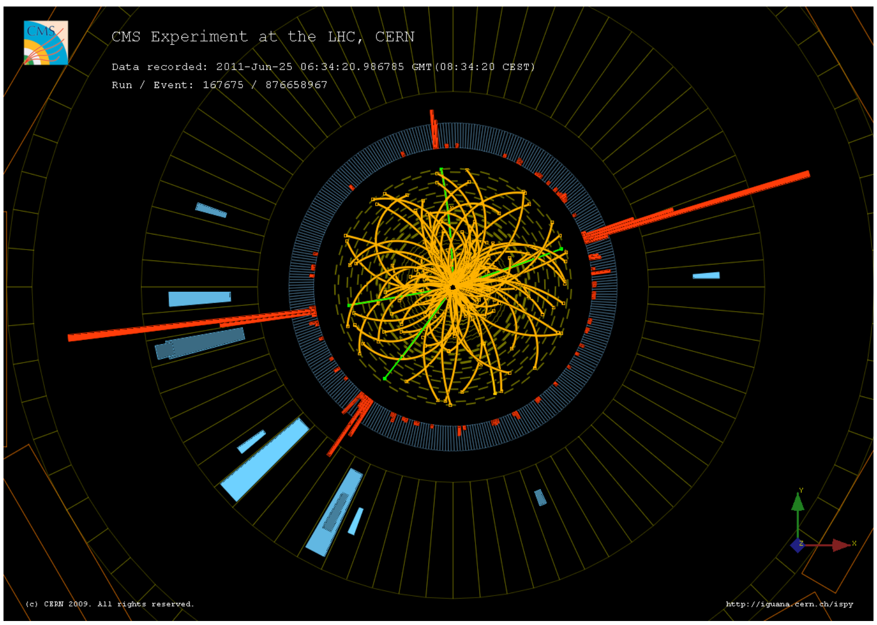
Phenomenology of Particle Physics II

ETH Zurich and University of Zurich

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

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.

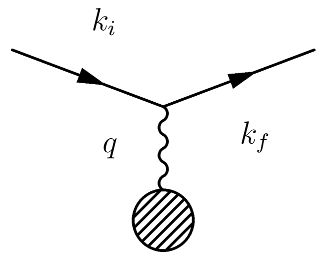


Figure 9.1: Probing a charge distribution

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.

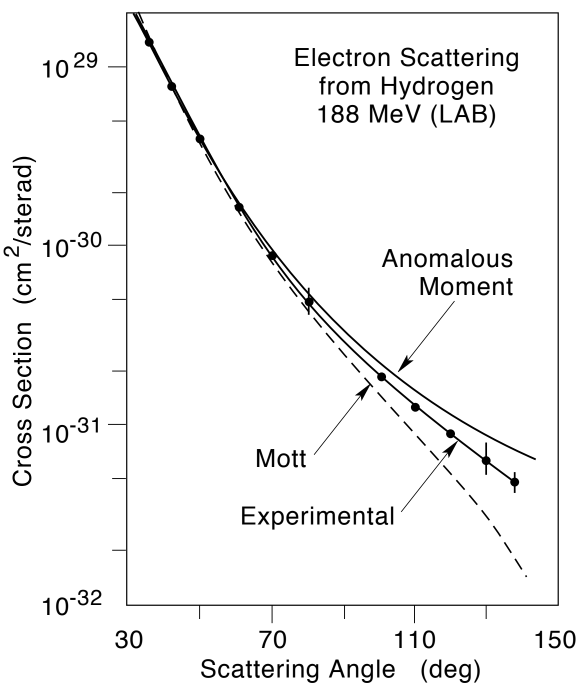


Figure 9.2: Mott cross section (dashed line) and compared to the experimental data form electron-hydrogen scattering. The measurement disagrees with the point-linke cross section at large scattering angles.

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,

$$\begin{aligned}\overline{|\mathcal{M}_{fi}|^2} &= \frac{e^4}{q^4} L_{e^-}^{\mu\nu} L_{\mu\nu}^{\mu^-} \\ &= \frac{8e^4}{q^4} 2M^2 E'E \left(\cos^2(\theta/2) - \frac{q^2}{2M^2}\sin^2(\theta/2)\right),\end{aligned}$$

where E' is the energy of the scattered electron, and the transferred momentum,

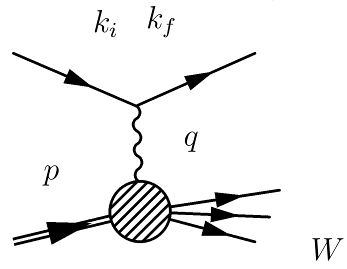
$$q^2 \approx -2k \cdot k' \approx -4EE' \sin^2(\theta/2),$$

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. 50). 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,

$$e^-p \to e^-\Delta^+ \to e^-p\pi^0$$

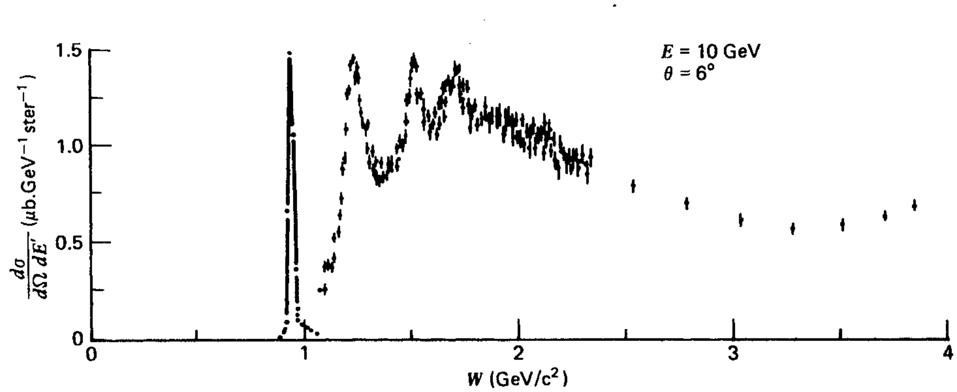


Figure 9.3: Differential cross section as a function of the invariant mass W.

To calculate the $e^{-}p$-scattering, one makes the substitution $L_{\mu^-}^{\mu\nu} \to W_p^{\mu\nu}$, where,

$$W_p^{\mu\nu} = -W_1 g^{\mu\nu} + \frac{W_2}{M^2} p^\mu p^\nu + \frac{W_4}{M^2} q^\mu q^\nu + \frac{W_5}{M^2}(p^\mu q^\nu + q^\mu p^\nu), \tag{9.4}$$

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 $\partial_\mu j_p^\mu = 0$, 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) :

$$W_p^{\mu\nu} = W_1 \left(-g^{\mu\nu} + \frac{q^\mu q^\nu}{q^2}\right) + \frac{W_2}{M^2}\left(p^\mu - \frac{p \cdot q}{q^2} q^\mu\right)\left(p^\nu - \frac{p \cdot q}{q^2} q^\nu\right). \tag{9.5}$$

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,

$$x = -\frac{q^2}{2p\cdot q} = \frac{Q^2}{2M\nu} \text{ : the Bjorken scaling } x\text{-variable,} \qquad 0 \le x \le 1,$$
$$y = \frac{p\cdot q}{p\cdot k_i}, \qquad 0 \le y \le 1.$$

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

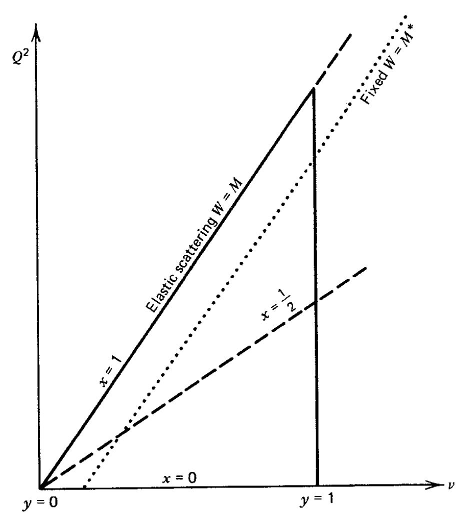


Figure 9.4: Allowed kinematical region of the Q^2-ν-plane.

kinematic region in the Q^2-ν-plane.

Using the hadron tensor, Eq. (9.5), the scattering matrix element is,

$$L^{e^-}_{\mu\nu} W_p^{\mu\nu} = 4EE' \left(W_2(Q^2,\nu)\cos^2(\theta/2) + W_1(Q^2,\nu)\sin^2(\theta/2) \right)$$

Including the flux and phase-space factors (Sect. 2.2.4, p. 13 & 3.2.3, p. 23) 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 $\lambda \approx 1\,\mathrm{fm}$, one can "see" the proton as a single particle, whereas for, $\lambda \ll 1\,\mathrm{fm}$, the

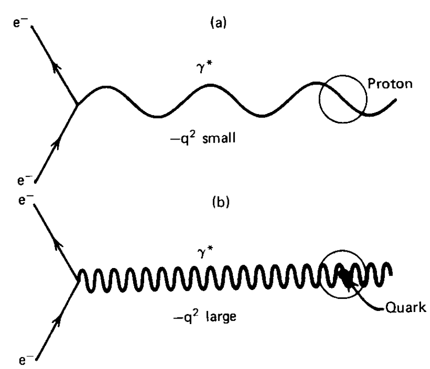
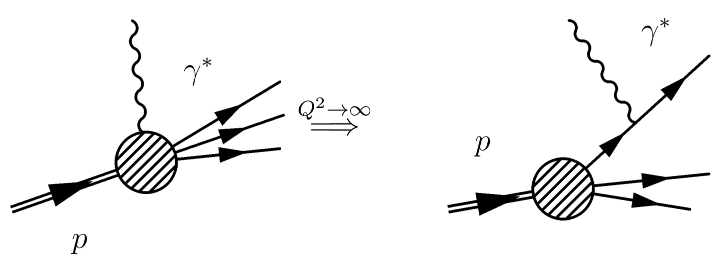


Figure 9.5: Relationship between resolution and transferred momentum.

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.



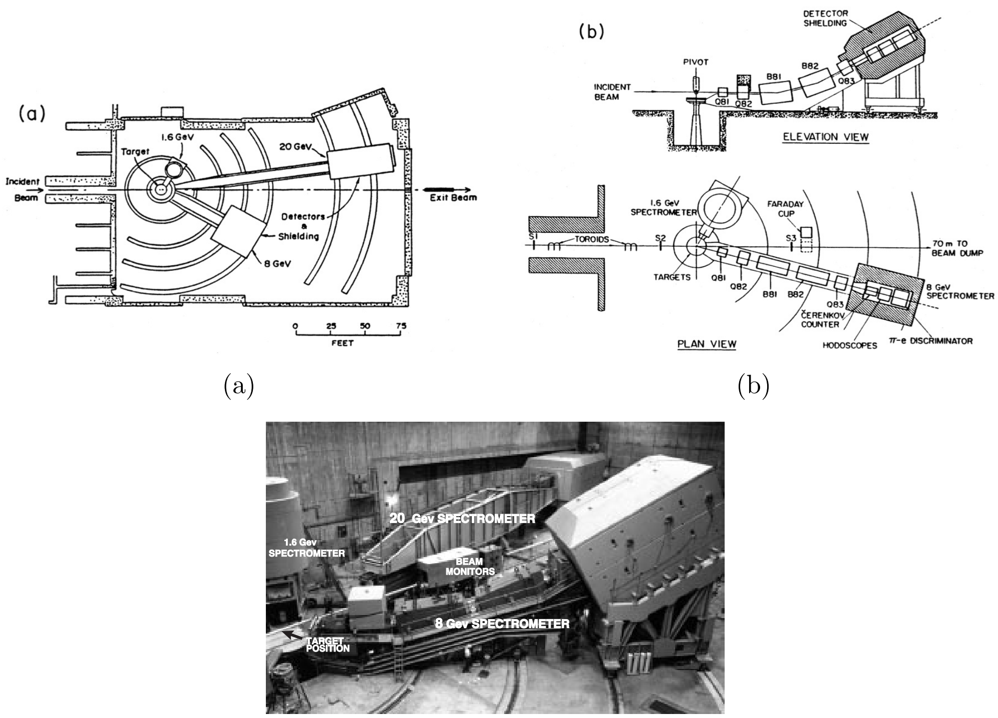
In this limit, one defines then the functions,

$$F_1(x) := \lim_{Q^2\to\infty} MW_1(Q^2,\nu), \tag{9.7}$$

$$F_2(x) := \lim_{Q^2\to\infty} \nu W_2(Q^2,\nu). \tag{9.8}$$

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,

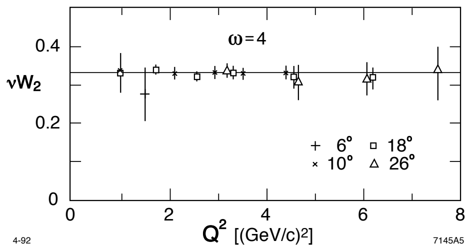
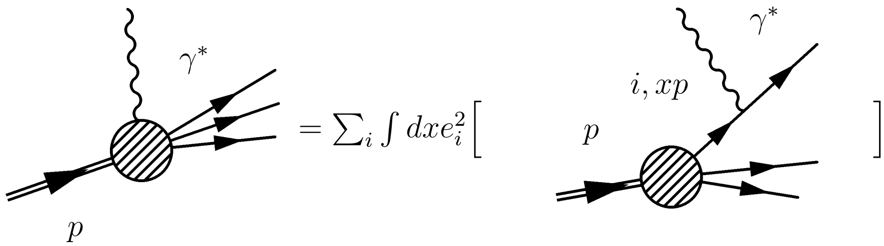


Figure 9.7: Experimental evidence for Bjorken scaling as measured at the SLAC-MIT experiment ($\omega = 1/x$).



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.

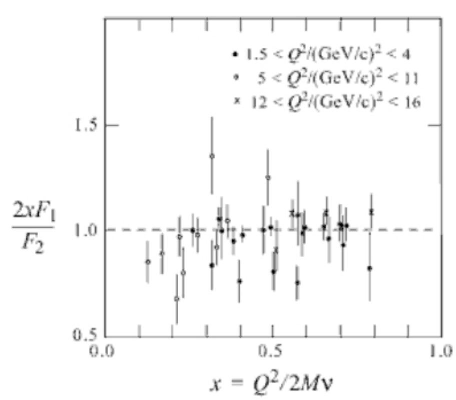


Figure 9.8: Experimental evidence for the Callan-Gross relation.

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. 131). 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,

$$\begin{aligned} u &:= u_v + u_s = u^p = d^n \\ d &:= d_v + d_s = d^p = u^n. \end{aligned}$$

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.

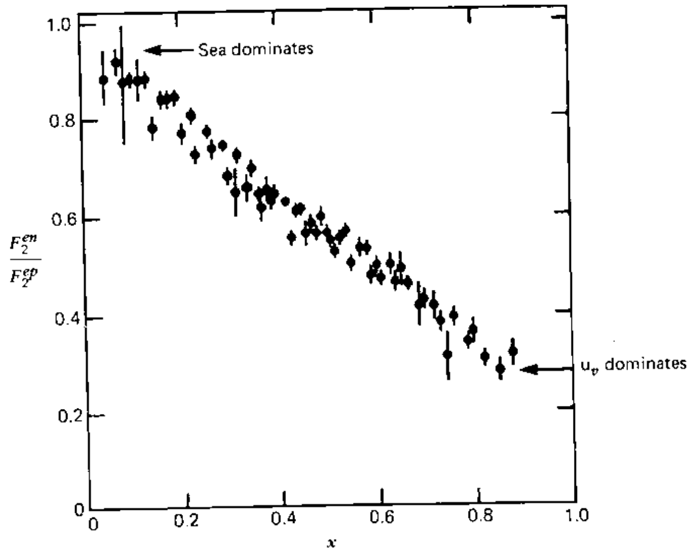


Figure 9.9: Ratio of the proton and neutron structure functions as a function of the Bjorken x-variable.

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,

$$\int_0^1 dx F_2^{ep} = \frac{4}{9}\varepsilon_u + \frac{1}{9}\varepsilon_d = 0.18,$$
$$\int_0^1 dx F_2^{en} = \frac{1}{9}\varepsilon_u + \frac{4}{9}\varepsilon_d = 0.12.$$

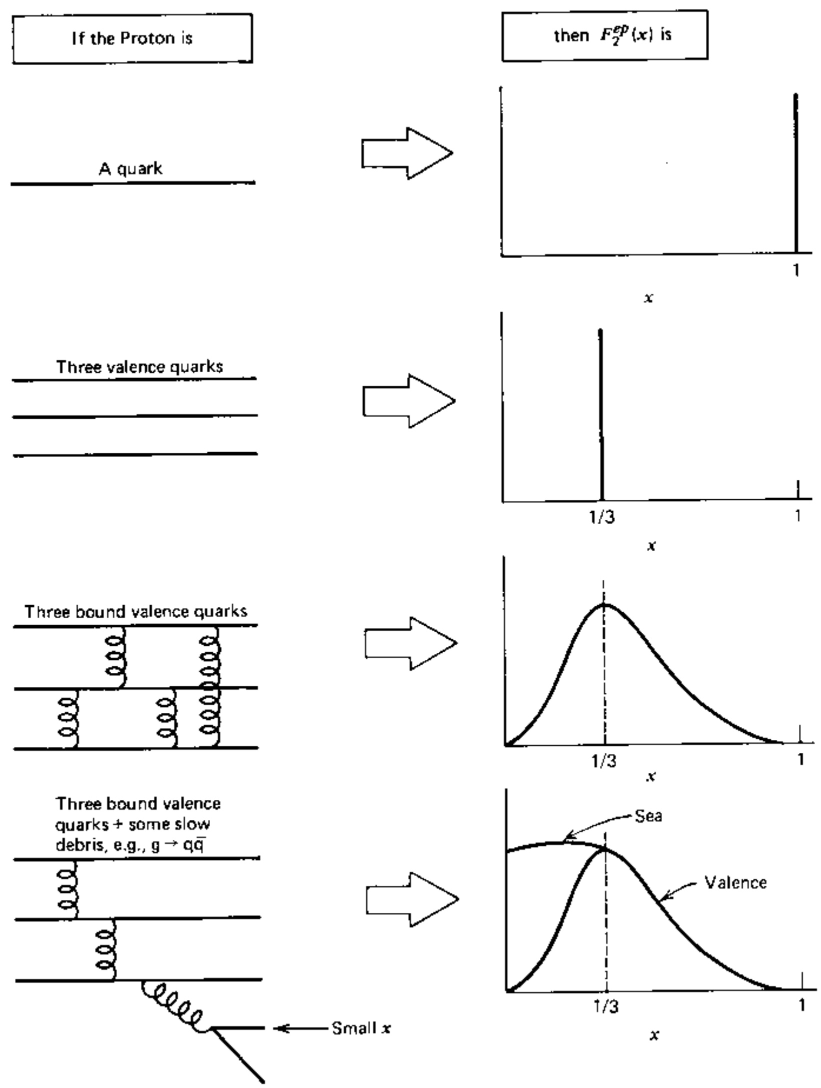


Figure 9.10: Structure functions F_2^{ep} in different scenarios of the proton structure.

Therefore,

$$\varepsilon_u = 0.36$$
$$\varepsilon_d = 0.18,$$

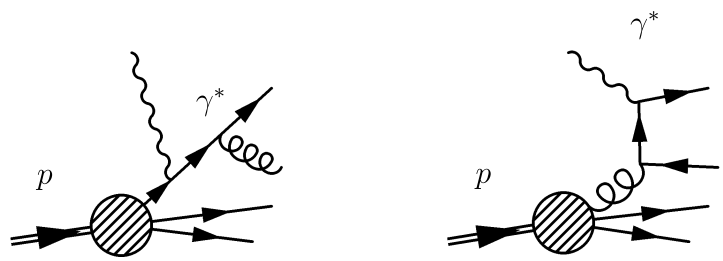
and the fraction of the proton momentum not carried by quarks is,

$$\varepsilon_g = 1 - \varepsilon_u - \varepsilon_d = 0.46.$$

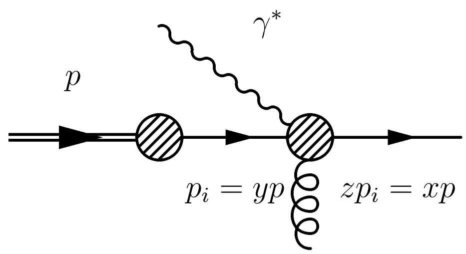
Almost half of the proton momentum is carried by electrically uncharged partons. By repeating the scattering experiments with neutrinos instead of electrons, one observes that these uncharged partons do not interact weakly either. The parton carrying the missing momentum is now known as the **gluon**, 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 $E_e = 30\,\text{GeV}$ and $E_p = 900\,\text{GeV}$ for electrons and protons respectively.

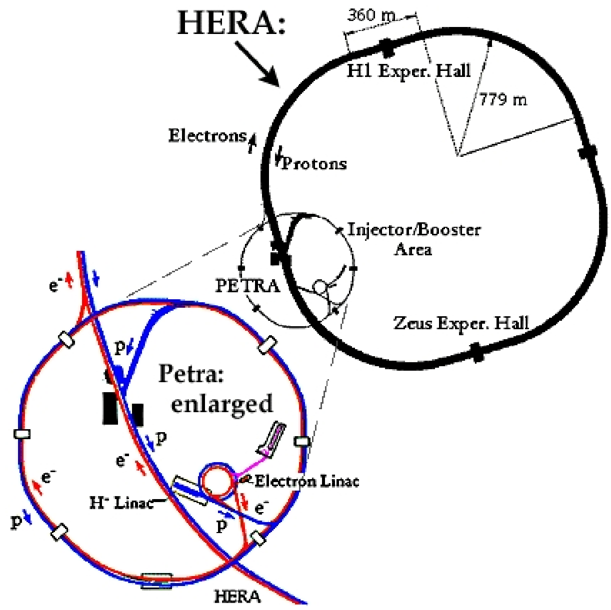


Figure 9.11: Schematics of the HERA accelerator at DESY.

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.

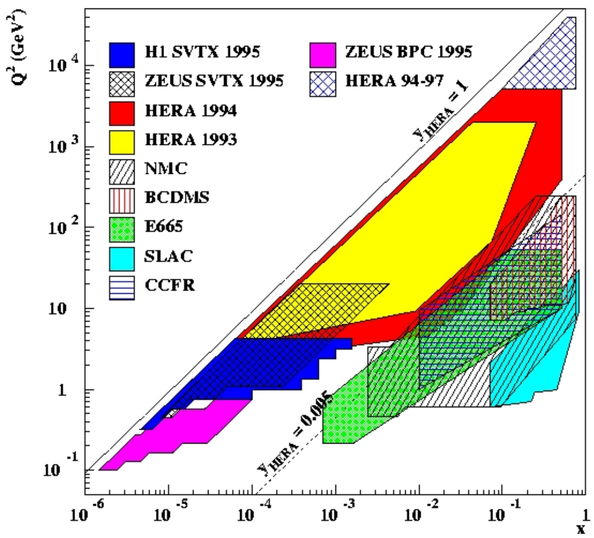


Figure 9.12: Coverage of the Q^2-x-kinematic region at HERA.

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,

$$\begin{aligned} y_e &= 1 - \frac{E_e'}{2E_e}(1 - \cos\theta_e) \\ Q_e^2 &= 2E_e E_e'(1 + \cos\theta_e). \end{aligned}$$

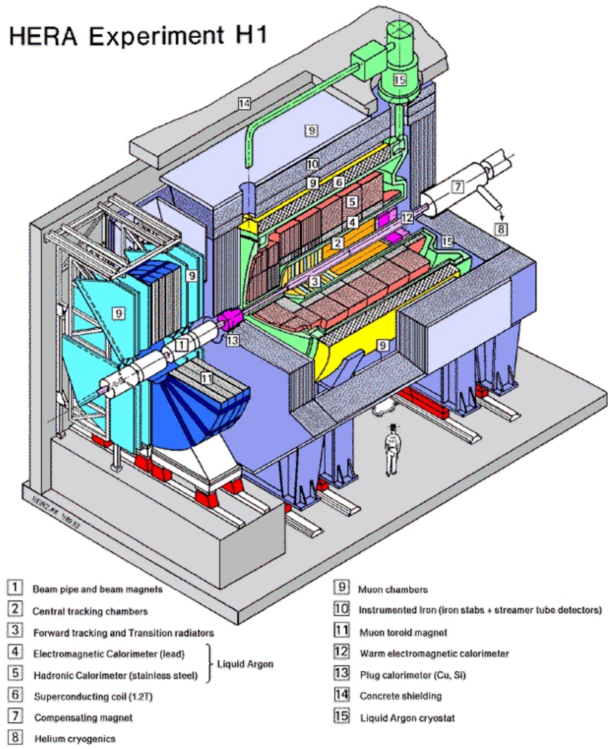
Fig. 9.16 shows the kinematic region measured at ZEUS while Fig. 9.17 shows the experimental results for the structure function F_2 as well as the NLO QCD fits. For low values of x, the scaling violation appears very clearly. It is due to the inclusion of the processes containing gluons.

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 $1.9\,\mathrm{GeV}^2$ and $10\,\mathrm{GeV}^2$ (note the scale reduction!). One can notice similarities with the expectation shown in Fig. 9.10.

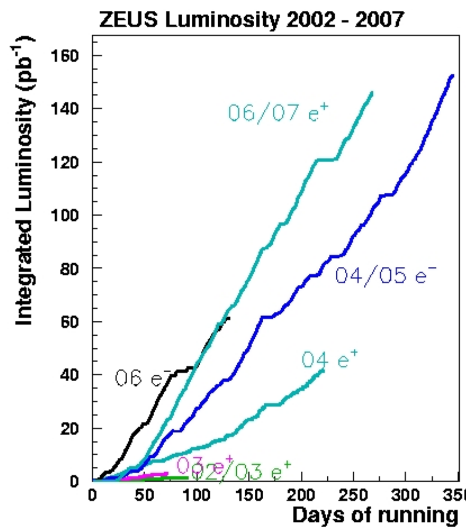
9.6 Parton model revisited

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).

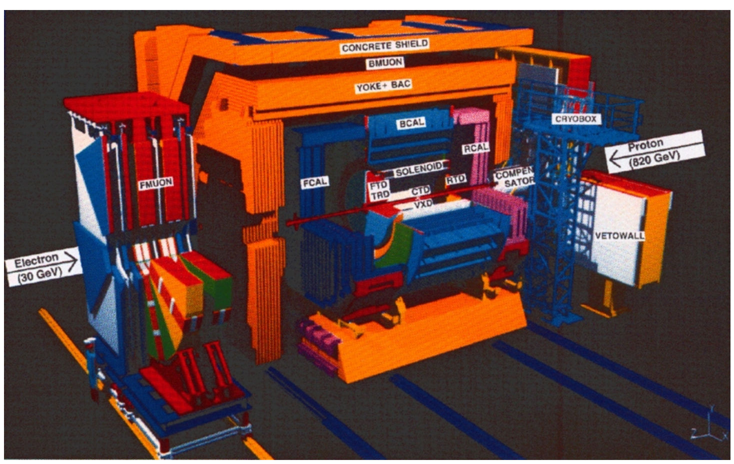
As we have seen the proton is a bound state of three quarks with strong binding. "Strong binding" says that the quark binding energy is much larger than the light quark masses: $E_{\mathrm{bind}} \gg m_q$. Compare this to the weak binding of the hydrogen atom electron: $E_{\mathrm{bind}} \ll m_e$.



(a)



(b)



(c)

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 ($|\vec{p}| \gg m_p$):

$$p^\mu = \begin{pmatrix} \sqrt{|\vec{p}|^2 + m_p^2} \\ \vec{p} \end{pmatrix} \simeq \begin{pmatrix} |\vec{p}| \\ \vec{p} \end{pmatrix}.$$

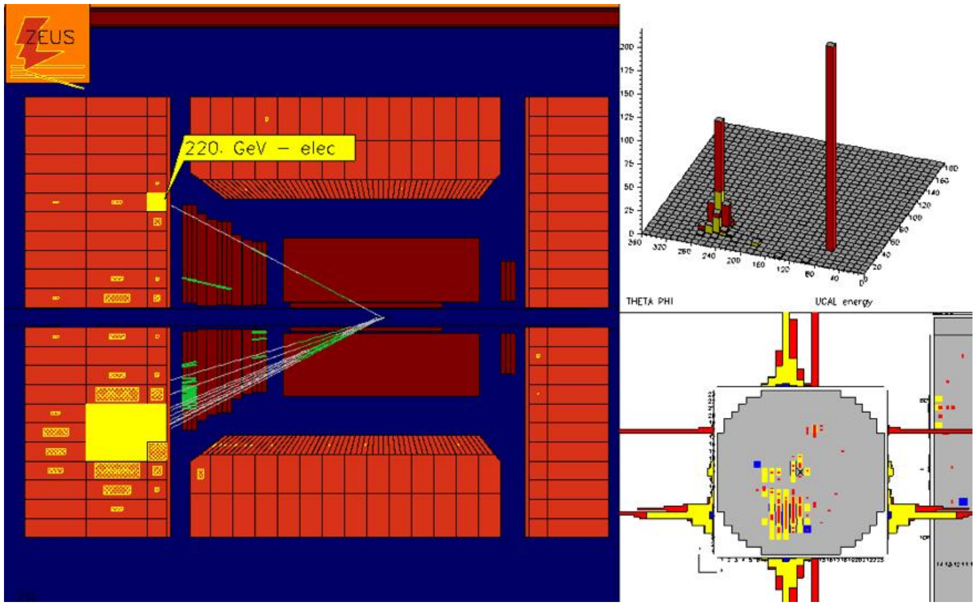
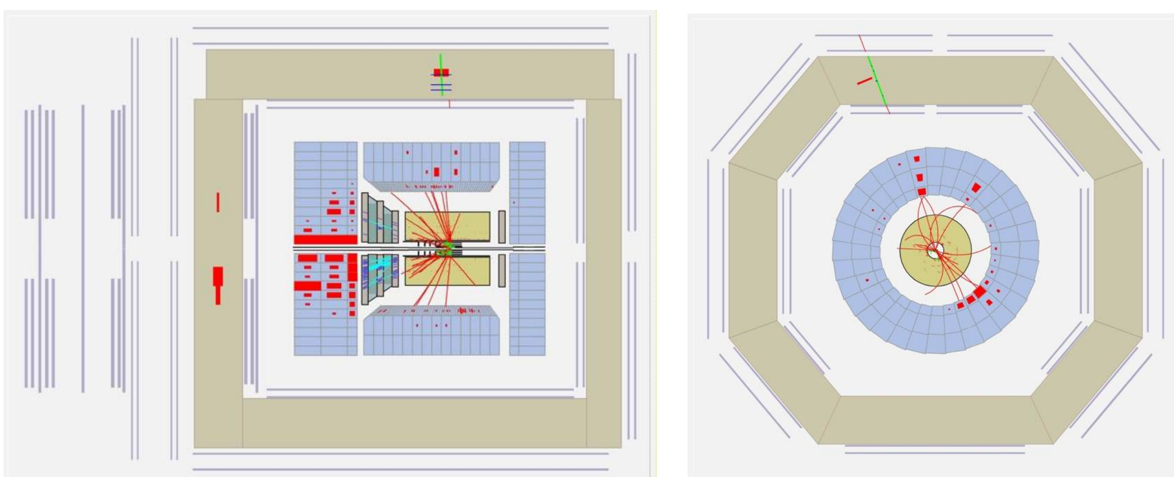


Figure 9.14: DIS event recorded by the ZEUS experiment.



(a) (b)

Figure 9.15: *Two jet event at ZEUS* (a) Side view. (b) Transverse view.

In Sect. 7.4.2 (p. 148) we discussed asymptotic freedom, namely the fact that for $Q^2 \gg \Lambda^2_{\rm QCD}$ 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

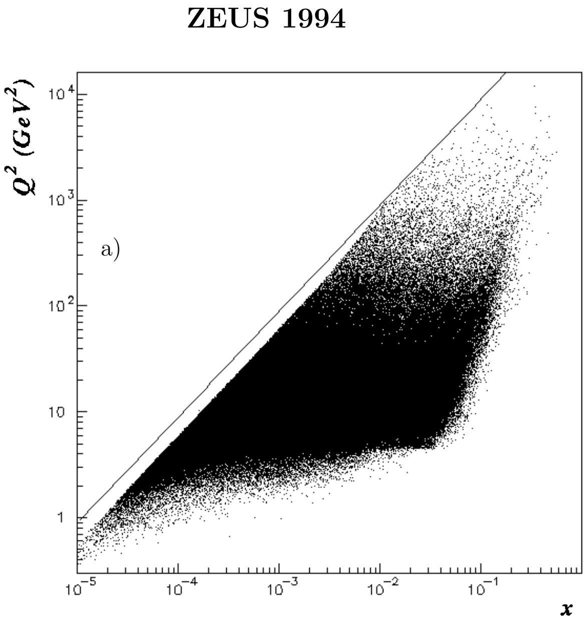


Figure 9.16: Kinematic phase-space measured by the ZEUS experiment.

processes (see Sect. 9.3.2 for diagrams) with the doubly differential cross section[2]

$$\frac{d^2\sigma}{dxdQ^2} = \sum_q \int_0^1 d\xi f_q(\xi) \frac{d^2\hat{\sigma}^{lq}}{dxdQ^2} \tag{9.18}$$

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.

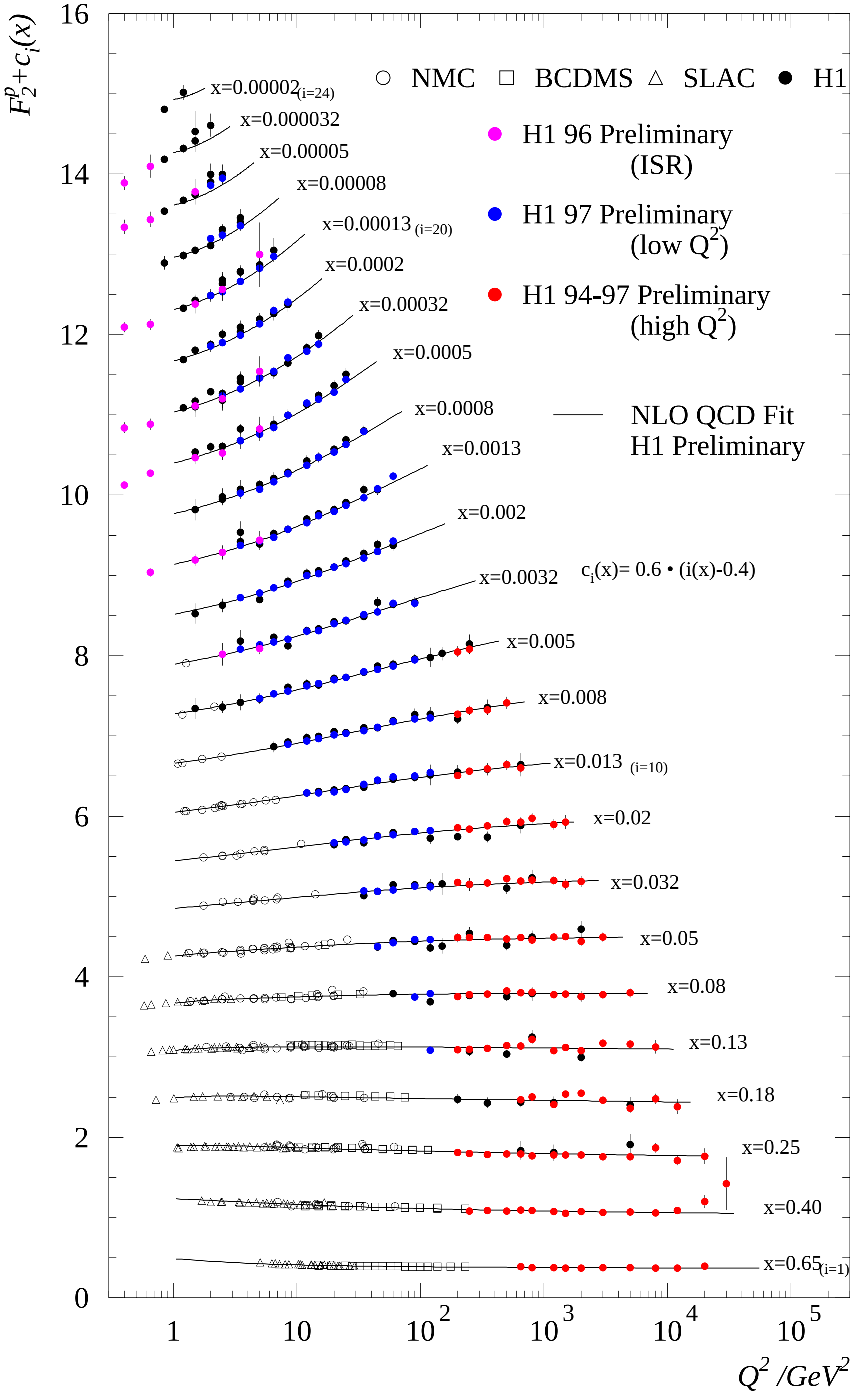


Figure 9.17: Proton structure function F_2^p measured by H1 and other experiments for various values of Q^2 and x. Scaling violations appear for $x < 10^{-2}$.

