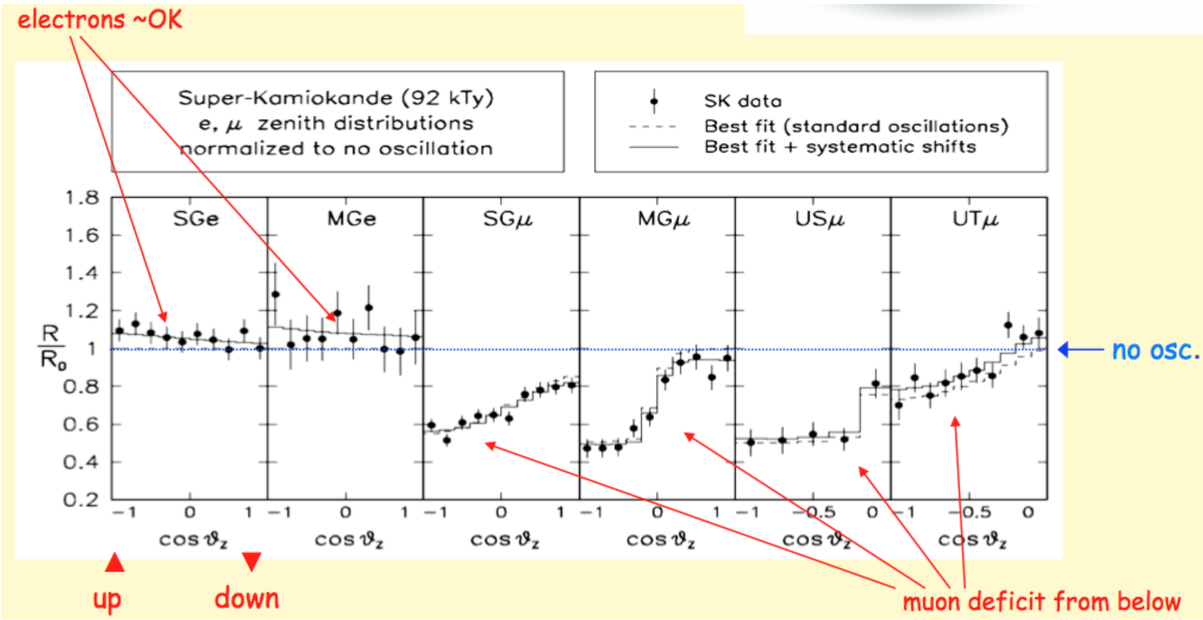
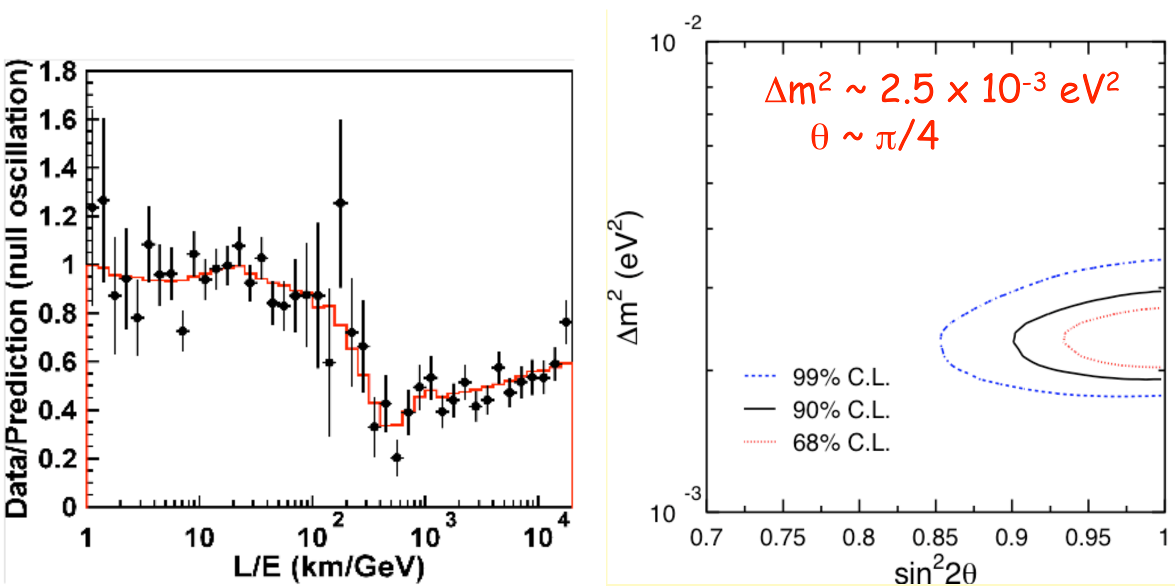


Figure 12.17: *Super-Kamiokande results on atmospheric neutrinos.*



(a) (b)

Figure 12.18: *Super-Kamiokande results on oscillation period (a) and constraints on the parameters Δm^2 and θ.*

