



Fermi National Accelerator Laboratory

FERMILAB-Conf-91/275-E

Hadron Collider Physics

L. Pondrom

*Physics Department, University of Wisconsin
1150 University Avenue, Madison, Wisconsin 53706*

October 1991

* *A Series of Lectures presented at CINVESTAV, Mexico, D.F., July, 1991.*



Operated by Universities Research Association Inc. under contract with the United States Department of Energy

HADRON COLLIDER PHYSICS

A Series of Lectures
given by

LEE G. PONDROM

*Physics Department, University of Wisconsin,
1150 University Avenue
Madison, Wisconsin 53706, USA*

at

CINVESTAV, Mexico, D.F.
July, 1991

ABSTRACT

An introduction to the techniques of analysis of hadron collider events is presented in the context of the quark-parton model. Production and decay of W and Z intermediate vector bosons are used as examples. The structure of the Electroweak theory is outlined. Three simple FORTRAN programs are introduced, to illustrate Monte Carlo calculation techniques.

Introduction

These lectures are intended for a classroom rather than a technical seminar. A review of the present experimental situation in hadron collider physics has been published elsewhere.¹ Here we will attempt to explain the ideas behind the interpretation of hadron collider events. To do this we rely heavily on some standard texts in the field: "Collider Physics" by Barger and Phillips;² "Quarks and Leptons" by Halzen and Martin;³ "Gauge Theories of the Strong, Weak, and Electromagnetic Interactions" by Quigg;⁴ "Gauge Theories in Particle Physics" by Aitchison and Hey;⁵ and "Concepts of Particle Physics" by Gottfried and Weisskopf.⁶ The periodic publications of the Particle Data Group⁷ furnish the latest experimental results and in addition establish conventions to be followed in defining parameters of the theory, naming the particles, etc.

In the notes that follow we will borrow freely from these texts, sometimes citing specific sources, but sometimes not. We admit our debt to the above authors at the outset, hoping that lapses in citation will henceforth be forgiven.

The selection of topics and the organization has been chosen to cover the subject in five one hour lectures. Although much material has been omitted by necessity, the objective of these notes is to be comprehensible and self contained.

There are homework problems scattered throughout the notes, which should be worked out by the interested reader. Calculating physical quantities is always enlightening, and is necessary for progress to be made in the science.

Lecture I

1.1 e^+e^- Annihilation

Quantum electrodynamics is a good place to start. The interactions of charged leptons and photons can be accurately described by the Feynman rules. These rules have been extended to cover the Electroweak theory and QCD as well, so the more familiar QED furnishes a natural framework to introduce the notation and kinematics.

We begin with the four dimensional notation for photons and leptons, the Dirac equation, and the application of the Feynman rules to the calculation of cross sections. All of the texts mentioned above have adopted the four dimensional metric tensor defined by Bjorken and Drell:⁸

$$g_{\mu\nu} = g^{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \quad (1.1)$$

so that the space components of a lower index four vector have the opposite signs to those of an upper index four vector:

$$x^\mu = \{x^0, \vec{x}\}, \quad x_\mu = \{x_0, -\vec{x}\} \quad x_\mu x^\mu = t^2 - \vec{x}^2 \quad (1.2)$$

We will adopt natural units throughout: $\hbar = c = 1$. In these units length, time, and 1/energy have the same dimensions. For cross sections, $(1 \text{ GeV})^{-2} = 0.389 \text{ mb}$ is handy, while for decay rates $\hbar = 6.58 \times 10^{-25} \text{ GeV-sec}$ does the trick. There is no distinction among the units for mass, momentum, and energy.

The gradient $\partial_\mu = \frac{\partial}{\partial x^\mu}$ transforms like a lower index vector. A simple Lorentz transformation to a prime frame moving with velocity $+v$ along the 3 axis is

$$x^{\nu'} = a^\nu_\lambda x^\lambda \quad \text{with} \quad a^\nu_\lambda = \begin{pmatrix} \gamma & 0 & 0 & -v\gamma \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -v\gamma & 0 & 0 & \gamma \end{pmatrix} \quad \text{and} \quad \gamma = \frac{1}{\sqrt{1-v^2}}. \quad (1.3)$$

The Dirac equation in momentum space in this metric for a free positive energy electron reads

$$(\not{k} - m)u(p) = 0, \quad \not{k} \equiv \gamma^\mu p_\mu. \quad (1.4)$$

In coordinate space the form is $(i\gamma^\mu \partial_\mu - m)\psi(x) = 0$, with $p_\mu = i\partial_\mu$.

The adjoint equation is

$$\bar{u}(p)(\not{k} - m) = 0. \quad (1.5)$$

where $\bar{u} = u^\dagger \gamma^0$ and $\gamma^{\mu\dagger} = \gamma^0 \gamma^\mu \gamma^0$.

The corresponding negative energy spinor equations are

$$(\not{k} + m)v(p) = 0, \text{ and } \bar{v}(p)(\not{k} + m) = 0. \quad (1.6)$$

The anticommutator of the gamma matrices is

$$\gamma^\mu \gamma^\nu + \gamma^\nu \gamma^\mu = 2g^{\mu\nu}. \quad (1.7)$$

Invariant variables first introduced by Mandelstam are very useful for describing the kinematics of reactions of the type $A + B \rightarrow C + D$. QED formulas exhibit useful symmetries when written as functions of these variables. In terms of the four momenta of the particles, they are defined as follows:

$$\begin{aligned} s &= (p_A + p_B)^2 \\ t &= (p_A - p_C)^2 \\ u &= (p_A - p_D)^2 \end{aligned} \quad (1.8)$$

with the constraint $s + t + u = m_A^2 + m_B^2 + m_C^2 + m_D^2$. The rapidity of a particle is defined relative to the direction of motion of the AB rest frame by the logarithmic relation:

$$y = \frac{1}{2} \ln \left(\frac{E + p_{\parallel}}{E - p_{\parallel}} \right). \quad (1.9)$$

Rapidities add algebraically when going from one Lorentz frame to another, as long as the motion is parallel to the original AB direction. Therefore rapidity differences, or shapes of rapidity distributions, are invariant under such Lorentz transformations. The maximum value of the rapidity for a particle of mass m in a system with total energy \sqrt{s} is

$$y_{\max} = \ln \left(\frac{\sqrt{s}}{m} \right). \quad (1.10)$$

Since at high energies the particle mass can be neglected, the pseudorapidity is often used

$$\eta = \frac{1}{2} \ln \left(\frac{1 + \cos \theta}{1 - \cos \theta} \right) = -\ln \left(\tan \frac{\theta}{2} \right) \quad (1.11)$$

Pseudorapidity depends only on the polar angle. This definition is invalid for very small angles where the quantity in Eq (1.11) is larger than the allowed maximum of Eq (1.10). At the SPS and Tevatron colliders, pseudorapidity is never a good approximation for W and Z bosons. They are too heavy, and so is the as yet undiscovered top quark. But it works well for everything else. Because of the invariance of rapidity differences, many detectors are segmented in constant bites in $\Delta\eta$.

In order to perform calculations, we must know how to relate cross sections and decay rates to invariant matrix elements, which in turn are calculated from the Feynman rules. The decay rate formula will be introduced in Lecture III. The differential

cross section per unit solid angle for the process $A + B \rightarrow C + D$ in the AB rest frame is given by the formula

$$\frac{d\sigma}{d\Omega} = \frac{1}{64\pi^2} \frac{1}{s} \frac{p_f}{p_i} |\mathcal{M}|^2 \quad (1.12)$$

where \mathcal{M} is the invariant matrix element for the reaction, and $s = (p_A + p_B)^2$ is the Mandelstam variable.

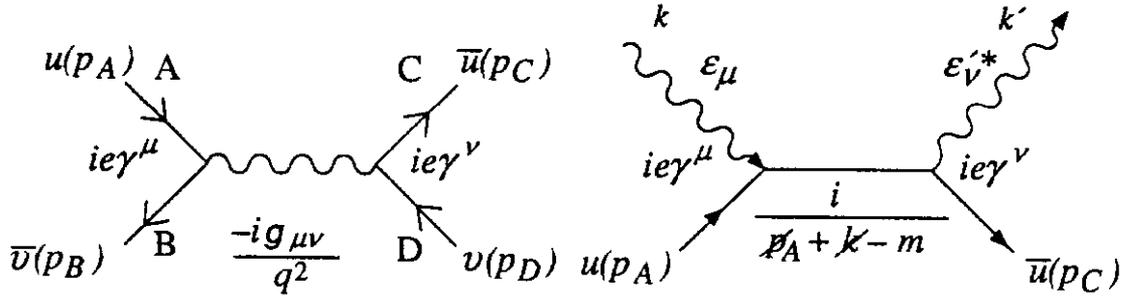


Figure 1.1. Feynman rules for $e^+e^- \rightarrow \mu^+\mu^-$.

Figure 1.2. Compton scattering diagram and Feynman rules.

Figure 1.1 shows the Feynman rules applied to the basic reaction $e^+e^- \rightarrow \mu^+\mu^-$. The invariant matrix element can be written down by applying the factors shown in the Fig.—the spinors, the vertex factors, and the photon propagator—as follows:

$$-i\mathcal{M} = ie\bar{v}(p_B)\gamma^\nu u(p_A) \left(\frac{-ig_{\mu\nu}}{q^2} \right) ie\bar{u}(p_C)\gamma^\mu v(p_D) \quad (1.13)$$

where $s = q^2 = (\text{mass})^2$ of the intermediate photon. The Compton scattering diagram involves real photons and an electron propagator, as shown in Fig. 1.2. In this case the photon is characterized by a four momentum and a polarization four vector. Similar formulas will be obtained for real W 's and Z 's. The use of the Feynman rules in this case gives the amplitude:

$$-i\mathcal{M} = \bar{u}(p_C) \left[\epsilon_\nu^* i e \gamma^\nu \frac{i(\not{p}_A + \not{k} + m)}{(p_A + k)^2 - m^2} i e \gamma^\mu \epsilon_\mu \right] u(p_A) \quad (1.14)$$

The amplitude for the crossed diagram must be added coherently to Eq (1.14) to obtain the correct expression for the cross section.

The annihilation process $e^+e^- \rightarrow \mu^+\mu^-$ is of fundamental importance in electron storage rings. In the high energy limit where all of the lepton masses can be ignored, the differential cross section is

$$\frac{d\sigma}{d\Omega} = \frac{\alpha^2}{2s} \left(\frac{t^2 + u^2}{s^2} \right) = \frac{\alpha^2}{4s} (1 + \cos^2 \theta), \quad (1.15)$$

where $\alpha = \frac{e^2}{4\pi} = \frac{1}{137}$. The total cross section is:

$$\sigma = \frac{4}{3}\pi \frac{\alpha^2}{s}. \quad (1.16)$$

HOMEWORK PROBLEM: Use the Feynman rules to derive the differential and total cross sections for $e^+e^- \rightarrow \mu^+\mu^-$, Eqs. 1.15 and 1.16 above. To perform the spin sums, use the trace theorems given in References 3 and 4. Note that the trace factors into two terms which have the same form. The Mandelstam variables of Eq. 1.8 have been simplified by the zero mass approximation.