$m_p \simeq 3m_q$, 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

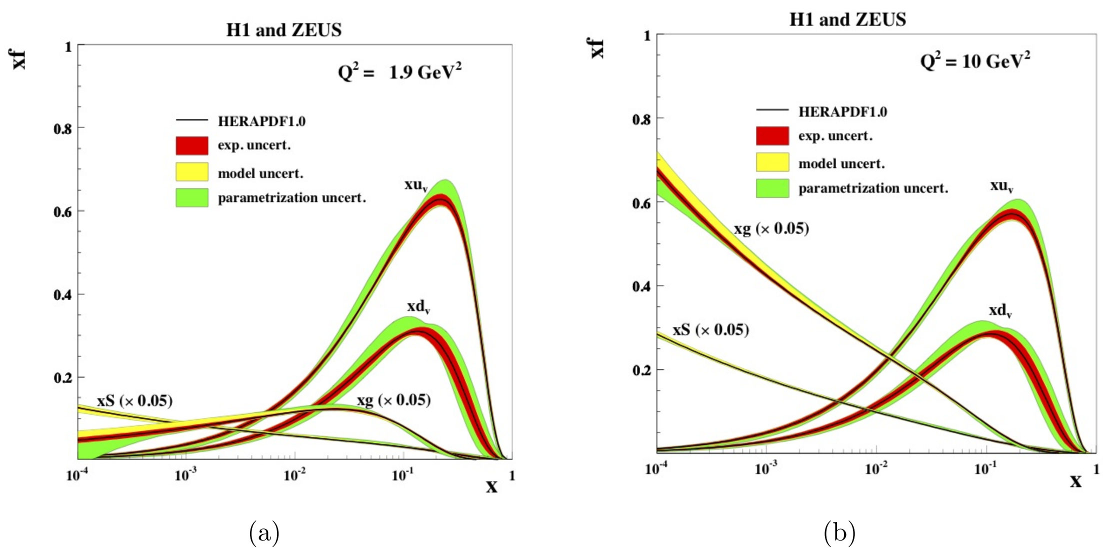


Figure 9.18: *Parton distribution functions of the proton* (a) $Q^2 = 1.9\,\mathrm{GeV}^2$. (b) $Q^2 = 10\,\mathrm{GeV}^2$. The sea and gluon PDFs are reduced by a factor 20.

smeared out significantly at any given instant of time, as shown in Fig. 9.19(c).

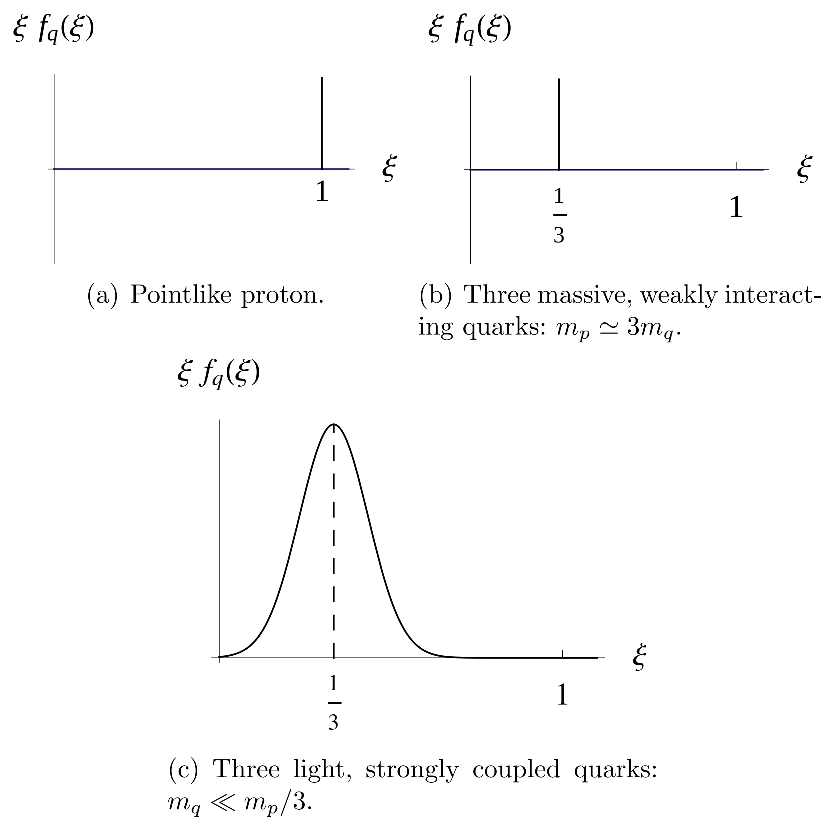


Figure 9.19: *Quark momentum density* $\xi f_q(\xi)$.

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.

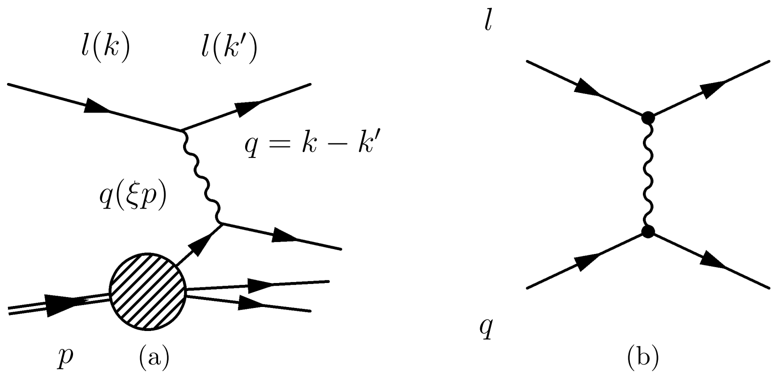


Figure 9.20: *(a) Kinematics of simple parton model and (b) Feynman diagram for lepton-quark scattering.*

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. 92). 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

$$F_2(x,Q^2) = \sum_q e_q^2 x f_q(x)$$
$$F_L(x,Q^2) = F_2(x,Q^2) - 2xF_1(x,Q^2) = 0$$

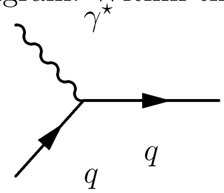
where F_L is called longitudinal structure function. We recognize that $F_2(x,Q) = F_2(x)$ ceases to be a function of two variables, but under the assumed conditions depends only on one variable, a phenomenon generally referred to as scaling. Furthermore, $F_L = 0 \Leftrightarrow 2xF_1 = F_2$ is the 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

$$\begin{aligned} f_q(x) &= q(x) \quad (q = u, d, s, c, \dots, \bar{u}, \dots) \\ f_g(x) &= g(x) \quad \text{(gluons)}. \end{aligned}$$

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



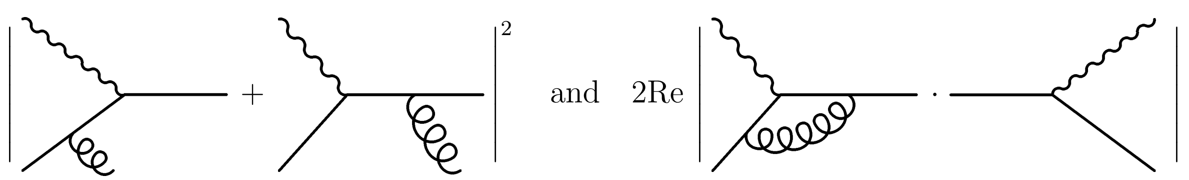
$$= \frac{4\pi\alpha e_q^2}{\hat{s}} \delta(x-\xi) =: \hat{\sigma}_0 \delta(x-\xi) \tag{9.19}$$

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. 26).

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:

$$\frac{\hat{\sigma}^{\gamma^\star q \to qg}}{\hat{\sigma}_0} = \frac{\alpha_s}{2\pi} P_{qq}(z) \int_{\mu^2}^{Q^2} \frac{dp_T^2}{p_T^2} = \frac{\alpha_s}{2\pi} P_{qq}(z) \log \frac{Q^2}{\mu^2}$$

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. 160). 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:

$$\frac{F_2(x, Q^2)}{x} = \sum_q \int_x^1 \frac{d\xi}{\xi} q(\xi) e_q^2 \left\{ \delta\Big(1 - \frac{x}{\xi}\Big) + \frac{\alpha_s}{2\pi} \left[P_{qq}\Big(\frac{x}{\xi}\Big) \log \frac{Q^2}{\mu^2} + \text{finite} \right] + \mathcal{O}(\alpha_s^2) \right\} \tag{9.21}$$

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:

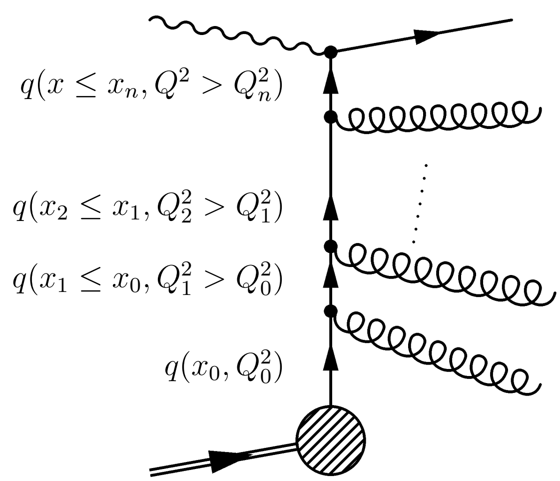
$$\mu_F^2 \frac{d}{d\mu_F^2} q(x) = 0.$$

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. 102) or QCD (Sect. 7.4.2, p. 148).

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 [2, 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,

$$P_{qq}(z) = C_F \left(\frac{1+z^2}{(1-z)_+} + \frac{3}{2}\delta(1-z) \right),$$

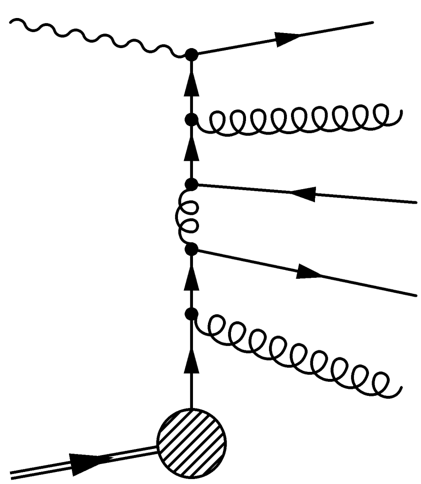
which takes into account the virtual corrections occuring at $z = 1$. We use the '+'-presciption, coming from the reguarisation procedure and defined by,

$$\int_0^1 dz \frac{f(z)}{(1-z)_+} = \int_0^1 dz \frac{f(z)-f(1)}{1-z}. \tag{9.24}$$

The factor in front of the δ-function can be inferred from the quark number conservation, which can be stated as,

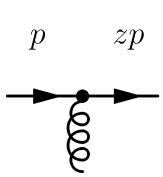
$$\int_0^1 dz P_{qq}(z) = 0. \tag{9.25}$$

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,



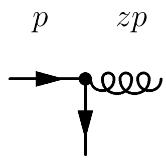
By inspection, one can find out that there are four different splitting processes at $O(\alpha_s)$:

- $q \to q$:



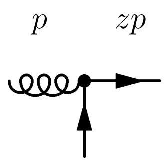
$$\rightsquigarrow P_{qq}(z) = C_F \left(\frac{1+z^2}{(1-z)_+} + \frac{3}{2}\delta(1-z) \right), \tag{9.26}$$

- $q \to g$:



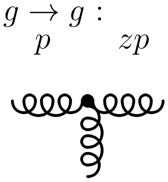
$$\leftrightsquigarrow P_{gq}(z) = C_F \left(\frac{1+(1-z)^2}{z} \right), \tag{9.27}$$

- $g \to q$:



$$\leftrightsquigarrow P_{qg}(z) = T_F \left(z^2 + (1-z)^2 \right), \tag{9.28}$$

- $g \to g$:



$$\leftrightsquigarrow P_{gg}(z) = 2C_A \left(\frac{z}{(1-z)_+} + \frac{1-z}{z} \right) + \left(\frac{11}{6} C_A - \frac{3}{2} T_F n_f \right) \delta(1-z). \tag{9.29}$$

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,

$$\frac{\partial q_i^V}{\partial \log \mu_F^2} = \frac{\alpha_s}{2\pi} P_{qq} \otimes q_i^V \tag{9.34}$$

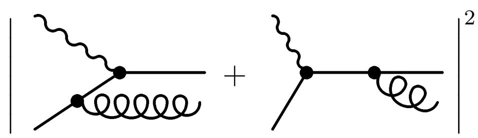
$$\frac{\partial q_i^F}{\partial \log \mu_F^2} = \frac{\alpha_s}{2\pi} P_{qq} \otimes q_i^F \tag{9.35}$$

$$\frac{\partial q^S}{\partial \log \mu_F^2} = \frac{\alpha_s}{2\pi} \left(P_{qq} \otimes q^S + 2n_f P_{qg} \otimes g\right) \tag{9.36}$$

$$\frac{\partial g}{\partial \log \mu_F^2} = \frac{\alpha_s}{2\pi} \left(P_{gq} \otimes q^S + P_{gg} \otimes g\right). \tag{9.37}$$

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,

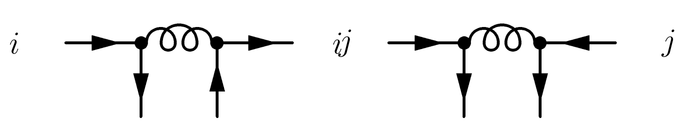


This translates in the expressions for the structure functions,

$$\frac{1}{x} F_2(x, Q^2) = \int_x^1 \frac{d\xi}{\xi} \left\{ \sum_q q\left(\xi, Q^2\right) \left[\delta\left(1 - \frac{x}{\xi}\right) + \frac{\alpha_s}{2\pi} C_{2,q}\left(\frac{x}{\xi}\right)\right] + g(\xi, Q^2) \frac{\alpha_s}{2\pi} C_{2,g}\left(\frac{x}{\xi}\right)\right\} \tag{9.38}$$

$$F_L(x, Q^2) = \mathcal{O}(\alpha_s) \neq 0 \tag{9.39}$$

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. 151) 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. 152,

$$\boxed{q(n,Q^2) = q(n,Q_0^2)\exp\left\{\frac{2}{\beta_0}P_{qq}(n)\log\frac{\log(Q^2/\Lambda^2)}{\log(Q_0^2/\Lambda^2)}\right\}}. \tag{9.47}$$

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,

$$q^S(2, Q^2 \to \infty) \to \frac{3n_f}{16 + 3n_f} \overset{n_f=5}{=} \frac{15}{31}$$
$$g(2, Q^2 \to \infty) \to \frac{16}{16 + 3n_f} \overset{n_f=5}{=} \frac{16}{31}.$$

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).

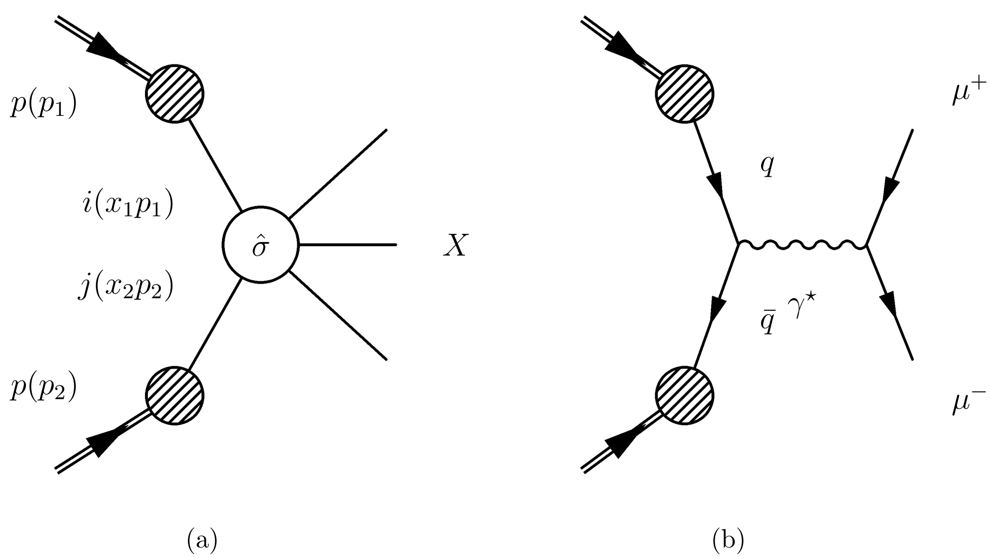


Figure 9.21: *(a) Hadron-hadron collision in naive parton model and (b) Drell-Yan process.*

As an example consider the Drell-Yan process, $pp \to \mu^+\mu^-$, shown in Fig. 9.21(b). The parton model cross section reads

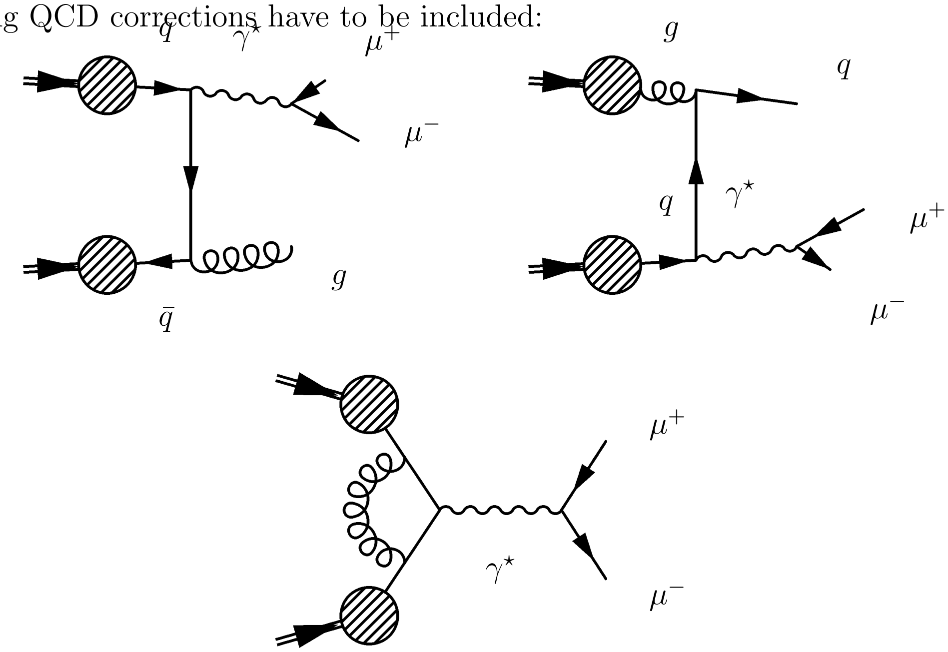
$$\sigma^{\mathrm{DY}} = \sum_q \int dx_1 dx_2 \left[q(x_1)\bar{q}(x_2) + q(x_2)\bar{q}(x_1)\right] \hat{\sigma}_{q\bar{q} \to \mu^+\mu^-} \tag{9.50}$$

where

$$\hat{\sigma}_{q\bar{q}\to\mu^+\mu^-} = \underbrace{\frac{4\pi\alpha^2}{3s_{q\bar{q}}}\frac{1}{3}}_{\hat{\sigma}_0^{\mathrm{DY}}} e_q^2\,\delta(1 - x_1x_2s_{pp}/M^2_{\mu^+\mu^-}) \tag{9.51}$$

which we basically already calculated before (Sect. 5.10, p. 92). 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 $(p_{\mu^+} + p_{\mu^-})^2 =: M^2_{\mu^+\mu^-} = x_1x_2s_{pp}$.

The following QCD corrections have to be included:



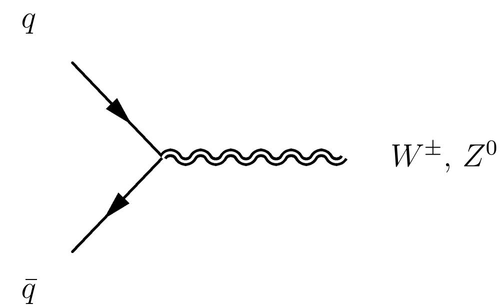
where the first two diagrams are because of parton evolution and the third diagram is a virtual correction. Setting $z = x_1x_2s_{pp}/M^2_{\mu^+\mu^-}$, the QCD corrected Drell-Yan cross section reads

$$\begin{aligned}\sigma^{\mathrm{DY}} = \hat{\sigma}_0^{\mathrm{DY}} \sum_q e_q^2 \int dx_1dx_2 \Big\{ & q(x_1)\bar{q}(x_2)\delta(1-z) + \frac{\alpha_s}{2\pi}C_{q\bar{q}}(z) \\ & + [q(x_1) + \bar{q}(x_1)]\, g(x_2)\frac{\alpha_s}{2\pi}C_{qg}(z) + (x_1 \leftrightarrow x_2)\Big\}\end{aligned}$$

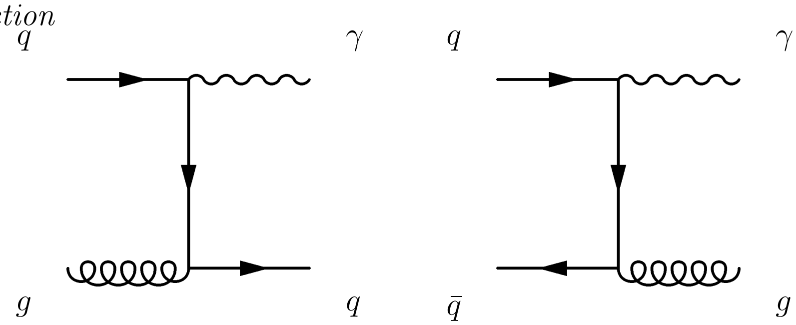
where $q(x_i)$ etc. are the QCD evolved parton distributions.

In the following some standard reactions are listed.

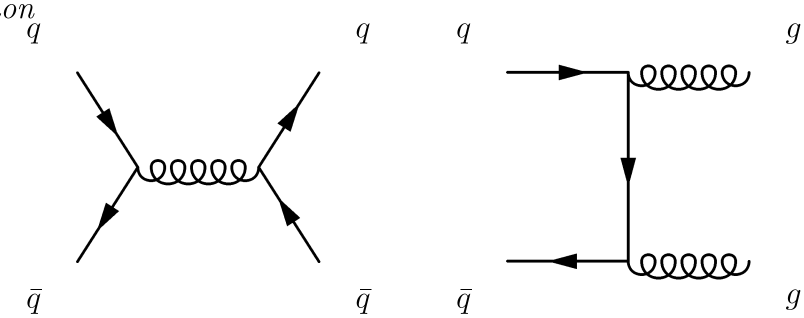
- *$W^\pm$, Z^0 production*



- $\gamma + \text{jet}$ *production*



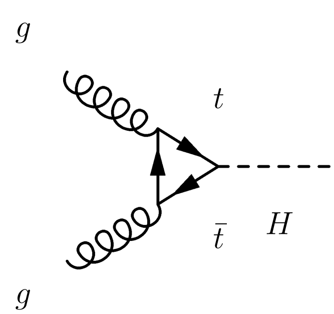
- *2-jet production*



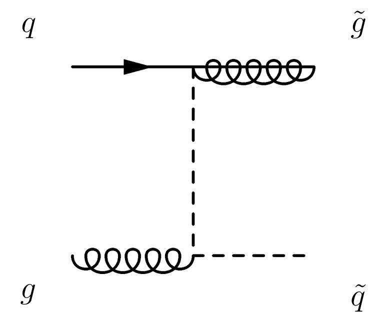
  Further processes leading to 2-jet events are $qg \to qg$, $gg \to gg$, $gg \to q\bar{q}$ and $qq \to qq$.

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. 50).

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.

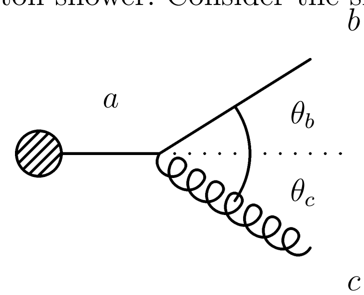
  These numbers illustrate that a computation even on the amplitude level is time-consuming.

- *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

$$t = 2E_b E_c(1 - \cos\theta) = z(1-z)E_a^2\theta^2 \tag{9.53}$$

$$\frac{\theta_b}{1-z} = \frac{\theta_c}{z} = \theta. \tag{9.54}$$

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:

$$\begin{aligned} d\phi_n &= \ldots \frac{d^3 p_a}{2E_a(2\pi)^3} \\ d\phi_{n+1} &= \ldots \frac{d^3 p_b}{2E_b(2\pi)^3} \frac{d^3 p_c}{2E_c(2\pi)^3}. \end{aligned}$$

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]

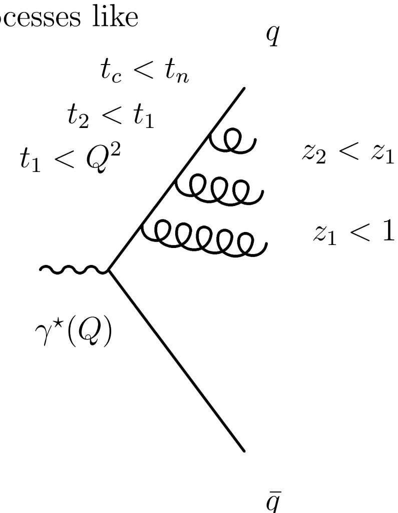
$$
\begin{aligned}
d\phi_{n+1} &= d\phi_n \frac{1}{2(2\pi)^3} \int E_b dE_b \theta_b d\theta_b d\phi \frac{dz}{1-z} \delta(z - E_b/E_a) dt \delta(t - E_a E_b \theta^2) \\
&= d\phi_n \frac{1}{4(2\pi)^3} dt dz d\phi
\end{aligned}
$$

(recall Eq. (9.52) and (9.53)).

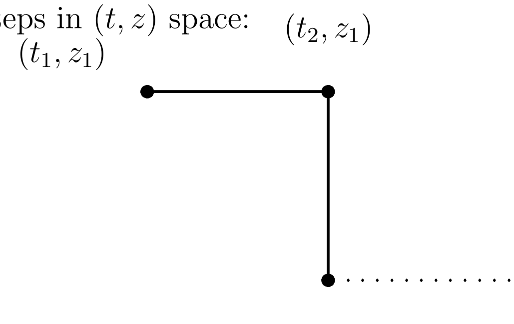
Since the matrixelement and the phase space factorize, so does the cross section:

$$
d\sigma_{n+1} = d\sigma_n \frac{dt}{t} dz \frac{d\phi}{2\pi} \frac{\alpha_s}{2\pi} C_F F(z).
$$

Therefore, multiple emission processes like



where t_c is a cutoff scale at which hadronization sets in, $t_c \gtrsim \Lambda^2_{\rm QCD}$, 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.

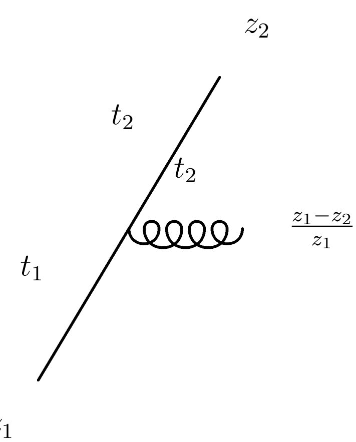


Figure 9.22: *Starting point for second Monte Carlo step.*

Chapter 10

Hadron collider physics

Literature:

- Ellis/Stirling/Webber [3]
- Dissertori/Knowles/Schmelling [4]
- Kane/Pierce [5]
- Review on QCD of the Particle Data Group [6]
- Technical Design Reports (TDRs) about the Physics Performance of ATLAS and CMS [7, 8]

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 ($< 114\,\mathrm{GeV}$).

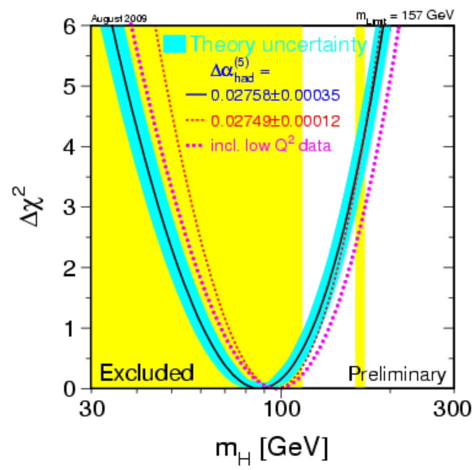


Figure 10.2: Most likely region for the Higgs mass (indicated by the minimum in the χ^2 value) as obtained from a fit of standard model predictions to LEP, SLD and Tevatron data

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.

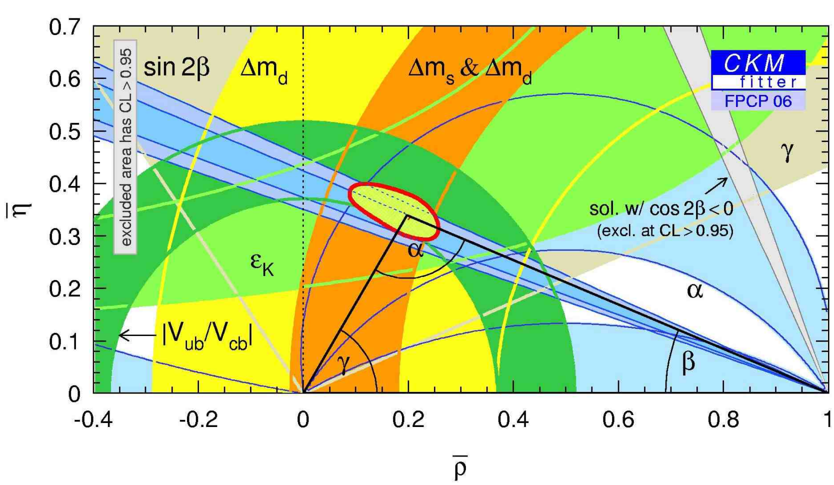


Figure 10.3: Experimental constraints on the parameters of the CKM matrix [6]

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,

$$p_T = p \sin\theta,$$
$$p_L = p \cos\theta.$$

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.

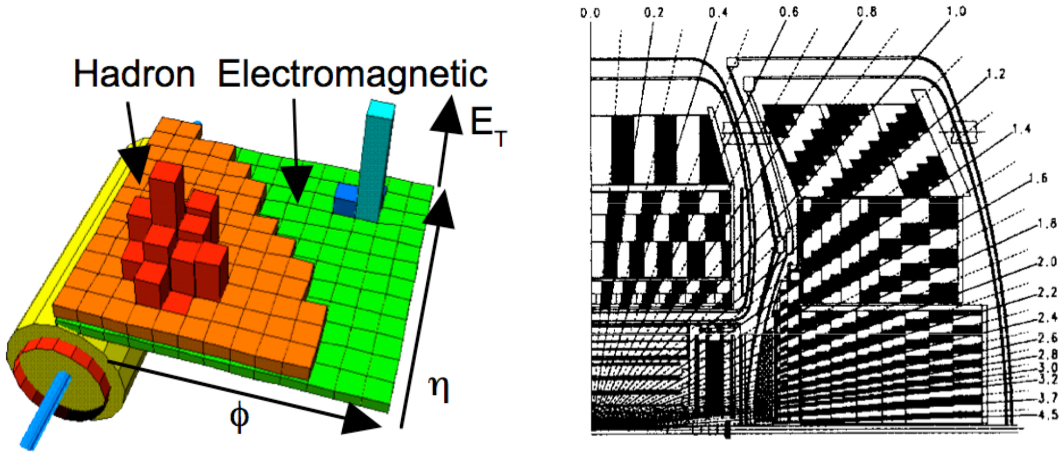
The interest of introducing rapidity and pseudorapidity lies in the fact that at hadron colliders the laboratory(detector)-frame in general does not coincide with the center of mass frame of the parton-parton collision, unless the two beams have the same energy and the parton momentum fractions fulfill $x_1 = x_2$. 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.

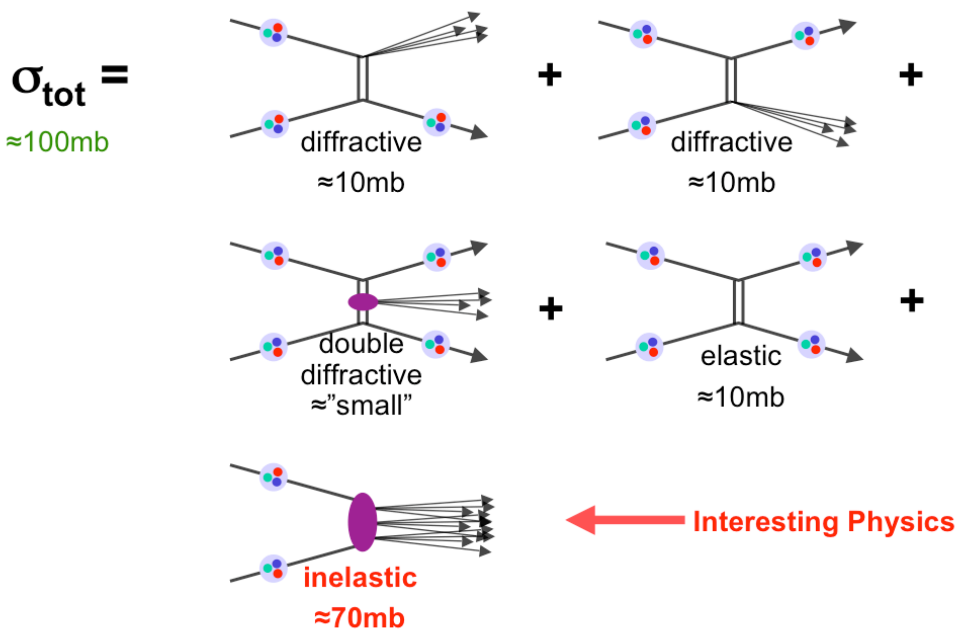


Figure 10.5: Pictorial representation of the components of the total proton-proton cross section. Here "interesting physics" refers to those processes relevant for the study of hard interactions.