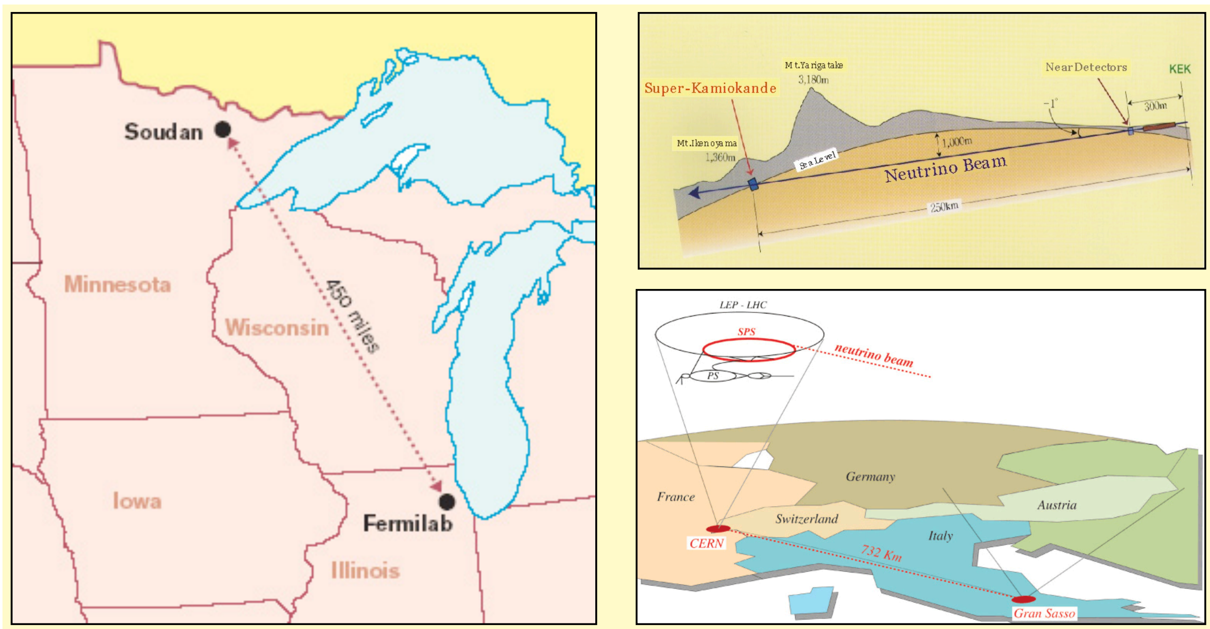


Figure 12.19: *Examples of long-baseline neutrino experiments.* Source: [71].

the Cherenkov technique at Super-Kamiokande (K2K and T2K) or by a steel/scintillator detector in the case of MINOS. Both experiments are supplemented by near detectors to control the flux of muon neutrinos for normalization. Once more the dominant probability is $P(\nu_\mu \to \nu_\tau) = \sin^2 2\theta_{23} \sin^2(\Delta m^2 L/4E)$ such that the results can be compared to the atmospheric results. Combining the corresponding exclusion plots, one finds the oscillation parameters to be consistent among the experiments (see Fig. 12.21). The OPERA detector searches for dominant oscillations via τ appearance. This is done using a hybrid of emulsion layers and scintillator trackers: If the tracker indicates a candidate event, the layers are scanned to document tau decays (see Fig. 12.22).

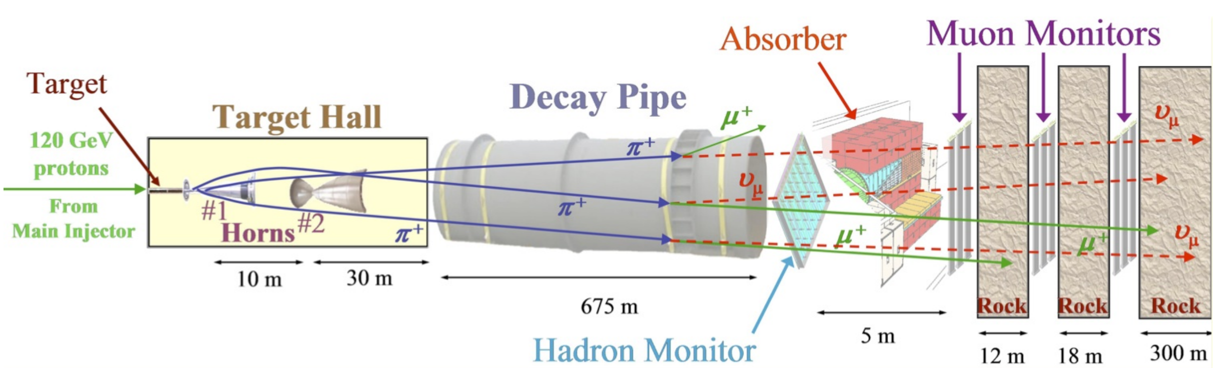


Figure 12.20: *Muon-neutrino beam production at hadron accelerators.*

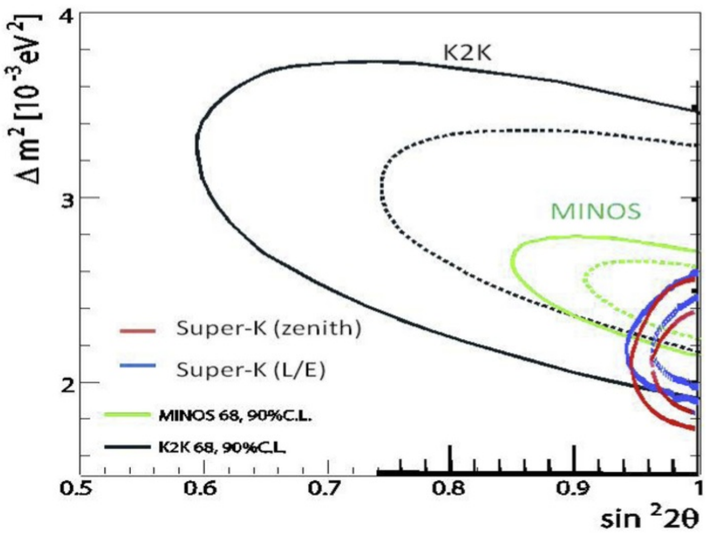


Figure 12.21: *Long-baseline neutrino experiments combination and consistency check with atmospheric results.*