1.2 The Standard Model

The standard model is built with six leptons and six quarks (and their antiparticles) given in Table 1. The quark varieties are called flavors, and each flavor comes in three colors. There are four gauge fields: γ , Z , and W of the electroweak sector, and g for QCD. The photon and gluon are massless, while the Z and W are very heavy. There are eight gluons carrying the colors of the $\mathbf{3} \times \mathbf{3}^*$ octet representation of color SU(3). Leptons are restricted to electroweak interactions. The quarks share the electroweak forces, and in addition couple to gluons through the color charge to give QCD. Neutrinos are assumed to be massless. In fact, at $\sqrt{s} \cong 1$ TeV all leptons and quarks except for the t and perhaps the b may be considered massless.

Table 1: “Standard Model” Ingredients

Leptons			Charge	Quarks			Charge
ν_e	ν_μ	ν_τ	0	u	c	t	$+\frac{2}{3}$
e^-	μ^-	τ^-	-1	d	s	b	$-\frac{1}{3}$

A quark-antiquark pair can be produced instead of a lepton pair in the final state after e^+e^- annihilation. The cross section for this process in its simplest form is just like Eq. 1.16, weighted by the square of the quark charge and multiplied by 3 for color, which gives a formula for the ratio R :

$$R = \frac{\sigma(e^+e^- \rightarrow q\bar{q})}{\sigma(e^+e^- \rightarrow \mu^-\mu^+)} = 3 \sum_f Q_f^2. \quad (1.17)$$

$R = 2$ below charm threshold; $R = 3\frac{1}{3}$ between charm and bottom thresholds; and $R = 3\frac{2}{3}$ above bottom threshold. The top quark threshold would cause a step in R of $1\frac{1}{3}$ units. There are QCD corrections to these numbers, and in the resonance regions near the J/ψ , upsilons, and Z there is sharp structure in the ratio, but the trend in R , shown in Fig. 1.3 (Ref. 7), follows expectations. The lower part of Fig. 1.3 shows R above the upion region, where a search has been made for the top quark threshold to no avail. The rise above 50 GeV is due to the tail of the Z pole, which is centered at (91.161 ± 0.031) GeV?

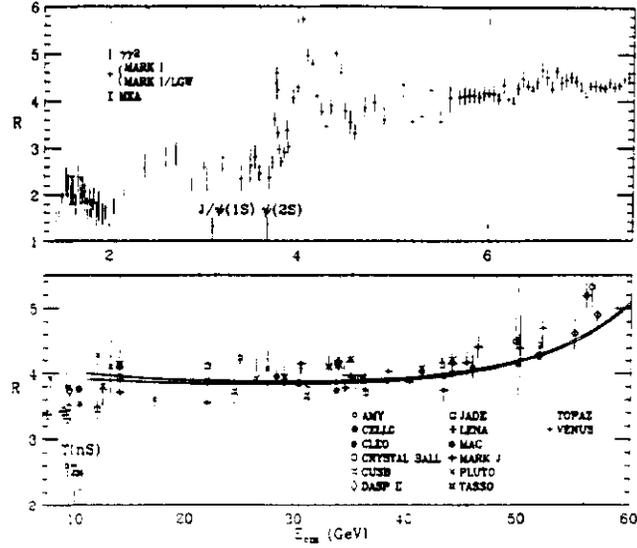


Figure 1.3. $R = (e^+e^- \rightarrow \text{hadrons})/(e^+e^- \rightarrow \mu^+\mu^-)$

The Electroweak theory specifies how the Z propagator and the photon propagator interfere. The formula for the cross section obtained by the coherent sum of amplitudes for these two intermediate states

$$e^+e^- \rightarrow q\bar{q} = \left| \begin{array}{c} \text{diagram with } \gamma \text{ and } Z \text{ exchange} \\ \text{charge } Q_f \end{array} \right|^2$$

is given by

$$\frac{d\sigma}{d\Omega} = \frac{\alpha^2}{4s} \left[Q_f^2(1 + \cos^2 \theta) - 2Q_f\chi_1 \{VV_f(1 + \cos^2 \theta) - 2a_f \cos \theta\} + \chi_2 \{(V_f^2 + a_f^2)(1 + V^2)(1 + \cos^2 \theta) - 8VV_f a_f \cos \theta\} \right] \quad (1.18)$$

The first term is the same as Eq. 1.15. The second term, involving the real part of the Breit-Wigner amplitude χ_1 , is the interference between the photon and the Z . Here the vector-vector interference is symmetric, while the vector-axial vector part is linear in $\cos \theta$, the angle between the electron and the outgoing fermion, and results in a front-back asymmetry. The third term, multiplied by the square of the Breit-Wigner amplitude χ_2 , is the Z amplitude squared. Here the vector and axial vector parts of the Z coupling to fermions contribute to the $(1 + \cos^2 \theta)$ term, while the vector-axial vector interference of the Z itself gives an asymmetric part. As we shall see, the Z coupling is almost pure axial vector, so the χ_2 asymmetric term is small. The χ_1 term can be large, but it changes sign as the energy increases through the Z pole, and vanishes on resonance. The variation of the coefficient of $\cos \theta$ as

the energy passes through the Z resonance is shown in Fig. 1.4. These properties are clear from the formulas below:

$$\chi_1 = \frac{1}{16 \sin^2 \theta_W \cos^2 \theta_W} \frac{s(s - M_Z^2)}{(s - M_Z^2)^2 + \Gamma^2 M_Z^2} \quad (1.19)$$

$$\chi_2 = \frac{1}{256 \sin^4 \theta_W \cos^4 \theta_W} \frac{s^2}{(s - M_Z^2)^2 + \Gamma^2 M_Z^2} \quad (1.20)$$

$$V = -1 + 4 \sin^2 \theta_W \quad (\text{leptonic vector term}) \quad (1.21)$$

$$a_f = 2T_{3f} \quad (a_\ell = -1) \quad (1.22)$$

$$V_f = 2T_{3f} - 4Q_f \sin^2 \theta_W; \quad T_3 = \begin{cases} +\frac{1}{2} & \text{for } \nu_e, \nu_\mu, u, c, t \\ -\frac{1}{2} & \text{for } e^-, \mu^-, \tau^-, d, s, b \end{cases} \quad (1.23)$$

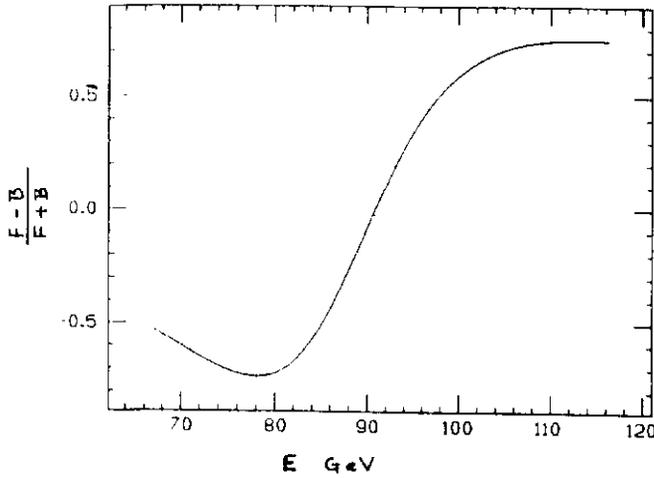


Figure 1.4. Asymmetry of $e^+e^- \rightarrow \mu^+\mu^-$ as a function of energy.

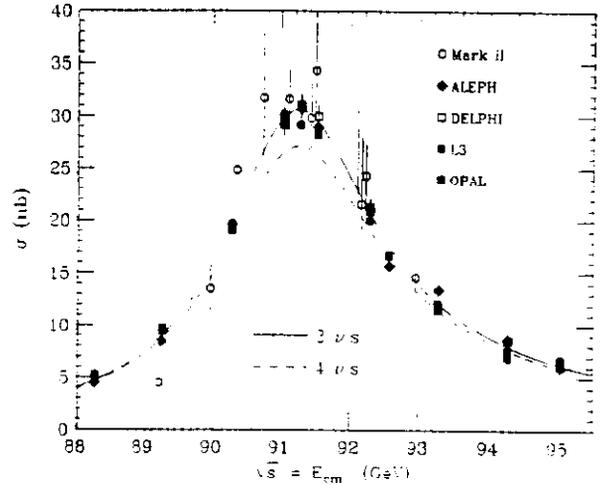


Figure 1.5. Z resonance line shape

In the following lectures we will attempt to justify these assignments. The Weinberg angle θ_W is a parameter of the theory. The first measurements of θ_W were performed by studying the neutral current cross sections in neutrino scattering, with the result $\sin^2 \theta_W = 0.23$, which results in a vector coupling constant for leptons of 0.08, about 10% of the axial vector coupling.

Figure 1.5 (Ref. 7) shows the cross section for e^+e^- annihilation into hadronic final states—*i.e.*, all quark final states which then dress themselves into jets of hadrons—as the energy is varied through the Z resonance. The plot shows the consistency of the cross section data with the hypothesis of three neutrino families, and gives reality to the above formulas, characterizing the Z as a massive intermediate state. In the next lecture we will see how these formulas apply to hadron colliders.

Lecture II

2.1 The Drell-Yan Process in Hadronic Collisions

In the quark-parton model, hadrons are assumed to consist of free point particles with negligible mass, dubbed partons by Feynman. These partons have been identified as fractionally charged quarks and neutral gluons. This picture is valid in an impulse approximation: the hadron is probed for a very short time span, corresponding via the uncertainty principle to a very large energy transfer, during which only one parton participates in the reaction, and the other constituents of the hadron act as spectators. The incident projectile, which is a lepton in lepton-hadron collisions (deep inelastic scattering) and another parton in hadron-hadron collisions, strikes a parton which carries a momentum fraction $0 < x < 1$ of the total momentum of the target. The collision occurs between point particles, analogous to e^+e^- annihilation, except that the kinematics are determined by x_1 and x_2 , while for e^+e^- $x_1 = x_2 = 1$. The probability distributions of parton momenta in the hadron are called structure functions⁹. More will be said about them later.