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 $\sim$25 soft collisions per bunch crossing, each generating $\sim$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

$$25 \cdot 9 \cdot 2|y_{max}| \approx 2250$$

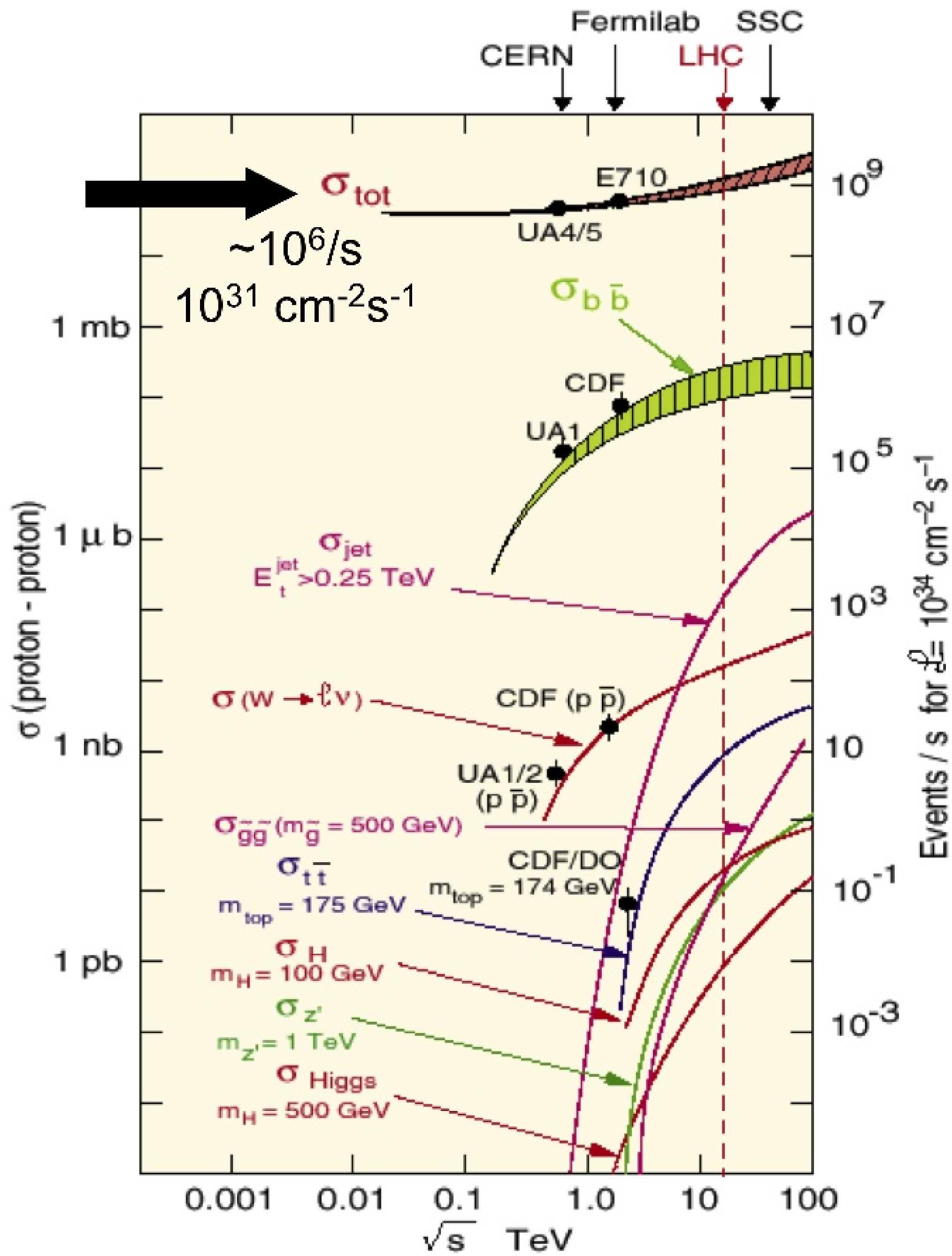
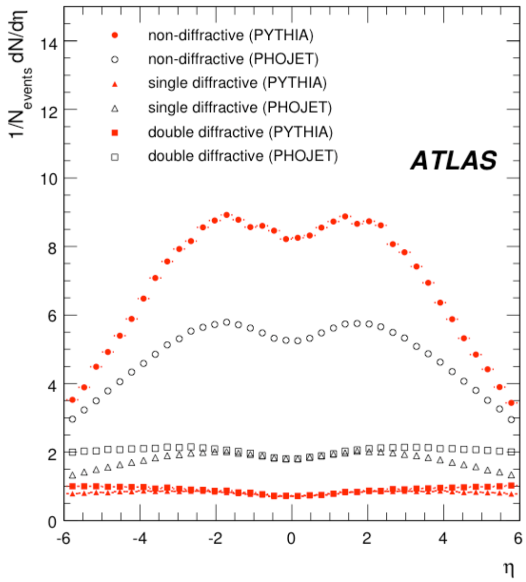


Figure 10.6: Cross sections for different processes in proton-proton scattering

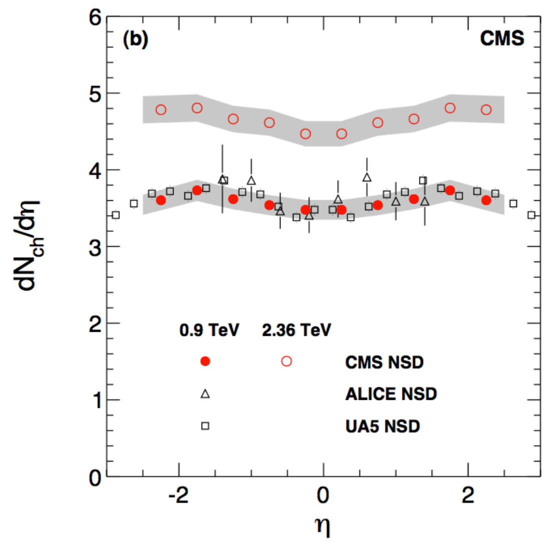
pions produced that will deposit a total energy of,

$$2250 \cdot \underbrace{700\,[\text{MeV}]}_{\langle p_T \rangle} \approx 1.6\,[\text{TeV}]$$

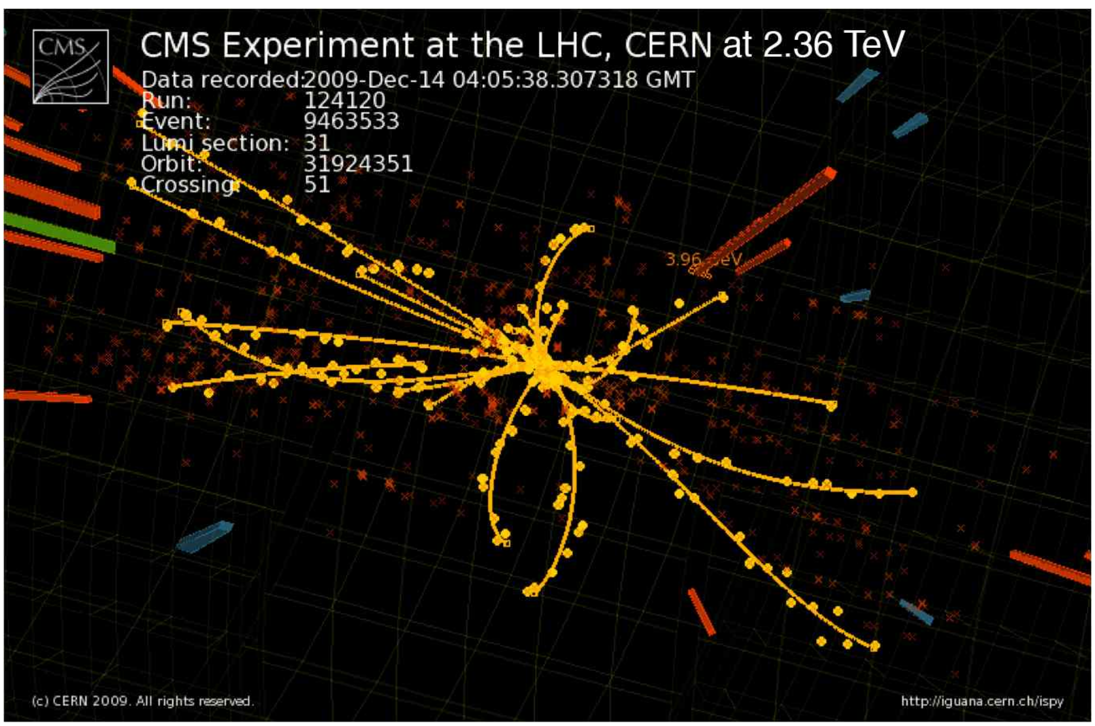
in the calorimeters for each bunch crossing, resulting in an important background noise,



(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.

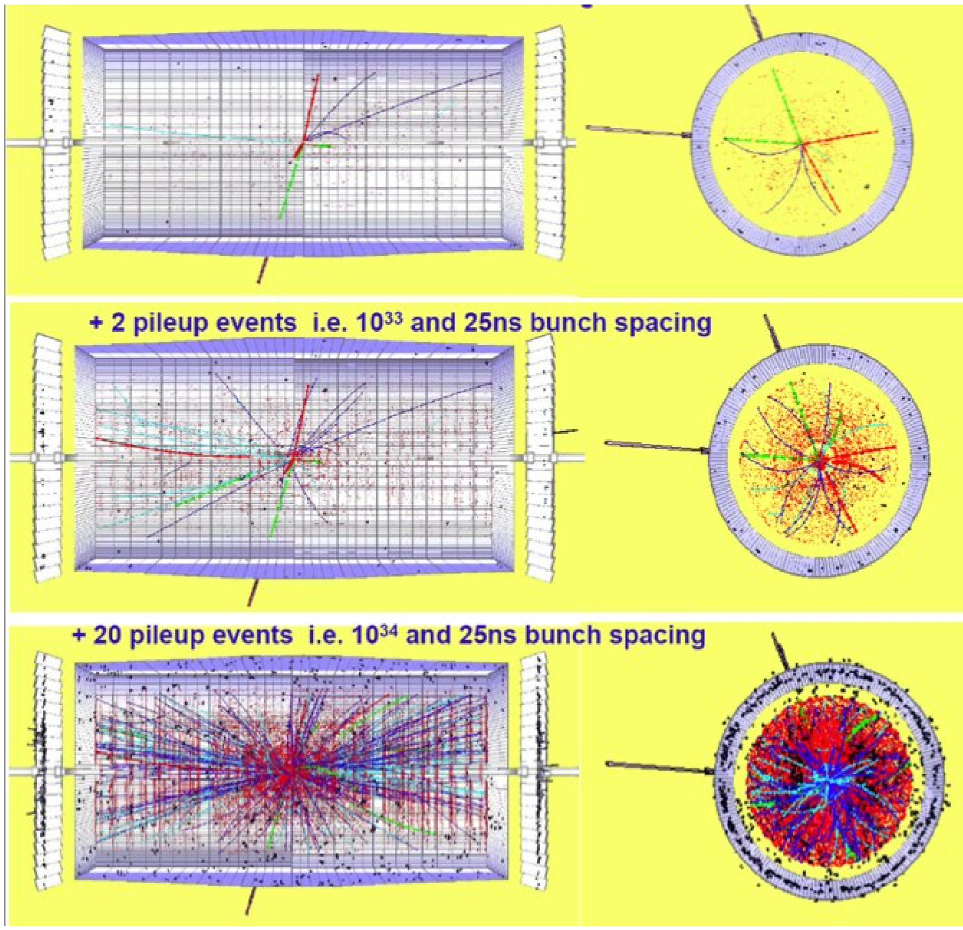


Figure 10.8: Pile-up events at different luminosities.

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.

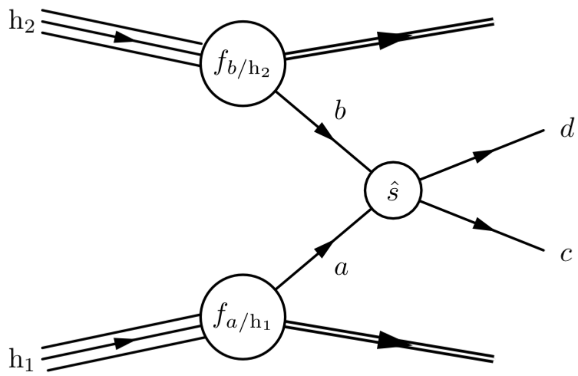


Figure 10.9: Basic Feynman graph for the description of a hard scattering process in a hadron-hadron collision.

At high energies ($\gg \Lambda_{\mathrm{QCD}}$), 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. 32. 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 $M \sim 1\,\mathrm{TeV} \gg m_p$, 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 $E_{\mathrm{beam}} = \sqrt{s}/2 \gg m_p$, 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,

$$e^y = \sqrt{\frac{x_1}{x_2}},$$

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,

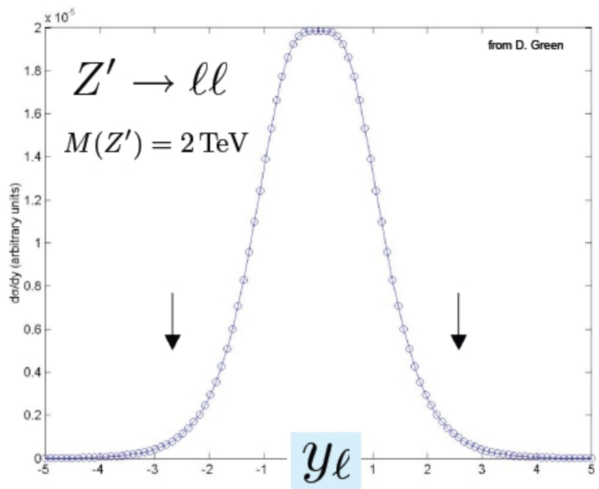


Figure 10.10: Rapidity distribution for Z' production

$$x_{1,2} = \frac{M}{\sqrt{s}} e^{\pm y}, \tag{10.5}$$

i.e. to produce a Z' with rapidity y, one of the partons must have had a momentum fraction $Me^{+y}/\sqrt{s}$, while the other had the momentum fraction $Me^{-y}/\sqrt{s}$. 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,

$$xf_i(x, Q_0^2) = A_i \underbrace{x^{\alpha_i}}_{\text{low-}x} \underbrace{(1-x)^{\beta_i}}_{\text{large-}x},$$

   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
$\boldsymbol{\ell N \to c\bar{c}X}$ F_2^c (EMC; H1, ZEUS)*	$\gamma^* c \to c$	c $(x \gtrsim 0.01;\ x \lesssim 0.01)$
$\boldsymbol{\nu N \to \mu^+\mu^- X}$ (CCFR)*	$W^* s \to c$ $\hookrightarrow \mu^+$	$s \approx \frac{1}{4}(\bar{u} + \bar{d})$
$\boldsymbol{pN \to \gamma X}$ (WA70*, UA6, E706, ...)	$qg \to \gamma q$	g at $x \simeq 2p_T^\gamma/\sqrt{s} \to$ $x \approx 0.2 - 0.6$
$\boldsymbol{pN \to \mu^+\mu^- X}$ (E605, E772)*	$q\bar{q} \to \gamma^*$	$\bar{q} = ...(1-x)^{\eta_S}$
$\boldsymbol{pp, pn \to \mu^+\mu^- X}$ (E866, NA51)*	$u\bar{u}, d\bar{d} \to \gamma^*$ $u\bar{d}, d\bar{u} \to \gamma^*$	$\bar{u} - \bar{d}$ $(0.04 \lesssim x \lesssim 0.3)$
$\boldsymbol{ep, en \to e\pi X}$ (HERMES)	$\gamma^* q \to q$ with $q = u, d, \bar{u}, \bar{d}$	$\bar{u} - \bar{d}$ $(0.04 \lesssim x \lesssim 0.2)$
$\boldsymbol{p\bar{p} \to WX(ZX)}$ (UA1, UA2; CDF, D0)	$ud \to W$	u, d at $x \simeq M_W/\sqrt{s} \to$ $x \approx 0.13;\ 0.05$
$\to \boldsymbol{\ell^\pm}$ **asym** (CDF)*		slope of u/d at $x \approx 0.05 - 0.1$
$\boldsymbol{p\bar{p} \to t\bar{t}X}$ (CDF, D0)	$q\bar{q}, gg \to t\bar{t}$	q, g at $x \gtrsim 2m_t/\sqrt{s} \simeq 0.2$
$\boldsymbol{p\bar{p} \to}$**jet** $+ \boldsymbol{X}$ (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 $\sim 100\,\text{GeV}$. 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).

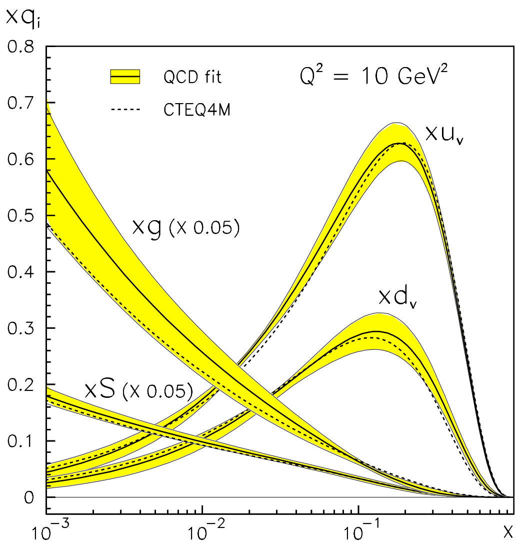
As can be seen in Fig. 10.12, the kinematic regime of the LHC is much broader than the one currently tested experimentally. Much of the relevant x range is covered by HERA, but for much smaller values of Q^2. The question arises if the DGLAP evolution is sufficient to evolve the PDFs to the full LHC 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

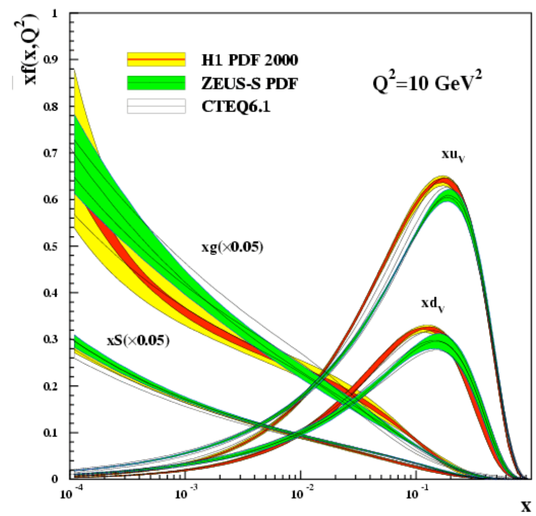
At the LHC, an important component of the inelastic cross-section after soft scattering is jet production, i.e. events where colored partons with significant transverse momentum are produced in the final state. Fig. 10.5 shows this component, labeled σ_{jet} in the case of a minimal jet energy of 250 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.

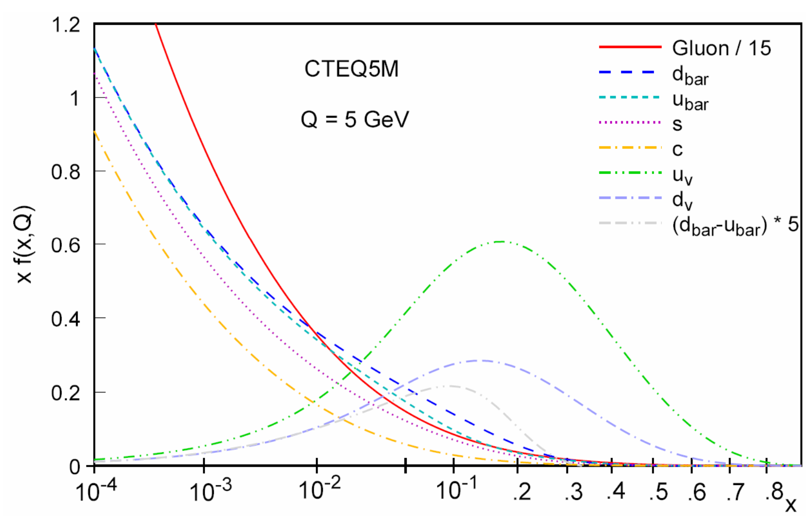
[4] It is implicitly understood that one integrates over the z variable or the x and y variables respectively, see Sect. 9.8.



(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.

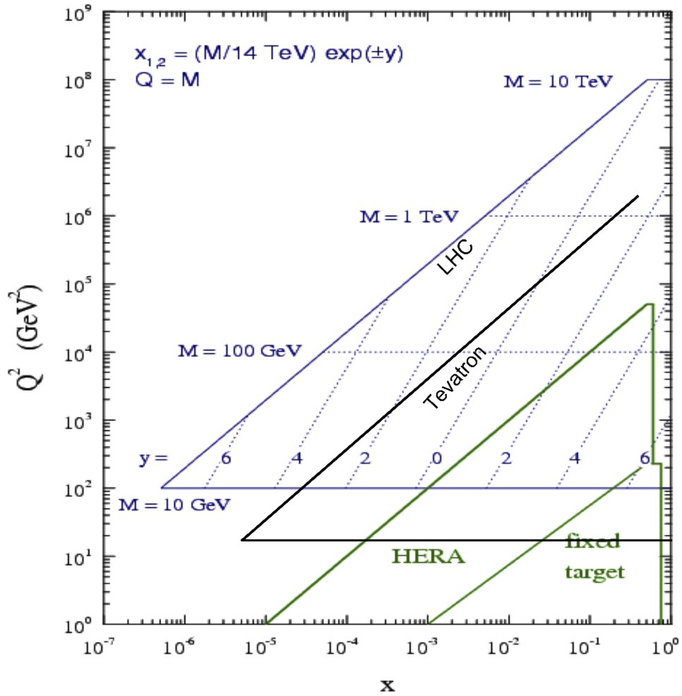


Figure 10.12: Q^2-x range of LHC, Tevatron and HERA.

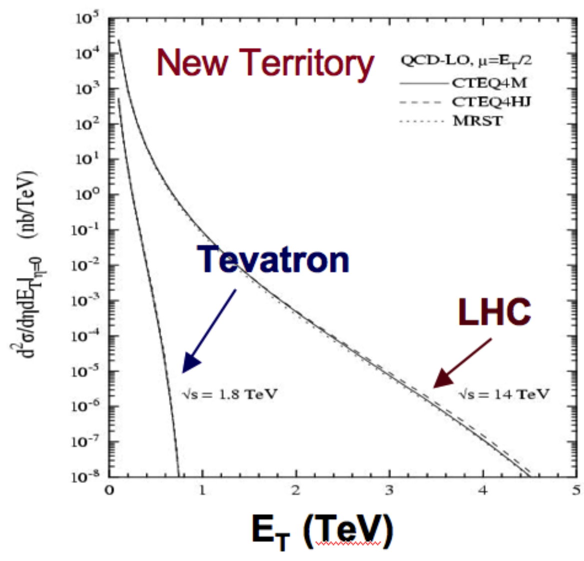


Figure 10.13: Differential cross section for jet production at zero rapidity as a function of the transverse momentum for the Tevatron and the LHC.

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.

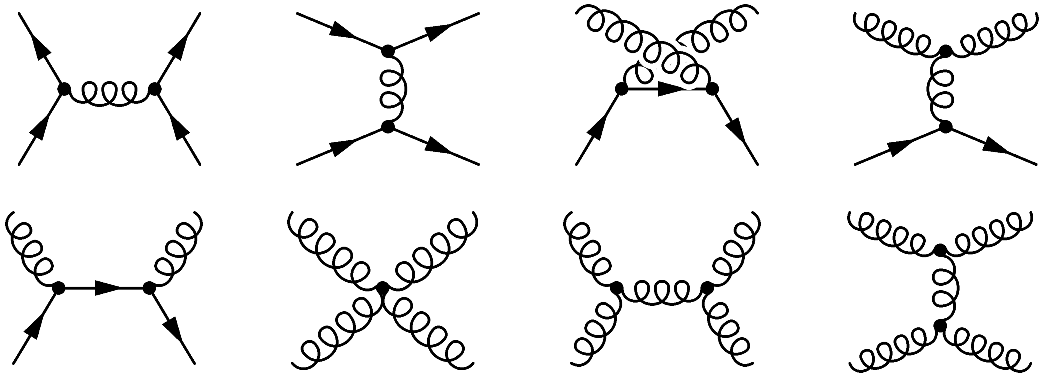
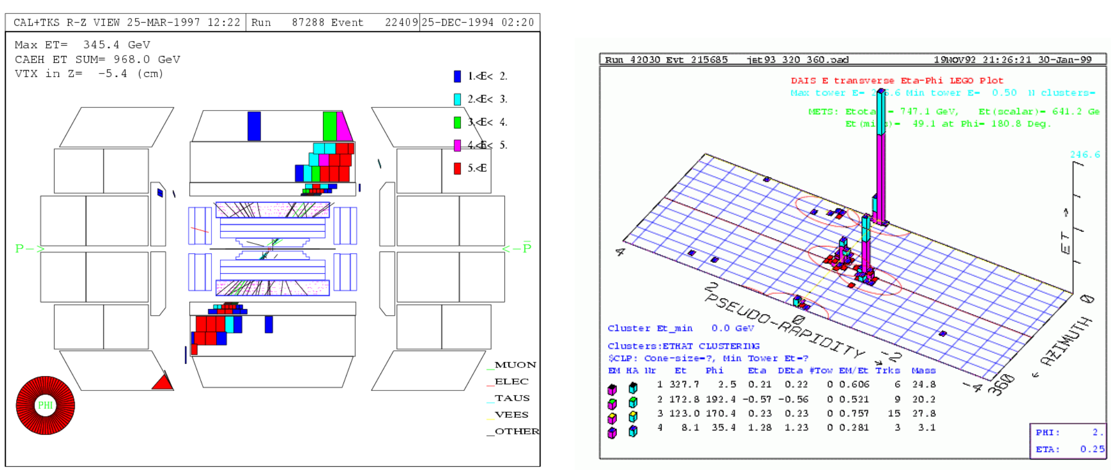


Figure 10.14: Elementary processes at hadron colliders.

10.5.1 Jet algorithms

Section 8.2, p. 159, 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 [9], 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 $\{p_i\}_{i=1,\ldots,n}$ with coordinates (η_i, ϕ_i). One then defines the metric,

$$\begin{aligned} d_{ij} &= \min(p_{T,i}^2, p_{T,j}^2)\frac{\Delta R_{ij}}{D^2} \qquad i > j \\ d_{ii} &= p_{T,i}^2, \end{aligned} \tag{10.6}$$

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 $p \in \mathbb{Z}$,

$$\begin{aligned} d_{ij} &= \min(p_{T,i}^{2p}, p_{T,j}^{2p})\frac{\Delta R_{ij}}{D^2} \qquad i > j \\ d_{ii} &= p_{T,i}^{2p}. \end{aligned} \tag{10.7}$$

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).

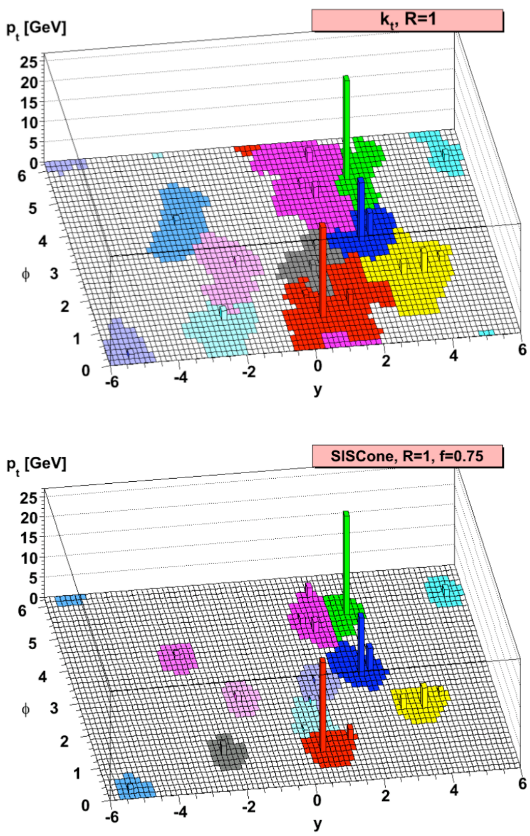


Figure 10.16: Jet areas for different values of the p-parameter in the modified k_T jet algorithm and for the SISCone jet algorithm for the same input data.

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.

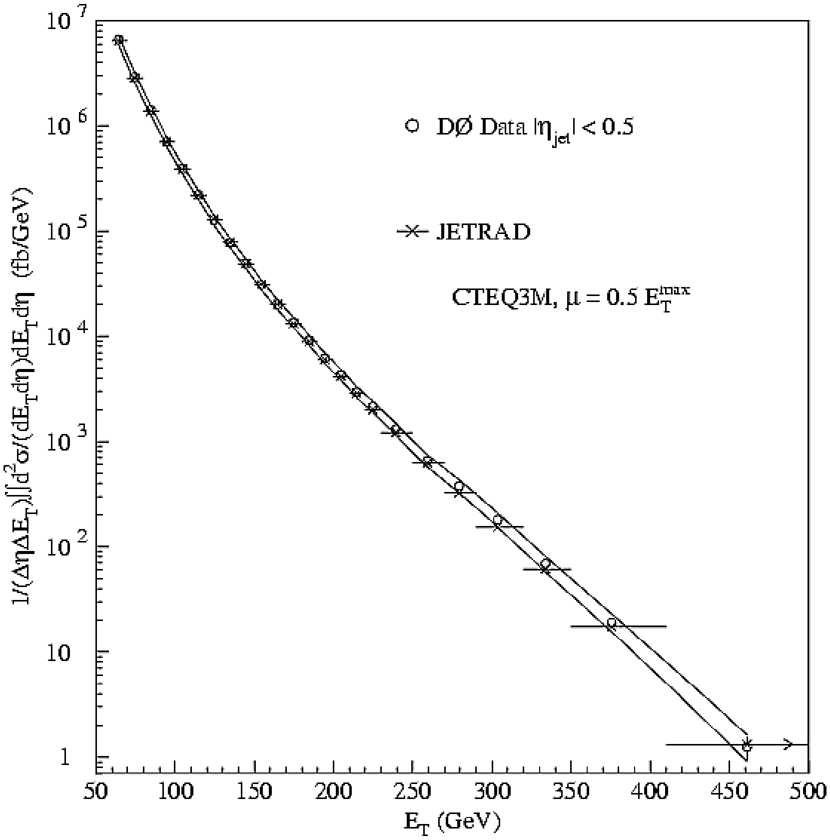


Figure 10.17: Differential jet cross-section as a function of E_T. Monte Carlo simulation (JETRAD) and DØdata.

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 $\sqrt{s} = 10\,\text{TeV}$ at CMS. We see that the jet energy scale and the PDFs induce the largest uncertainties.

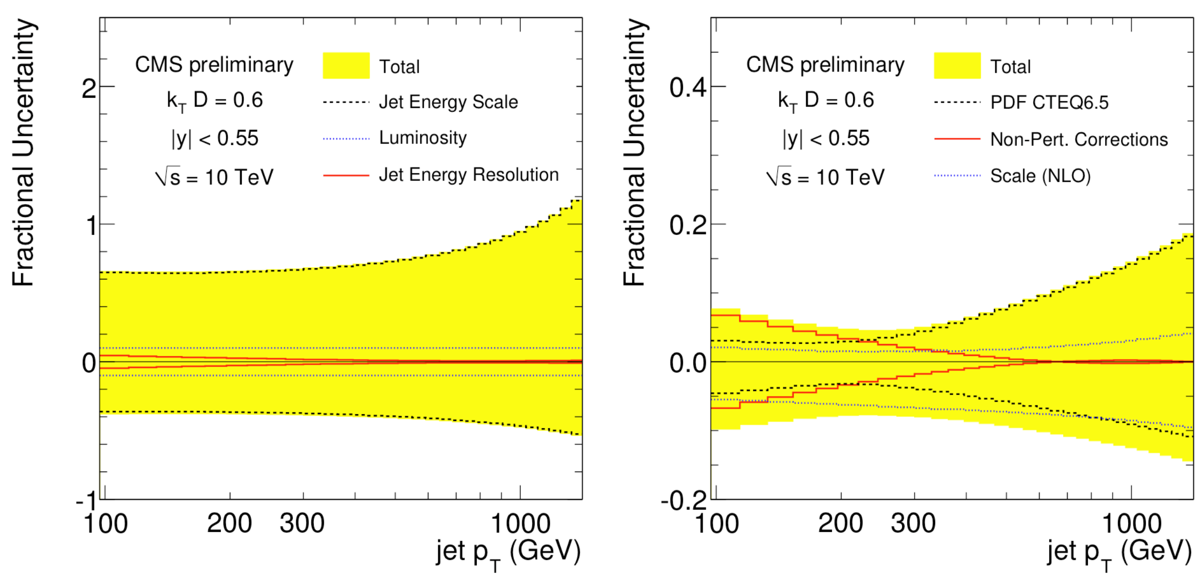
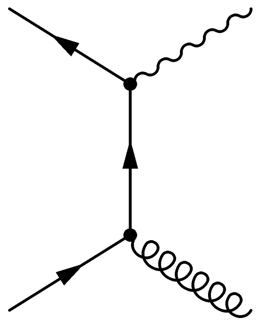


Figure 10.18: Experimental and theoretical part of the fractional uncertainty as a function of the jet transverse energy E_T.

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.

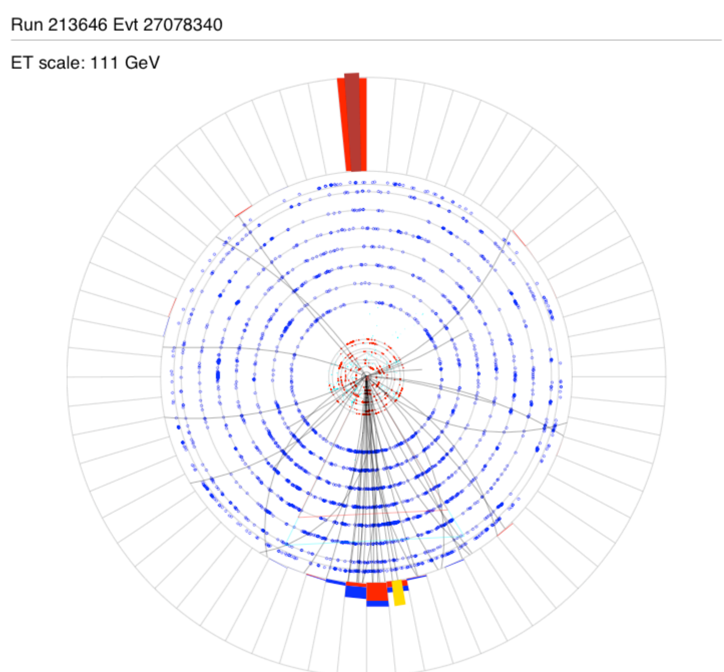
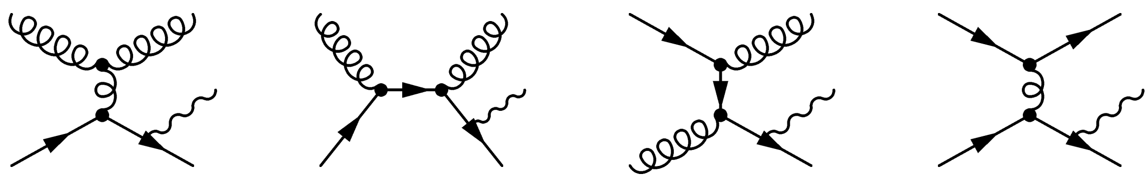


Figure 10.19: Event display of DØwith a jet and a photon used to calibrate the jet energy scale.

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.

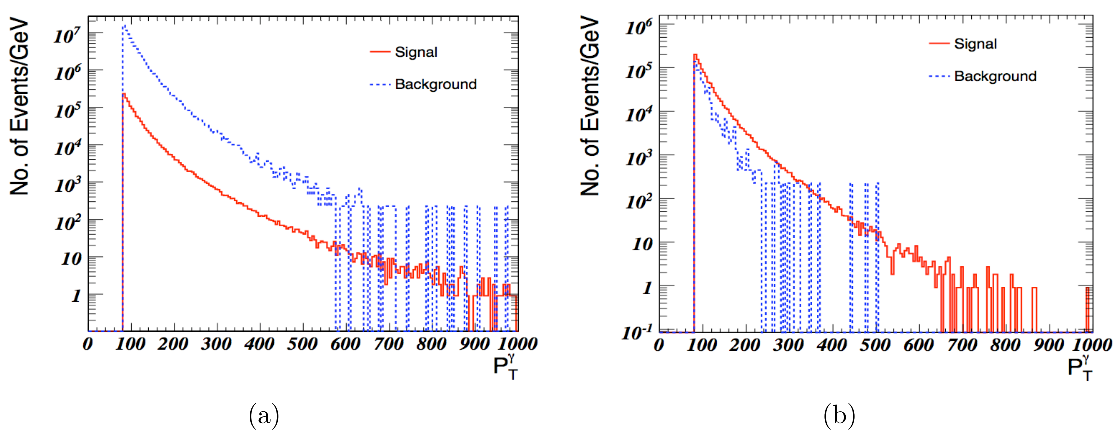


Figure 10.20: *Isolation.* Signal and background (a) before and (b) after isolation cut.

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).

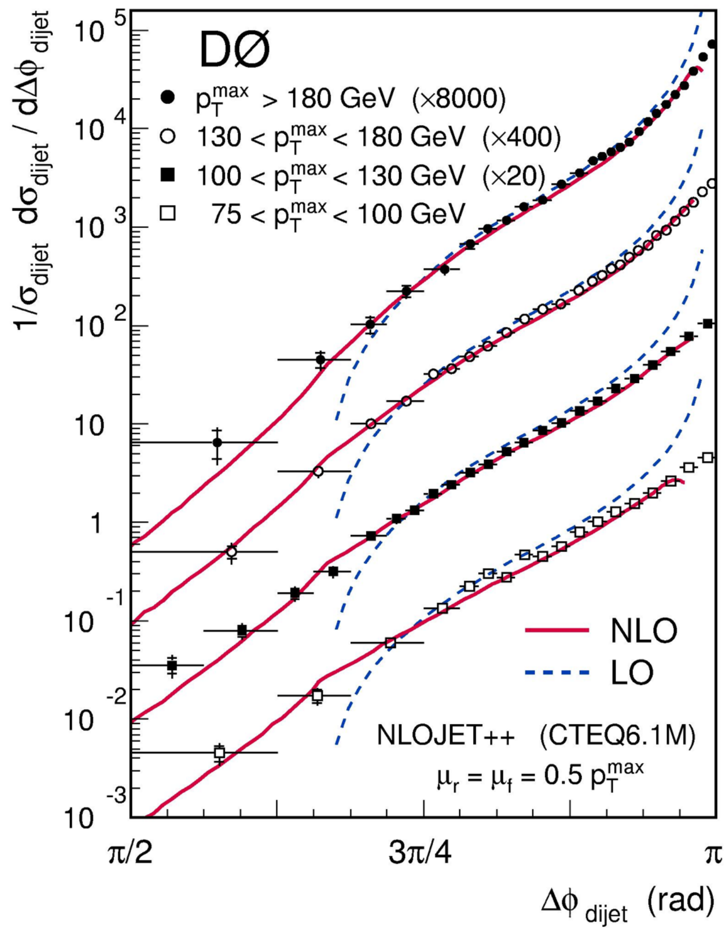


Figure 10.21: Measured dijet angular distributions at DØcompared to LO and NLO QCD calculations.