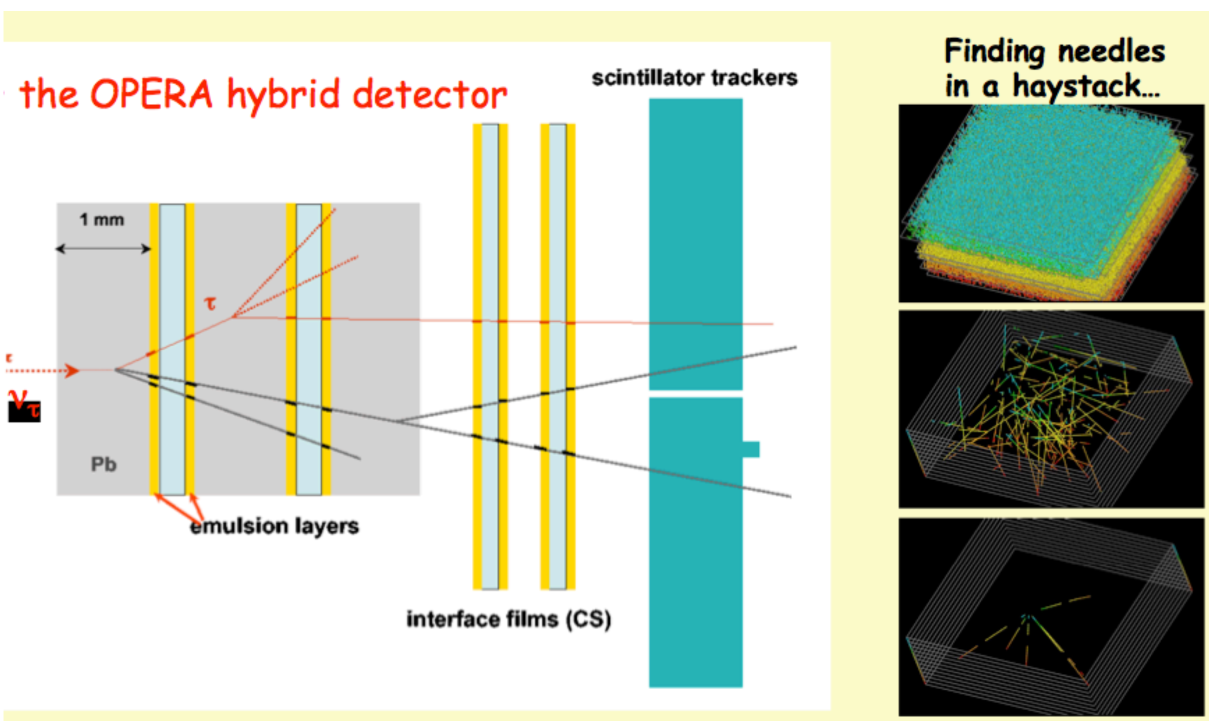


Figure 12.22: *Sketch of the OPERA detector (LHS) and of a reconstructed event (RHS).*

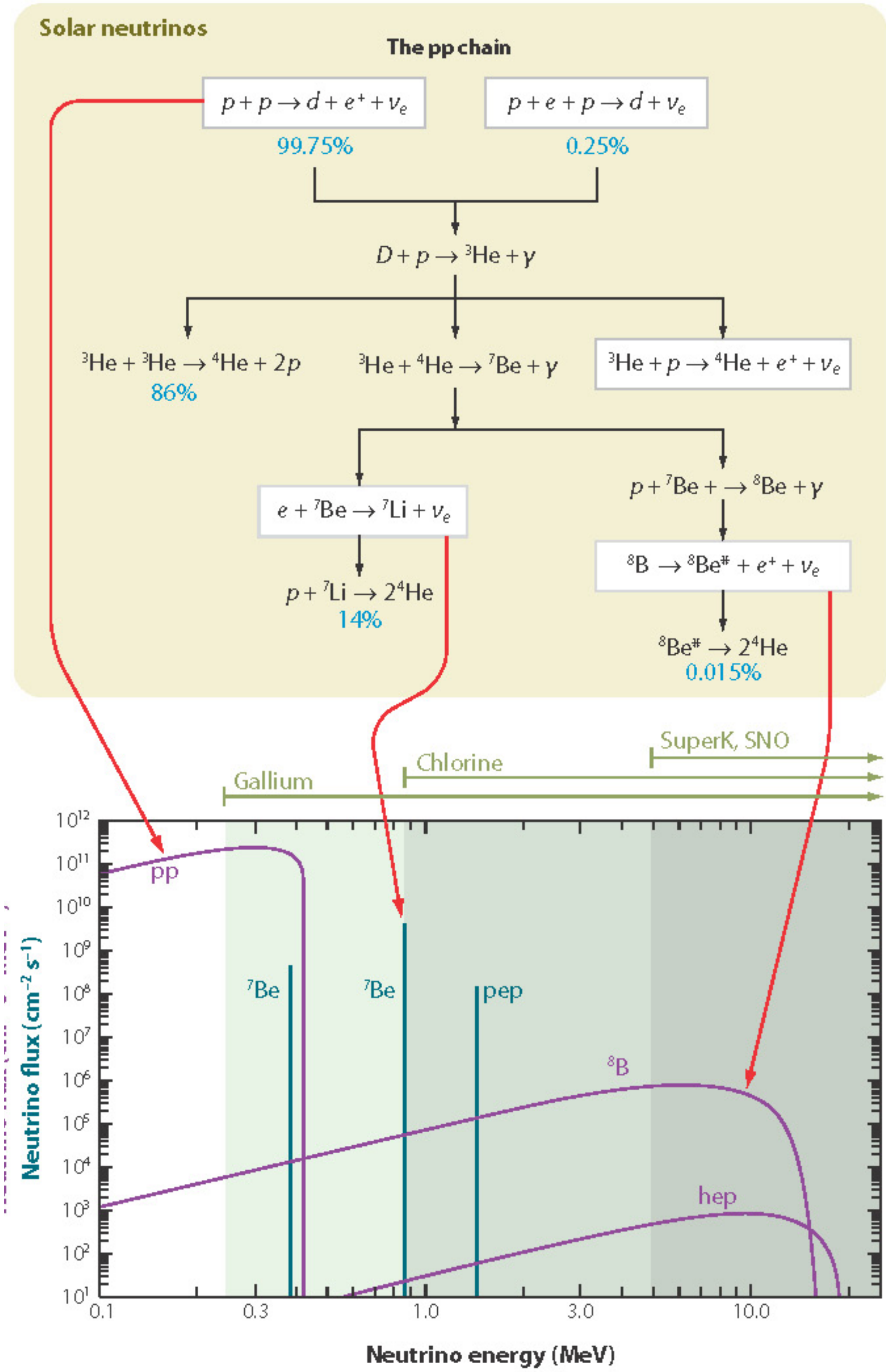


Figure 12.23: *Production of solar neutrinos in the pp cycle.*

Solar neutrinos. We now turn to experiments sensitive to the small mass splitting δm^2. Solar neutrino production proceeds via the *pp* (and CNO) cycles (see Fig. 12.23), where the energy spectrum of the neutrinos varies with the stage of their production. There are different ways to detect "solar neutrinos". In the radiochemical method, one counts the decays of unstable final-state nuclei. Advantageous is the low energy threshold of this method. Problematic is, though, the loss/integration of the energy and time information. Possible reactions for detection are