Equation 1.18 for the process $e^+e^- \rightarrow q\bar{q}$ can be time reversed without any changes to describe the inverse reaction $\bar{q}q \rightarrow e^+e^-$, which in the parton model is the subprocess for the observed reaction $\bar{p} + p \rightarrow e^+e^- + X$, and is called the Drell-Yan mechanism, shown in Fig 2.1. We proceed with the kinematics of the reaction in the $\bar{p}p$ rest frame, which coincides with the laboratory frame at the SPS and Tevatron colliders. The parton model differs from e^+e^- annihilation in two important ways:

- a) the $\bar{q}q$ rest frame does not in general correspond to the $\bar{p}p$ rest frame, so that the lepton pair can be moving in the laboratory along the beam direction (Transverse motion also happens, but can be neglected in first approximation.); and
- b) although the $\bar{p}p$ system is neutral, the $\bar{q}q$ system need not be, which permits the production of not only the Z boson, but single W^+ or W^- as well. Production of W^\pm , at least until LEP 200 begins operation, is the sole province of hadron colliders.

The $\bar{q}q$ collision is collinear, with total momentum

$$\vec{p}_q + \vec{p}_{\bar{q}} = \vec{k} \quad (2.1)$$

and energy

$$E_q + E_{\bar{q}} = E_k. \quad (2.2)$$

Now introduce the momentum fractions x_1 and x_2 :

$$\vec{p}_q = x_1 \vec{P}, \text{ and } \vec{p}_{\bar{q}} = x_2 (-\vec{P}) \quad (2.3)$$

where \vec{P} and $-\vec{P}$ are the proton/antiproton incident momenta, and x_1, x_2 are dimensionless variables. Momentum conservation then gives the momentum fraction of the intermediate state (or the lepton pair) as

$$k = (x_1 - x_2)P \text{ or } x = x_1 - x_2. \quad (2.4)$$

Since $E_k = (x_1 + x_2)P$, $M^2 = E_k^2 - \vec{k}^2 = 4x_1x_2P^2$, resulting in

$$x_1x_2 = M^2/s, \quad (2.5)$$

where $s = 4P^2$. The rapidity of the produced state y is given in terms of x_1 and x_2 by

$$x_{1,2} = \frac{M}{\sqrt{s}} e^{\pm y} \quad (2.6)$$

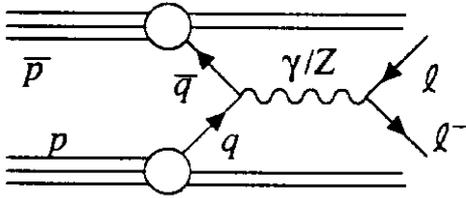


Figure 2.1. Drell-Yan mechanism in the Parton Model

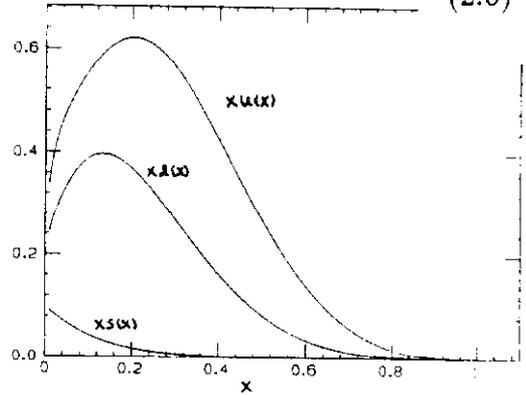


Figure 2.2. Field and Feynman structure functions u, d, s for the proton.

The Drell-Yan cross section for an intermediate photon, i.e. well below the Z resonance, can be written down from the definitions of the structure functions and the basic formula of Eq. 1.16. Thus

$$d^2\sigma(\bar{p}p \rightarrow \ell^+\ell^-X) = \frac{4\pi}{3} \frac{\alpha^2}{M^2} \frac{1}{3} \sum_a Q_a^2 [q_a(x_1)\bar{q}_a(x_2)dx_1dx_2] \quad (2.7)$$

where M^2 is the square of the mass of the lepton pair, and Q_a is the quark charge. The factor of $1/3$ is for color, because only color neutral $\bar{q}q$ pairs are allowed in the sum, and the structure functions are identical for the three quark colors. For simplicity the quark is assumed to reside in the proton and the antiquark in the antiproton, although it is possible for the two to be reversed. Antiquarks occur in the proton only as vacuum fluctuations of $\bar{q}q$ pairs. The proton structure functions are u and d valence quarks, plus $\bar{u}u$, $\bar{d}d$, and $\bar{s}s$ sea quarks. The gluons play no direct role in either deep inelastic scattering or the Drell-Yan mechanism—a reason why the gluon structure functions are not as well known as those of the quarks. Quark structure functions for the proton obtained by Field and Feynman¹⁰ are shown in Fig. 2.2.

With a change of variables Eq. 2.7 can be written in terms of $dM^2 dx$:

$$\frac{d^2\sigma}{dM^2 dx} = \frac{4\pi\alpha^2}{9M^2} \frac{1}{s(x^2 + 4M^2/s)^{1/2}} \sum_a Q_a^2 q_a(x_1) \bar{q}_a(x_2) \quad (2.8)$$

HOMEWORK PROBLEM: Show that this expression follows from Eq. 2.7.

2.2 The Structure Functions

Deep inelastic lepton scattering experiments (e 's, μ 's, and ν 's) have been the sources for the structure functions, which are obtained by fits to the data. The structure functions cannot be derived from first principles, and must be calculated empirically. Examples of deep inelastic scattering in the parton model together with the kinematics are shown in Fig. 2.3. The cross section is linear in the structure functions (only one hadron), so by assuming charge conjugation invariance— u in the proton equals \bar{u} in the antiproton—the Drell-Yan cross section can be written down directly in terms of measured proton characteristics. This is the interpretation of Eq. 2.8. Unfortunately this procedure does not quite work, as the cross section so calculated is about 2/3 of the experimental value. To fix things up, a “ K factor” of 1.5 is applied to the right hand side of Eq. 2.8. This factor can be calculated in QCD.

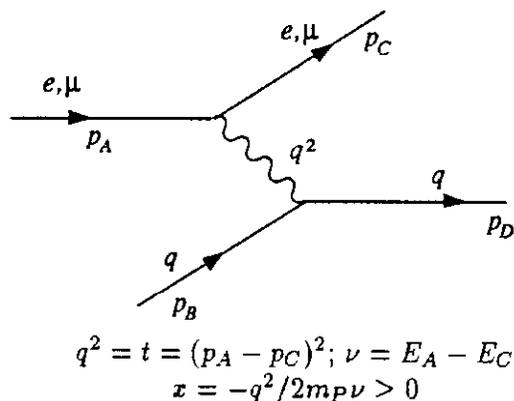


Figure 2.3. Deep inelastic scattering in the parton model.

The motion of bound nucleons in a spinless nucleus is spherically symmetrical, and related to the Fourier transform of a nucleon wave function. Do the quark structure functions in the proton arise from similar considerations? The answer is yes, but there is a fundamental difference between nucleons and nuclei.

Consider a nuclear analog of the proton with its three valence quarks (u, u, d)— He^3 (p, p, n). The negative binding energy of He^3 , about 6 MeV, is much smaller than the masses of the constituents, so we may safely write the mass formula for the nucleus as:

$$M = 3m \quad (2.9)$$

Now boost the nucleus to very high momentum:

$$P = \gamma M = 3\gamma m \quad (2.10)$$

Now each “parton” has 1/3 the momentum of the nucleus, and hence $x = 1/3$. The nucleon motion smears this delta function out a little bit, the width being equal to the nucleon motion in the transverse plane.

The proton is a very different object. We are told that the quarks are confined—they cannot escape as free particles—and that the quark masses are negligibly small. Hence a formula like Eq. 2.9 is not valid. Hadrons, unlike any other materials, weigh more than the sum of their constituent masses. Nothing else does this - atoms, nuclei, lead bricks, etc all weigh less than their parts. The energy which forms the proton mass in its rest frame is kinetic, so we write the analog of Eq. 2.9 as follows:

$$M = \sum_i E_i \quad (2.11)$$

where E_i is the kinetic (or total) energy of the massless i th constituent. To conserve momentum, there must be at least two partons, and the maximum value of E_i for any parton is

$$(E_i)_{\max} = M/2 \quad (2.12)$$

Now boost the hadron to high momentum by multiplying both sides of Eq. 2.11 by the Lorentz factor:

$$P = \gamma M = \gamma \sum_i E_i \quad (2.13)$$

Since $\sum_i p_{i\parallel} = 0$ in the proton rest frame, it can be added to the right hand side of Eq. 2.13 without penalty:

$$P = \gamma \sum_i (E_i + p_{i\parallel}) \quad (2.14)$$

The longitudinal momentum of the i th parton in the boosted frame is $P_i = \gamma(E_i + p_{i\parallel})$, and the momentum fraction $x = \frac{E_i + p_{i\parallel}}{M}$, where both terms are proton rest frame quantities. Note that since $E_{\max} = p_{\parallel\max} = M/2$, and p_{\parallel} can be plus or minus, the range for x is $0 < x < 1$. This is Feynman’s x . The Lorentz invariant transverse momentum can be ignored.

Thus confined massless partons are the key to understanding the structure functions. The binding of the partons and the fact that the number of partons is not fixed play a role, in fact it is the binding that gives the shape of the x distribution, but these considerations are not fundamental to obtaining a momentum fraction anywhere between zero and one. Free massless particles will do that very nicely. If we knew how to solve the Dirac equation and obtain wave functions for the bound massless partons, then their distributions in momentum space would permit the calculation of $u(x)$. So far we are not able to solve this bound state problem.¹¹

2.3 Hadronic Z production

The Drell-Yan formula Eq. 2.7 can be readily adopted to the case where the intermediate photon and Z interfere. Equation (1.18) when integrated over the solid angle gives the cross section for $\bar{q}_a + q_a \rightarrow \ell^+ \ell^- X$ as:

$$\hat{\sigma}(\bar{q}_a q_a \rightarrow \ell^+ \ell^- X) = \frac{4\pi\alpha^2}{3\hat{s}} [Q_a^2 - 2Q_a V V_a \chi_1 + (V_a^2 + a_a^2)(1 + V^2)\chi_2] \quad (2.15)$$

where \hat{s} is the square of the mass of the lepton pair, Q_a is the quark charge, and (V_a, a_a) are the vector and axial vector coupling coefficients for the quarks. The Z Breit-Wigner terms are χ_1 and χ_2 . The details are given in Eqs. 1.19-1.23. It follows then that in the parton model the cross section for $\bar{p}p \rightarrow \ell^+ \ell^- X$ is given by (neglecting the sea quark contributions):

$$d^2\sigma(\bar{p}p \rightarrow \ell^+ \ell^- X) = \frac{4\pi\alpha^2 K}{3 M^2} \frac{K}{3} \sum_a [Q_a^2 - 2Q_a V V_a \chi_1 + (V_a^2 + a_a^2)(1 + V^2)\chi_2] q_a(x_1) \bar{q}_a(x_2) dx_1 dx_2. \quad (2.16)$$