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.

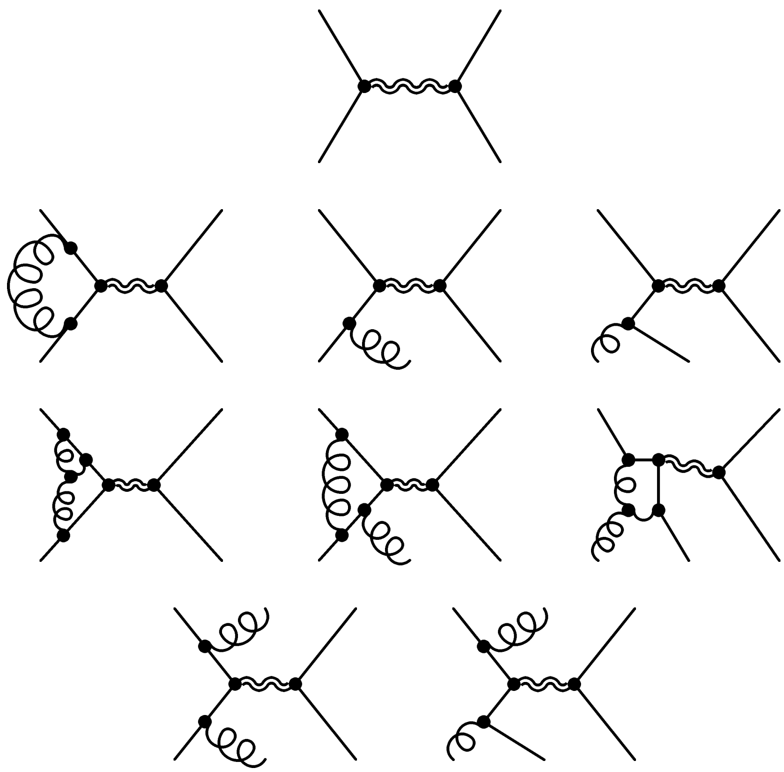
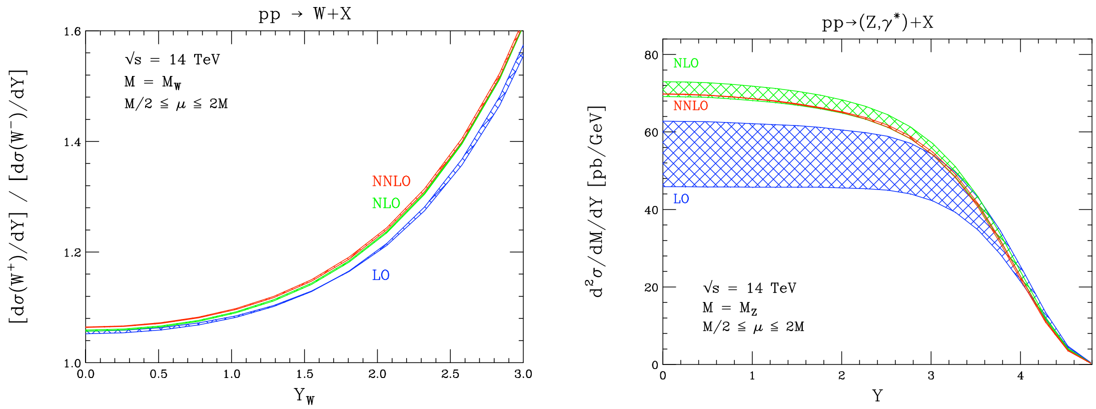


Figure 10.22: LO, NLO and NNLO Feynman diagrams for Z-production

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. 50) in a typical range of $70 - 110$ GeV. Fig. 10.24 shows the topology of a typical Z event.

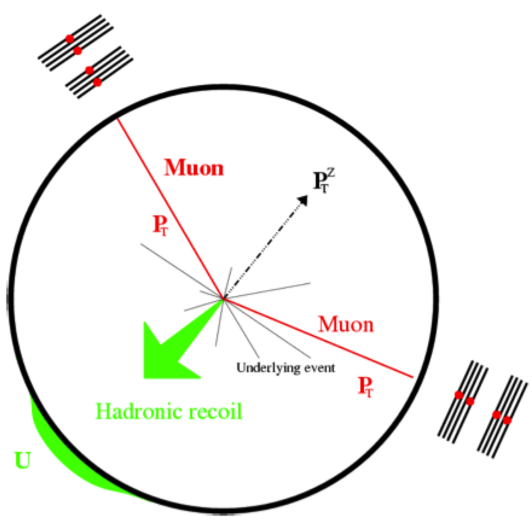


Figure 10.24: Typical dileptonic signature for a Z event.

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.

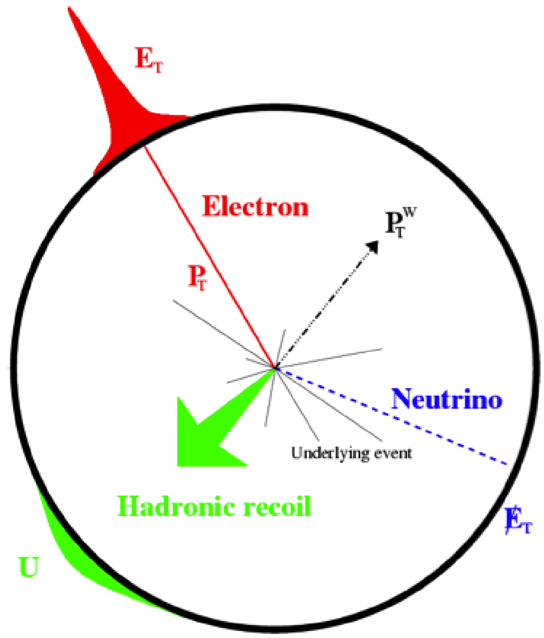


Figure 10.25: Typical signature for a W event.

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.

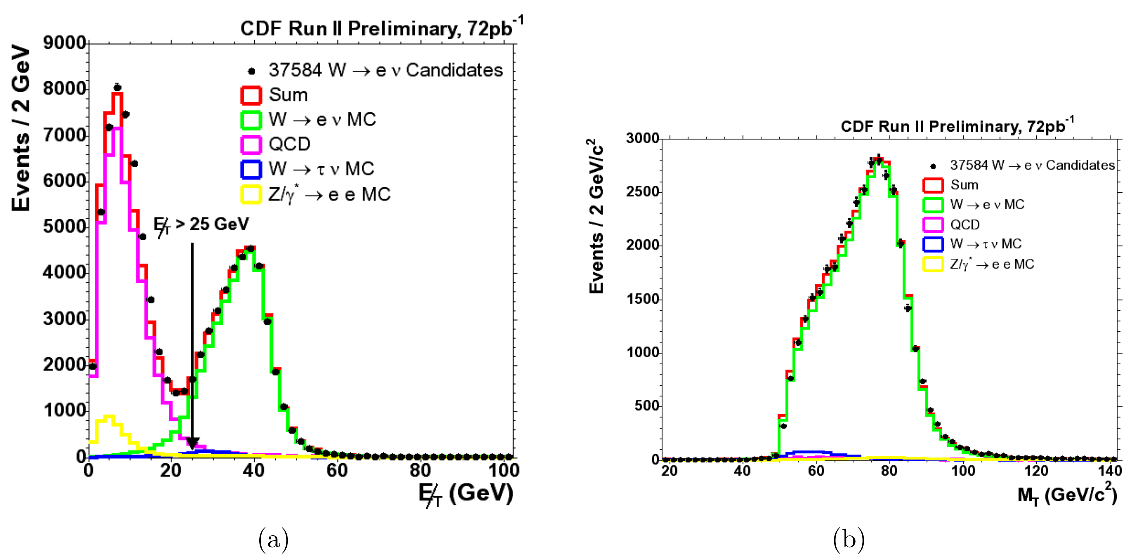
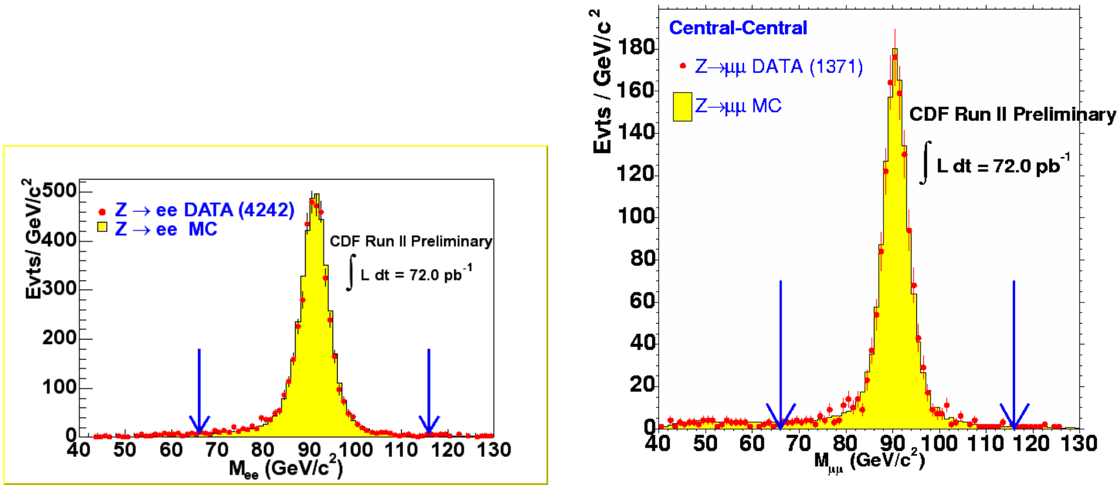


Figure 10.26: *Histogram of missing E_T associated with W production.* (a) Raw data. (b) Data after $\not{E}_T > 25\,\mathrm{GeV}$ cut.

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 $\mathcal{L}_{int} = 72\,\mathrm{pb}^{-1}$ 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):

$$\sigma_W = 2775 \pm 10(stat) \pm 53(sys) \pm 167(lum)[\mathrm{pb}]$$
$$\sigma_Z = 254.9 \pm 3.3(stat) \pm 4.6(sys) \pm 15.2(lum)[\mathrm{pb}].$$

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 $10\,\mathrm{pb}^{-1}$ 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 $\not{E}_T$.

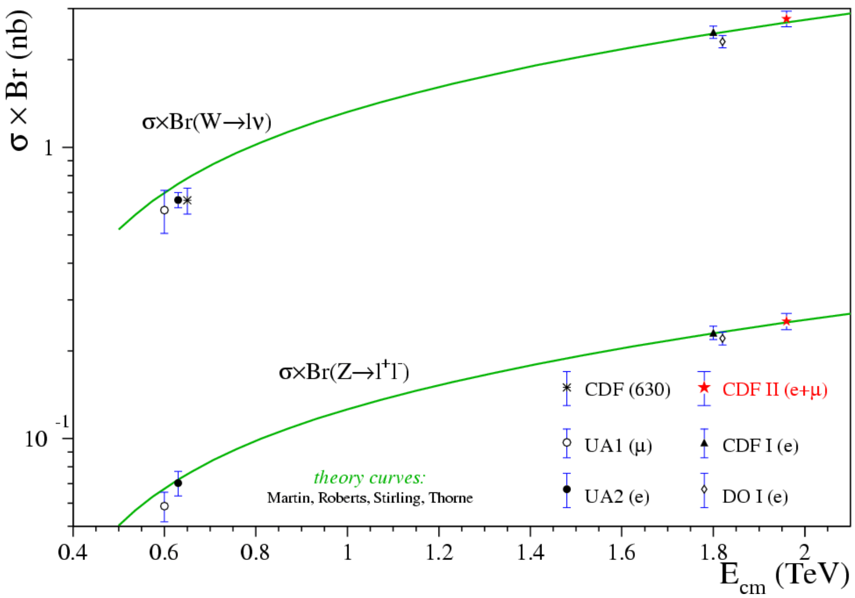


Figure 10.28: Evolution of the production cross-section for W and Z bosons.

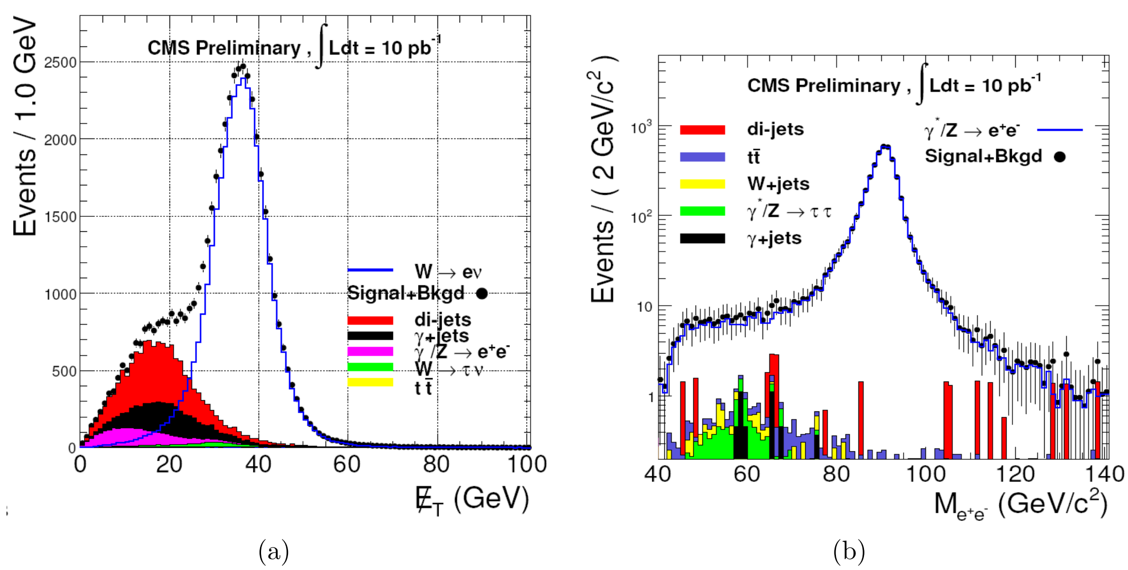
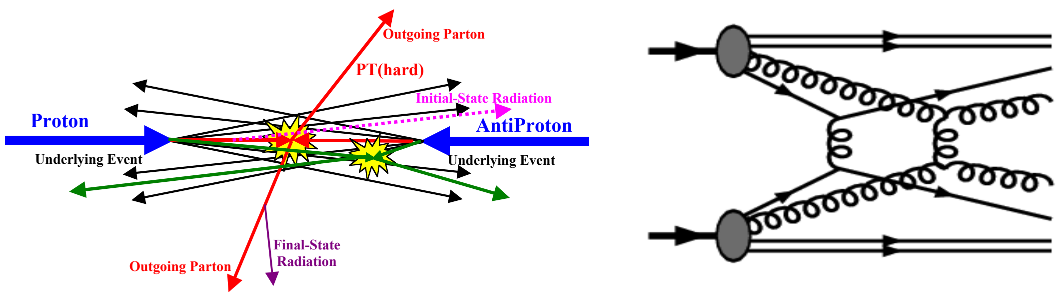


Figure 10.29: *Simulation of the different signal and backgrounds for CMS.* (a) $W^- \to e^-\bar{\nu}_e$. (b) $Z \to e^+e^-$.

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 $dp_{T,\mathrm{sum}}/d\eta d\phi$. 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.

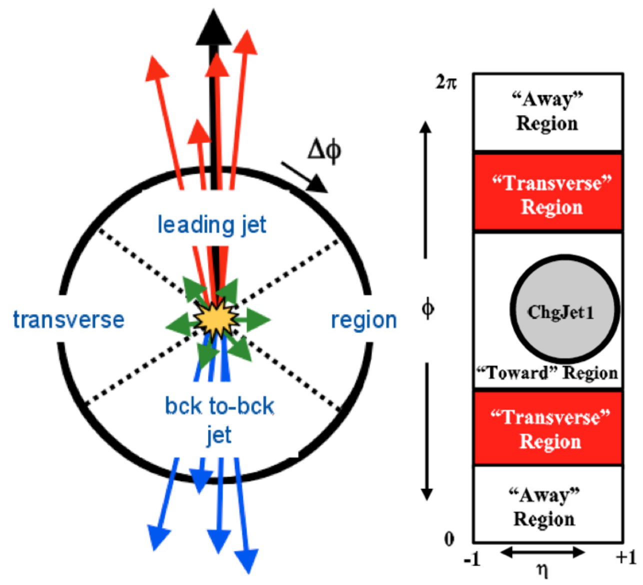


Figure 10.31: *Leading jet and transverse region.*

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 [10].

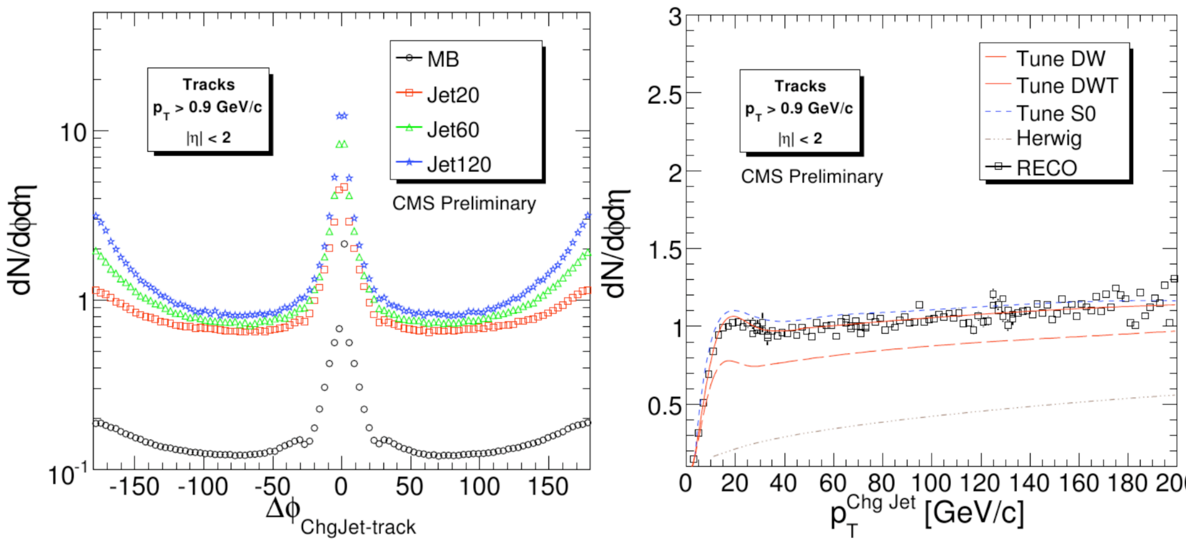


Figure 10.32: *Underlying event studies at CMS.*

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.

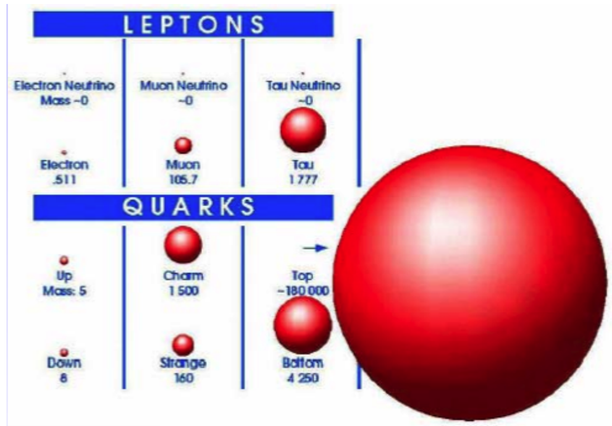
Possible measurements related to the top quark include the production cross section, the production via a heavy intermediate state Z' (resonance production), along with mass, spin and charge. A summary of top quark physics is given in Fig. 10.34.

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.

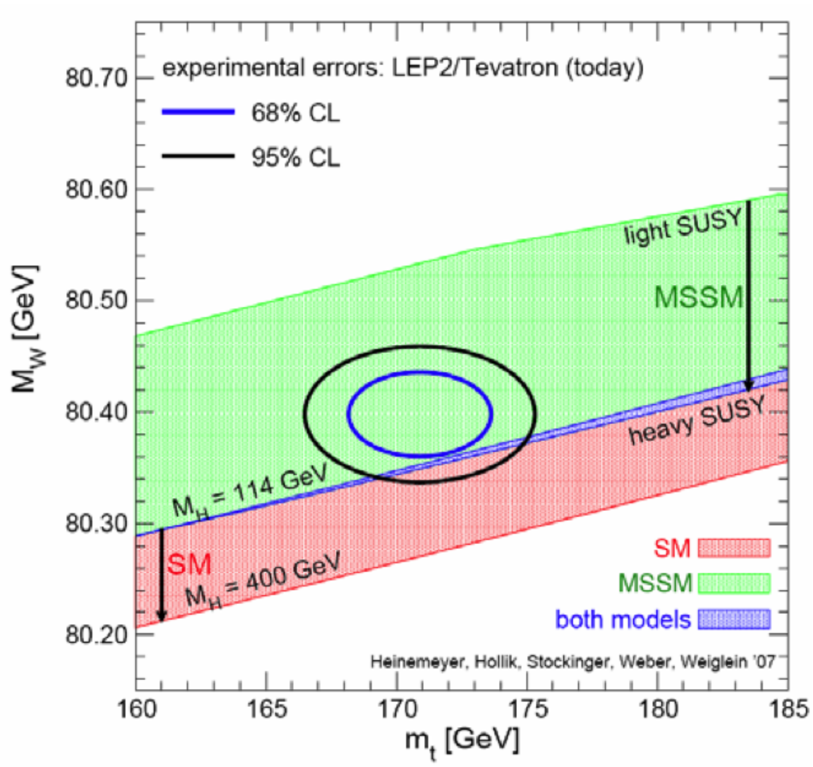
As mentioned before, the top almost immediately and exclusively decays into W^+b: $\mathrm{BR}(t \to W^+b) \sim 100\%$. According to the subsequent decays of the thus produced Ws one classifies the top decay channels as follows:

- *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).*

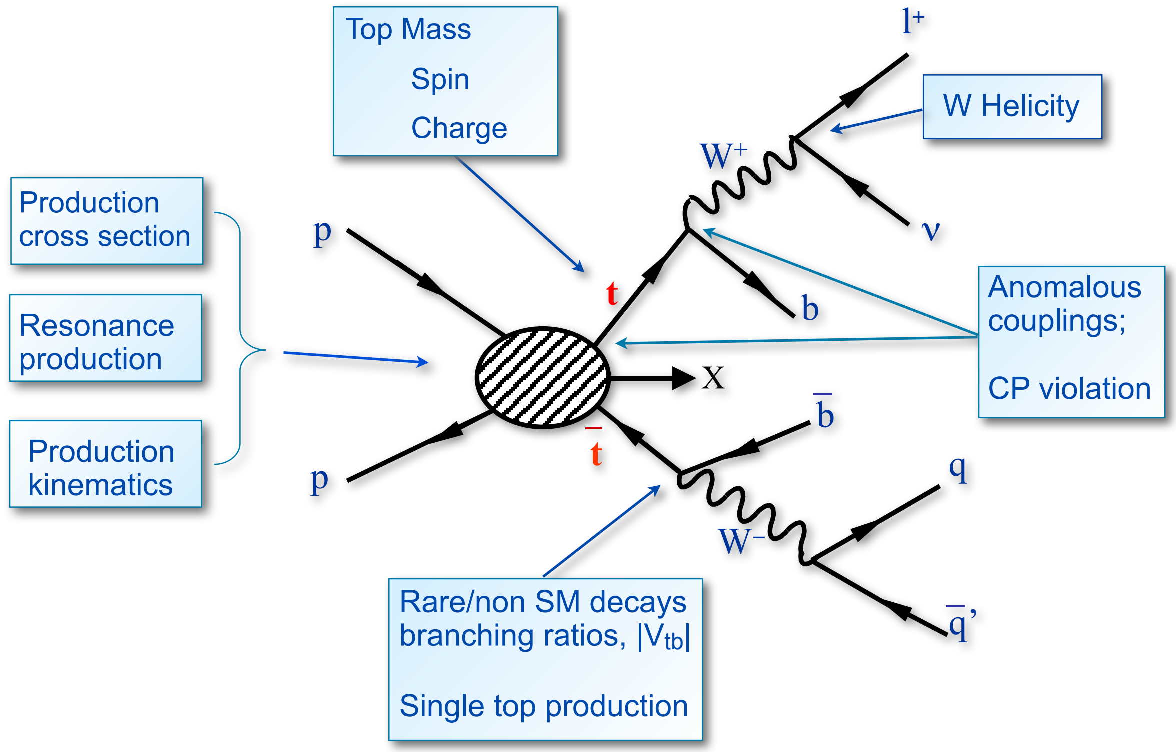


Figure 10.34: *Top physics summary.*

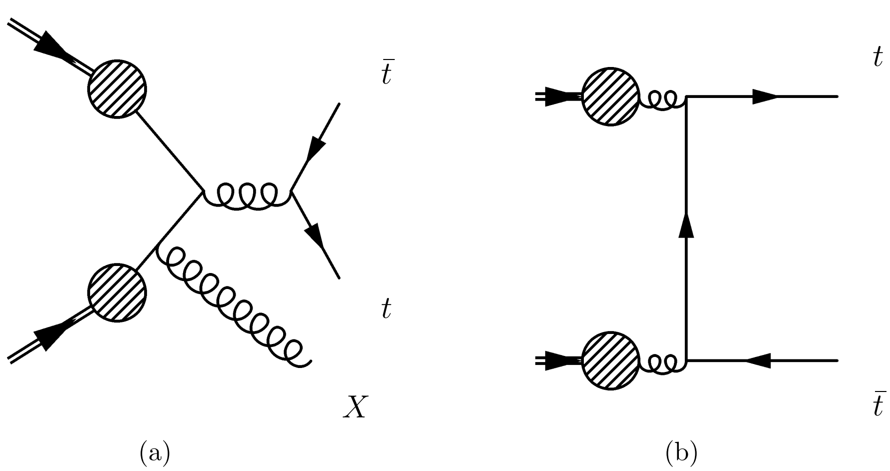
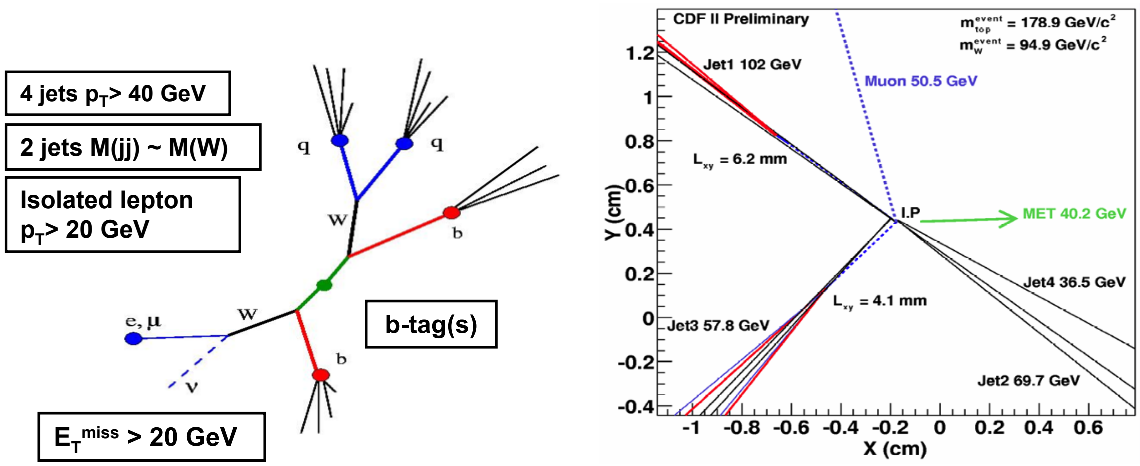


Figure 10.35: *Two possibilities for top production.* In (a) initial state radiation produces additional hadrons X while in (b) the top pair is produced by pair creation.

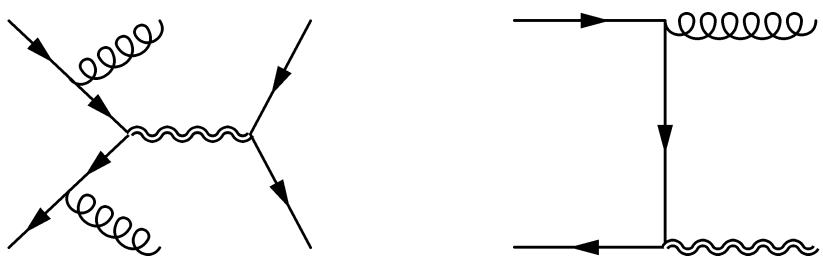


(a) (b)

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`.

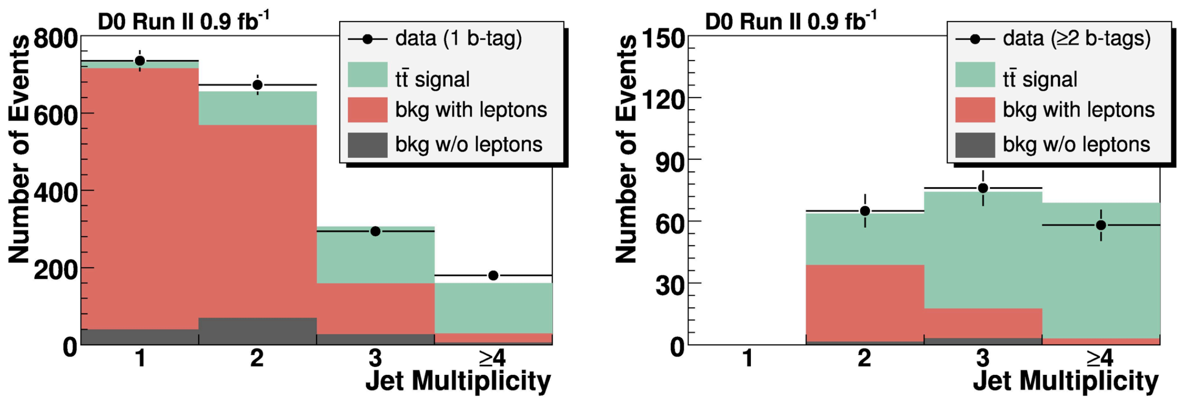


Figure 10.37: *Jet multiplicity and b tagging.*

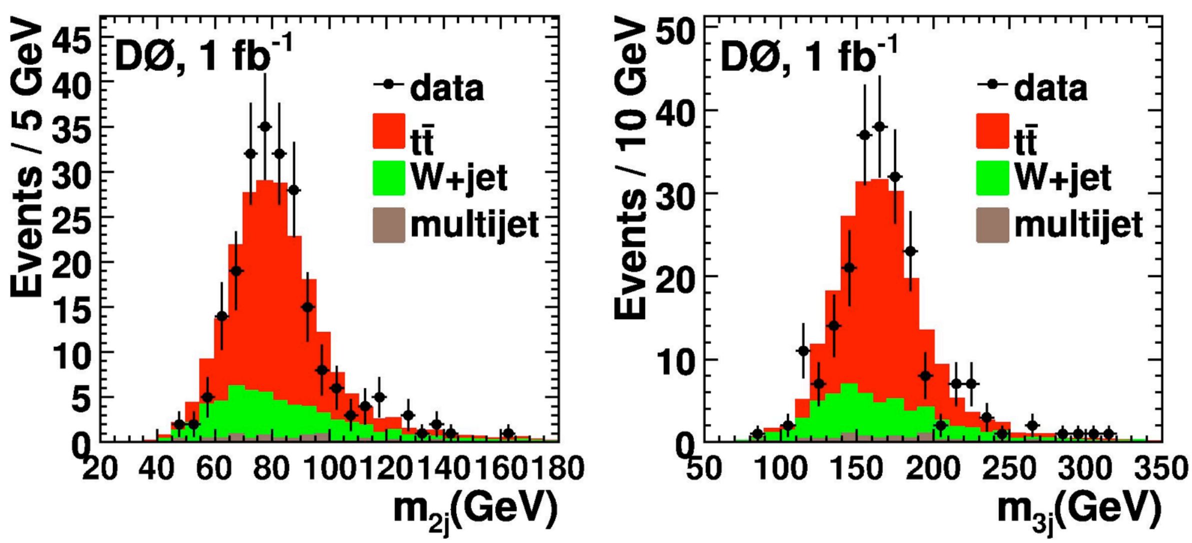
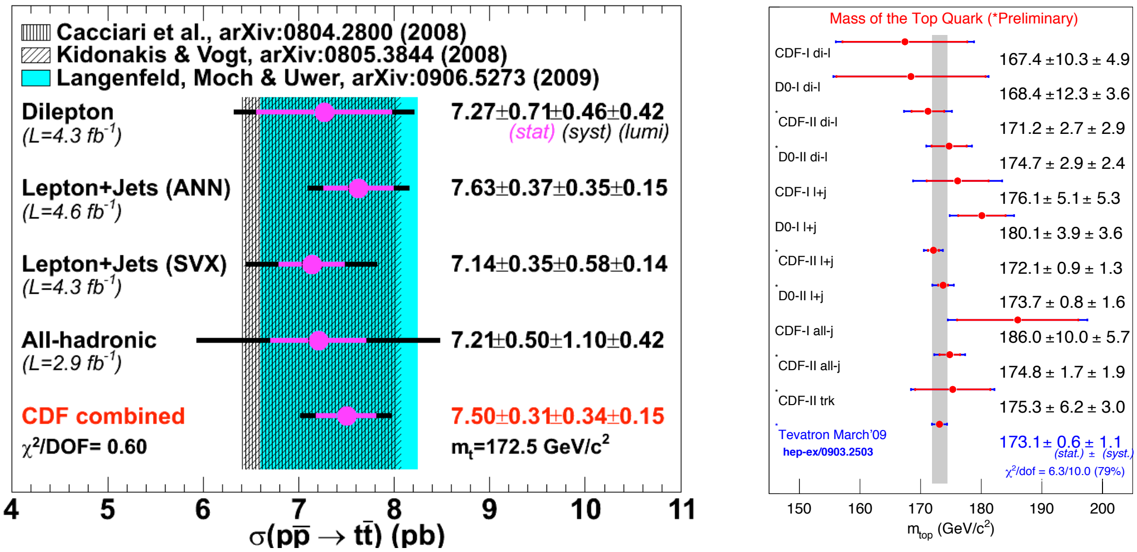


Figure 10.38: *Mass reconstruction.* Comparison between data and Monte Carlo two-jet (m_{2j}) and three-jet (m_{3j}) invariant mass distributions.



(a) (b)

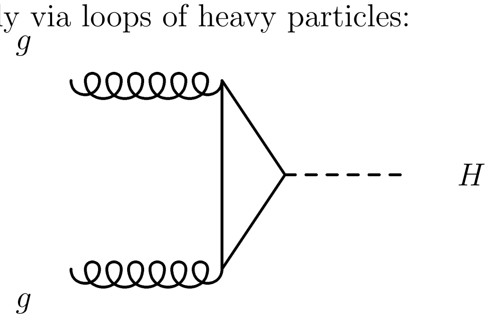
Figure 10.39: *Tevatron results for top production cross section and mass.*

By combining measurements for different decay channels (see Fig. 10.39(a)) the CDF experiment determined the top production cross section to be $\sigma^{p\bar{p}\to t\bar{t}}/\mathrm{pb} = 7.50 \pm 0.31 \pm 0.34 \pm 0.15$ (statistical, systematic, and integrated luminosity errors). This is compared to theoretical predictions for a top mass of $m_t = 172.5\,\mathrm{GeV}$ and $\sqrt{s} = 1.96\,\mathrm{TeV}$. Top mass results obtained by considering various channels are given in Fig. 10.39(b).