$$\begin{array}{lcr} {}^{37}\mathrm{Cl} + \nu_e \to {}^{37}\mathrm{Ar} + e^- & (\mathrm{CC}) & \text{Homestake} \\ {}^{71}\mathrm{Ga} + \nu_e \to {}^{71}\mathrm{Ge} + e^- & (\mathrm{CC}) & \text{GALLEX/GNO, SAGE.} \end{array}$$

The second detection possibility for solar neutrinos is elastic scattering:

$$\nu_x + e^- \to \nu_x + e^- \qquad (\mathrm{NC,CC}) \qquad \text{SK, SNO, Borexino}$$

where events are detected in real time with either a high energy threshold (Cherenkov, directional) or with a low threshold (scintillators). Thirdly, there is the possibility to detect solar neutrinos via interactions with deuterium, where the charged current events are detected in real time and the neutral current events are separated statistically and using neutron counters. The corresponding reactions read:

$$\begin{array}{lcr} \nu_e + \mathrm{d} \to p + p + e^- & (\mathrm{CC}) & \text{SNO} \\ \nu_x + \mathrm{d} \to p + n + \nu_x & (\mathrm{NC}) & \text{(Sudbury Neutrino Observatory).} \end{array}$$

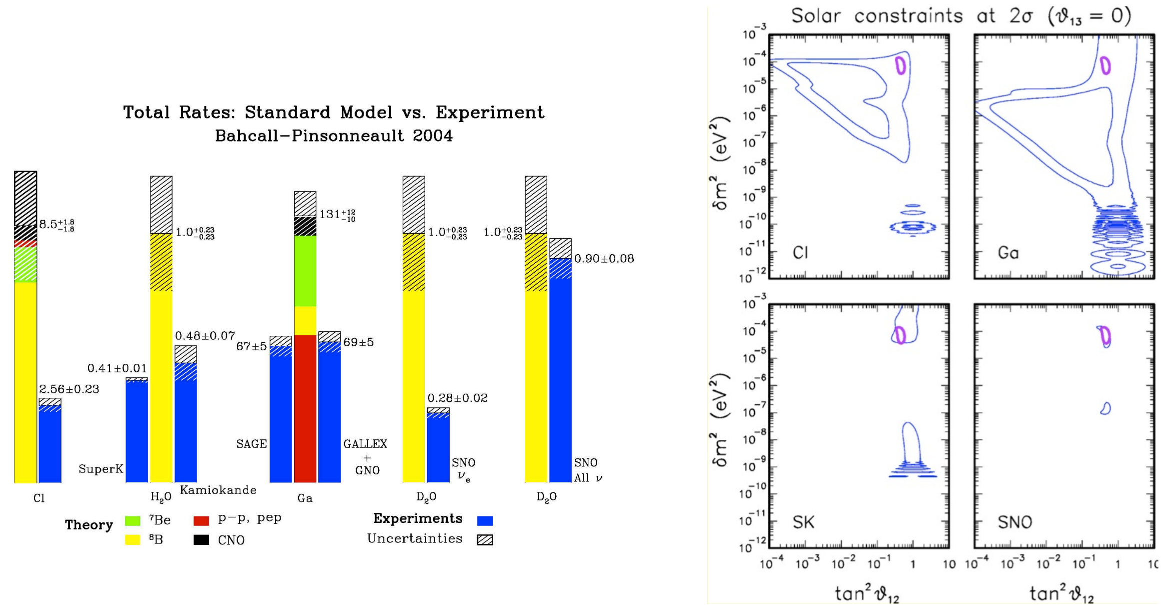
All CC-sensitive results on solar neutrinos indicated a ν_e deficit, when compared to solar model expectations (see Fig. 12.24(a)). Interpreting the results in terms of neutrino oscillations yielded solar constraints on δm^2 and θ_{12} (see Fig. 12.24(b)). A crucial role in this development was played by the Sudbury Neutrino Observatory. As we have seen, at SNO deuterium was used as target. In deuterium one can separate CC events (induced by ν_e only) from NC events (induced by $\nu_e,\ \nu_\mu,\ \nu_\tau$), and double check via elastic scattering events (due both to NC and CC). In terms of flux this means

$$\frac{\mathrm{CC}}{\mathrm{NC}} \simeq \frac{\Phi(\nu_e)}{\Phi(\nu_e) + \Phi(\nu_{\mu,\tau})}.$$

Therefore

$$\frac{\mathrm{CC}}{\mathrm{NC}} < 1 \Rightarrow \Phi(\nu_{\mu,\tau}) > 0 \Rightarrow P(\nu_e \to \nu_{\mu,\tau}) \neq 0$$

since solar neutrinos are produced exclusively as electron neutrinos. It was found that $\mathrm{CC/NC} \sim 1/3 < 1$ and the solar model turned out to be adequate. Note also that since $\mathrm{CC/NC} \sim P(\nu_e \to \nu_e) \sim 1/3 < 1/2$ this is also evidence of three-neutrino like mixing and of matter effects. A summary of neutrino mass differences and mixing parameters with their $n\sigma$ ranges from a global three-neutrino analysis is shown in Fig. 12.25.