On resonance $\hat{s} = M_Z^2$, the Q_a^2 term is negligible and the χ_1 term vanishes, so that

$$d^2\sigma(\bar{p}p \rightarrow Z \rightarrow \ell^+ \ell^-) = \frac{4\pi\alpha^2 K}{3 M_Z^2} \frac{K}{3} \sum_a (V_a^2 + a_a^2)(1 + V^2)\chi_2 q_a(x_1) \bar{q}_a(x_2) dx_1 dx_2 \quad (2.17)$$

The variables can be changed to rapidity and M^2 by using the identity $dx_1 dx_2 = dy dM^2/s$, where s refers to the $\bar{p}p$ system. The result is

$$\frac{d^2\sigma}{dy dM^2} = \frac{4\pi\alpha^2 K}{9 s M_Z^2} \sum_a (V_a^2 + a_a^2)(1 + V^2)\chi_2 q_a(x_1) \bar{q}_a(x_2) \quad (2.18)$$

Integrating over the Breit-Wigner and using the partial width formula $\Gamma(Z \rightarrow \ell^+ \ell^-) = \frac{G_F M_Z^3 (1 + V^2)}{24\sqrt{2}\pi}$ gives the result

$$\frac{d\sigma}{dy} = \frac{\pi K G_F}{3 \sqrt{2}} B_{\ell\ell} x_1 x_2 \left[\left(1 - \frac{8}{3}x_W + \frac{32}{9}x_W^2\right)u(x_1)u(x_2) + \left(1 - \frac{4}{3}x_W + \frac{8}{9}x_W^2\right)d(x_1)d(x_2) \right] \quad (2.19)$$

where $\sin^2 \theta = x_W$, and $B_{\ell\ell} = \Gamma_{\ell\ell}/\Gamma$. Here we have introduced the Fermi coupling constant G_F , the characteristic strength of the weak interactions, which is derived from the muon lifetime. This agrees with the equation on p. 252 of Ref. 2.

HOMEWORK PROBLEM: Derive Eq 2.19 from Eq 2.18. Hint:

$$\int_0^\infty \frac{\hat{s} d\hat{s}}{(\hat{s} - M_Z^2)^2 + \Gamma^2 M_Z^2} \approx \frac{\pi M_Z}{\Gamma}$$

Lecture III

3.1 Hadronic W Production

In the last lecture we learned that the hadronic reaction to produce a high mass lepton pair, $\bar{p}p \rightarrow \ell^+\ell^-X$, proceeds via the parton model through the subprocess $\bar{q}q \rightarrow \ell^+\ell^-$, which is the inverse of the e^+e^- annihilation into $\bar{q}q$ (two jets).

Since the $\bar{q}q$ pair can have a net charge, single W 's can also be produced in $\bar{p}p$ collisions by an extension of the Drell- Yan mechanism to include weak charged currents. The analog to Eq. 2.19 for $\bar{p}p \rightarrow W^+X \rightarrow \ell^+\nu X$ is:

$$\frac{d\sigma}{dy}(\bar{p}p \rightarrow W^+ \rightarrow \ell^+\nu) = \frac{2\pi G_F}{3\sqrt{2}} B_{\ell\nu} K x_1 x_2 u(x_1) d(x_2). \quad (3.1)$$

$B_{\ell\nu} = \Gamma_{\ell\nu}/\Gamma$ is the branching ratio for the decay of the W into a (ℓ, ν) lepton pair. Eq. 3.1 produces a W^+ by the simplest parton model process: a u quark in the proton annihilates a \bar{d} quark in the antiproton to produce a W^+ plus spectator debris. In a similar fashion a W^- would be created by the annihilation of a d in the proton and a \bar{u} in the antiproton. This model gives similar interpretations for neutral lepton pair (ℓ^+, ℓ^-) and charged lepton pair (ℓ^\pm, ν) production in $\bar{p}p$ collisions.

The missing neutrino in W^\pm decay complicates the reconstruction of the events in the laboratory. The transverse momentum of the neutrino can be inferred from measured quantities and momentum conservation, but the longitudinal momentum is unknown. In principle p_{\parallel} could be determined by conserving momentum in the beam direction, but it is usually not possible to do this because measurements cannot be made close to the direction of the colliding beams. As a result, there are two solutions for $\cos\theta$ the polar angle of the charged lepton in the W rest frame. Transverse momentum balance gives:

$$\vec{p}_{\nu\perp} = -\vec{p}_{\ell\perp}, \text{ or } \vec{p}_{\ell\perp} \cdot \vec{p}_{\nu\perp} = \frac{-M_W^2}{4} \sin^2\theta \quad (3.2)$$

Hence the two solutions for $\cos\theta$ are symmetric about $\pi/2$ in the W rest frame:

$$\cos\theta = \pm \left(1 - \frac{4p_{\ell\perp}^2}{M_W^2}\right)^{1/2} \quad (3.3)$$

These two values correspond to two W momenta in the laboratory. This is a kinematic ambiguity, characteristic of a zero constraint fit.

Figure 3.1 shows the W^+ rapidity distribution calculated with the W_PROD program described in the APPENDIX for $\sqrt{s} = 1800$ GeV (the Tevatron). Eq. 3.1 and the simple structure functions of Feynman and Field were used for these calculations. Figure 3.2 shows the charged lepton rapidity distribution, which, in contrast to that

of the W 's, is directly measurable. In proceeding from 3.2 to 3.1, there are two possible W momenta—a higher value, and a lower one. Over the central range ± 1 unit of rapidity these two values are equally probable at 1800 GeV, which means that it is impossible from this information to determine whether the charged lepton was emitted in the forward or the backward hemisphere in the W rest frame.

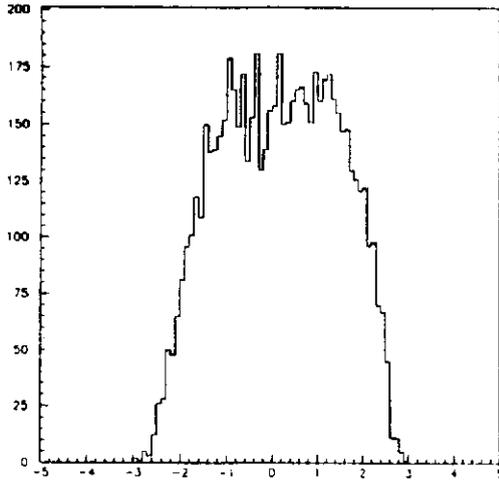


Figure 3.1. W^+ rapidity distribution at 1.8 TeV.

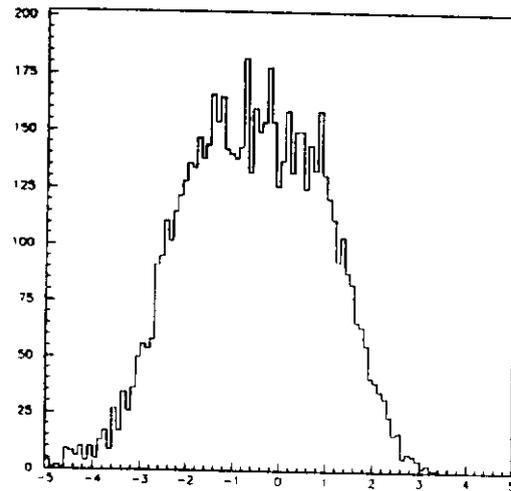


Figure 3.2. Lepton rapidity distribution at 1.8 TeV.

Figure 3.1 also shows a slight shift towards positive rapidity for W^+ . Positive rapidity is defined in the proton direction. Thus there is a preference in W production to preserve the flow of charge. In the parton model this preference is accommodated by an x dependence to the ratio $u(x)/d(x)$ of structure functions. Figure 2.2 shows that $d(x)$ damps out faster than $u(x)$ with increasing x . The lepton distributions in Fig. 3.2 are shifted slightly the other way—towards negative rapidity for ℓ^+ . This is because while the Drell-Yan mechanism preserves the flow of charge, the weak decay of the W preserves the direction quark \rightarrow lepton, and antiquark \rightarrow antilepton. Since antiprotons carry the antiquarks, and antileptons are positively charged, the flow of charge is reversed. The observed lepton rapidity distribution is a result of the competition between these two oppositely directed phenomena.

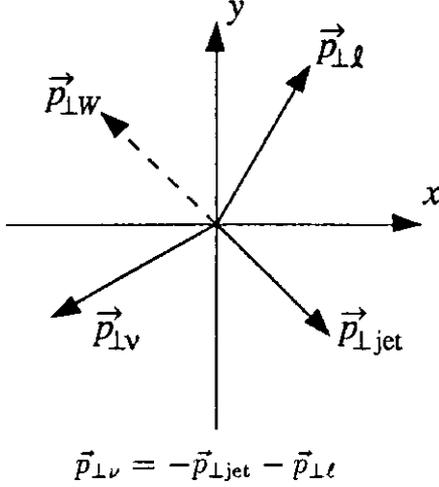


Figure 3.3. Vectors in the transverse plane for M_{\perp}^2 .

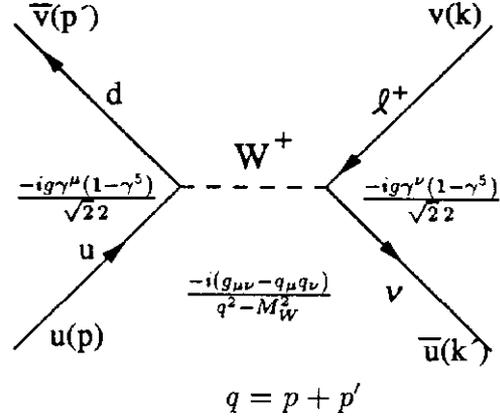


Figure 3.4. Feynman rules for $u + d \rightarrow W \rightarrow \ell + \nu$.

3.2 The Transverse Mass

The kinematic ambiguity resulting from a zero constraint fit to W decay assumes that the W mass is known. If measurement of the W mass is the experimental objective, then one constraint goes away, and the events are no longer reconstructable. It is possible, however, to work with transverse variables, and define a transverse mass by the equation:

$$M_{\perp}^2 = 2p_{\ell\perp} p_{\nu\perp} (1 - \cos \phi_{\ell\nu}) \quad (3.4)$$

where the measured and inferred vectors in the transverse plane are shown in Fig. 3.3. From the definition it is apparent that the maximum value of M_{\perp} is the true mass. There is another transverse mass sometimes used in kinematics, defined by the formula $m_{\perp}^2 = m^2 + p_{\perp}^2$, which has a minimum value of the true mass. Do not confuse the two.