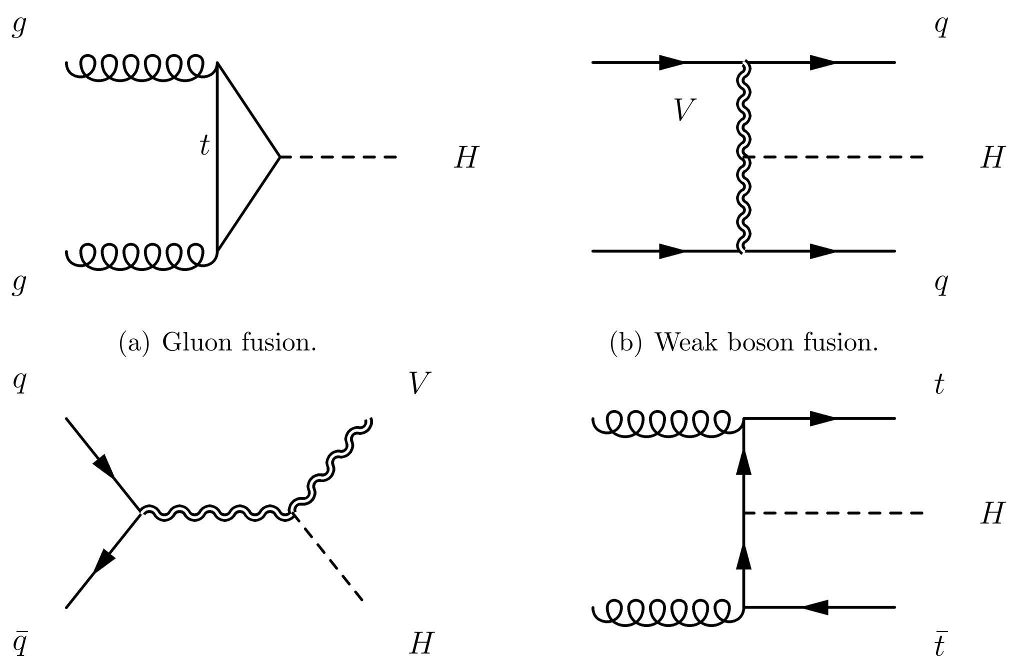
10.9 Searches for a SM Higgs and SUSY

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 $m_H \sim 100-200\,\mathrm{GeV}$), 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 $\sqrt{s} = 14\,\mathrm{TeV}$. 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 $m_H = 120 - 200\,\mathrm{GeV}$ 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.

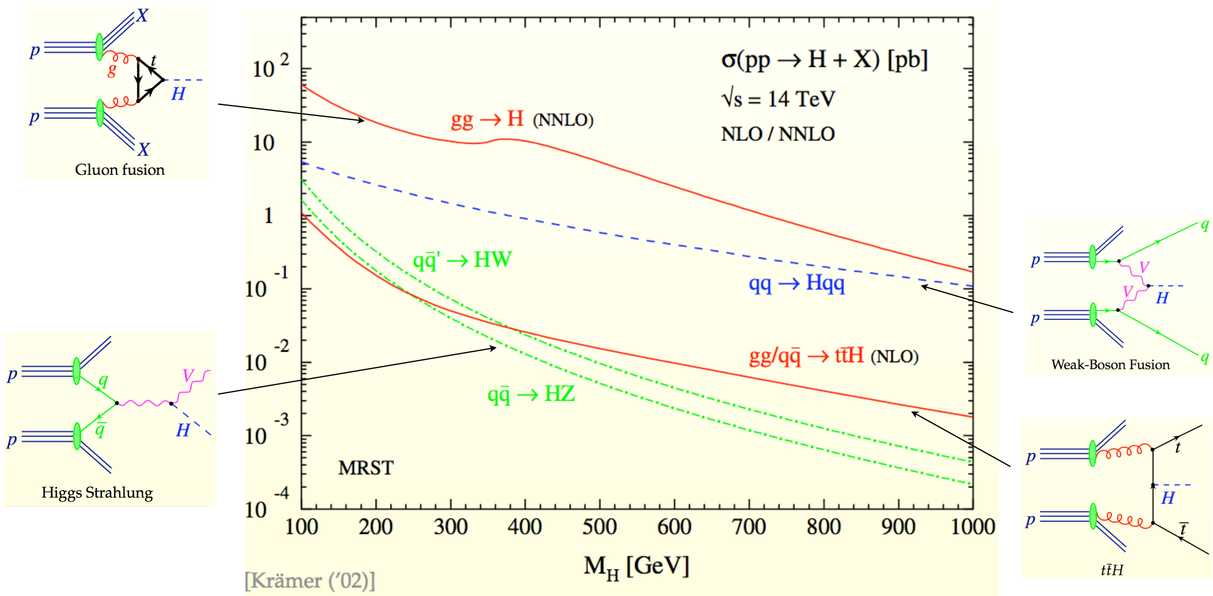
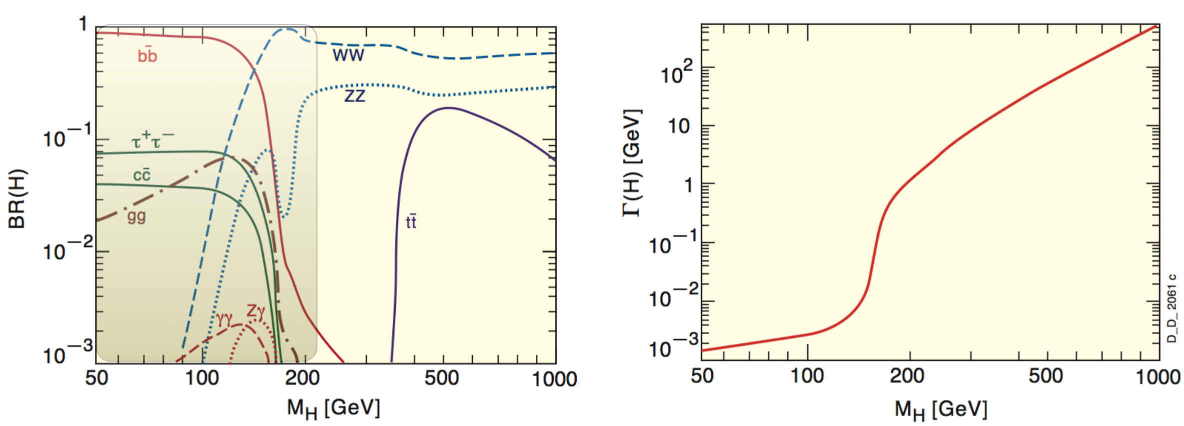


Figure 10.41: *Higgs production cross section as function of Higgs mass.*



(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.

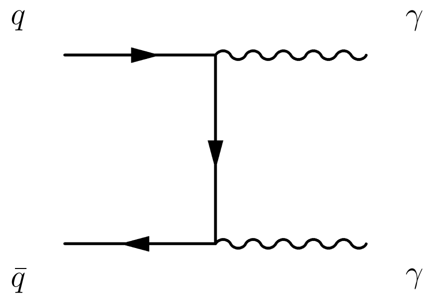
The final states can be classified according to whether mass reconstruction is possible: For the final states $\gamma\gamma$, $4l$, and $b\bar{b}$ the mass can be fully reconstructed. In these cases the background is obtained from the "sidebands" surrounding the signal box. For hadronic final states 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:

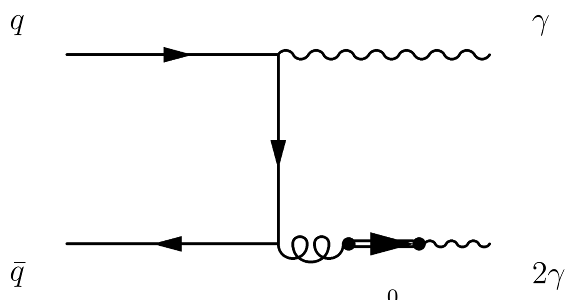
- $m_H \simeq 114 - 140\,\text{GeV}$: $\gamma\gamma(H \to \gamma\gamma)$;
- $m_H \simeq 140 - 175\,\text{GeV}$: $2l + \not{E}_T(H \to WW^{(\star)})$ and $4l(H \to ZZ^{(\star)})$;
- $m_H \simeq 175 - 600\,\text{GeV}$: $4l(H \to ZZ^{(\star)})$.

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.

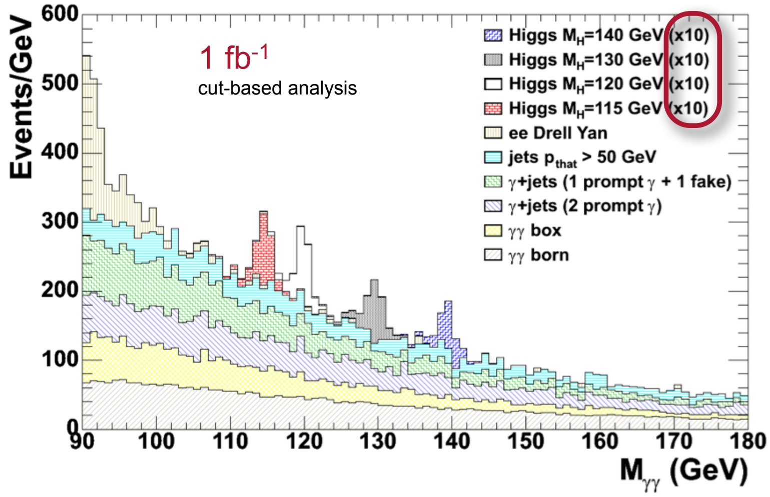
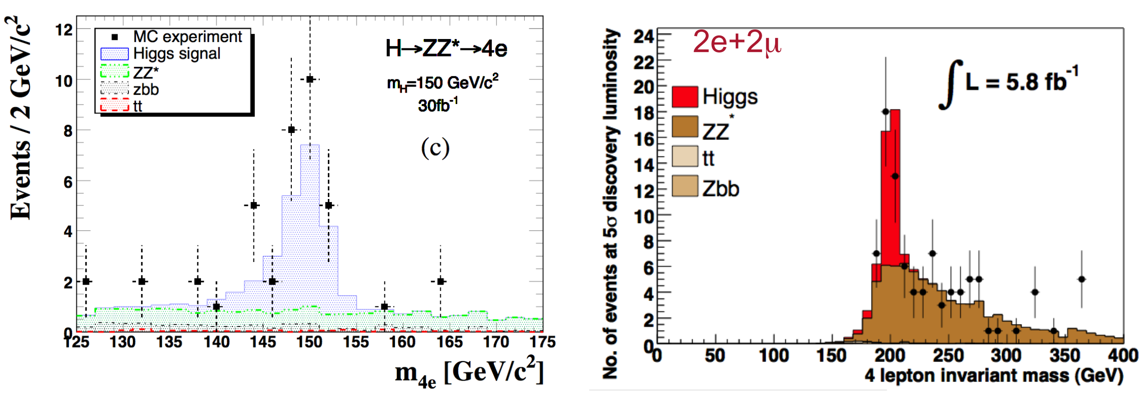
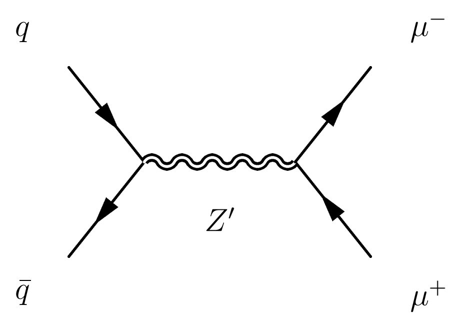


Figure 10.43: *Invariant mass a of photon pair for the Higgs decay $H \to \gamma\gamma$.* Note that the Higgs peaks are increased by a factor of 10.

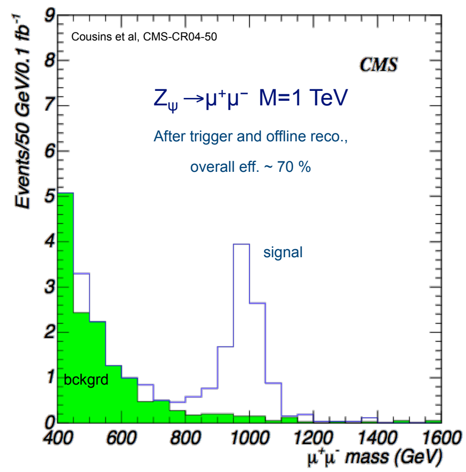


(a) (b)

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 $1-2\,\mathrm{TeV}$ would be easily detectable.

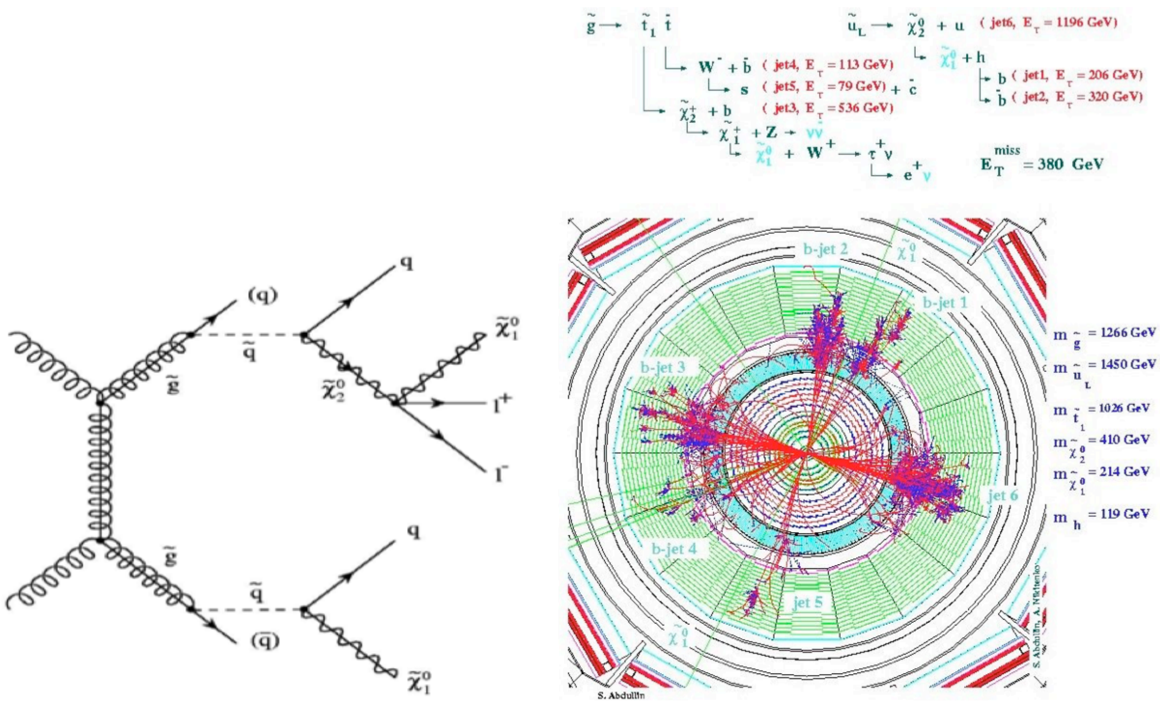
2. *An intermediate case.*
   SUSY (See below.)

3. *A difficult case.*
   As we have seen, a *light Higgs* with $m_H \sim 115-120\,\mathrm{GeV}$ 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.



(a) (b)

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

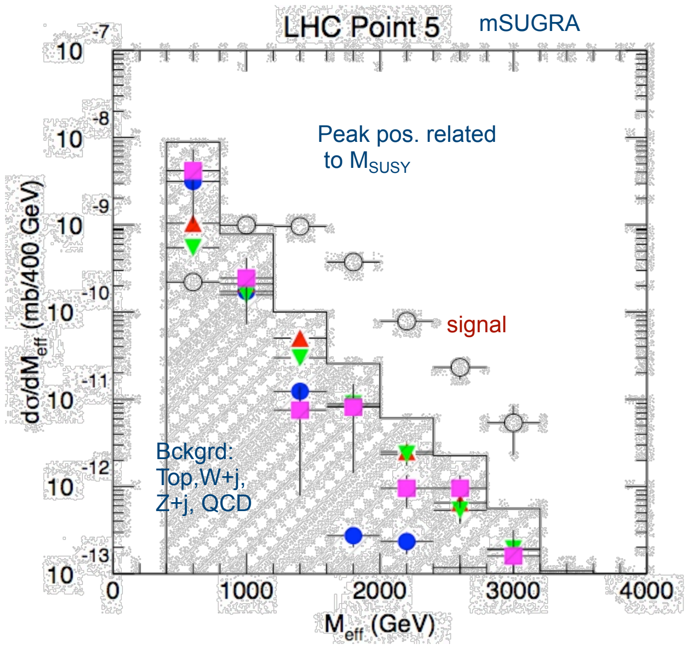


Figure 10.47: *Typical SUSY signal and backgrounds.*

selection would be based on the following criteria: $N_{\rm jet} > 4$, $E_T > 100$, 50, 50, 50 GeV, and $\not{E}_T > 100$ 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 [11]

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. $\mu^+ \to e^+ + \bar{\nu}_\mu + \nu_e$; $\nu_e + e^- \to \nu_e + e^-$.

- *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

$$\mathcal{M} = \frac{G_F}{\sqrt{2}} (\bar{\psi}_p \gamma^\mu \psi_n)(\bar{\psi}_e \gamma_\mu \psi_{\bar{\nu}}), \tag{11.1}$$

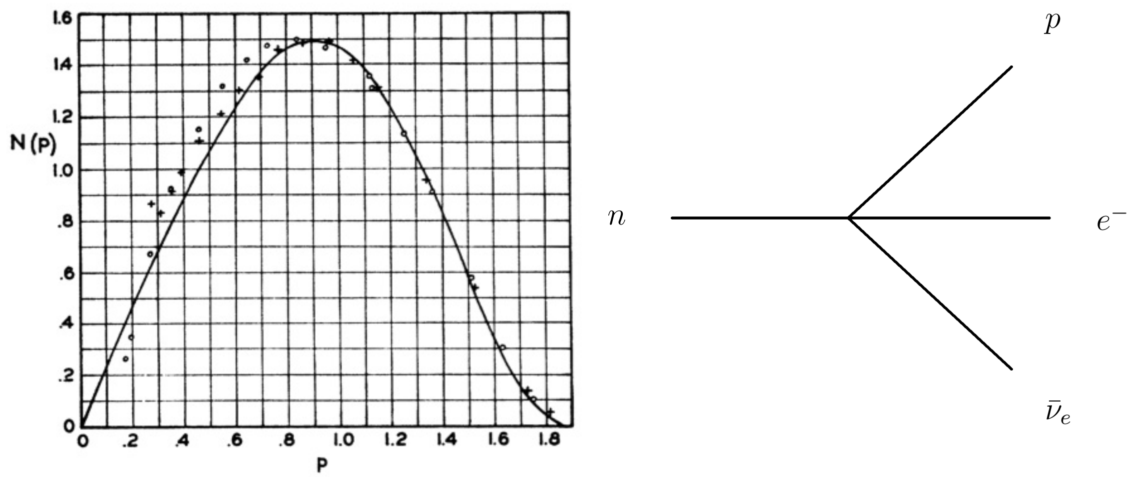
where $G_F \sim 1/(300\,\mathrm{GeV})^2$ is the Fermi constant. Note that the expression in Eq. (11.1) has vector structure and therefore does not violate parity. This point will be revisited later on.

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

$$\begin{aligned} L_l &= 1 && e^-,\ \mu^-,\ \nu_e,\ \ldots \\ L_{\bar{l}} &= -1 && e^+,\ \mu^+,\ \bar{\nu}_e,\ \ldots. \end{aligned}$$

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: [12, 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:

$$K^+ \begin{cases} \theta \to \pi^+\pi^0 \\ \tau \to \pi^+\pi^+\pi^- \end{cases}$$
$$P\,|\pi\pi\rangle = (-1)(-1)(-1)^l = +1$$
$$P\,|\pi\pi\pi\rangle = (-1)^3(-1)^{l_{\pi_1\pi_2}}(-1)^{l_{\pi_3}} = -1,$$

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:

$$\begin{array}{ll} \uparrow & \underbrace{{}^{60}\mathrm{Co}}_{J=5} \to \underbrace{{}^{60}\mathrm{Ni}^\star}_{J=4} + \underbrace{e^- + \bar{\nu}_e}_{J_z=1} \\ \vec{B} & \quad\Uparrow \qquad\quad \Uparrow \qquad \downarrow\Uparrow \quad \uparrow\Uparrow \end{array}$$

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]

$$\begin{array}{ll}\bar\psi\psi & \text{scalar}\\ \bar\psi\gamma^\mu\psi & \text{vector}\\ \bar\psi\sigma^{\mu\nu}\psi & \text{tensor}\\ \bar\psi\gamma_5\psi & \text{pseudoscalar}\\ \bar\psi\gamma^\mu\gamma_5\psi & \text{pseudovector.}\end{array}$$

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. [13, p. 64].

- $\gamma_5^2 = \mathbb{1}$;
- $\{\gamma_5, \gamma_\mu\} = 0$;
- $\gamma_5^\dagger = i\gamma^3\gamma^2\gamma^1\gamma^0 = \gamma_5$;
- $\mathrm{Tr}\gamma_5 = 0$;
- 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:

- $\mathrm{Tr}\gamma_5 = 0$;
- $\mathrm{Tr}(\gamma_5\gamma^\mu\gamma^\nu) = 0$;

- $\mathrm{Tr}(\gamma_5\gamma^\alpha\gamma^\beta\gamma^\gamma\gamma^\delta) = -4i\varepsilon^{\alpha\beta\gamma\delta}$
  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

$$\Gamma = \int_0^{m_\mu/2} dE_{p'} \frac{d\Gamma}{dE_{p'}} = \frac{G_F^2 m_\mu^5}{192\pi^3} = \frac{1}{\tau}.$$

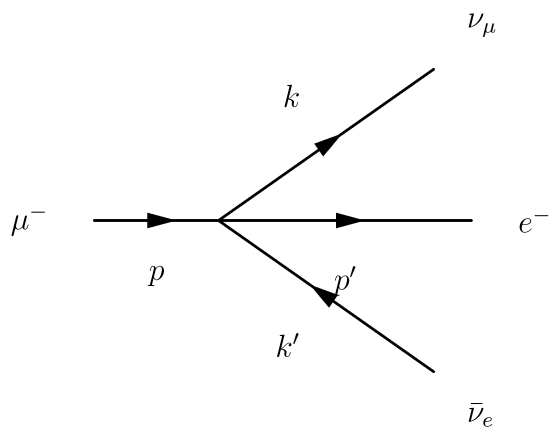


Figure 11.2: *Leptonic muon decay.*

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 $E_{\mathrm{CM}} \to \infty$:

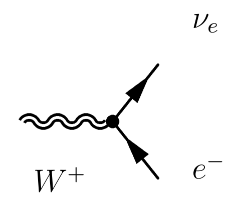
$$\sigma^{e^- + \nu_e \to e^- \nu_e} = \frac{4G_F^2}{\pi} E_{\mathrm{CM}}^2.$$

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.

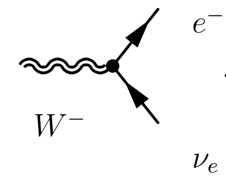
11.5 Weak isospin and hypercharge

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,

$$\tau_\pm = \frac{1}{2}(\tau_1 \pm i\tau_2),$$

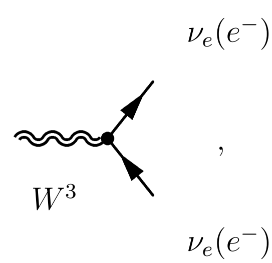
where the τ_i are the usual Pauli matrices. With this formalism, one can write the charged currents as,

$$j_\mu^+ = \bar{\chi}_L \gamma_\mu \tau_+ \chi_L \tag{11.9}$$

$$j_\mu^- = \bar{\chi}_L \gamma_\mu \tau_- \chi_L \tag{11.10}$$

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.98) and QCD (Sec. 7.4, p. 138), we expect the electroweak interaction to be mediated by gauge fields. In the case of QED, we had,

$$\mathcal{L}_{int}^{\mathrm{QED}} = -iej_\mu^{em}A^\mu,$$

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 :

$$\mathcal{L}_{int}^{\mathrm{EW}} = -igj_\mu^i W^{i\mu} - i\frac{g'}{2}j_\mu^Y B^\mu, \tag{11.16}$$

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,

$$
\begin{aligned}
-igj^3_\mu W^{3\mu} - i\frac{g'}{2}j^Y_\mu B^\mu = &- i\left(g\sin\theta_w j^3_\mu + g'\cos\theta_w \frac{j^Y_\mu}{2}\right)A^\mu \\
&- i\left(g\cos\theta_w j^3_\mu - g'\sin\theta_w \frac{j^Y_\mu}{2}\right)Z^\mu .
\end{aligned}
$$

The first term corresponds to the electromagnetic current, for which we had $j^{em}_\mu = j^3_\mu + \frac{1}{2}j^Y_\mu$, 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 $j^Y_\mu = 2(j^{em}_\mu - j^3_\mu)$, we get,

$$
j^{\mathrm{NC}}_\mu = \frac{g}{\cos\theta_w}(j^3_\mu - \sin^2\theta_w j^{em}_\mu). \tag{11.18}
$$

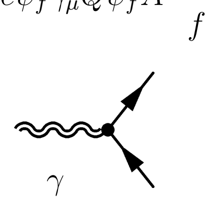
11.7 Electroweak Feynman rules

Vertices The Feynman rules for vertices stemming from,

$$
\mathcal{L}^{\mathrm{EW}}_{int} = \mathcal{L}^{\mathrm{QED}}_{int} + \mathcal{L}^{\mathrm{CC}}_{int} + \mathcal{L}^{\mathrm{NC}}_{int},
$$

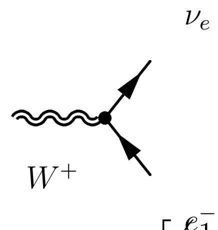
can be computed as follows,

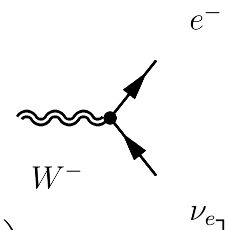
$$i\mathcal{L}_{int}^{\text{QED}} = -ie\bar{\psi}_f\gamma_\mu Q\psi_f A^\mu$$

$\Rightarrow$ 

$= -ieQ_f\gamma_\mu$

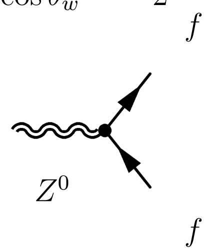
$$\begin{aligned} i\mathcal{L}_{int}^{\text{CC}} &= -i\frac{g}{\sqrt{2}}(\bar{\chi}_L\gamma_\mu\tau_+\chi_L)W^{+\mu} - i\frac{g}{\sqrt{2}}(\bar{\chi}_L\gamma_\mu\tau_-\chi_L)W^{-\mu} \\ &= -i\frac{g}{\sqrt{2}}\bar{\nu}\gamma_\mu\left(\frac{1-\gamma_5}{2}\right)eW^{+\mu} - i\frac{g}{\sqrt{2}}\bar{e}\gamma_\mu\left(\frac{1-\gamma_5}{2}\right)\nu W^{-\mu} \end{aligned}$$

$\Rightarrow$ 

$=$ 

$= -i\frac{g}{\sqrt{2}}\gamma_\mu\left(\frac{1-\gamma_5}{2}\right)$

$$\begin{aligned} i\mathcal{L}_{int}^{\text{NC}} &= -i\frac{g}{\cos\theta_w}\bar{\psi}_f\gamma_\mu\left[\left(\frac{1-\gamma_5}{2}\right)T_3 - \sin^2\theta_w Q\right]\psi_f Z^\mu \\ &= -i\frac{g}{\cos\theta_w}\bar{\psi}_f\gamma_\mu\frac{1}{2}(c_V^f - c_A^f\gamma_5)\psi_f Z^\mu \end{aligned}$$

$\Rightarrow$ 

$= -i\frac{g}{\cos\theta_w}\gamma_\mu\frac{1}{2}(c_V^f - c_A^f\gamma_5)$

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,

$$i\frac{\sum_\lambda (\varepsilon_\lambda^\mu)^* \varepsilon_\lambda^\nu}{p^2 - M^2}.$$

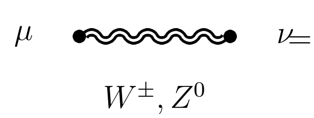
The polarisation sum $\Pi^{\mu\nu}$ must take the form,

$$\Pi^{\mu\nu} = \sum_\lambda (\varepsilon_\lambda^\mu)^* \varepsilon_\lambda^\nu = A g^{\mu\nu} + B p^\mu p^\nu.$$

Using the identities,

$$p_\mu p^\mu = M^2, \qquad p_\mu \Pi^{\mu\nu} = p_\nu \Pi^{\mu\nu} = 0, \qquad g_{\mu\nu}\Pi^{\mu\nu} = 3,$$

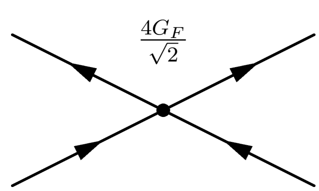
coming from the on-shell condition, the conservation of current and the count of polarization states (for a massive particle) respectively, we get $A = -1$ and $B = M^{-2}$, making us able to write,



$$= i\frac{-g^{\mu\nu} + p^\mu p^\nu / M^2}{p^2 - M^2}.$$

So unless momentum transfer is not of the order of $M \gtrsim 100\,\mathrm{GeV}$, the propagator gets suppressed drastically by the mass.

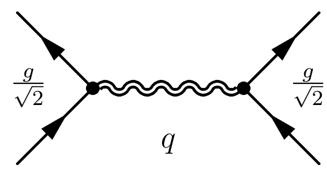
Relation of the Fermi $V - A$-interaction In $V - A$-theory, we have a 4 point vertex,



yielding the matrix element,

$$\mathcal{M}^{V-A} = \frac{4G_F}{\sqrt{2}} j^\mu j_\mu^\dagger.$$

The same process, viewed as the exchange of a low momentum ($q^2 \ll M_W^2$) vector boson,



corresponds to the matrix element,

$$\mathcal{M}^{\text{EW}} \approx \left(\frac{g}{\sqrt{2}} j^\mu\right) \frac{1}{M_W^2} \left(\frac{g}{\sqrt{2}} j_\mu^\dagger\right),$$

yielding the relation,

$$\boxed{G_F = \frac{\sqrt{2}g^2}{8M_W^2}}. \tag{11.21}$$

From this relation, the first estimates of the mass of the $W^\pm$ bosons were $50 - 100\,\text{GeV}$.

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:

$$\phi(x) = v + \eta(x) \tag{11.26}$$

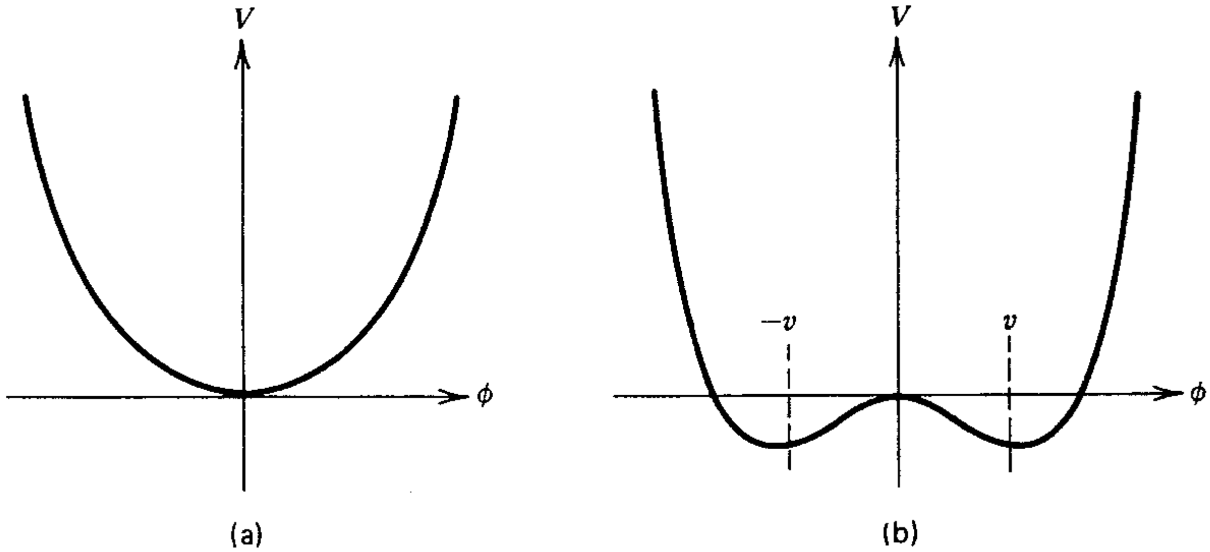


Figure 11.3: *The Potential* $V(\phi)$ *for (a)* $\mu^2 > 0$ *and (b)* $\mu^2 < 0$ *and* $\lambda > 0$. Source: [1, p. 322].

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:

$$\phi = \frac{1}{\sqrt{2}}(\phi_1 + i\phi_2) \tag{11.28}$$

$$\mathcal{L} = (\partial_\mu \phi)^*(\partial^\mu \phi) - \mu^2 \phi^* \phi - \lambda(\phi^* \phi)^2. \tag{11.29}$$

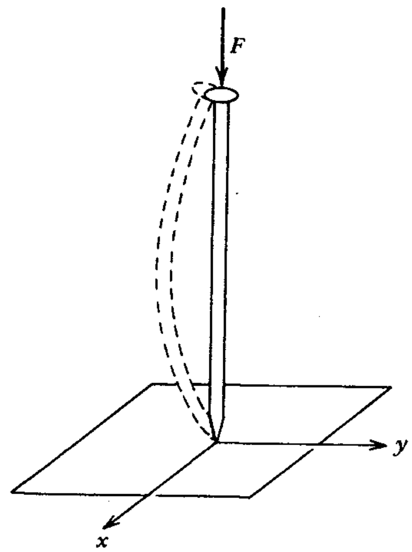


Figure 11.4: *Bending of an elastic bar.* Source: [1, p. 324].

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.

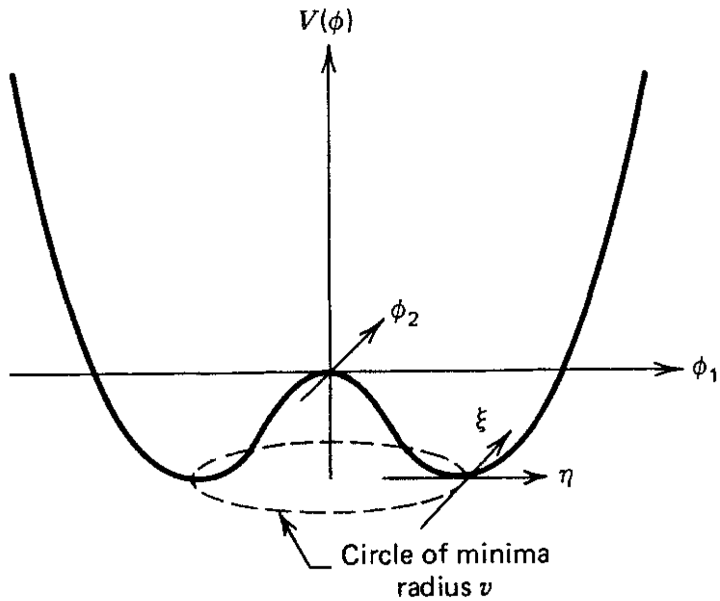


Figure 11.5: *The potential $V(\phi)$ for a complex scalar field for the case $\mu^2 < 0$ and $\lambda > 0$.* Source: [1, p. 325].