(a) (b)

Figure 12.24: *Electron neutrino deficit in solar neutrino measurements as compared to standard solar model (a) and parameter constraints from interpretation in terms of mixing (b).*

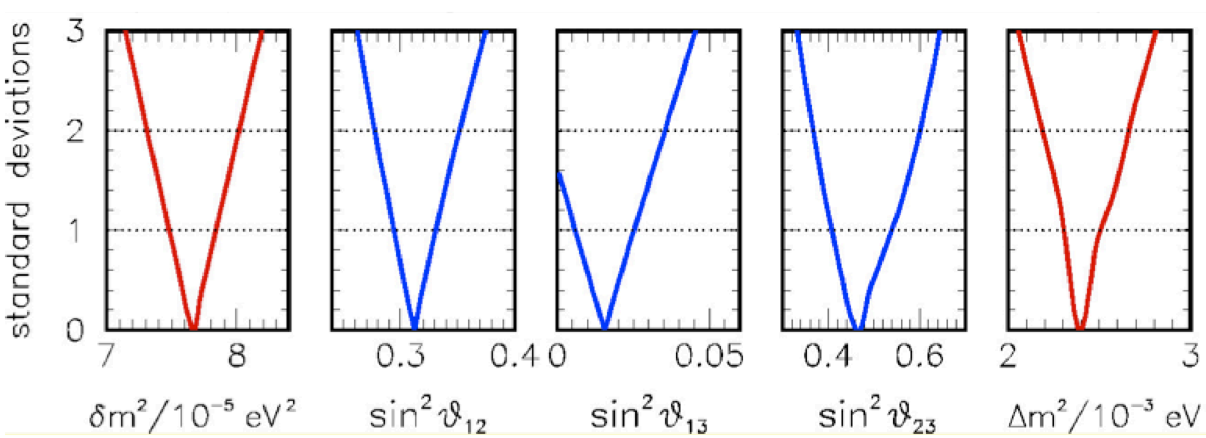


Figure 12.25: *Synopsis of neutrino mass splitting and mixing parameters.*

What are the next experimental steps in determining these parameters? First of all it is important to know θ_{13} more precisely. Since $\sin^2\theta_{13} = |U_{e3}|^2$, this is the small ν_e part of ν_3. Thus what is needed is an experiment with L/E sensitive to Δm ($L/E \sim 500\,\mathrm{km/GeV}$), and involving ν_e. One possibility is disappearance of $\bar{\nu}_e$ produced by a reactor while traveling $L \sim 1.5\,\mathrm{km}$. This process depends on θ_{13} alone (recall Eq. (12.21)):

$$P(\bar{\nu}_e\,\text{disappearance}) = \sin^2 2\theta_{13}\sin^2\frac{\Delta m^2 L}{4E}.$$

Another interesting possibility is the measurement of $P(\nu_\mu \to \nu_e)$ for ν_μ produced by accelerators with L several hundred kilometers. This process depends on θ_{13}, θ_{23}, on whether the hierarchy is normal or inverted and on whether CP is violated (δ).

12.4.3 Absolute masses

As we have seen, neutrino oscillations constrain neutrino mixings and mass splittings but not the absolute mass scale. E. g., one can choose the lightest neutrino mass as a free parameter. However, the lightest neutrino mass cannot be directly observed. There are three realistic observables to attack neutrino masses:

1. β decay. A non-vanishing neutrino mass can affect the spectrum endpoint in β decay.
2. Neutrinoless double beta decay. This is only possible for Majorana neutrinos, we will not discuss this possibility here.
3. Cosmology. Non-vanishing neutrino masses can affect large scale structures in the standard model of cosmology, constrained by CMB and other data. Again, we will not go into detail here.

One can use the high energy end of a beta decay spectrum like the one shown in Fig. 11.1(a) to search for neutrino masses. Since beta decay is essentially emission and decay of a W boson, the matrix element squared is proportional to G_F^2. Thus the decay rate reads $d\Gamma \propto G_F^2 \times$ (phase space factor). The energy spectrum can be written as

$$\frac{d\Gamma}{dE_e} \propto \begin{cases} G_F^2 p_e E_e (Q-E_e)^2 & (m_\nu = 0) \\ G_F^2 p_e E_e (Q-E_e)\sqrt{(Q-E_e)^2 + m_\nu^2} & (m_\nu > 0) \end{cases}$$

where Q is the high energy endpoint of the electron spectrum. Tritium is well suited for this experiment, since Q (18.57 keV) and half life (12.32 y) are low. The reaction reads as follows:

$${}^3\mathrm{H} \to {}^3\mathrm{He} + e^- + \bar{\nu}_e.$$

Figure 12.26 shows a close-up of the spectrum around its endpoint. Note that only a very small fraction of all events lies in the region sensitive to the neutrino mass. To detect its effect, good energy resolution is needed. In fact, E_0 is not Q, but the end point value corrected by a recoil contribution which can be assumed to be constant in the region of interest ($E_{\mathrm{rec}} = 1.72\,\mathrm{eV}$): $E_0 = Q - E_{\mathrm{rec}}$ (see [74, 75] for details).

There are three mass eigenstates whose eigenvalues cannot be individually resolved by this experiment: Beta-decay produces electron neutrinos; as we have seen, these are superpositions of the three mass eigenstates ν_i. Therefore, the experiment is sensitive to the sum of the masses m_i, weighted by the squared mixing coefficients $|U_{ei}|^2$:

$$m_\beta = \sqrt{\mathrm{c}_{13}^2 \mathrm{c}_{12}^2 m_1^2 + \mathrm{c}_{13}^2 \mathrm{s}_{12}^2 m_2^2 + \mathrm{s}_{13}^2 m_3^2}$$

which is called “effective electron neutrino mass”. Note that the mass eigenstate with the largest electron flavor component is ν_1, $|U_{e1}|^2 \simeq \cos^2\theta_{12} \simeq 0.7$, and it cannot be excluded that ν_1 is nearly massless (in the normal hierarchy, see Fig. 12.4). A historical summary of the mass limits obtained by the beta-decay method is shown in Fig. 12.27. Latest bounds are at the level of $2\,\mathrm{eV}$.