Equation 3.1 is the cross section per unit rapidity times the branching ratio, integrated over angles and over the Breit-Wigner formula for the W line width. We can now derive this relation from first principles, following the procedures of Lecture II for the Z production and decay, and borrowing the matrix element from muon decay given by Eq. 3.20 below. The Feynman rules for the reaction $u + \bar{d} \rightarrow \ell^+ + \bar{\nu}$ are shown in Fig. 3.4 The matrix element for this reaction is given by:

$$\mathcal{M} = \frac{g^2}{8} \frac{(\bar{v}(p')\gamma^{\mu}(1-\gamma^5)u(p))}{\hat{s} - M_W^2 + iM_W\Gamma} (\bar{u}(k')\gamma_{\mu}(1-\gamma^5)v(k)). \quad (3.5)$$

In writing down Eq. 3.5 we have added an imaginary term to the propagator to account for the line width of the W . The momentum terms in the numerator of the propagator do not contribute because of the Dirac equation. If the dimensionless constant g^2 and the Fermi constant G_F are related by

$$\frac{G_F}{\sqrt{2}} = \frac{g^2}{8M_W^2}, \quad (3.6)$$

then in the low energy limit $\hat{s} \ll M_W^2$ Eq. 3.5 has the same form as Eq. 3.20. If we substitute $g^2/8 = G_F M_W^2/\sqrt{2}$ into Eq. 3.5, introduce the Mandelstam variables $\hat{s} = (p + p')^2$ and $\hat{u} = (p - k)^2$, and evaluate $\sum_{\text{spins}} |\mathcal{M}|^2$, we get the result:

$$\sum_{\text{spins}} |\mathcal{M}|^2 = 64 \left(\frac{G_F M_W^2}{\sqrt{2}} \right)^2 \frac{\hat{u}^2}{(\hat{s} - M_W^2)^2 + M_W^2 \Gamma^2} \quad (3.7)$$

Dividing by four to average over the u and \bar{d} quark spins (the wrong helicities do not contribute to the cross section), and using the cross section formula Eq. 1.12 gives the expression

$$\frac{d\hat{\sigma}}{d\Omega} = \frac{1}{16\pi^2} \left(\frac{G_F M_W^2}{\sqrt{2}} \right)^2 \frac{\hat{s}(1 - \cos \theta)^2}{(\hat{s} - M_W^2)^2 + M_W^2 \Gamma^2} \quad (3.8)$$

Here the angular distribution factor $(1 - \cos \theta)^2$, mentioned above in the discussion of the asymmetry in lepton rapidity, follows from the helicity rules. The cross section for $\bar{p}p$, in terms of the structure functions, the K factor, and the 1/3 for color is

$$\frac{d\sigma}{d\Omega} = \frac{K}{3} \int dx_1 \int dx_2 u(x_1) d(x_2) \frac{d\hat{\sigma}}{d\Omega} \quad (3.9)$$

The differential cross section per unit W rapidity per unit solid angle can be obtained from this expression by substituting the Jacobian $dx_1 dx_2 = dy \frac{d\hat{s}}{s}$. The result is:

$$\frac{d\sigma}{dy d\Omega} = \frac{K}{48\pi^2} \frac{1}{s} \left(\frac{G_F M_W^2}{\sqrt{2}} \right)^2 \int \frac{d\hat{s} u(x_1) d(x_2) \hat{s}(1 - \cos \theta)^2}{(\hat{s} - M_W^2)^2 + M_W^2 \Gamma^2}, \quad (3.10)$$

where the structure functions are evaluated at $x_{1,2} = \left(\frac{\hat{s}}{s} \right)^{1/2} e^{\pm y}$. The integration over the Breit Wigner line shape then gives the cross section

$$\frac{d\sigma}{dy d\Omega} = \frac{K}{48\pi} \frac{G_F}{\sqrt{2}} \left(\frac{G_F M_W^2}{\sqrt{2}} \right) \frac{1}{\Gamma} x_1 x_2 u(x_1) d(x_2) (1 - \cos \theta)^2. \quad (3.11)$$

Finally, integration over the solid angle produces Eq. 3.1.

HOMEWORK PROBLEM: Go through the steps from Eq. 3.7 to Eq. 3.11, and then to Eq. 3.1. You will need the W partial width into $\ell\nu$:

$$\Gamma(W \rightarrow \ell\nu) = \frac{1}{6\pi} \frac{G_F M_W^3}{\sqrt{2}} \quad (3.12)$$

The derivation of Eq. 3.7 will be sketched below in the section on muon decay.

The angular distribution Eq. 3.10 can be transformed into an expression in terms of transverse mass. Using $M_{\perp}^2 = 4p_{\perp}^2$, and $p_{\perp} = \frac{\sqrt{\hat{s}}}{2} \sin \theta$ gives the Jacobian of this transformation

$$d \cos \theta = \frac{dM_{\perp}^2}{\sqrt{1 - M_{\perp}^2/\hat{s}}} \quad (3.13)$$

This relation exhibits a singularity in the denominator for $M_{\perp}^2 = \hat{s}$, which is called the Jacobian peak. Although the distribution is singular, the integral is well behaved. The term linear in $\cos \theta$ in the angular distribution $(1 - \cos \theta)^2$ averages to zero in the transverse mass expression, because the transverse variables are independent of the sign of $\cos \theta$. The transformed formula then becomes

$$\hat{s}(1 - \cos \theta)^2 d \cos \theta = \frac{2 - M_{\perp}^2/\hat{s}}{\sqrt{1 - M_{\perp}^2/\hat{s}}} dM_{\perp}^2. \quad (3.14)$$

This derivation follows the discussion on pages 259 and 260 of Ref. 2. Equation 3.14 can be substituted into Eq. 3.10 to give the transverse mass distribution smeared by the Breit-Wigner line shape. Since the transverse mass does not depend on the rapidity of the W , the expression can be evaluated at $y = 0$. In addition, as the energy is varied over the W line width, the structure functions $u(x)$ and $d(x)$ can be assumed constant in first approximation. The results of a Monte Carlo calculation based on these assumptions are shown in Fig. 3.5. The sharp peak at $M_{\perp} = M_W$ comes from the denominator in Eq. 3.13. The program which produced this curve is described in the APPENDIX.

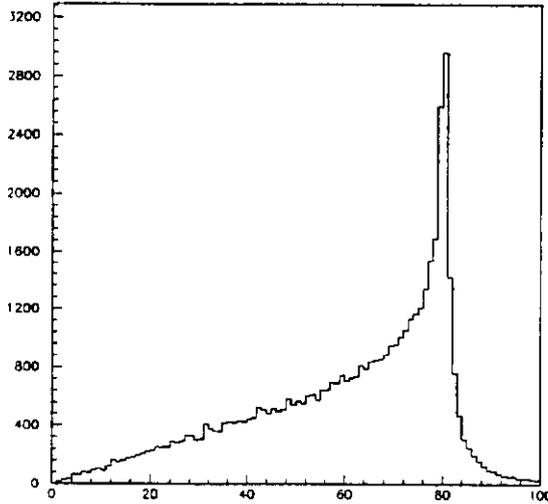


Figure 3.5. Transverse mass distribution for the perfect detector.

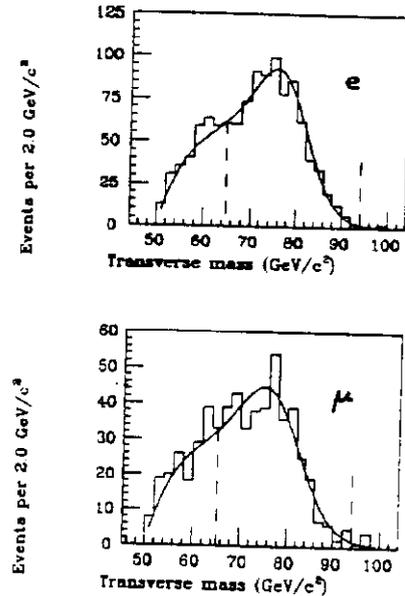


Figure 3.6. Transverse mass distributions for CDF data.

Experimental distributions for the W transverse mass are dominated by detector resolutions, which broaden the Jacobian peak considerably. The published results of CDF (12) are reproduced in Fig. 3.6. A cut on the lepton $p_{\perp} > 25$ GeV has been made on these data to reduce low M_{\perp} background. The smearing of the curve is well described by modelling the detector response, as is apparent from the agreement obtained between the data and the expected shape. Very little sensitivity to the W line width remains in the measured leading edge. The W line width has been measured by an indirect method.¹³ In this paper the cross section times branching ratio $\sigma(\bar{p} + p \rightarrow W^{\pm} + X)B(W \rightarrow e^{\pm}\nu)$ was divided by the corresponding expression for Z 's: $\sigma(\bar{p} + p \rightarrow Z + X)B(Z \rightarrow e^{+}e^{-})$. The cross section ratio, believed to be less sensitive to theoretical uncertainties than either cross section alone, was calculated from the standard model. Then using the standard model for the partial width $\Gamma(W \rightarrow e\nu)$, and the direct measurements of $B(Z \rightarrow e^{+}e^{-})$ from LEP and SLC gave the value $\Gamma = (2.19 \pm 0.2)$ GeV, in agreement with expectations.

3.3 Quark Decays of W 's

The inverse of W production by $\bar{q}q$ is the decay of the W into quarks: $W^{+} \rightarrow u + \bar{d}$. According to the standard model 64% of all W 's decay this way. The leptonic channels $W \rightarrow e\nu$ and $W \rightarrow \mu\nu$ represent about 22% of W decays, so the detection of the $\bar{q}q$ final state would be beneficial for two reasons: a) it would increase the detection efficiency by a factor of 3; and b) it would give a constrained fit to the W

mass, since there is no missing energy.

Unfortunately, although the $\bar{q}q$ pair is color neutral, the individual quarks are not, and are prohibited from appearing as free particles by confinement. As the quarks separate, the color field creates $\bar{q}q$ pairs, which combine to form color singlet mesons (and qqq triplets which form baryons a few percent of the time). These mesons appear in the laboratory as jets of particles. This process, which is called fragmentation, cannot be derived from first principles, but rather must be described phenomenologically. Two approaches have been successful: the independent fragmentation model of Feynman and Field¹⁴; and the correlated string model of the Lund group.¹⁵ The descriptions of the jets of particles are very similar in the two models. When there are differences between the predictions, the experimental data favor the Lund model.