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 $SU(2)_L \times U(1)_Y$:

$$D_\mu = \partial_\mu - ig\frac{1}{2}\vec{\tau} \cdot \vec{W}_\mu - ig'\frac{1}{2}YB_\mu. \tag{11.40}$$

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

$$v^2 = -\frac{\mu^2}{\lambda} \qquad \phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix}. \tag{11.43}$$

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

$$\boxed{M_W = \frac{1}{2} vg} \qquad \boxed{M_Z = \frac{1}{2} v \sqrt{g^2 + g'^2}}. \tag{11.44}$$

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:

$$G_F = \frac{\sqrt{2} g^2}{8 M_W^2} = \frac{1}{\sqrt{2} v^2} \rightarrow \boxed{v = 246\,\mathrm{GeV}}.$$

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):

$$\begin{aligned}\mathcal{L}_{m-} &= -G^e\left[\begin{pmatrix}\bar{\nu}_e & \bar{e}\end{pmatrix}_L\begin{pmatrix}\phi^+\\ \phi^0\end{pmatrix}e_R + \bar{e}_R\begin{pmatrix}\bar{\phi}^+ & \bar{\phi}^0\end{pmatrix}\begin{pmatrix}\nu_e\\ e\end{pmatrix}_L\right]\\ &= -\frac{G^e v}{\sqrt{2}}\left(\bar{e}_L e_R + \bar{e}_R e_L\right) - \frac{G^e}{\sqrt{2}}h\left(\bar{e}_L e_R + \bar{e}_R e_L\right),\end{aligned} \tag{11.46}$$

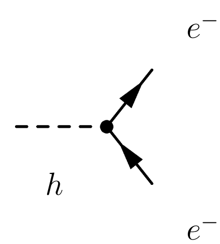
where G^e denotes the **Yukawa coupling** of the electron, and we used,

$$\begin{pmatrix}\phi^+\\ \phi^0\end{pmatrix} = \frac{1}{\sqrt{2}}\begin{pmatrix}0\\ v + h(x)\end{pmatrix}.$$

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 $m_e = 511\,\mathrm{keV}$ and $v = 246\,\mathrm{GeV}$, this vertex factor is very small for the electron. In the case of the top, $m_t = 172\,\mathrm{GeV}$ 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.

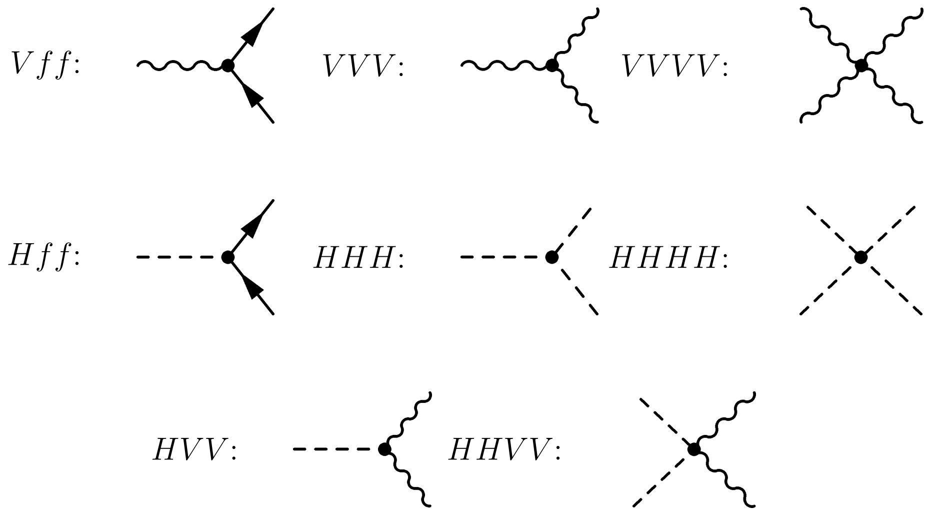
After the spontaneous symmetry breaking, we have,

$$\phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + h(x) \end{pmatrix},$$

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. 22),

$$\Gamma_H = \frac{1}{2M_H}\frac{1}{(2\pi)^2}\sum_f \int \frac{d^3p_1}{2E_1}\frac{d^3p_2}{2E_2}\delta^{(4)}(p_f - p_H)|\mathcal{M}_{fH}|^2,$$

where f denotes the final state : $b\bar{b}, t\bar{t}, W^+W^-, Z^0Z^0, \tau^+\tau^-, \dots$ 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,

$$R_2 = \int \frac{d^3p_1}{2E_1}\frac{d^3p_2}{2E_2}\delta^{(4)}(p_f - p_H) = \frac{\pi}{2M_H^2}\sqrt{\lambda\left(M_H^2, m_f^2, m_f^2\right)} = \frac{\pi}{2}\sqrt{1 - \frac{4m_f^2}{M_H^2}},$$

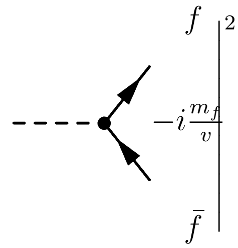
and hence,

$$\Gamma_H = \frac{1}{16\pi M_H}\sum_f \sqrt{1 - \frac{4m_f^2}{M_H^2}}|\mathcal{M}_{fH}|^2 = \sum_f \Gamma_{H\to f}. \tag{11.58}$$

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,

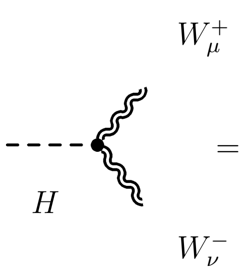
$$\Gamma_{H\to l^+l^-} = \frac{1}{8\pi^2 v^2} m_f^2 M_H \left(1 - \frac{4m_f^2}{M_H^2}\right)^{\frac{3}{2}} \tag{11.59}$$

$$\Gamma_{H\to q\bar{q}} = \frac{3}{8\pi^2 v^2} m_q^2 M_H \left(1 - \frac{4m_q^2}{M_H^2}\right)^{\frac{3}{2}}. \tag{11.60}$$

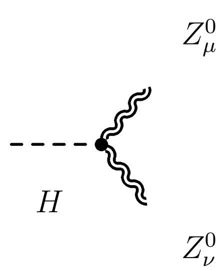
We remark at this point that the dominant decay mode (corresponding to the largest partial width) is always into the heaviest kinematically allowed fermion. In the case of a light Higgs boson ($M_H < 2M_{W,Z}$), the dominant channels would be into $b\bar{b}$ and $\tau^+\tau^-$.

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,

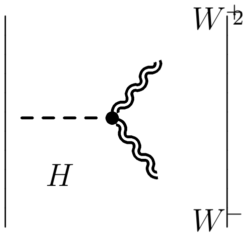


$$= igM_W g_{\mu\nu},$$



$$= \frac{igM_Z}{\cos\theta_w} g_{\mu\nu},$$

Summing the moduli squared over the polarizations, we get,



$$\sum_\lambda \left| \cdots \right|^2 = g^2 M_W^2 \left(-g^{\mu\rho} + \frac{p_1^\mu p_2^\rho}{M_W^2}\right)\left(-g_{\mu\rho} + \frac{p_{1\mu}p_{2\rho}}{M_W^2}\right)$$

$$= \frac{g^2 M_H^4}{4M_W^2}\left(1 - 4\frac{M_W^2}{M_H^2} + 12\frac{M_W^4}{M_H^4}\right),$$

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 $M_H \approx 1\,\text{TeV}$, $\Gamma_H \approx M_H$ 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 $Q \approx \Lambda^{\text{QCD}}$).

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 $A_{\text{fb}}^{0,b}$, 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 $M_W \sim 83\,\text{GeV}$ and $M_Z \sim 93\,\text{GeV}$. Therefore, to produce W and Z bosons, a particle collider was needed capable of producing particles with mass $\sim 100\,\text{GeV}$. A the time, two candidates were available at CERN. The ISR with $\sqrt{s} = 61\,\text{GeV}$ 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 $\sqrt{s} = 540\,\text{GeV}$. It had a luminosity of $5 \cdot 10^{27}\,\text{cm}^{-2}\text{s}^{-1}$, 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.

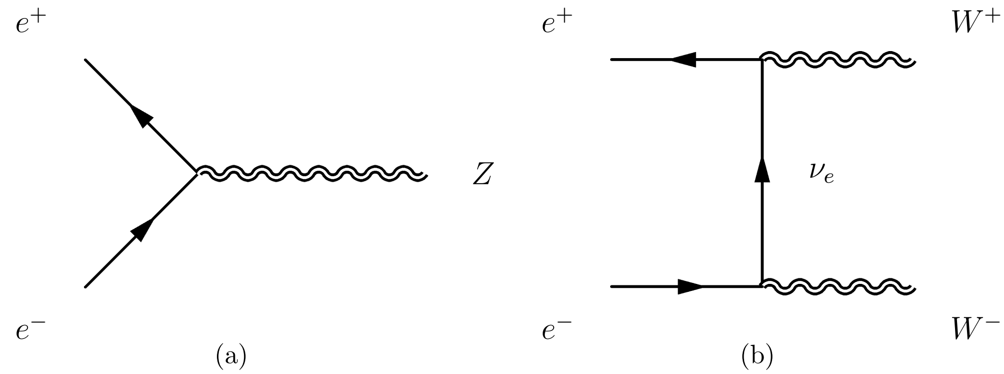


Figure 11.6: *Z (a) and W (b) boson production at electron-positron colliders.*

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:

$$\frac{\sigma(\bar{p}p \to WX \to e\nu X)}{\sigma_T(p\bar{p})} \simeq 10^{-8}.$$

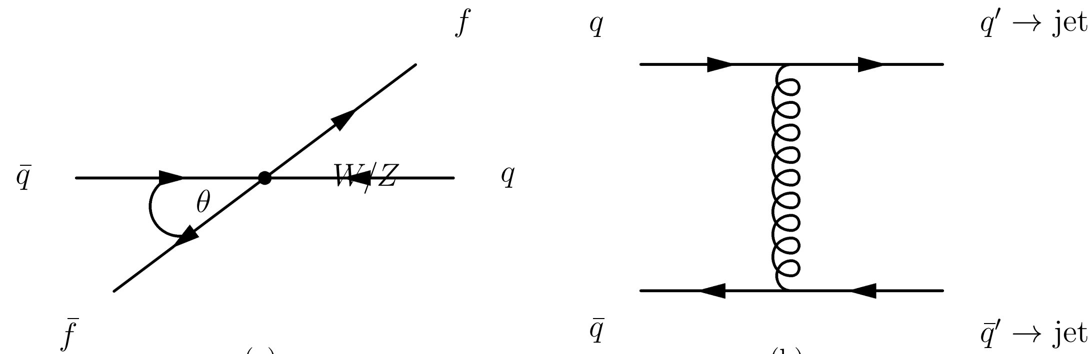
Therefore, it is preferred to look for W and Z decays into leptons, where the background is smaller:[8]

$$\begin{aligned} W^\pm &\to e^\pm \overset{(-)}{\nu}_e,\ \mu^\pm \overset{(-)}{\nu}_\mu,\ \tau^\pm \overset{(-)}{\nu}_\tau \\ Z^0 &\to e^+e^-,\ \mu^+\mu^-, \tau^+\tau^-. \end{aligned}$$

11.13.2.1 W discovery and mass measurement

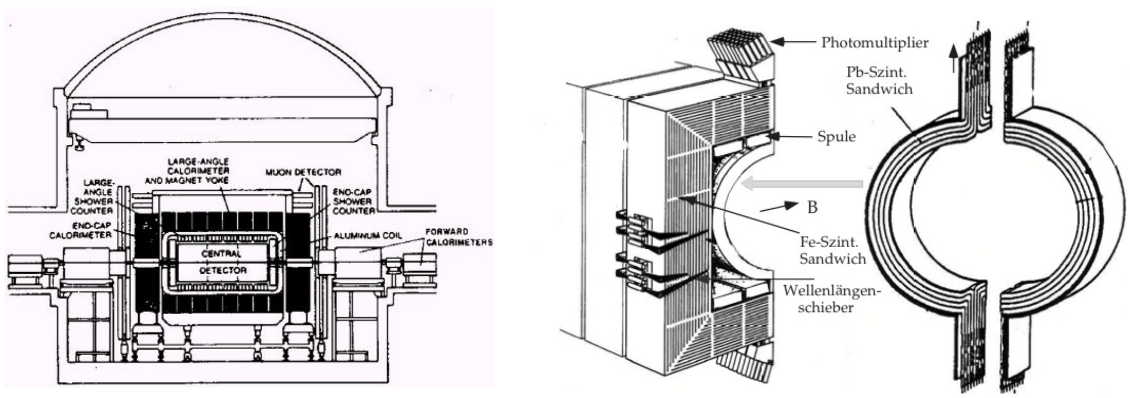
The UA1 experiment at the S$p\bar{p}$S collider was an hermetic particle detector optimized for the $W^\pm \to e^\pm\nu_e/\bar{\nu}_e$ measurement. It featured for the first time the general design principles of collider detectors (see also Sect. 4.3.3): tracking devices inside a magnetic

[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: [14, p. 305].

field, followed by electromagnetic calorimeters, hadron calorimeters and muon chambers (see Fig. 11.8(a)). Since $M_W \sim 80\,\text{GeV}$, the electromagnetic calorimeter resolution is optimized for $40\,\text{GeV}$ electrons to $\pm 500\,\text{MeV}(1\%)$. 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.

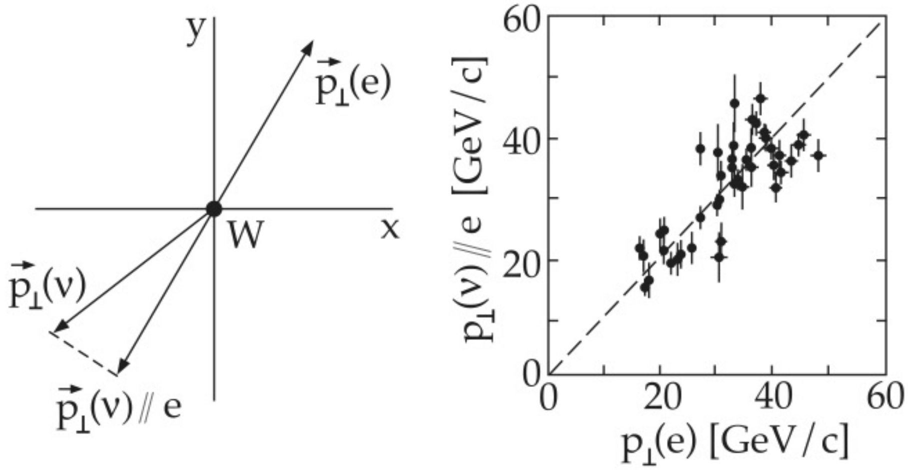


Figure 11.9: *Transverse momenta in a leptonic W decay.* On the LHS one sees a sketch of the electron and neutrino transverse momenta. $\vec{p}_\perp \| e$ is the component of the neutrino transverse momentum parallel to $\vec{p}_\perp(e)$. The correlation between these momenta is shown in the RHS Subfig. Source: [14, 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 ($dN/d\cos\theta$ = 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

$$\frac{dp_\perp}{d\cos\theta} = \frac{M_W}{2}\frac{\cos\theta}{\sin\theta} = \frac{M_W}{2}\frac{\sqrt{1-\sin^2\theta}}{\sin\theta} = \left(\frac{M_W}{2}\right)^2\frac{\sqrt{1-\frac{4p_\perp^2}{M_W^2}}}{p_\perp}.$$

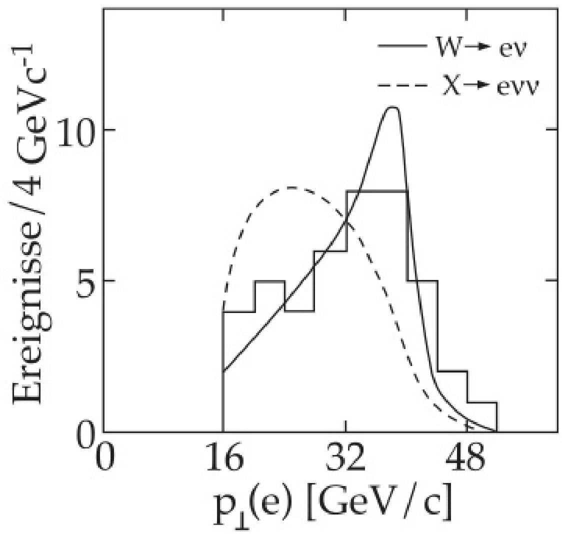


Figure 11.10: *Momentum distribution of the electron perpendicular to the beam (43 events).* The histogram shows the data while the continuous and dashed lines show the Monte Carlo expectation for a two-body decay and three-body decay scenarios, respectively. Source: [14, p. 306].

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

$$\frac{d\Gamma}{d\Omega} = \frac{1}{64\pi^2 M_W}|\mathcal{M}_{fi}|^2.$$

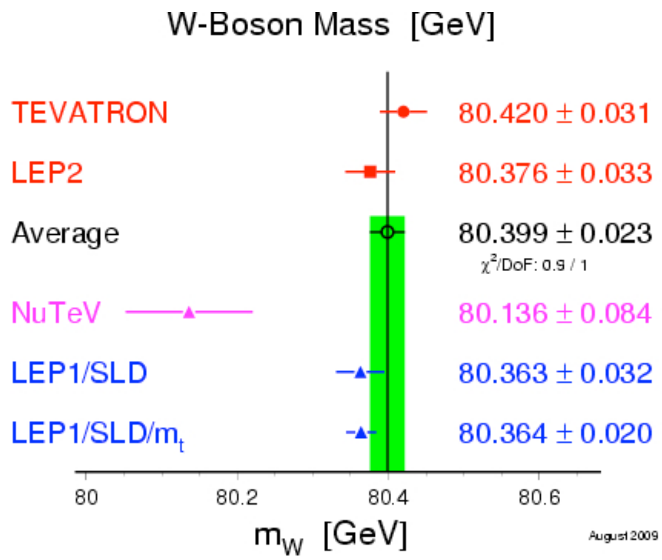


Figure 11.11: *Summary of the current W boson mass measurements.* Source: [17].

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 ($e,\ \mu,\ \tau$) 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

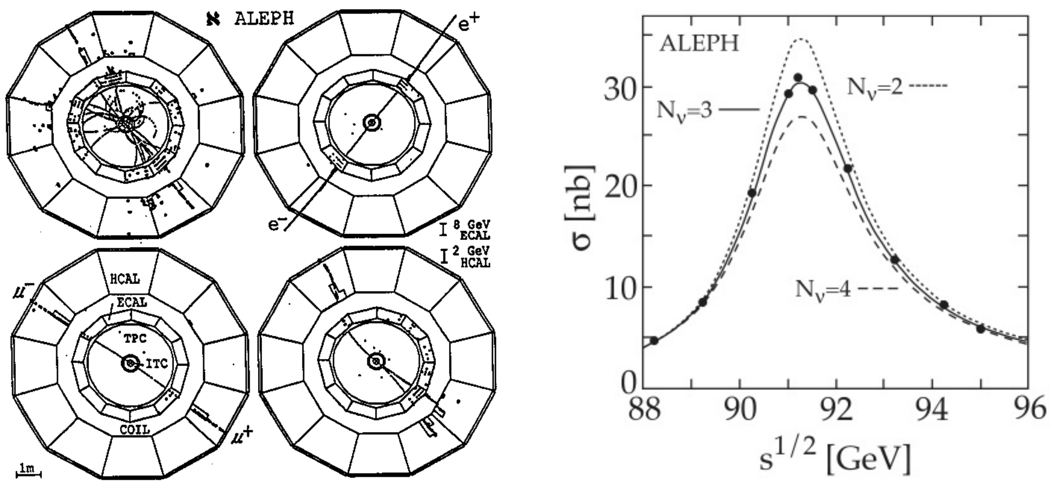
$$\sum_{m_f < M_Z/2}^{\text{fermions}} \left([c_V^f]^2 + [c_A^f]^2\right)$$

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: [15, 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: [14, 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

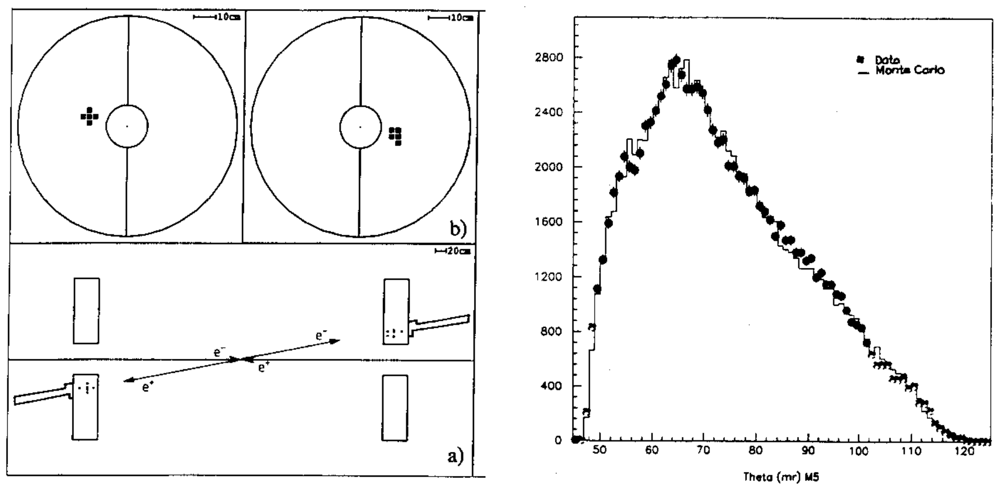


Figure 11.13: *Luminosity measurement in ALEPH using the Bhabha scattering.* On the left a small angle electron-positron scattering event is shown. (a) shows a cut including the beam direction and (b) is a view along the beam of the two luminosity calorimeters. A comparison of measured and simulated polar angle of the scattered electron is shown on the right. Source: [15, 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.

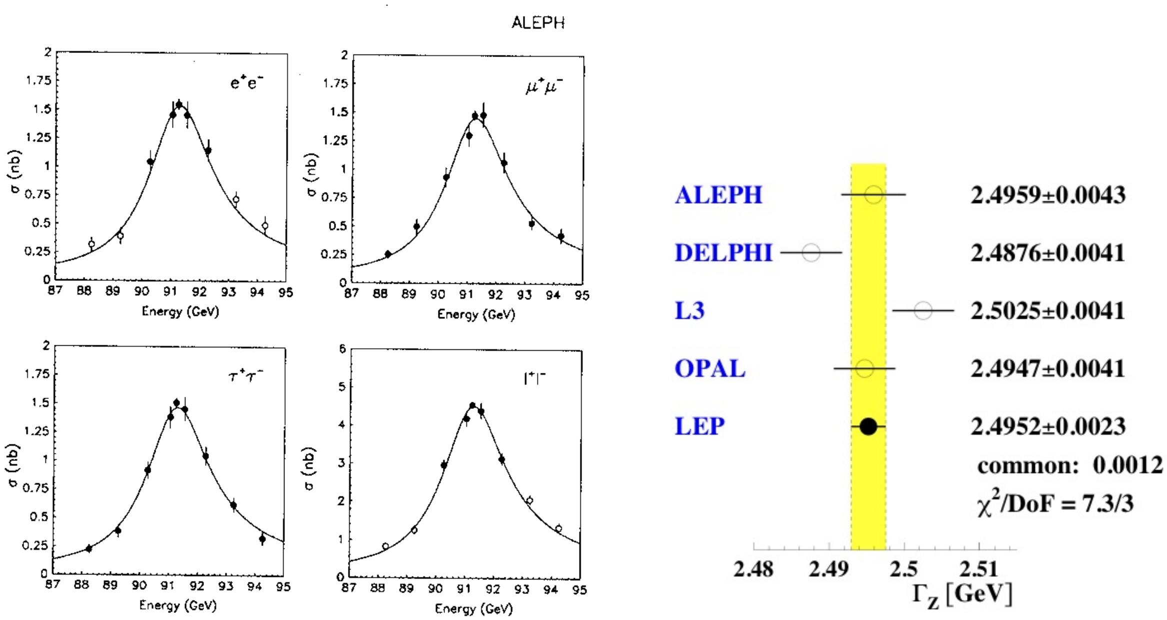
Selecting leptonic events, one can perform the same measurement as the one shown for the hadronic case (see Fig. 11.14(a); note that the cross sections are considerably smaller). This delivers the partial widths $\Gamma(l\bar{l})$ and thus allows for a test of lepton universality. Remembering our discussion 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: [15, 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

$$F_1(s) = Q_f^2 - 2v_e v_f Q_f \mathrm{Re}\chi + (v_e^2 + a_e^2)(v_f^2 + a_f^2)|\chi|^2$$
$$F_2(s) = -2a_e a_f Q_f \mathrm{Re}\chi + 4v_e a_e v_f a_f |\chi|^2$$

[10] `http://www-theorie.physik.unizh.ch/~pfmonni/PPPII_FS10/sheet8.pdf`

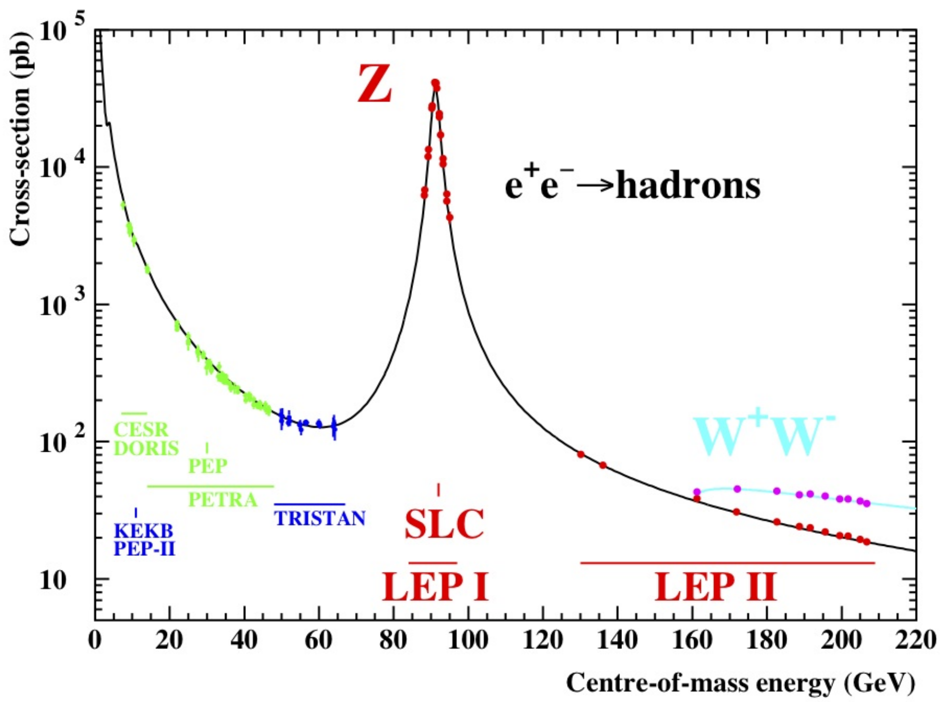


Figure 11.15: *Summary of the* $e^+e^- \to$ *hadrons cross section measurements as a function of the center of mass energy* $\sqrt{s}$*.*

with

$$\chi = \frac{s}{s - M_z^2 + iM_Z\Gamma_T(Z)}$$

the Breit-Wigner term (compare Eq. (11.68)) and

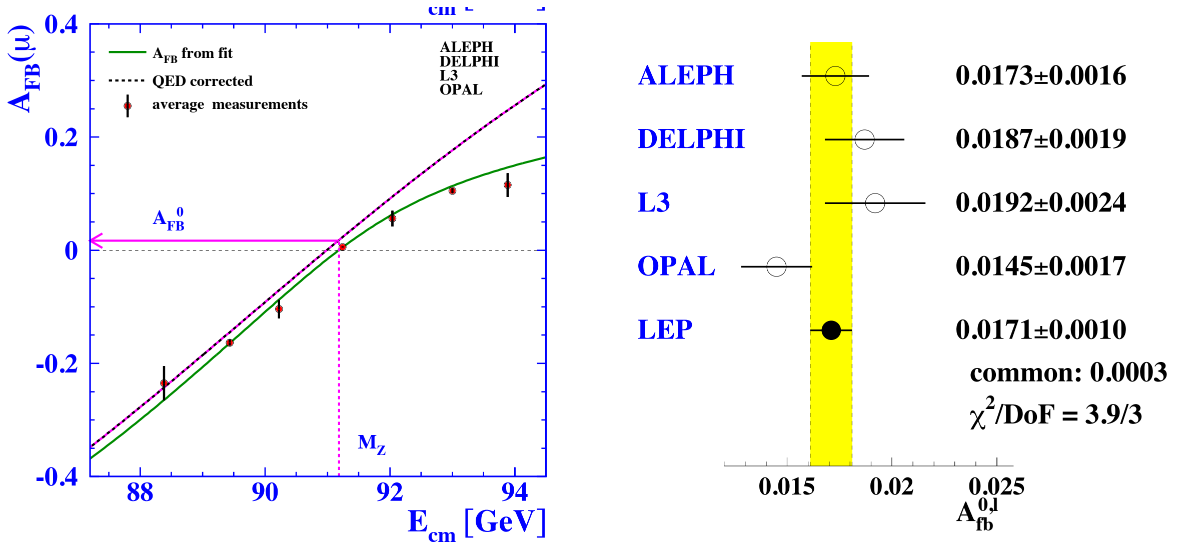
$$v_f \equiv \frac{c_V^f}{2\sin\theta_w\cos\theta_w}$$
$$a_f \equiv \frac{c_A^f}{2\sin\theta_w\cos\theta_w}.$$

To get a quantitative estimate of the forward-backward asymmetry, we define the following quantity

$$A_{\mathrm{FB}} = \frac{\mathcal{I}(0,1) - \mathcal{I}(-1,0)}{\mathcal{I}(0,1) + \mathcal{I}(-1,0)} \tag{11.70}$$

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 $\sqrt{s} = M_Z$.

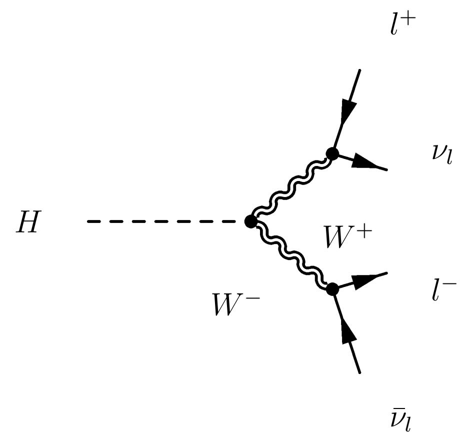


Figure 11.17: *Feynman diagram for the decay of a heavy Higgs into a W^+W^- pair.*

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

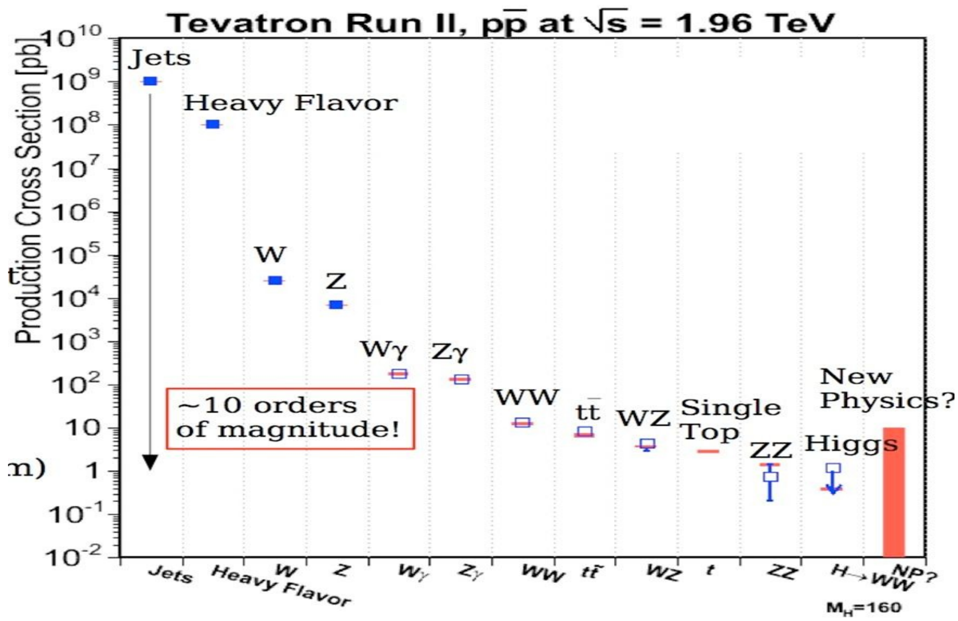


Figure 11.18: *Various production cross sections at Tevatron.* Note that the scale is logarithmic.

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: $\Delta\phi < 2\,\text{rad}$. 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 $\mathcal{L}_{\text{int}} \sim 5\,\text{fb}^{-1}$ 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/`

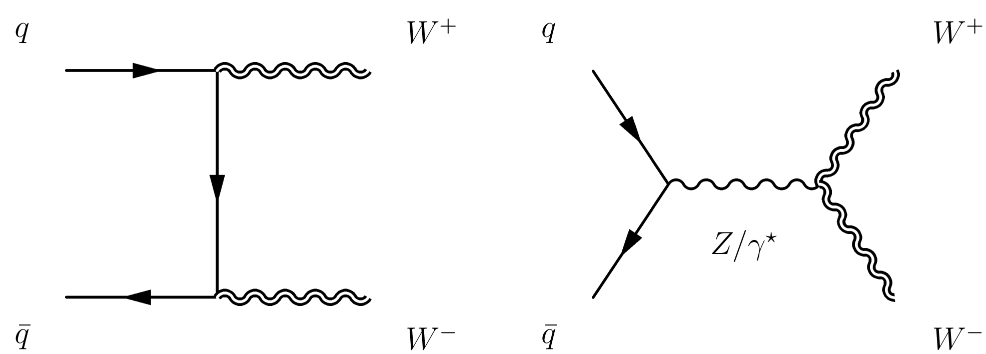


Figure 11.19: *Examples of* W^+W^- *production diagrams at hadron colliders not involving Higgs boson exchange.*

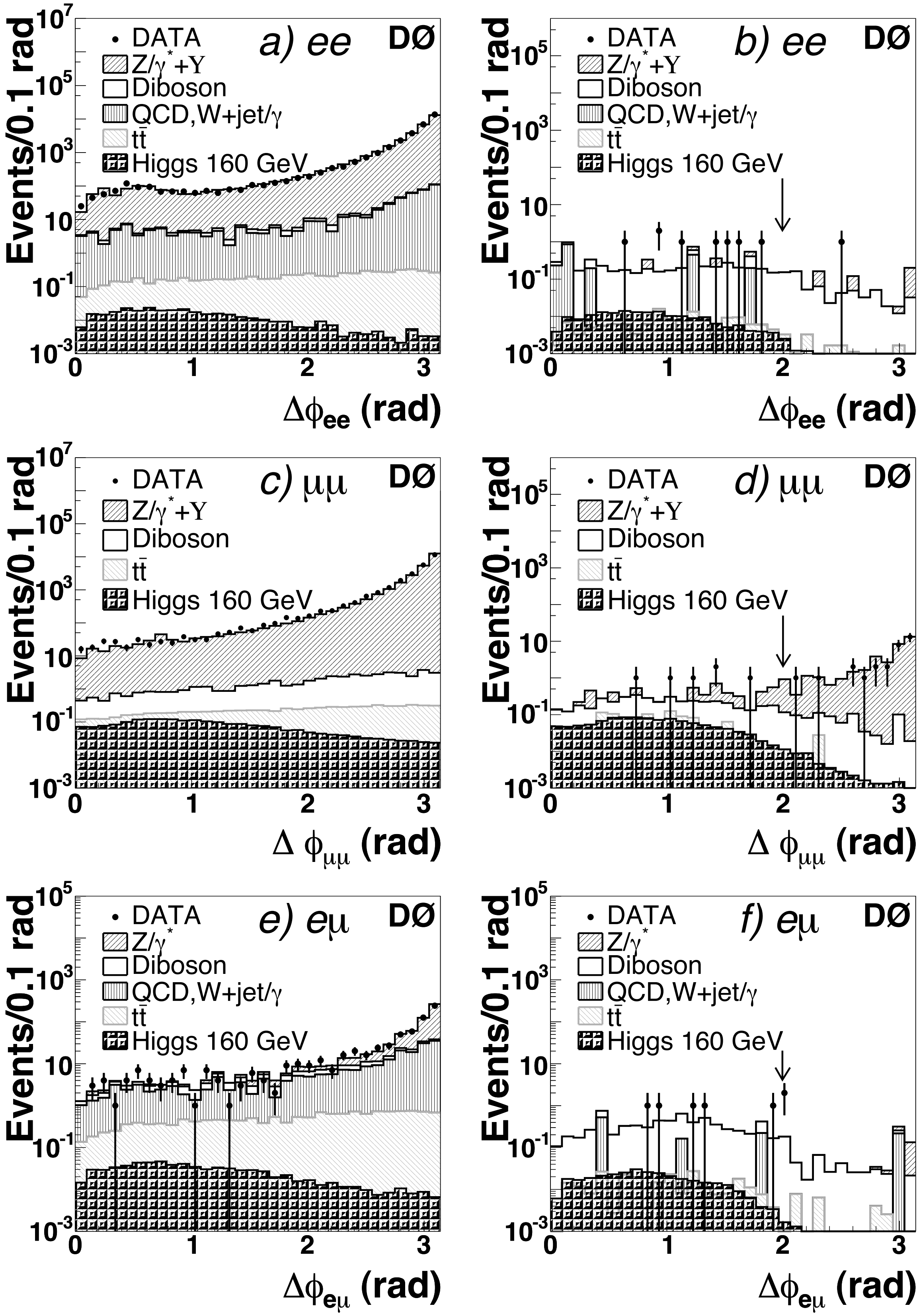


Figure 11.20: *Distribution of the opening angle $\Delta\phi_{ll'}$ after applying the initial transverse momentum cuts (a), (c), (e) and after all cuts, except for the $\Delta\phi$ cut (b), (d), (f).* Source: [18].