The significant improvement in the neutrino mass sensitivity at the Troitsk and the Mainz experiments (compared to the older ones) is due to so-called MAC-E-Filters (Magnetic Adiabatic Collimation with an Electrostatic Filter) [74, p. 17]. Figure 12.28 shows the main features of the MAC-E-Filter. β electrons emitted by the tritium source in the LHS solenoid into the forward hemisphere are guided magnetically on a cyclotron motion along the magnetic field lines into the spectrometer, resulting in an accepted solid angle of nearly 2π. On their way into the center of the spectrometer the magnetic field B drops adiabatically by several orders of magnitude keeping the ratio of cyclotron energy and magnetic field constant: $E_\perp/B = \mathrm{const.}$ Therefore, nearly all cyclotron energy $E_\perp$ is transformed into longitudinal motion giving rise to a broad beam of electrons flying almost parallel to the magnetic field lines. Finally, the parallel beam of electrons is energetically analyzed by applying an electrostatic barrier. The KATRIN experiment, currently under construction, is expected to improve the mass limit by one order of magnitude to about $0.2\,\mathrm{eV}$.

Neutrino physics is a vast field, accordingly important topics like Majorana neutrinos, neutrino-less double-beta decay, cosmological bounds on the neutrino mass and future perspectives in neutrino physics are not discussed here (see lecture on neutrino physics by Prof. Rubbia[5]).

[5] `http://neutrino.ethz.ch/Vorlesung/HS2009/`

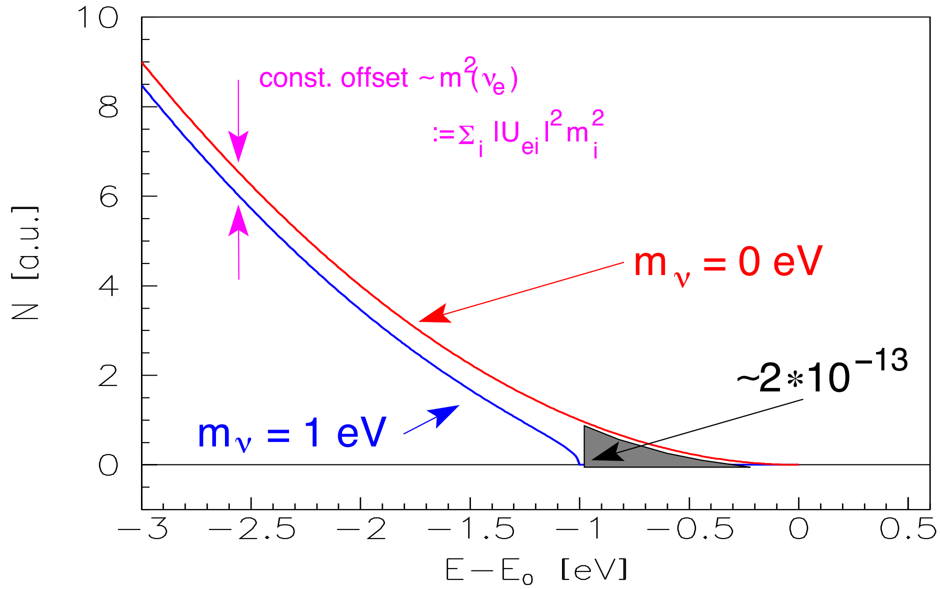


Figure 12.26: *Close-up of the high-energy end of the beta decay spectrum.* In the case of tritium the shaded area corresponds to a fraction of about $2 \cdot 10^{-13}$ events. Source: [74, p. 12].

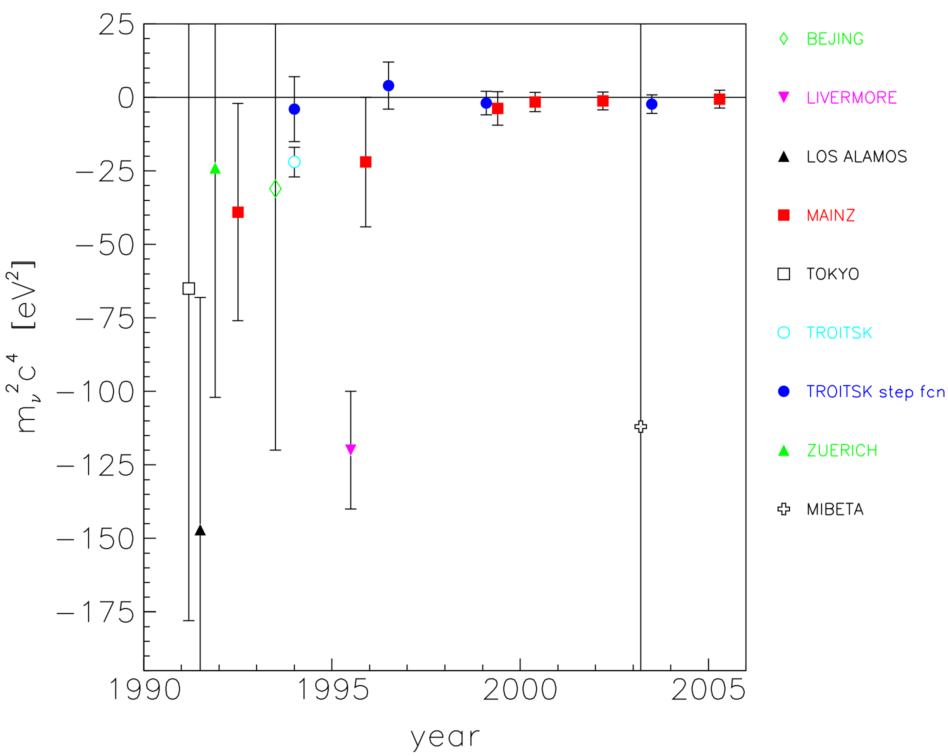


Figure 12.27: *Recent results of tritium beta decay experiments on the effective electron neutrino mass.* Source: [74, p. 15].