Fragmentation models are based on extensive empirical data regarding the production of hadrons. It has been known for a long time that, as s increases, the mean transverse momentum (p_{\perp}) remains roughly constant, and the average number of charged particles $\langle n \rangle$ increases only logarithmically with s . Thus the extra energy goes neither into more particles, nor into more p_{\perp} . These observations lead naturally to the introduction of the scaling variable $x = 2p_{\parallel}/\sqrt{s}$, or the related variable y of Eq. 1.9, and the expression of the cross section for hadron production as a Gaussian in p_{\perp} times a function of y . In fact, the y dependence is almost constant, as can be seen from the logarithmic dependence of the multiplicity. Since the end points in rapidity increase logarithmically with s (Eq. 1.10), the multiplicity grows like $\log(s)$ if dn/dy is a constant. In this very simplified picture the multiplicity per unit rapidity is energy independent—like teeth in a comb. The length of the comb increases like $\log(s)$.

The physical characteristics of a jet—its multiplicity, average p_{\perp} , cone size, *etc.*—depend upon the transverse energy of the jet. Longitudinal motions arising from the structure functions have no effect. The definition of a jet in terms of $(\Delta\eta, \Delta\phi)$ is independent of the polar angle. See Fig. 3.7.

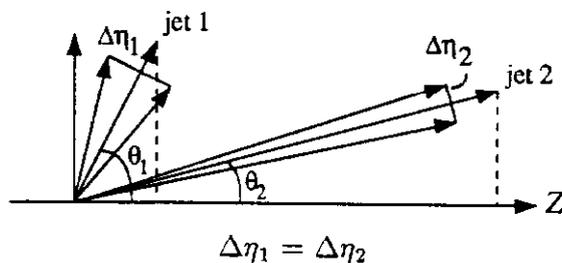


Figure 3.7. Two jets with same E_T but different thetas.

In the APPENDIX a program which generates jets using the simpler technique of Ref. 14 is outlined, and some of the results are shown.

The fragmentation process greatly complicates the experimental task of identifying the $W \rightarrow \bar{q}q$ decay for two reasons: a) Quark fragmentation to produce jets is not

perfectly understood; and b) The response of a detector to jet energy is not very good. These effects combine to give a rather poor mass resolution in the expression

$$M^2 = (E_1 + E_2)^2 - (\vec{p}_1 + \vec{p}_2)^2 \quad (3.15)$$

where one and two refer to the jets. Resolution smearing results in a poor signal to noise ratio, where the noise comes from a continuum of QCD dijet events. The UA2 group at CERN has studied the dijet mass range from 50 to 130 GeV in search for an enhancement in the W/Z mass region!¹⁶ (The Z is 11 GeV heavier than the W , but the mass resolution is unable to split the line.) The resulting distribution is shown in Fig. 3.8, and the background subtracted signal is shown in Fig. 3.9.

There are lots of jets in hadronic reactions, but very few line sources, which is one reason for the interest in these results. Another reason is that W 's and Z 's are expected to be signatures of new physics beyond the standard model, and so the detection of W 's with good efficiency through a kinematically constrained channel is important for the future.

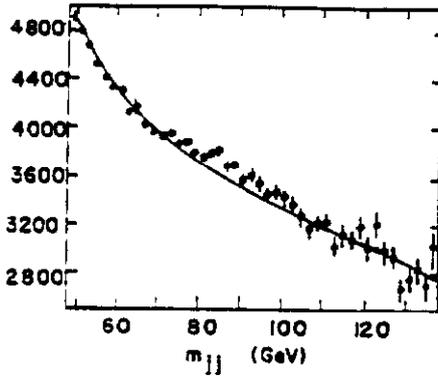


Figure 3.8. UA2 jet-jet mass between 50 and 130 GeV.

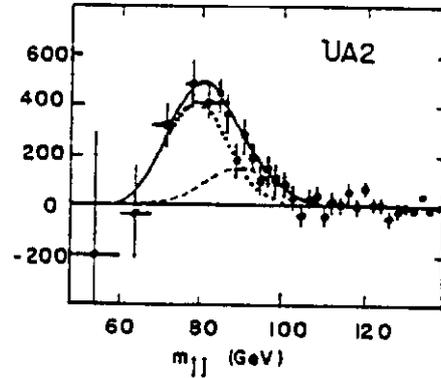


Figure 3.9. Data from 3.7 after background subtraction.

There is one complication in the quark coupling, which we have ignored in Eq. 3.1. Namely, the flavor eigenstates (u, d) and (c, s) are not the states that couple at the weak vertex. This mixing phenomenon was first noted by Cabibbo before the c quark was discovered. Cabibbo introduced a mixing angle θ_c and wrote the weak current (in modern notation) as

$$J^\mu = (\bar{u} \ \bar{c}) \frac{\gamma^\mu(1 - \gamma^5)}{2} U \begin{pmatrix} d \\ s \end{pmatrix} \quad (3.16)$$

where the mixing matrix U is

$$U = \begin{pmatrix} \cos \theta_c & \sin \theta_c \\ -\sin \theta_c & \cos \theta_c \end{pmatrix} \quad (3.17)$$

Experimentally $\sin \theta_c = 0.22$. All of our amplitudes involving W 's coupling to u and d quarks should be multiplied by $\cos \theta_c = 0.975$. This idea has been extended to six quarks and a 3×3 unitary mixing matrix by Kobayashi and Maskawa.

3.4 Weak Interactions

We now turn to the phenomenology of leptonic weak interactions in the pre W/Z era. We want to show how particle decays and neutrino interactions can be described by a point four fermion interaction without an intermediate boson; how the divergence of the neutrino cross section with increasing energy can be avoided by the introduction of W^\pm of unknown (but large) mass; and how the neutral current weak interactions were first observed in neutrino scattering. Then later we will learn that the standard model of Weinberg and Salam is able to predict the masses of the W and Z bosons in terms of the Fermi coupling constant G_F , the fine structure constant α , and the Weinberg angle θ_W obtained from the strength of the weak neutral current. But first we need to review the traditional theory of weak interactions.

The decay rate formula in terms of the invariant matrix element, the companion to the cross section Eq. 1.12 describes the decay $A \rightarrow 1 + 2 + 3 + \dots + n$ is as follows:

$$d\Gamma = \frac{1}{2E_A} |\mathcal{M}|^2 \frac{d^3\vec{p}_1}{(2\pi)^3 2E_1} \frac{d^3\vec{p}_2}{(2\pi)^3 2E_2} \dots \frac{d^3\vec{p}_n}{(2\pi)^3 2E_n} (2\pi)^4 \delta^{(4)}(p_A - p_1 - p_2 - \dots - p_n) \quad (3.18)$$

This is Eq. 4.36 in Reference 3. To describe the matrix element for muon decay via the Feynman rules, define the momenta

$$\begin{aligned} \mu^- &\rightarrow e^- \quad \bar{\nu}_e \quad \nu_\mu \\ p &\rightarrow p' + k + k' \end{aligned} \quad (3.19)$$

and the matrix element

$$\mathcal{M} = \frac{G_F}{\sqrt{2}} [\bar{u}(k)\gamma^\mu(1 - \gamma^5)u(p)] [\bar{u}(p')\gamma_\mu(1 - \gamma^5)v(k')]. \quad (3.20)$$

This can be expressed as the dot product of two four dimensional currents:

$$\mathcal{M} = \frac{4G_F}{\sqrt{2}} J^\mu J_\mu^\dagger. \quad (3.21)$$

In our metric, the parity violating operator $\gamma^5 = i\gamma^0\gamma^1\gamma^2\gamma^3$, and has these properties:

$$\gamma^{5\dagger} = \gamma^5; (\gamma^5)^2 = 1; \gamma^\mu\gamma^5 + \gamma^5\gamma^\mu = 0. \quad (3.22)$$

Left handed and right handed spinors are defined by the following projection operators:

$$\frac{1}{2}(1 - \gamma^5)u \equiv u_L; \quad \frac{1}{2}(1 + \gamma^5)u \equiv u_R. \quad (3.23)$$

Evaluation of Eq. 3.18 for this decay gives the energy distribution of electrons as a function of $\epsilon = E/E_{\max} = 2p'/m_\mu$:

$$\frac{d\Gamma}{d\epsilon} = \frac{G_F^2 m_\mu^5}{96\pi^3} \epsilon^2 (3 - 2\epsilon) \quad (3.24)$$

The integral of this expression from zero to one gives the total decay rate, or the reciprocal of the muon lifetime, as

$$\Gamma(\mu^- \rightarrow e^- \bar{\nu}_e \nu_\mu) = \frac{G_F^2 m_\mu^5}{192\pi^3} = \frac{1}{\tau_\mu} \quad (3.25)$$

This is a very useful formula. It describes the decay of any spin 1/2 fermion into three massless spin 1/2 fermions.

HOMEWORK PROBLEM: Derive the formula Eq. 3.25 by evaluating the $\sum_{\text{spins}} |\mathcal{M}|^2$ and integrating over the phase space. Hint: The \sum_{spins} factors into the contraction of two tensors which have the same form. Drop the final lepton mass terms. Check your intermediate result with Eq. 12.35 in Ref. 3:

$$\frac{1}{2} \sum_{\text{spins}} |\mathcal{M}|^2 = 64 G_F^2 (p \cdot k')(k \cdot p'). \quad (3.26)$$

This equation is handy for working out the W production problem started in Eq. 3.5.

Equation 3.25 defines the Fermi coupling constant G_F . The numbers are tabulated in Ref. 7:

$$\begin{aligned} m_\mu &= 0.105658387 (34) \text{ GeV} \\ 1/\tau_\mu &= (0.455160 (8)) \times 10^6 \text{ sec}^{-1} \\ &= (2.99592 (5)) \times 10^{-19} \text{ GeV} \end{aligned} \quad (3.27)$$

which combine to give $G_F = (1.16380 (2)) \times 10^{-5} \text{ GeV}^{-2}$. This number differs by 2/1000 from the Fermi constant quoted in Ref. 7:

$$G_F = (1.16637 (2)) \times 10^{-5} \text{ GeV}^{-2}. \quad (3.28)$$

The difference arises from two small corrections to the muon decay rate: a) a phase space term of about 2/10000 due to the finite electron mass; and b) radiative corrections which alter the electron energy distribution, and when integrated affect the rate by 4.2/1000, and hence G_F by half as much. The details of these calculations are discussed by Marshak, Riazuddin, and Ryan.¹⁷

Lecture IV

4.1 Neutrino Scattering

In his Nobel address Mel Schwartz¹⁸ described how the idea of attempting an experiment with high energy neutrinos came about. T.D. Lee raised the question at a coffee hour at Columbia: How does one study the weak interaction at high energy? Various ideas involving weak interactions of hadrons and charged leptons were raised and discarded, because of the unfavorable signal to noise. Neutrinos seemed the best choice to Schwartz, since they had only weak interactions. If enough neutrinos could be produced, and if backgrounds from other sources could be suppressed, then the signal in a massive detector would be observable, the signal to noise ratio would be tolerable, and T.D. Lee's question could be answered.