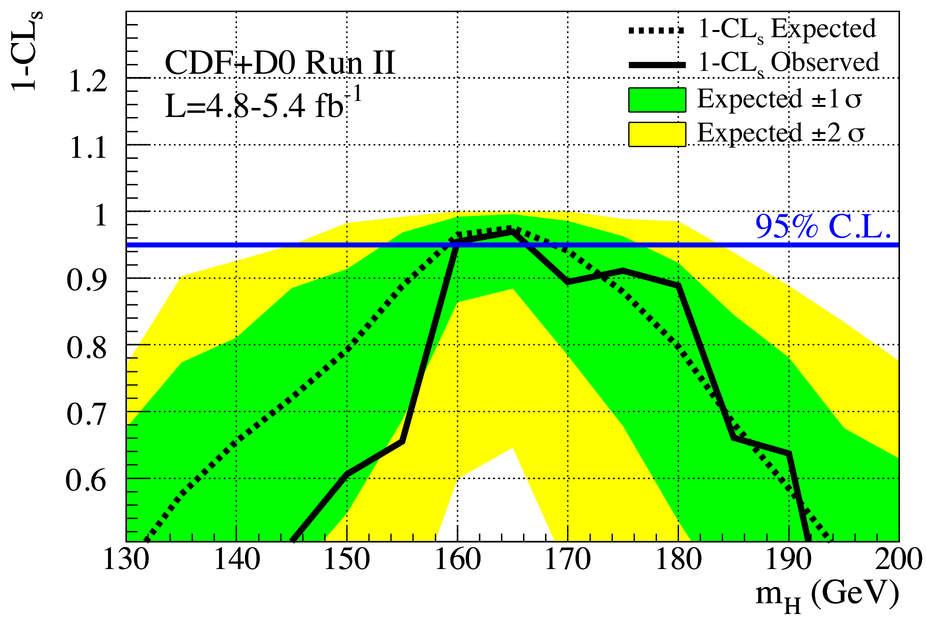


Figure 11.21: *Higgs mass range exclusion with combined Tevatron results.* Source: [19].

Chapter 12

Flavor physics

Quarks and leptons can be ordered in flavour doublets, each column being called a family,

$$\begin{array}{lcccc} \text{Quarks:} & \begin{pmatrix} u \\ d \end{pmatrix} & \begin{pmatrix} c \\ s \end{pmatrix} & \begin{pmatrix} t \\ b \end{pmatrix} & \begin{array}{c} Q = \frac{2}{3} \\ Q = -\frac{1}{3} \end{array} , \\ \text{Leptons:} & \begin{pmatrix} \nu_e \\ e \end{pmatrix} & \begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix} & \begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix} & \begin{array}{c} Q = 0 \\ Q = -1 \end{array} . \end{array}$$

These arrangements correspond to an approximate flavor $SU(6)$ symmetry. The isospin $SU(2)$ of p, n (Sect. 7, p. 125) or the flavor $SU(3)$ symmetry of u, d, s (Sect. 7.3, p. 131) 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,

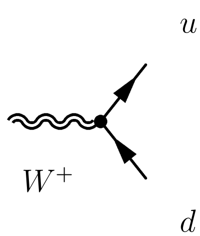
$$\begin{pmatrix} u \\ d' \end{pmatrix} \quad \begin{pmatrix} c \\ s' \end{pmatrix},$$

and we assume that the weak eigenstates $|d'\rangle$ and $|s'\rangle$ are linear combinations of the mass eigenstates $|d\rangle$ and $|s\rangle$,

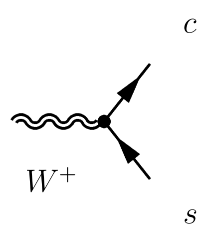
$$\begin{aligned} |d'\rangle &= \cos\theta_c\,|d\rangle + \sin\theta_c\,|s\rangle \\ |s'\rangle &= -\sin\theta_c\,|d\rangle + \cos\theta_c\,|s\rangle\,, \end{aligned} \tag{12.1}$$

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,

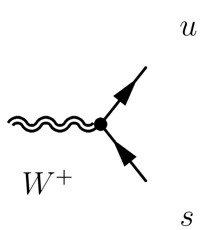


$\propto \cos\theta_c$

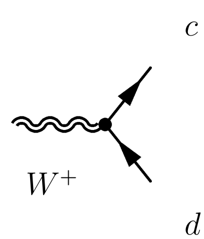


$\propto \cos\theta_c$,

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 $V \in U(3)$.

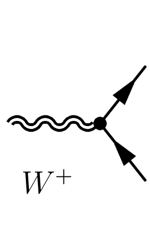
Recall that for a matrix $V \in U(N)$:

- 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.

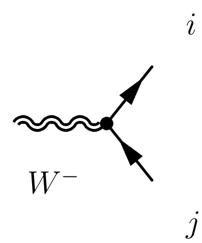
Thus V contains $N^2 - (2N-1) = (N-1)^2$ independent real parameters. On the other hand, a matrix $O \in O(N)$ is determined by $\frac{1}{2}N(N-1)$ independent real parameters (Euler angles).

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 [20]

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,

$$
\begin{aligned}
|\nu_\alpha, t\rangle &= \mathrm{e}^{-i|\vec{p}|t} \left(U \begin{bmatrix} \mathrm{e}^{-i\frac{m_1^2 t}{2E}} & 0 \\ 0 & \mathrm{e}^{-i\frac{m_2^2 t}{2E}} \end{bmatrix} U^\dagger \right)_{\alpha\beta} |\nu_\beta, t\rangle \\
&\approx \mathrm{e}^{-i|\vec{p}|t} \left(U \begin{bmatrix} 1 - \frac{i m_1^2 t}{2E} & 0 \\ 0 & 1 - \frac{i m_2^2 t}{2E} \end{bmatrix} U^\dagger \right)_{\alpha\beta} |\nu_\beta, t\rangle \, ,
\end{aligned}
$$

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 $L = t$),

$$
\frac{\Delta m^2 L}{4E} \approx 1.27\frac{\Delta m^2[\mathrm{eV}^2]L[\mathrm{m}]}{E[\mathrm{MeV}]}. \tag{12.17}
$$

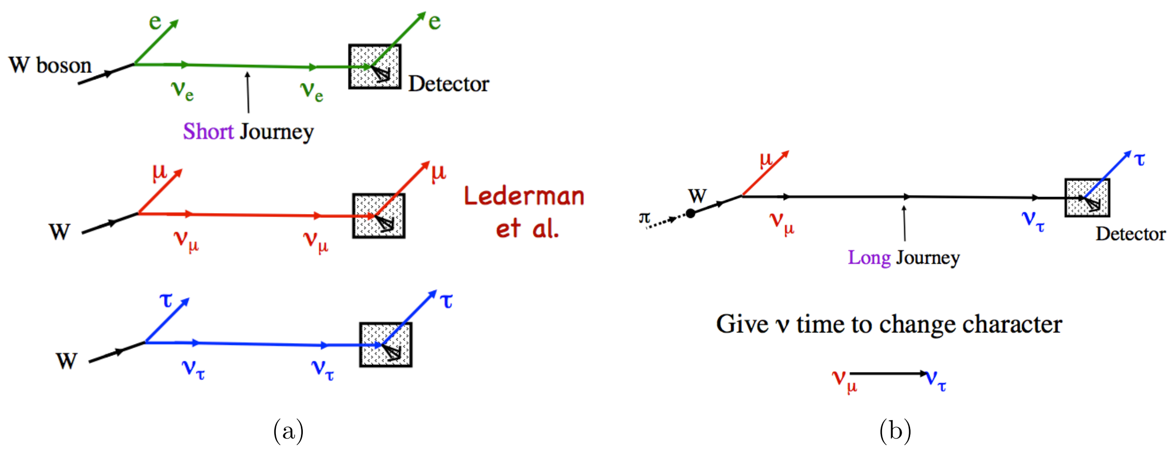


Figure 12.1: *Neutrino production and detection.* During a sufficiently long journey, the neutrinos may change character (b). Source: B. Kayser.

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 [24]). 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 [21, 22, 23].
[3] http://operaweb.lngs.infn.it/IMG/pdf/OPERA_press_release_May_2010_english-5.pdf

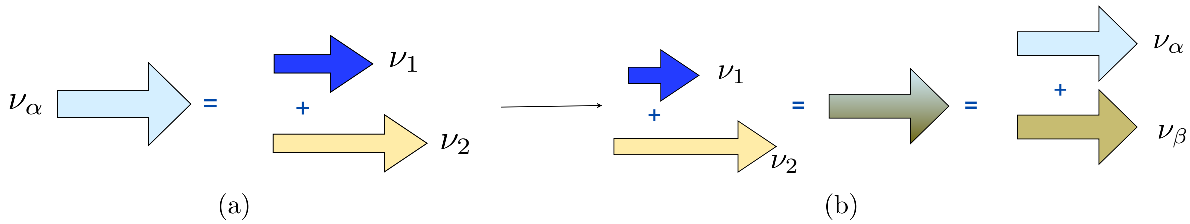


Figure 12.2: *Neutrino mixing in the two-neutrino case.*

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.

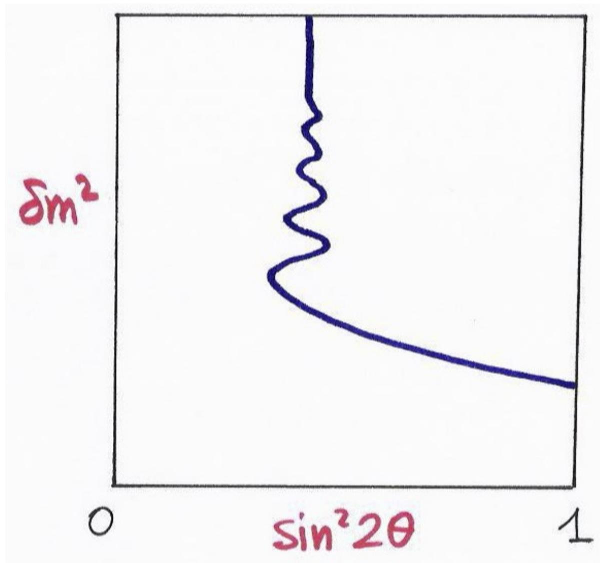


Figure 12.3: *Oscillation experiment sensitivity.* Source: [25].

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

$$\begin{aligned} \delta m^2 &= m_2^2 - m_1^2 > 0 \\ |\Delta m^2| &= |m_3^2 - m_{1,2}^2| \end{aligned}$$

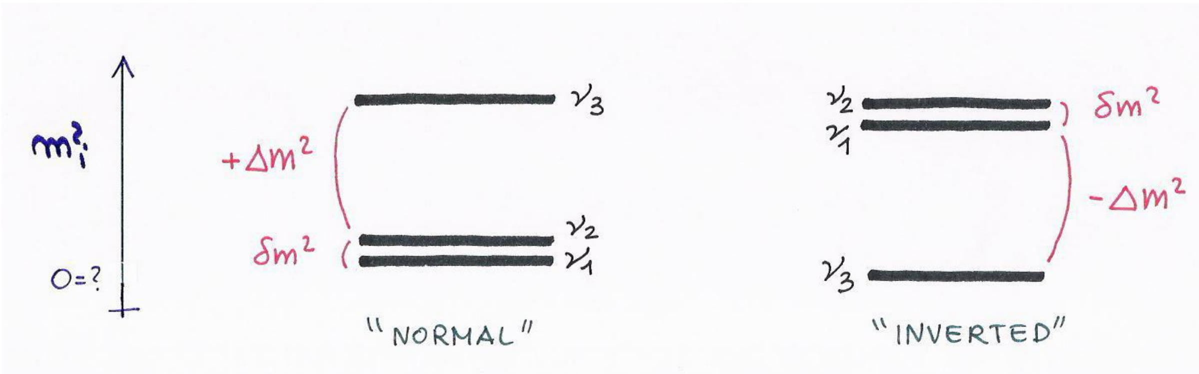


Figure 12.4: *Normal and inverted mass hierarchies for the three-neutrino case.* Source: [25].

(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 [26]:

$$P(\nu_e \to \nu_e) = 1 - \sin^2 2\theta_{13} \sin^2 \frac{\Delta m^2 L}{4E} \tag{12.21}$$

$$P(\nu_e \to \nu_\mu) = \sin^2 2\theta_{13} \sin^2 \theta_{23} \sin^2 \frac{\Delta m^2 L}{4E} \tag{12.22}$$

$$P(\nu_\mu \to \nu_\tau) = \sin^2 2\theta_{13} \sin^2 \theta_{23} \sin^2 \frac{\Delta m^2 L}{4E} \tag{12.23}$$

$$P(\nu_e \to \nu_\tau) = \sin^2 2\theta_{13} \cos^2 \theta_{23} \sin^2 \frac{\Delta m^2 L}{4E} \tag{12.24}$$

$$P(\nu_\mu \to \nu_\tau) = \cos^4 \theta_{13} \sin^2 2\theta_{23} \sin^2 \frac{\Delta m^2 L}{4E}. \tag{12.25}$$

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

$$P(\nu_\alpha \to \nu_\beta) = \delta_{\alpha\beta} - 4\sum_{i<j} \mathrm{Re} J_{\alpha\beta}^{ij} \sin^2\left(\frac{\Delta m_{ij}^2 L}{4E}\right) - 2\sum_{i<j} \mathrm{Im} J_{\alpha\beta}^{ij} \sin\left(\frac{\Delta m_{ij}^2 L}{2E}\right) \tag{12.26}$$

where $\Delta m_{ij} = m_i^2 - m_j^2$ and $J_{\alpha\beta}^{ij} = U_{\alpha i} U_{\beta i}^* U_{\alpha j}^* U_{\beta j}$. 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.

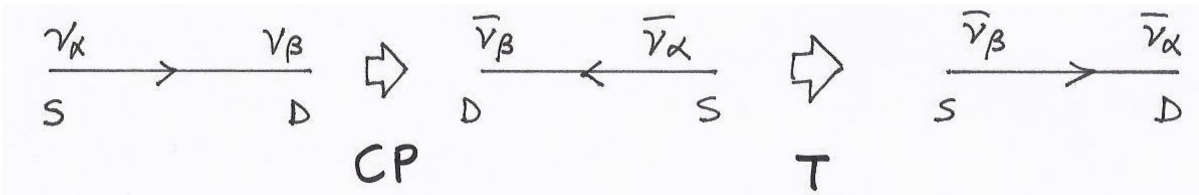


Figure 12.5: *Action of CP and T transformations on the $\nu_\alpha \to \nu_\beta$ process from source (S) to detector (D).* Source: [25].

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 [27, 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. [27].

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.

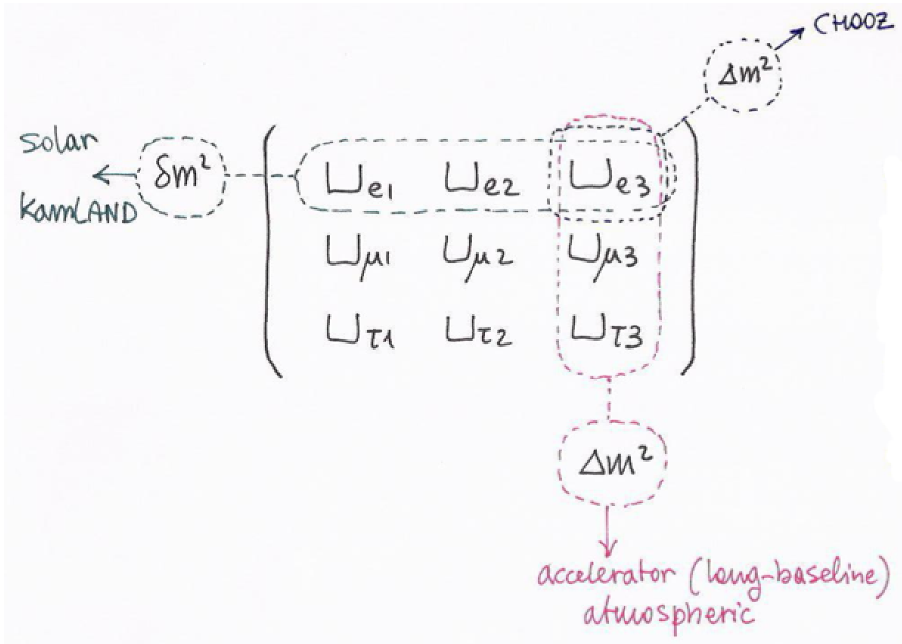


Figure 12.6: *Summary of experimental sensitivities to the neutrino mixing matrix.* Source: [25].

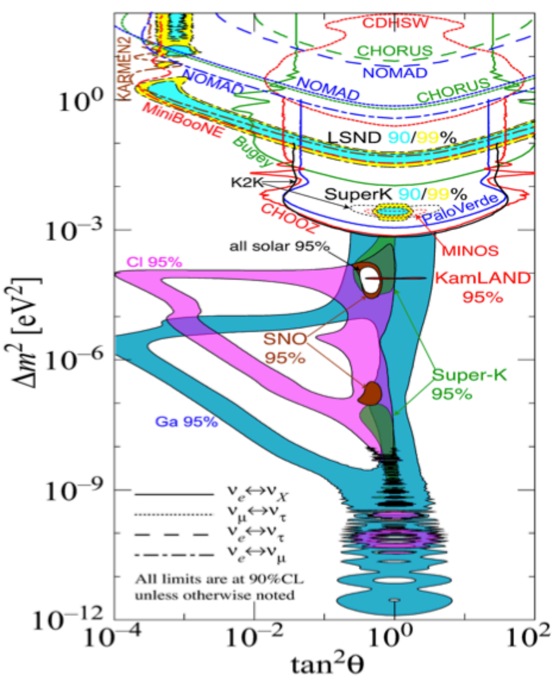


Figure 12.7: *Summary of neutrino oscillation experiments.* Source: Particle Data Group 2009.

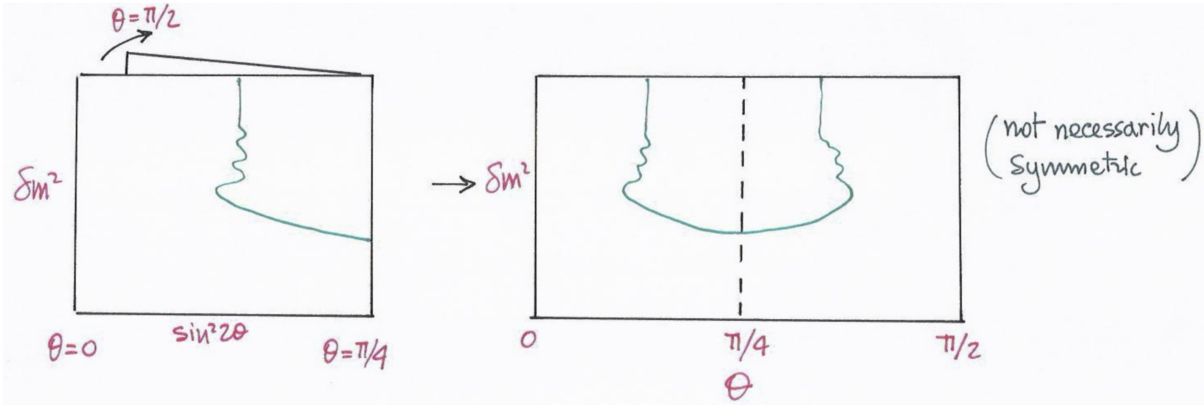


Figure 12.8: *Oscillation experiment sensitivity as a function of θ, rather than* $\sin^2 2\theta$. Source: [25].

Their three-neutrino interpretation is summarized in Fig. 12.9; the numerical values (with one digit accuracy) read:

$$\begin{aligned}
\delta m^2 &\sim 8 \cdot 10^{-5}\,\mathrm{eV}^2 \\
\Delta m^2 &\sim 3 \cdot 10^{-3}\,\mathrm{eV}^2 \\
m_\nu &< \mathcal{O}(1)\,\mathrm{eV} \\
\mathrm{sign}(\Delta m^2) &= ? \\
\sin^2\theta_{12} &\sim 0.3 \\
\sin^2\theta_{23} &\sim 0.5 \\
\sin^2\theta_{13} &\sim \text{few}\,\% \\
\delta(CP) &= ?.
\end{aligned}$$

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 $\sim 6 \cdot 10^{20}\,\mathrm{s}^{-1}$, the

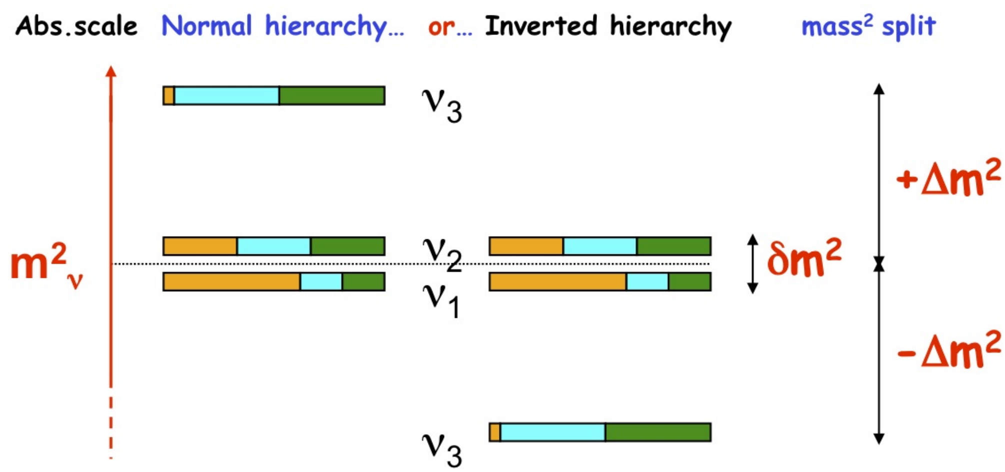


Figure 12.9: *Summary of the current knowledge on neutrino oscillations.* Source: [25].

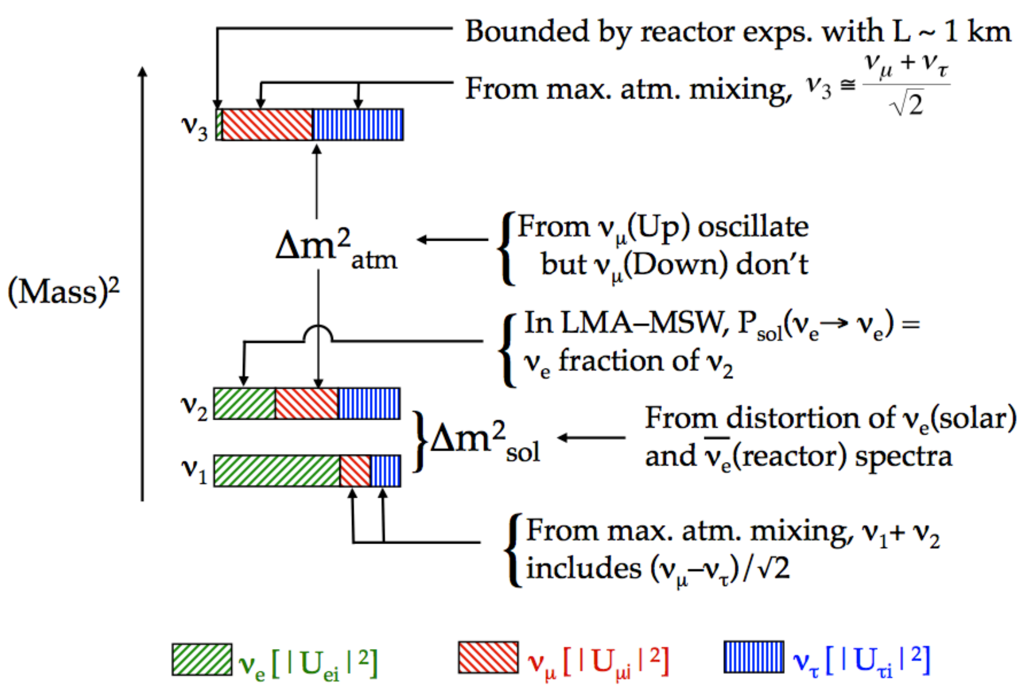


Figure 12.10: *Origin of the current knowledge on neutrino oscillations.* Source: B. Kayser.

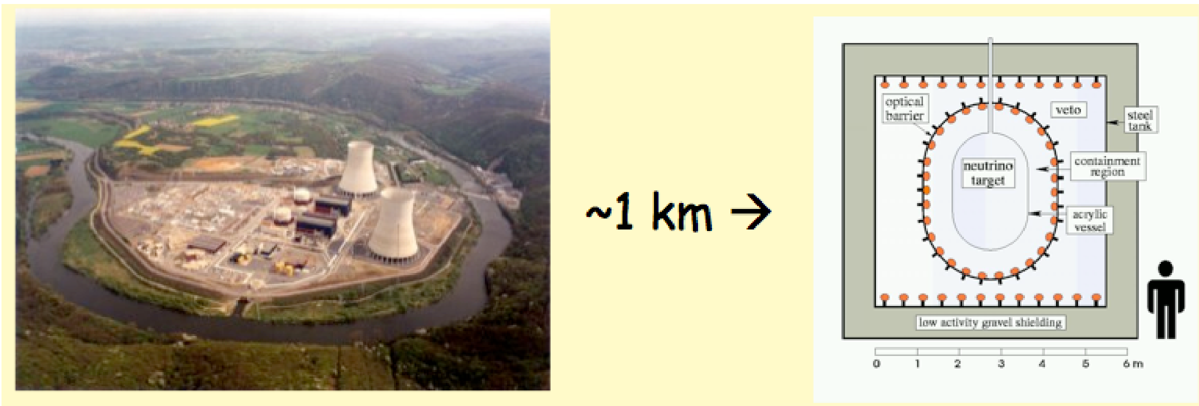


Figure 12.11: *Setup of short-baseline reactor experiments.* Source: [25].

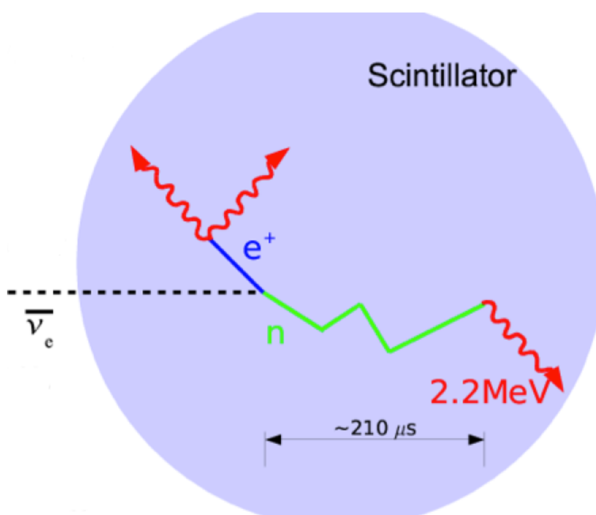
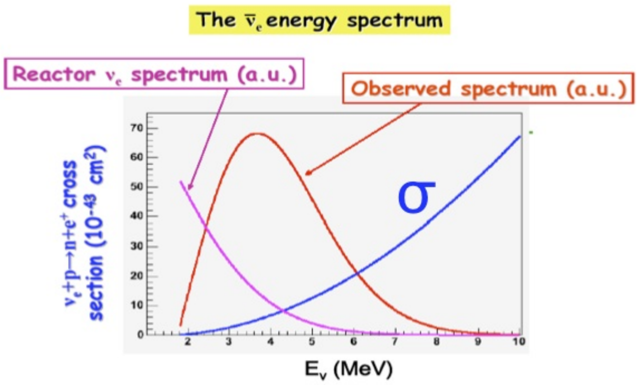


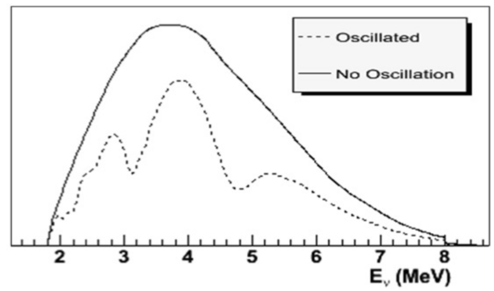
Figure 12.12: *Neutrino detection via inverse beta decay.*

energy being of the order of MeV. Detection is accomplished by inverse β-decay: $\bar{\nu}_e + p \rightarrow e^+ + n$; $n + p \rightarrow d + \gamma$, 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.

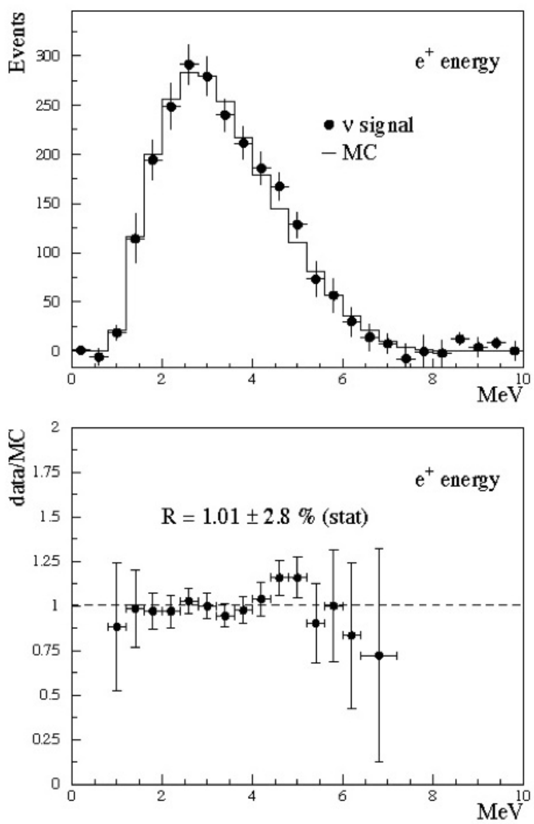
Atmospheric neutrinos: the Super-Kamiokande breakthrough. Figure 12.14(a) shows the zenith angle dependence of the number of events in the 50 kt Super-Kamiokande detector. One observes that there is a deficit in μ-like events in the up-going direction, whereas the electron-like events follow more or 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 $\nu_\mu \rightarrow \nu_e$ is non-existing or subdominant (in agreement with CHOOZ) and that the channel $\nu_\mu \rightarrow \nu_\tau$ is dominant. Recall that the one-mass scale



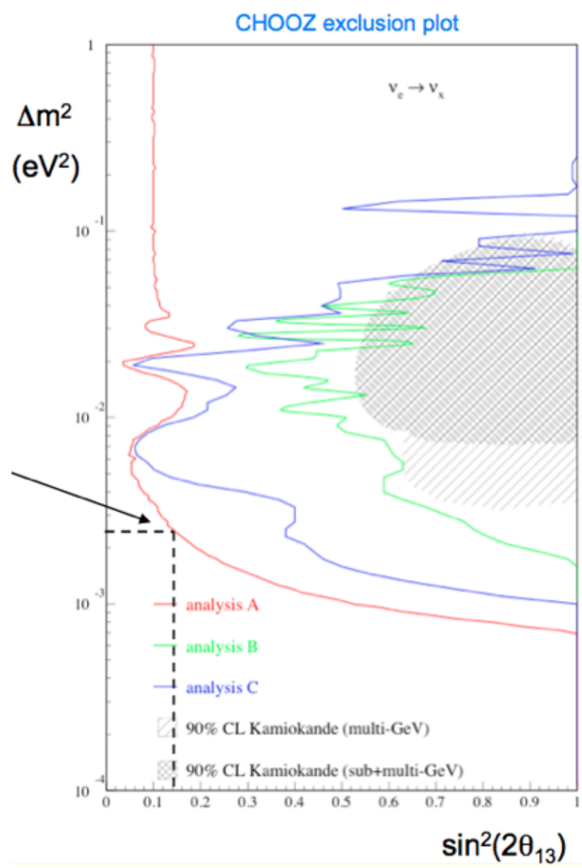
(a)



(b)

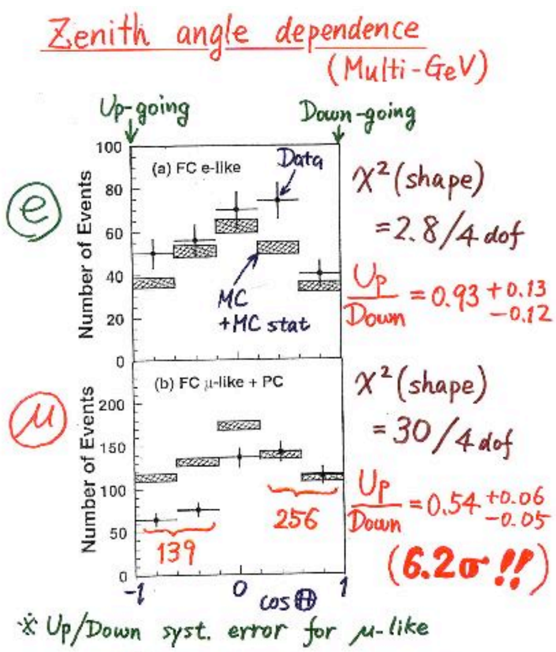


(c)

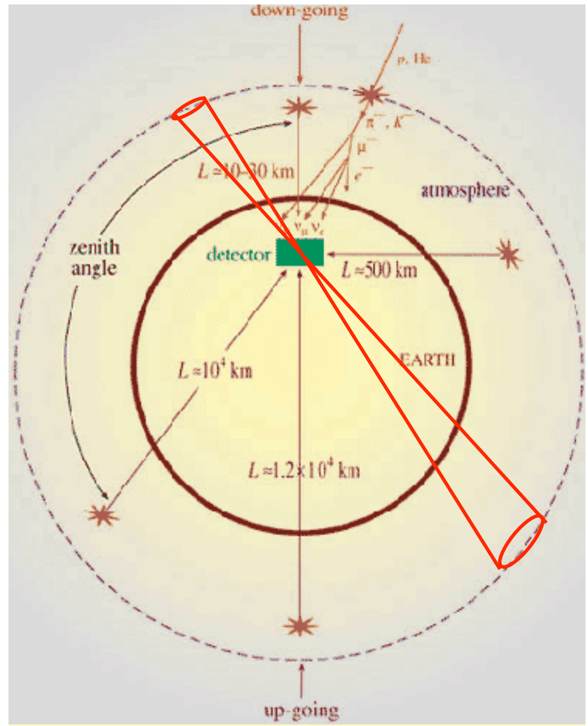


(d)

Figure 12.13: *Results of the short-baseline reactor experiment CHOOZ.*



(a)



(b)

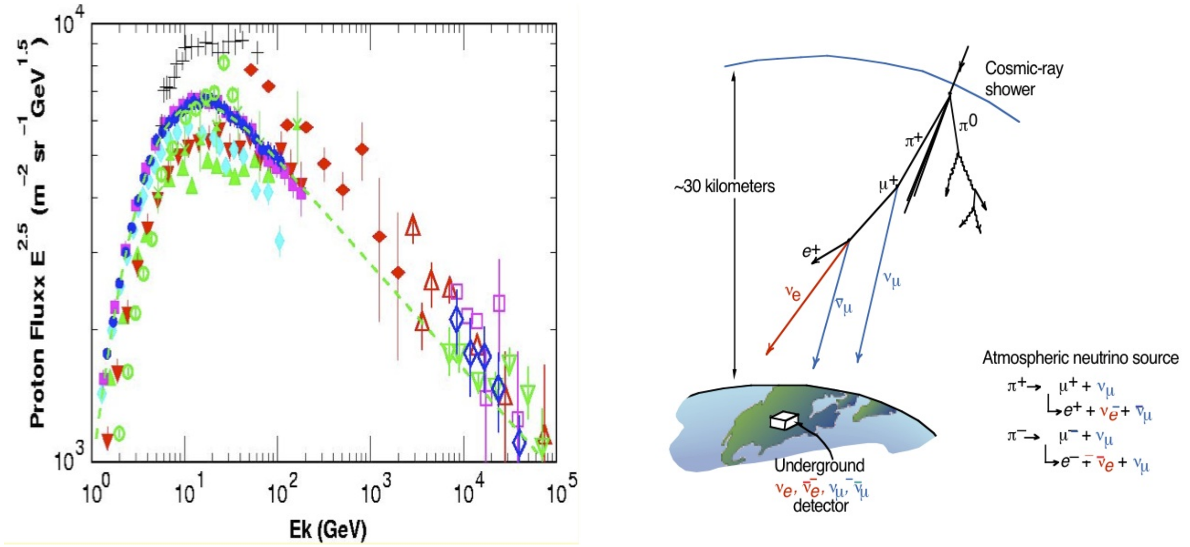
Figure 12.14: *Zenith angle dependence of μ-like events in the Super-Kamiokande experiment.* Source: T. Kajita at Neutrino '98, Takayama.