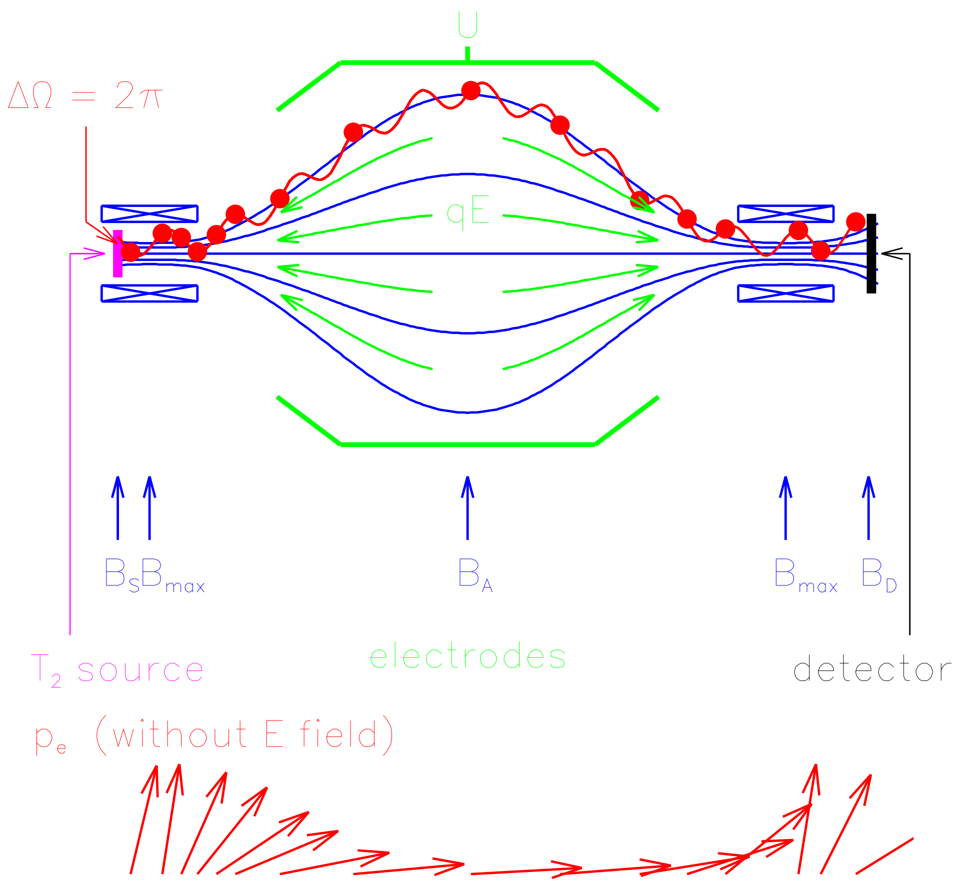


Figure 12.28: *Sketch of the MAK-E-Filter.* Source: [74, p. 17].

Bibliography

[1] F. Halzen and A. Martin. *Quarks & Leptons*. Wiley, 1984.

[2] I. Aitchison and A. Hey. *Gauge Theories in Particle Physics*. Taylor & Francis, 2002.

[3] A. Seiden. *Particle Physics*. Addison Wesley, 2004.

[4] O. Nachtmann. *Phänomene und Konzepte der Elementarteilchenphysik*. Vieweg, 1986.

[5] R. Hagedorn. *Relativistic Kinematics*. W. A. Benjamin, 1963.

[6] E. Byckling and K. Kajantie. *Particle Kinematics*. Wiley, 1973.

[7] F. Hinterberger. *Physik der Teilchenbeschleuniger und Ionenoptik*. Springer, 2008.

[8] C. Amsler. *Kern- und Teilchenphysik*. UTB, 2007.

[9] http://pdg.lbl.gov/2006/hadronic-xsections/hadron.html.

[10] http://www-cdf.fnal.gov/physics/new/qcd/run2/ue/chgjet/etaphi.html.

[11] http://en.wikipedia.org/wiki/Pseudorapidity.

[12] http://www.phys.ufl.edu/∼rfield/cdf/chgjet/dijet.eps.

[13] G. Arnison et al. *Physics Letters B*, Jan 1983.

[14] M. E. Peskin and D. V. Schroeder. *An Introduction to Quantum Field Theory*. Westview, 1995.

[15] E. Freitag and R. Busam. *Complex analysis*. Springer, 2005.

[16] G. Gabrielse. Measurements of the electron magnetic moment. In B. L. Roberts and W. J. Marciano, editors, *Lepton Dipole Moments: The Search for Physics Beyond the Standard Model*, volume 20 of *Advanced Series on Directions in High Energy Physics*. World Scientific, 2009.

[17] http://upload.wikimedia.org/wikipedia/commons/thumb/b/b6/penning_trap.svg/800px-penning_trap.svg.png.

[18] L. D. Landau and E. M. Lifschitz. *Quantenmechanik*. Harri Deutsch, 1986.

[19] G. Gabrielse. Determining the fine structure constant. In B. L. Roberts and W. J. Marciano, editors, *Lepton Dipole Moments: The Search for Physics Beyond the Standard Model*, volume 20 of *Advanced Series on Directions in High Energy Physics*. World Scientific, 2009.

[20] R.P. Feynman. *QED: The strange theory of light and matter*. Princeton University Press, 1988.

[21] A. Ali and P. Soeding. *High energy electron positron physics*. World Scientific, 1988.

[22] P. Dittmann and V. Hepp. *Particles and Fields*, 10, 1981.

[23] H. Lipkin. *Lie groups for pedestrians*. Dover, 2002.

[24] J. M. Lee. *Introduction to smooth manifolds*. Springer, 2006.

[25] V. Barnes et al. *Physical Review Letters*, Jan 1964.