Since the principal sources of high energy neutrinos are pi and K decays, muon neutrinos predominate over electron neutrinos in the beams. A typical beam has a few percent electron neutrinos. The neutrino/antineutrino ratio can be changed by charge selecting the pions and kaons. The basic reactions have the form $\nu_\mu + d \rightarrow \mu^- + u$, or the conjugate $\bar{\nu}_\mu + u \rightarrow \mu^+ + d$. The Feynman diagram, shown in Fig 4.1a, is similar to beta decay. The matrix element for the process is:

$$\mathcal{M} = \frac{G_F}{\sqrt{2}} \left(\bar{u}(k') \gamma^\mu (1 - \gamma^5) u(p) \right) \left(\bar{u}(p') \gamma_\mu (1 - \gamma^5) u(k) \right). \quad (4.1)$$

The sum over spins leads to an isotropic subprocess:

$$\frac{1}{2} \sum_{\text{spins}} |\mathcal{M}|^2 = 64 G_F^2 (k \cdot p)(k' \cdot p') = 16 G_F^2 \hat{s}^2 \quad (4.2)$$

Note the similarity to Eq. 3.26. The factor of 1/2 is for the quark spin—the ν_μ has only one spin state. Then the cross section in the (ν_μ, d) rest frame according to Eq. 1.12 is

$$\frac{d\hat{\sigma}(\nu_\mu d \rightarrow \mu^- u)}{d\Omega} = \frac{1}{4\pi^2} G_F^2 \hat{s} \quad \text{and} \quad \hat{\sigma} = \frac{1}{\pi} G_F^2 \hat{s}. \quad (4.3)$$

The conjugate reaction can be readily calculated by substituting \hat{t} for \hat{s} in Eq. 4.2, which results in an angular dependence

$$\frac{d\hat{\sigma}}{d\Omega} = \frac{G_F^2}{16\pi^2} \hat{s} (1 + \cos \theta)^2. \quad (4.4)$$

This can be understood by the helicity rules, which forbid the $(\bar{\nu}_\mu, u)$ reaction at $\theta = \pi$, while (ν_μ, d) has no problem reversing both fermion momenta. The total cross section is

$$\hat{\sigma}(\bar{\nu}_\mu u \rightarrow \mu^+ d) = \frac{G_F^2 \hat{s}}{3\pi}. \quad (4.5)$$

The subprocess $\hat{s} = xs$, where x is the momentum fraction carried by the struck quark. If E_ν is the incident neutrino energy in the laboratory, and the target nucleon (N) is at rest, then $s = 2M_p E_\nu$. The (ν_μ, N) cross section depends of course on integrals over the structure functions, but the result is proportional to the incident neutrino energy. Experimentally

$$\begin{aligned} \sigma(\nu_\mu + N \rightarrow \mu^- + X) &= 0.67 \times 10^{-38} \text{ cm}^2 E_\nu \\ \text{and } \sigma(\bar{\nu}_\mu + N \rightarrow \mu^+ + X) &= 0.34 \times 10^{-38} \text{ cm}^2 E_\nu, \end{aligned} \quad (4.6)$$

where E_ν is in GeV.

The fact that the ratio of these numbers is two rather than three is a reflection of the antiquark content of the nucleon. If the nucleon were made of three valence quarks, the ratio predicted by Eqs. 4.3 and 4.5 would obtain. See Section 12.8 of reference 3 for more details.

The analogous purely leptonic reaction $\nu_\mu + e^- \rightarrow \nu_e + \mu^-$ has a very small cross section. Eq. 4.3 is correct, with $\hat{s} = s = 2m_e E_\nu$ at high energy, which gives the cross section

$$\sigma(\nu_\mu e^- \rightarrow \nu_e \mu^-) = 1.7 \times 10^{-41} \text{ cm}^2 E_\nu. \quad (4.7)$$

This is .0025 times the cross section of Eq. 4.6. The enhancement of the nucleon cross section is an example of quark binding and the structure functions. The masses of a u quark and an electron are not that much different, and yet the quark cross section is 400 times larger, because the quarks live inside a heavy object, so that M_p rather than m_q appears in the formula.

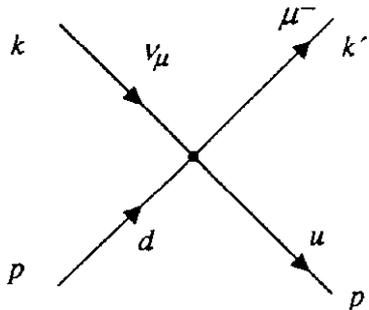


Figure 4.1a. Feynman diagram for $\nu_\mu + d \rightarrow \mu^- + u$.

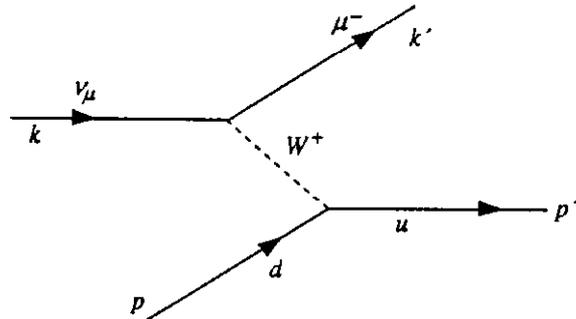


Figure 4.1b. Same with W^+t channel exchange.

A cross section which increases linearly with E_ν cannot continue indefinitely, so there must be something wrong with the four fermion point interaction. (No roll over has been observed for neutrino energies up to 200 GeV.) It has been recognized for a long time that the introduction of a massive intermediate state as shown in Fig. 4.1b would give a cross section which becomes a constant as $s \rightarrow \infty$. Now we began these lectures with massive intermediate bosons, putting the cart before the horse from a historical standpoint, so the exchanged object in Fig. 4.1b is not big news. The vector boson propagator and the dimensionless vertex factors for charged

weak currents were introduced in Fig. 3.4, and used in Eqs. 3.5 ff. To summarize the structure, the vertex factor is

$$\frac{-ig}{\sqrt{2}}\gamma^\mu\frac{1-\gamma^5}{2} \quad (4.8)$$

which corresponds to $ie\gamma^\mu$ in QED. The matrix element for the neutrino reaction given by the diagram in Fig. 4.1b is then

$$-i\mathcal{M} = \left(\bar{u}(k') \left(\frac{-ig\gamma^\mu(1-\gamma^5)}{\sqrt{2}}\frac{1-\gamma^5}{2}u(p) \right) W_{\mu\nu} \left(\bar{u}(p') \frac{-ig\gamma^\nu(1-\gamma^5)}{\sqrt{2}}\frac{1-\gamma^5}{2}u(k) \right) \right) \quad (4.9)$$

where the W propagator is

$$W_{\mu\nu} = \frac{-i(g_{\mu\nu} + q_\mu q_\nu/M_W^2)}{q^2 - M_W^2}. \quad (4.10)$$

We know from Eq. 3.6 that, given G_F , g and M_W are not independently specified. The standard model of Glashow, Weinberg, and Salam¹⁹ relates g to the electric charge e and the Weinberg angle θ_W . This model, by combining electromagnetic and weak forces in the same scheme, requires the introduction of weak neutral currents.

4.2 Weak Neutral Currents

The discovery of weak neutral currents makes an interesting story, which has been reviewed by Galison.²⁰ A brief description of the early history, and the influence of what is now called the Standard Model on experimenters is given also by Steinberger.²¹ Much of our knowledge of weak interactions comes from the study of decays of particles, but nature conspires in these decays to hide the weak neutral currents from view. For example, a neutral current decay like $\Sigma^0 \rightarrow \Lambda\nu\bar{\nu}$ ($u \rightarrow u\nu\bar{\nu}$ at the quark level) is energetically possible and violates no selection rules, but is swamped by the electromagnetic channel $\Sigma^0 \rightarrow \Lambda\gamma$. Only flavor changing neutral currents like $K^+ \rightarrow \pi^+\nu\bar{\nu}$, which are forbidden by electromagnetic selection rules, are not overwhelmed. But alas flavor changing neutral currents are forbidden in first order by the GIM mechanism.²² So far no first order neutral current weak decays have been experimentally observed.

Weak neutral currents were discovered in neutrino reactions. Examples include elastic scattering like $\nu_\mu + p \rightarrow \nu_\mu + p$, and $\nu_\mu + e^- \rightarrow \nu_\mu + e^-$, or inelastic scattering like $\nu_\mu + p \rightarrow \nu_\mu + X$. These have all been observed and extensively studied. The cross section is not that small, being about 20% of the total neutrino cross section, *i.e.*, $\frac{(\nu_\mu \rightarrow \nu_\mu)}{(\nu_\mu \rightarrow \mu^-) + (\nu_\mu \rightarrow \nu_\mu)} = 0.2$. One might ask why it took over ten years of experimental work to find these effects? This question is pursued in Ref. 20. In a nutshell, it was a combination of two factors:

- a) Until the early 1970's, when the renormalization of the new theory had been proven, there was not a strong conviction that the weak neutral currents should really be there, and that the experimenters should search for them with vigor; and
- b) Since neither the initial nor the final lepton can be observed, the signature of a neutral current reaction is a small amount of recoil energy, which can be lost in neutron background. Recent neutrino experiments have fine grained targets to observe the recoil energy, and employ timing and spatial distributions of events to discriminate against neutron background.