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).

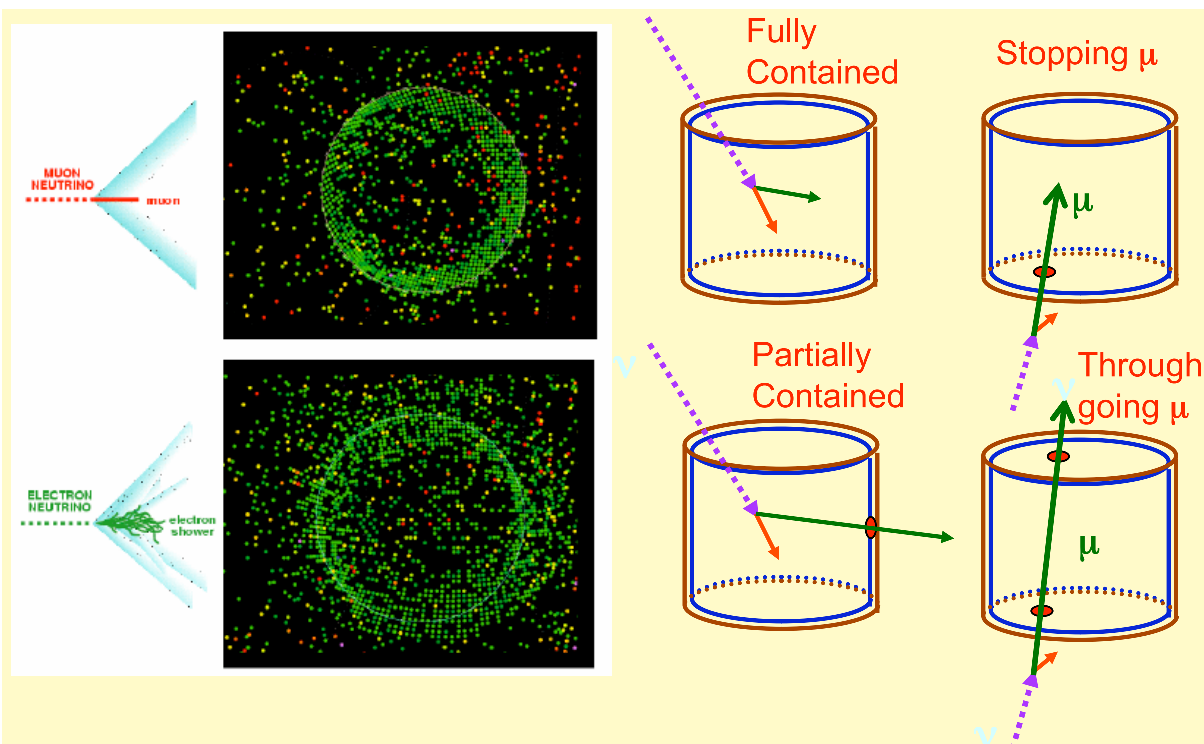


Figure 12.16: *Detection in Super-Kamiokande.* Parent neutrinos are detected via charged-current interactions in the water target.

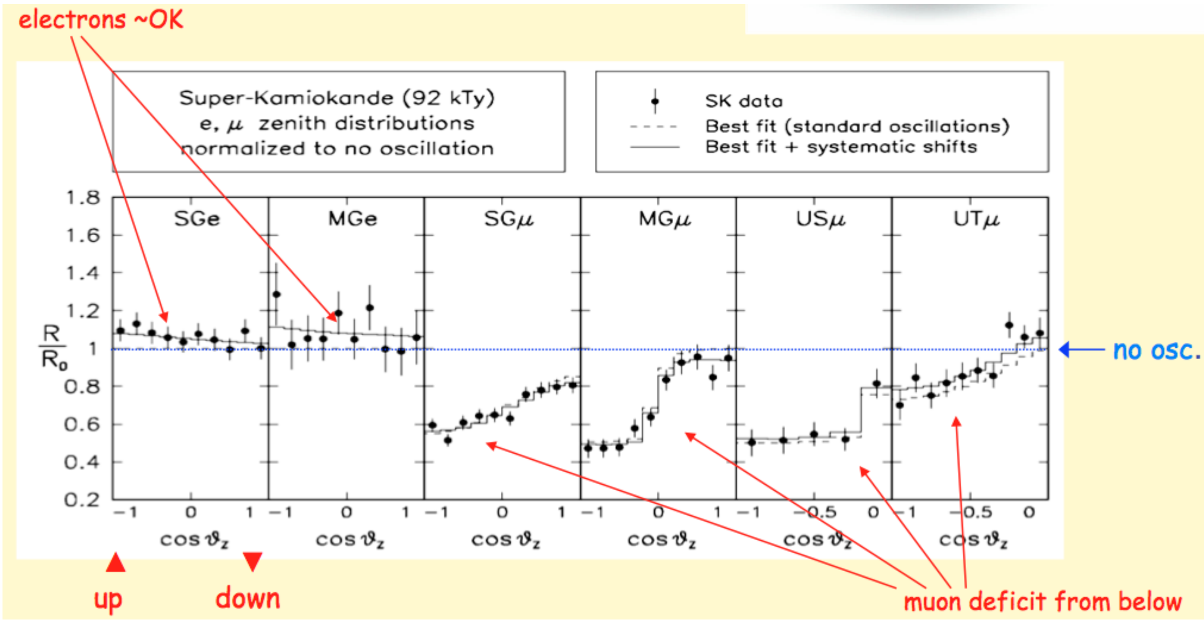
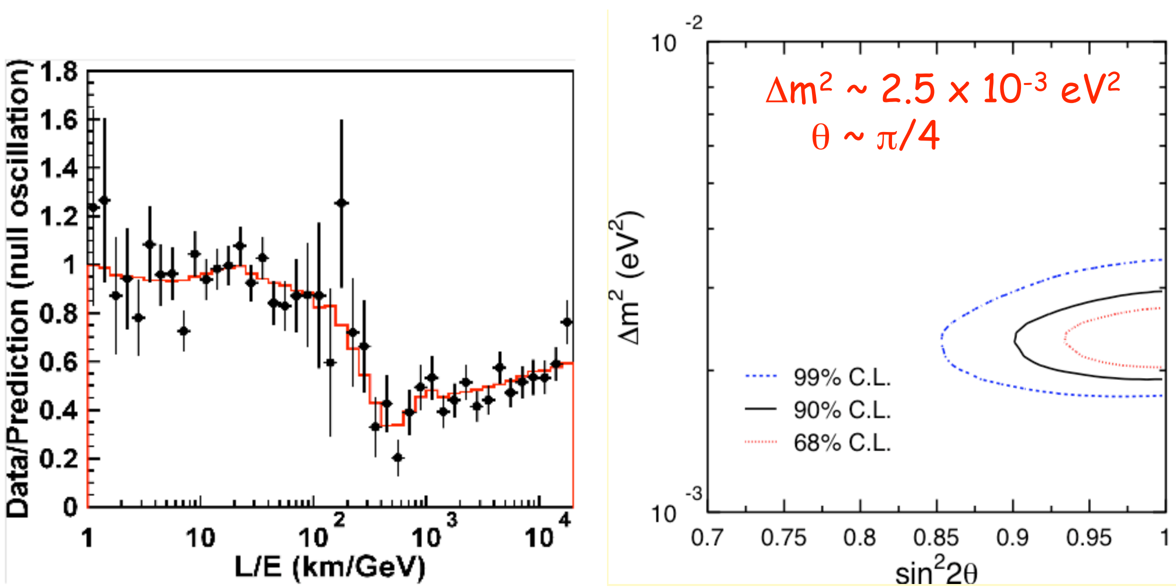


Figure 12.17: *Super-Kamiokande results on atmospheric neutrinos.*



(a) (b)

Figure 12.18: *Super-Kamiokande results on oscillation period (a) and constraints on the parameters Δm^2 and θ.*

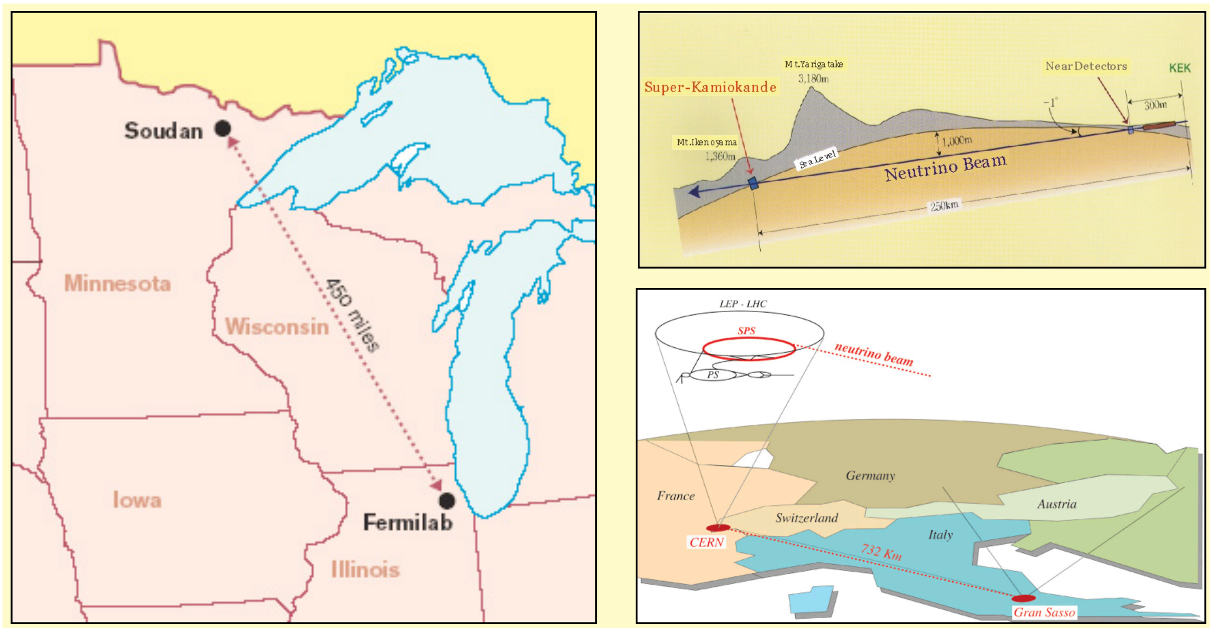


Figure 12.19: *Examples of long-baseline neutrino experiments.* Source: [25].

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).

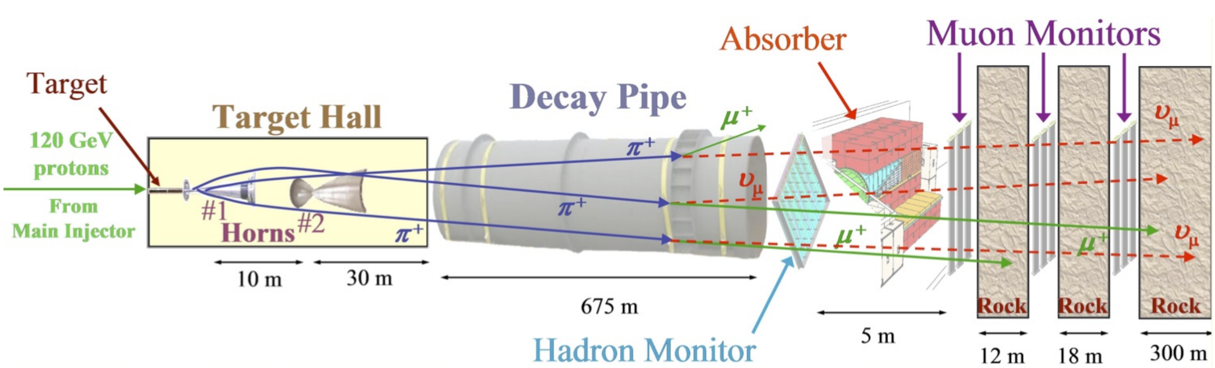


Figure 12.20: *Muon-neutrino beam production at hadron accelerators.*

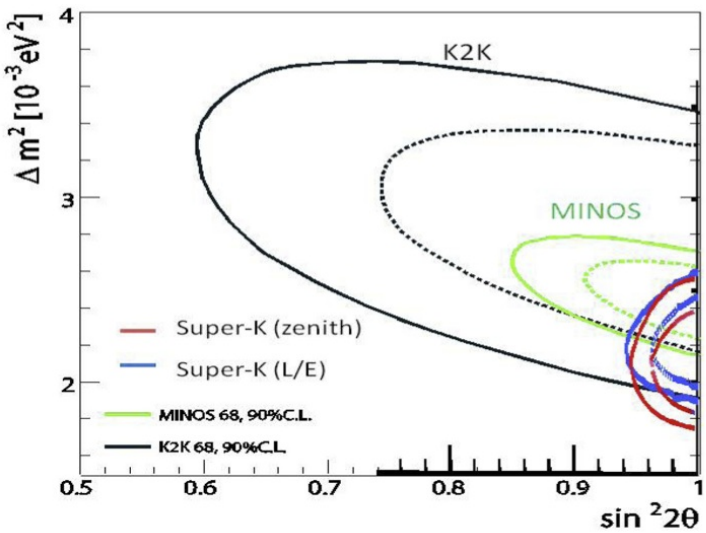


Figure 12.21: *Long-baseline neutrino experiments combination and consistency check with atmospheric results.*

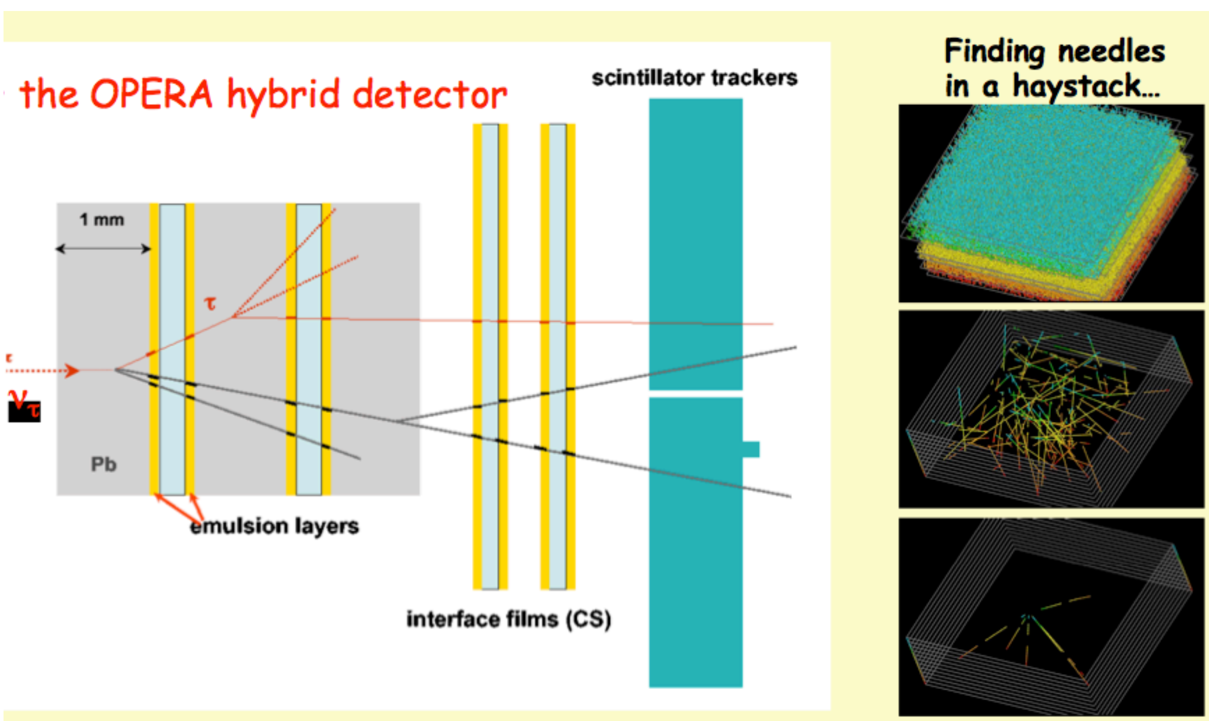


Figure 12.22: *Sketch of the OPERA detector (LHS) and of a reconstructed event (RHS).*

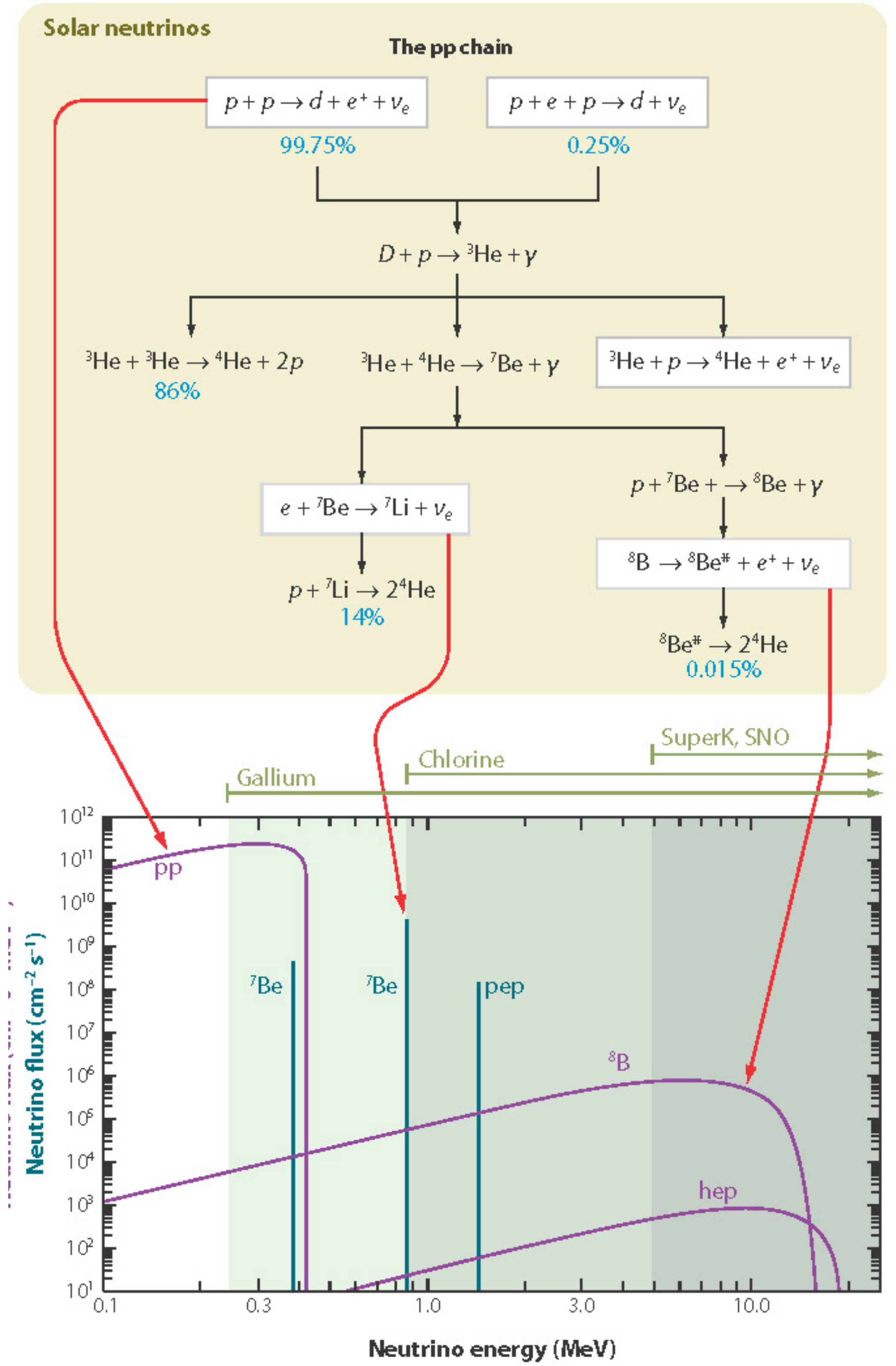


Figure 12.23: *Production of solar neutrinos in the pp cycle.*

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

$$\begin{array}{lcr} {}^{37}\mathrm{Cl} + \nu_e \to {}^{37}\mathrm{Ar} + e^- & \text{(CC)} & \text{Homestake} \\ {}^{71}\mathrm{Ga} + \nu_e \to {}^{71}\mathrm{Ge} + e^- & \text{(CC)} & \text{GALLEX/GNO, SAGE.} \end{array}$$

The second detection possibility for solar neutrinos is elastic scattering:

$$\nu_x + e^- \to \nu_x + e^- \qquad \text{(NC,CC)} \qquad \text{SK, SNO, Borexino}$$

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:

$$\begin{array}{lcr} \nu_e + \mathrm{d} \to p + p + e^- & \text{(CC)} & \text{SNO} \\ \nu_x + \mathrm{d} \to p + n + \nu_x & \text{(NC)} & \text{(Sudbury Neutrino Observatory).} \end{array}$$

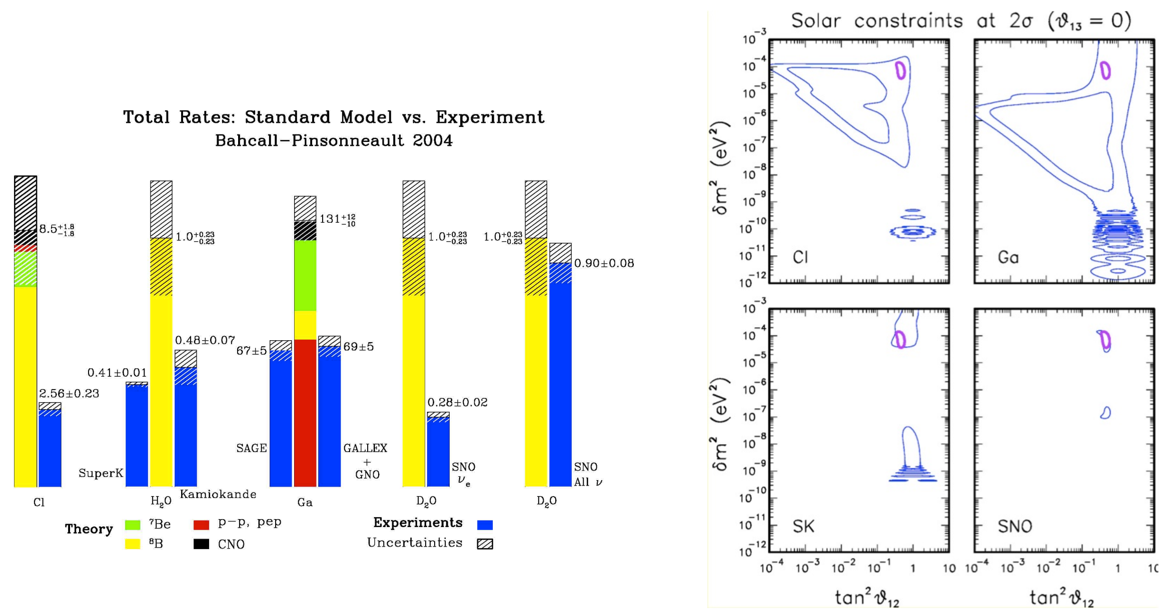
All CC-sensitive results on solar neutrinos indicated a ν_e deficit, when compared to solar model expectations (see Fig. 12.24(a)). Interpreting the results in terms of neutrino oscillations yielded solar constraints on δm^2 and θ_{12} (see Fig. 12.24(b)). A crucial role in this development was played 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 CC/NC $\sim 1/3 < 1$ and the solar model turned out to be adequate. Note also that since 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).*

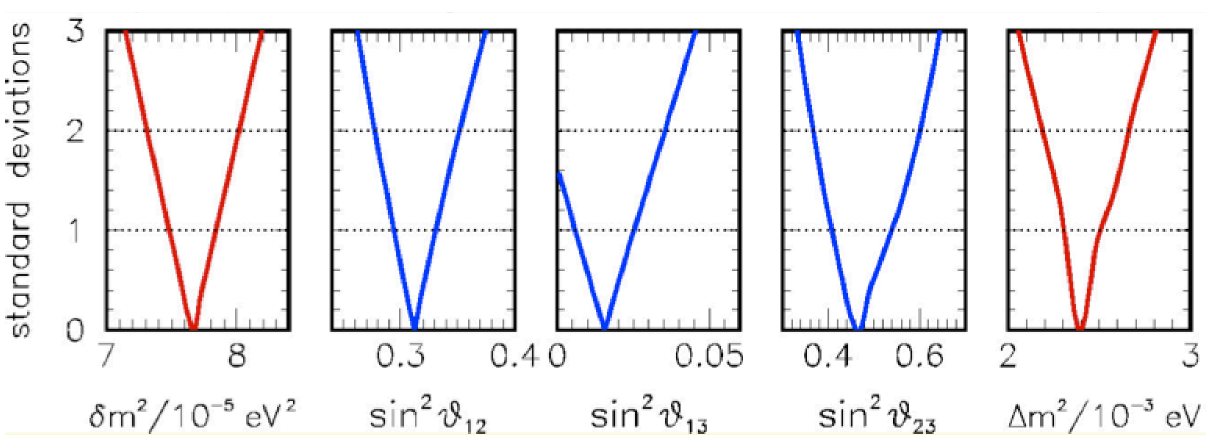


Figure 12.25: *Synopsis of neutrino mass splitting and mixing parameters.*

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 ($E_{\rm rec} = 1.72\,\text{eV}$): $E_0 = Q - E_{\rm rec}$ (see [28, 29] 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 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) [28, 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: $E_\perp/B = \text{const}$. 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 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/`

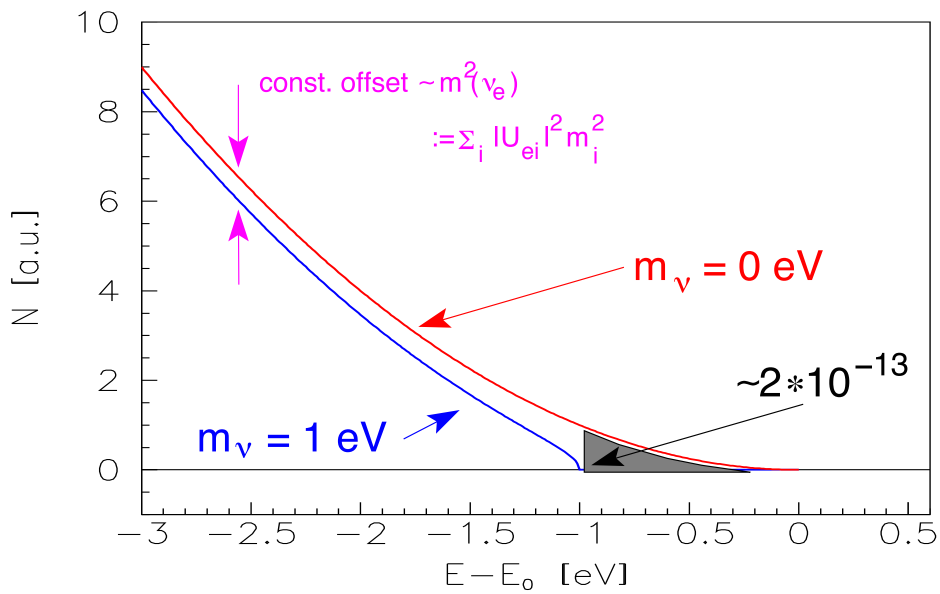


Figure 12.26: *Close-up of the high-energy end of the beta decay spectrum.* In the case of tritium the shaded area corresponds to a fraction of about $2 \cdot 10^{-13}$ events. Source: [28, p. 12].

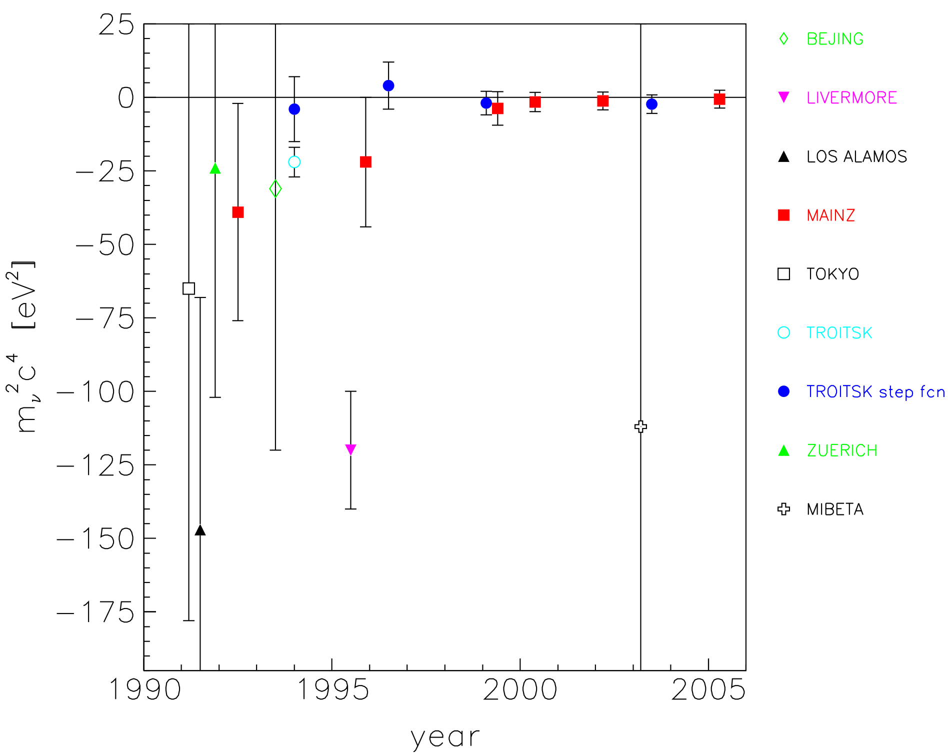


Figure 12.27: *Recent results of tritium beta decay experiments on the effective electron neutrino mass.* Source: [28, p. 15].

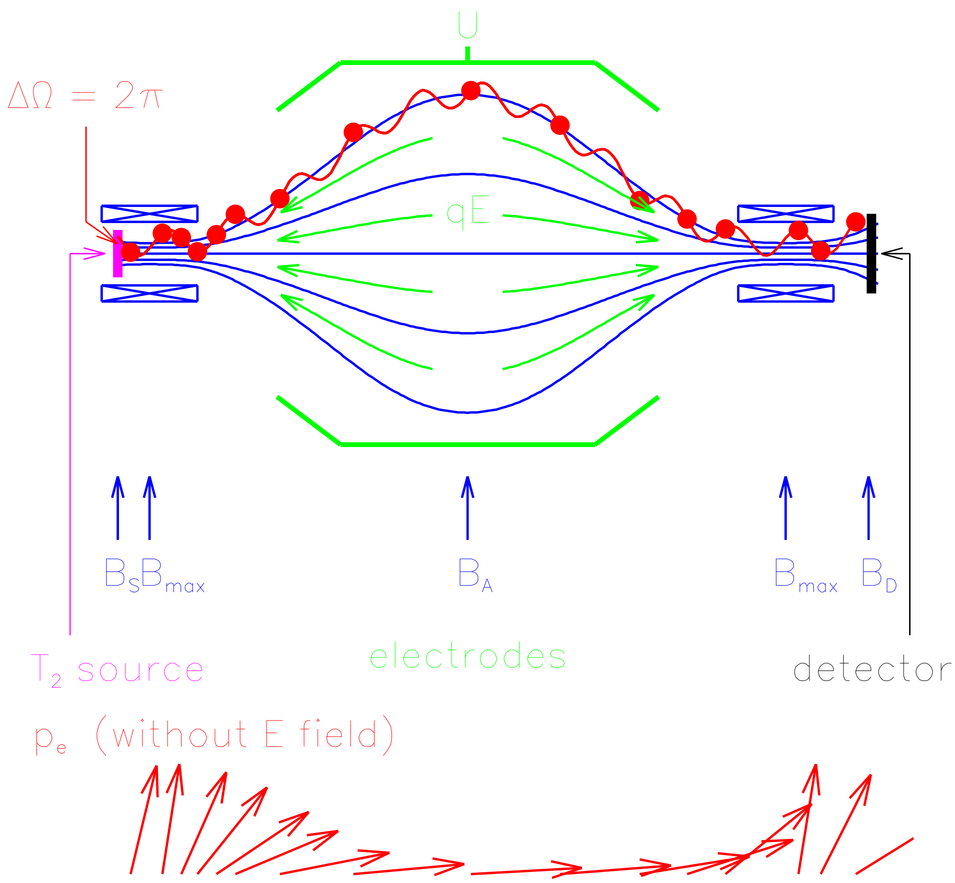


Figure 12.28: *Sketch of the MAK-E-Filter.* Source: [28, p. 17].

Bibliography

[1] F. Halzen and A. Martin. *Quarks & Leptons*. Wiley, 1984.

[2] A. Grozin. *Lectures on* QED *and* QCD. World Scientific, 2007.

[3] R. K. Ellis, W. J. Stirling, and B. R. Webber. *QCD and Collider Physics*. Cambridge University Press, 1996.

[4] G. Dissertori, I. K. Knowles, and M. Schmelling. *Quantum Chromodynamics: High Energy Experiments and Theory*. Oxford University Press, 2nd edition, 2009.

[5] G. Kane and A. Pierce. *Perspectives on LHC physics*. World scientific, 2008.

[6] Particle data group. http://pdg.lbl.gov/.

[7] ATLAS Collaboration. *arXiv:0901.0512 [hep-ex]*, 2009.

[8] CMS Collaboration. *J. Phys. G*, 34:995, 2007.

[9] D. E. Soper S. D. Ellis. Successive combination jet algorithm for hadron collisions. *Phys. Rev. D*, 48, 1993.

[10] E. Cattaruzza, A. Del Fabbro, and D. Treleani. 10.1103/physrevd.72.034022. *Phys. Rev. D*, 2005.

[11] M. Böhm, A. Denner, and H. Joos. *Gauge theories of the strong and electroweak interaction*. Teubner, 2001.

[12] J. R. Reitz. 10.1103/physrev.77.10. *Phys. Rev.*, 1950.

[13] M. E. Peskin and D. V. Schroeder. *An Introduction to Quantum Field Theory*. Westview, 1995.

[14] C. Amsler. *Kern- und Teilchenphysik*. UTB, 2007.

[15] J. Steinberger. `http://ccdb4fs.kek.jp/cgi-bin/img/allpdf?199012067`. 1990.

[16] The ALEPH Collaboration et al. *hep-ex/0509008*, 2005.

[17] `http://lepewwg.web.cern.ch/LEPEWWG/`.

[18] 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.

[19] http://tevnphwg.fnal.gov/results/SMHPubWinter2010/index.html.

[20] M. Fukugita and T. Yanagida. *Physics of neutrinos*. Springer, 2003.

[21] E. Lisi. http://indico.cern.ch/getFile.py/access?sessionId=4&resId=0&materialId=0&confId=80223. *Lectures at CHIPP PhD school*, 2010.

[22] E. Lisi. http://indico.cern.ch/getFile.py/access?sessionId=6&resId=0&materialId=2&confId=80223. *Lectures at CHIPP PhD school*, 2010.

[23] E. Lisi. http://indico.cern.ch/getFile.py/access?sessionId=6&resId=1&materialId=1&confId=80223. *Lectures at CHIPP PhD school*, 2010.

[24] B. Kayser. On the quantum mechanics of neutrino oscillation. *Phys. Rev. D*, 24(1), 1981.

[25] E. Lisi. Lectures at CHIPP PhD school. 2010.

[26] http://neutrino.ethz.ch/Vorlesung/HS2009/attachments/003_nt.pdf.

[27] K. Zuber. *Neutrino Physics*. Institute of Physics, 2004.

[28] C. Weinheimer. http://arxiv.org/pdf/0912.1619. 2009.

[29] E. W. Otten and C. Weinheimer. *Rep. Prog. Phys.*, 71, 2008.

Index