[26] Particle data group. http://pdg.lbl.gov/.

[27] G. Dissertori, I. K. Knowles, and M. Schmelling. *Quantum Chromodynamics: High Energy Experiments and Theory*. Oxford University Press, 2nd edition, 2009.

[28] R. K. Ellis, W. J. Stirling, and B. R. Webber. *QCD and Collider Physics*. Cambridge University Press, 1996.

[29] S. Bethke. *arXiv:hep-ex/0606035*, 2007.

[30] S. Bethke. *arXiv:0908:1135*, 2009.

[31] JADE collaboration. *Phys. Lett. B*, 123:460, 1983.

[32] Yu. L. Dokshitzer. *J. Phys. G*, 17:1537, 1991.

[33] N. Brown and W. J. Stirling. *Phys. Lett. B*, 252:657, 1990.

[34] Yu. L. Dokshitzer et al. *JHEP*, 9708:1, 1997.

[35] M. Cacciari and G. P. Salam. *arXiv:0707.1378*, 2007.

[36] G. P. Salam and G. Soyez. *JHEP*, 05:86, 2007.

[37] M. Cacciari and G. P. Salam. *Phys. Lett. B*, 641:57, 2006.

[38] B. Gary. CTEQ Summer School, Madison WI, June 26, 2009. *http://www.phys.psu.edu/∼cteq/schools/summer09/talks/gary-cteq-madison-2009-talk.pdf*.

[39] Wu-Ki Tung. Perturbative QCD and the parton structure of the nucleon. 2001.

[40] S. Bethke. Eleven years of QCD at LEP. *EPJ direct*, 2002.

[41] G. C. Blazey et al. Run II jet physics: Proceedings of the run II QCD and weak boson physics workshop. *arXiv:hep-ex/0005012*, 2000.

[42] G. Hanson et al. SLAC-LBL Collaboration. *Phys. Rev. Lett.*, 35:1609, 1975.

[43] TASSO Collaboration. *Phys. Lett. B*, 86(2):243, 1979.

[44] MARK-J Collaboration. *Phys. Rev. Lett.*, 43(12):830, 1979.

[45] PLUTO Collaboration. *Phys. Lett. B*, 86:418, 1979.

[46] JADE collaboration. *Phys. Lett. B*, 91:142, 1980.

[47] ALEPH Collaboration. *Phys. Rept.*, 294:1, 1998.

[48] L3 Collaboration. *Phys. Lett. B*, 248:227, 1990.

[49] G. Dissertori, A. Gehrmann-De Ridder, T. Gehrmann, E.W.N. Glover, G. Heinrich, and H. Stenzel. *arXiv:0712.0327*, 2008.

[50] G. Dissertori, A. Gehrmann-De Ridder, T. Gehrmann, E.W.N. Glover, G. Heinrich, G. Luisoni, and H. Stenzel. *arXiv:0906.3436*, 2009.

[51] G. Dissertori, A. Gehrmann-De Ridder, T. Gehrmann, E.W.N. Glover, G. Heinrich, and H. Stenzel. *arXiv:0806.4601*, 2008.

[52] S. Kluth. *arXiv:hep-ex/0603011*, 2006.

[53] A. Grozin. *Lectures on* QED *and* QCD. World Scientific, 2007.

[54] G. Kane and A. Pierce. *Perspectives on LHC physics.* World scientific, 2008.

[55] ATLAS Collaboration. *arXiv:0901.0512 [hep-ex]*, 2009.

[56] CMS Collaboration. *J. Phys. G*, 34:995, 2007.

[57] D. E. Soper S. D. Ellis. Successive combination jet algorithm for hadron collisions. *Phys. Rev. D*, 48, 1993.

[58] E. Cattaruzza, A. Del Fabbro, and D. Treleani. 10.1103/physrevd.72.034022. *Phys. Rev. D*, 2005.

[59] M. Böhm, A. Denner, and H. Joos. *Gauge theories of the strong and electroweak interaction.* Teubner, 2001.

[60] J. R. Reitz. 10.1103/physrev.77.10. *Phys. Rev.*, 1950.

[61] J. Steinberger. http://ccdb4fs.kek.jp/cgi-bin/img/allpdf?199012067. 1990.

[62] The ALEPH Collaboration et al. *hep-ex/0509008*, 2005.

[63] http://lepewwg.web.cern.ch/LEPEWWG/.

[64] DØ collaboration. Search for the Higgs boson in $H \to WW^{(*)}$ decays in $p\bar{p}$ collisions at $\sqrt{2} = 1.96\,\mathrm{TeV}$. *arXiv:hep-ex/0508054v1*, 2005.

[65] http://tevnphwg.fnal.gov/results/SMHPubWinter2010/index.html.

[66] M. Fukugita and T. Yanagida. *Physics of neutrinos*. Springer, 2003.

[67] E. Lisi. http://indico.cern.ch/getFile.py/access?sessionId=4&resId=0&materialId=0&confId=80223. *Lectures at CHIPP PhD school*, 2010.

[68] E. Lisi. http://indico.cern.ch/getFile.py/access?sessionId=6&resId=0&materialId=2&confId=80223. *Lectures at CHIPP PhD school*, 2010.

[69] E. Lisi. http://indico.cern.ch/getFile.py/access?sessionId=6&resId=1&materialId=1&confId=80223. *Lectures at CHIPP PhD school*, 2010.

[70] B. Kayser. On the quantum mechanics of neutrino oscillation. *Phys. Rev. D*, 24(1), 1981.

[71] E. Lisi. Lectures at CHIPP PhD school. 2010.

[72] http://neutrino.ethz.ch/Vorlesung/HS2009/attachments/003_nt.pdf.

[73] K. Zuber. *Neutrino Physics*. Institute of Physics, 2004.

[74] C. Weinheimer. http://arxiv.org/pdf/0912.1619. 2009.

[75] E. W. Otten and C. Weinheimer. *Rep. Prog. Phys.*, 71, 2008.

Index