The neutral current vertex factor is written in terms of the Weinberg angle and the coupling constant g introduced in Fig. 3.4 and in Eq. 4.8:

$$\frac{-ig}{\cos \theta_W} \gamma^\mu \frac{(V_f - a_f \gamma^5)}{4} \quad (4.11)$$

where V_f and a_f were introduced in Eqs (1.22) and (1.23). The neutral current reaction

$$\begin{aligned} \nu_\mu + e^- &\rightarrow \nu_\mu + e^- \\ k + p &\rightarrow k' + p' \end{aligned}$$

can be described by extending the Feynman rules to cover Z exchange, and writing the matrix element analog to Eq. 4.9:

$$-i\mathcal{M} = \frac{-g^2}{8 \cos^2 \theta_W} \left(\bar{u}(k') \gamma^\mu (1 - \gamma^5) u(k) \right) Z_{\mu\nu} \left(\bar{u}(p') \gamma^\mu \frac{(V_e - a_e \gamma^5)}{2} u(p) \right) \quad (4.12)$$

where the Z propagator $Z_{\mu\nu}$ is

$$Z_{\mu\nu} = \frac{-i(g_{\mu\nu} + q_\mu q_\nu / M_Z^2)}{q^2 - M_Z^2}. \quad (4.13)$$

The matrix element can be simplified by approximating $Z_{\mu\nu} \rightarrow ig_{\mu\nu}/M_Z^2$ in the low energy limit, by substituting $\cos^2 \theta_W = (M_W^2/M_Z^2)$, a relation obtained in the standard model, and using Eq. 3.6 to eliminate M_W . The result is

$$\mathcal{M} = \frac{G_F}{\sqrt{2}} \left(\bar{u}(k') \gamma_\mu (1 - \gamma^5) u(k) \right) \left(\bar{u}(p') \gamma^\mu \frac{(V_f - a_f \gamma^5)}{2} u(p) \right) \quad (4.14)$$

Performing the average $\frac{1}{2} \sum_{\text{spins}} |\mathcal{M}|^2$ and substituting into the cross section formula Eq. 1.12 leads to the total cross section for (ν_μ, e) scattering:

$$\sigma(\nu_\mu e^- \rightarrow \nu_\mu e^-) = \frac{G_F^2 m_e E_\nu}{2\pi} \left(\frac{a_e^2 + a_e V_e + V^2}{3} \right). \quad (4.15)$$

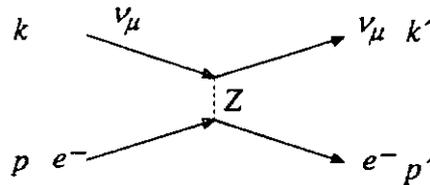
This equation involves the Weinberg angle in the vector coupling. The ratio of this cross section to its companion charged current cross section at the same neutrino energy is

$$\frac{\sigma(\nu_\mu e^- \rightarrow \nu_\mu e^-)}{\sigma(\nu_\mu e^- \rightarrow \mu^- \nu_e)} = \frac{4}{3}(1 - 4 \sin^2 \theta_W + \frac{16}{3} \sin^4 \theta_W). \quad (4.16)$$

A measurement of this ratio gives the Weinberg angle. Although these formulas avoid the complications from the structure functions which clutter up the companion formulas for $\nu_\mu + N \rightarrow \nu_\mu + X$, the nucleon reactions are much more probable, as discussed above. The systematic uncertainties inherent in the parton model are smaller than the statistical errors in the purely leptonic channel, so that the best current values for the Weinberg angle from neutrino scattering come from the ratio of neutral current to charged current neutrino nucleon scattering.

A global analysis of these data gave the result $\sin^2 \theta_W = 0.233 \pm 0.006(23)$.

HOMEWORK PROBLEM: Derive Eq. 4.15 from the Feynman rules for the neutral current t channel exchange reaction



Hint: Read the text above, and follow the steps outlined therein.

4.3 The Standard Model

We have now assembled the following ingredients of the standard model:

- a) The charged current weak interaction which is responsible for beta decay is described by a vertex with two fermions and a massive charged intermediate boson rather than by the four fermion vertex of the Fermi theory. The ratio of the coupling constant to the W mass is known from muon decay (Eq. 3.6), but the W mass is arbitrary, although large compared to experimentally observed momentum transfers.
- b) There are flavor conserving weak neutral currents. A third weak intermediate boson, the Z , is postulated to couple neutral leptonic currents. The Z coupling diagram resembles that of the photon in QED, except that the Z is presumed also to be very massive. Since the cross sections for charged current and neutral current interactions are comparable, the W and Z masses are comparable as well. A real mixing parameter, the Weinberg angle, was introduced in Eq. 4.11

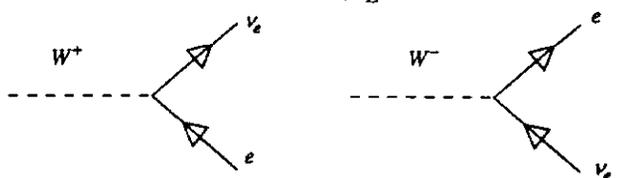
in the neutral current coupling. An important consequence of the standard model is that the ratio $(M_W/M_Z) = \cos \theta_W$, so that the Z must be heavier than the W .

The standard model was constructed to generate weak and electromagnetic interactions from a common source. The resulting structure relates the coupling constants, and gives predictions for the W and Z masses in terms of the Weinberg angle. The theory begins with four massless vector fields, two charged and two neutral. Three of these fields transform like a vector under $SU(2)$, which is a symmetry group of the Lagrangian, and the fourth field is an $SU(2)$ scalar. The left handed fermions—leptons and quarks—are $SU(2)$ doublets. This group is called weak isospin. The right handed fermions, which do not couple to the charged intermediate bosons, are assigned weak isospin zero. All of the fermions are initially massless. The weak isospin and the electric charge are related to a number Y , the weak hypercharge, by a formula identical to the Gell-Mann Nishijima relation of strong isospin:

$$Q = T_3 + \frac{1}{2}Y \quad (4.17)$$

The fourth IVB field is coupled to the weak isospin Y . The relationship between Q and Y is arbitrary, but works conveniently.

Following Chapter 13 of Ref. 3, we label the spinors of the leptons by their particle symbol, write the left handed states (defined by Eq. 3.23) in a weak isospin doublet:

$$\chi_L = \begin{pmatrix} \nu \\ e \end{pmatrix}_L \quad (4.18)$$


The currents shown in Fig. 4.2 are then written

$$J_\mu = \bar{\chi}_L \gamma_\mu \tau_+ \chi_L \text{ and } J_\mu^+ = \bar{\chi}_L \gamma_\mu \tau_- \chi_L \quad (4.19)$$

where the Pauli matrices $\tau_\pm = \frac{1}{2}(\tau_x \pm \tau_y)$ have been introduced to interchange the spinors $e \leftrightarrow \nu$ as required. The corresponding neutral current is

$$J_\mu^{(3)} = \bar{\chi}_L \gamma_\mu \frac{1}{2} \tau_3 \chi_L, \quad (4.20)$$

so we may write

$$J_\mu^{(i)} = \bar{\chi}_L \gamma_\mu \frac{1}{2} \tau^{(i)} \chi_L \quad (4.21)$$

where $J_\mu^+ = J_\mu^{(1)} + iJ_\mu^{(2)}$. The electromagnetic current (in units of the elementary charge) involves both left and right handed parts:

$$J_\mu^{em} = -\bar{e} \gamma_\mu e = -\bar{e}_R \gamma_\mu e_R - \bar{e}_L \gamma_\mu e_L \quad (4.22)$$

A weak hypercharge current can be defined in terms of Eq. 4.20 and 4.22 with the aid of the charge rule 4.17:

$$J_\mu^{em} = J_\mu^{(3)} + \frac{1}{2} J_\mu^Y \quad (4.23)$$

We now introduce the four massless vector fields, three coupled to the weak isospin current, called $W_\mu^{(i)}$, and one coupled to the weak hypercharge, B_μ . To do this two dimensionless coupling constants g and g' are inserted to give the basic electroweak interaction:

$$-igJ^{(i)\mu}W_\mu^{(i)} - \frac{ig'}{2} (J^Y)^\mu B_\mu \quad (4.24)$$

This equation involves charged current terms for $i = 1, 2$, coupled via g , the same constant introduced in Fig. 3.4, and two neutral current terms $-igJ^{(3)\mu}W_\mu^{(3)} - \frac{ig'}{2} (J^Y)^\mu B_\mu$.

So far so good, but we have a small problem since nature does not seem to need four massless electroweak vector fields. One of them, the photon field, is massless, but the other three are very heavy. The eigenstates of the coupling in Eq. 4.24 cannot be the mass eigenstates. The mass is generated by machinery known as spontaneous symmetry breaking and the Higgs mechanism, which will be sketched in Lecture V. For the moment, let us accept the fact that it is possible to transform the two massless neutral fields into one which remains massless:

$$A_\mu = B_\mu \cos \theta_W + W_\mu^{(3)} \sin \theta_W; \quad (4.25)$$

and one which becomes massive:

$$Z_\mu = W_\mu^{(3)} \cos \theta_W - B_\mu \sin \theta_W. \quad (4.26)$$

While this is going on, the two charged fields $W^\pm = (W^{(1)} \pm iW^{(2)})/\sqrt{2}$ also acquire a mass. The mixing angle θ_W in Eqs. 4.25 and 4.26 is the Weinberg angle.

Equations 4.25 and 4.26 can be inverted and substituted into Eq. 4.24 to give the neutral current expression:

$$\begin{aligned} -igJ^{(3)\mu}W_\mu^{(3)} - \frac{ig'}{2} (J^Y)^\mu B_\mu = & -i \left\{ \left(gJ^{(3)\mu} \sin \theta_W + \frac{g'}{2} (J^Y)^\mu \cos \theta_W \right) A_\mu \right. \\ & \left. + \left(gJ^{(3)\mu} \cos \theta_W - \frac{g'}{2} (J^Y)^\mu \sin \theta_W \right) Z_\mu \right\} \quad (4.27) \end{aligned}$$

The first term on the right hand side is the electromagnetic interaction. To get the correct coupling we must identify

$$g \sin \theta_W = g' \cos \theta_W = e \quad (4.28)$$

This identification can be combined with Eq. 3.6 to give a formula for the W mass in terms of the Fermi coupling constant G_F , the fine structure constant $\alpha = e^2/4\pi$, and the Weinberg angle θ_W :

$$M_W^2 = \frac{\pi\alpha}{\sqrt{2}G_F \sin^2 \theta_W} \quad (4.29)$$

This is historically an important formula, since as we have seen the neutrino weak neutral current experiments supplied a number for the Weinberg angle, and of course G_F and α were also known, so that a prediction could be made of the W mass before its discovery at the CERN SPS collider.²⁴ Substituting $\sin^2 \theta_W = 0.233$ into Eq. 4.29 and taking the square root gives $M_W = 77.3$ GeV, which is about 3 GeV below the current best experimental value for the W mass. This difference is explained by radiative corrections to Eq. 4.29, which will be discussed—but not derived—in Lecture V. Equation 4.29 clearly served as an invaluable guide to the experimenters, telling them where to look. Before this result was obtained, it was only known that the W had to be massive compared to experimentally studied weak interaction energies, which left a lot of room to search.

The companion Z coupling term can be derived by substituting Eq. 4.28 back into Eq. 4.27 and retaining only the part proportional to Z_μ :

$$-i \left(gJ^{(3)\mu} \cos \theta_W - \frac{g'}{2} (J^Y)^\mu \sin \theta_W \right) Z_\mu = \frac{-ie}{\sin \theta_W \cos \theta_W} \left(J^{(3)\mu} - \sin^2 \theta_W J^{em\mu} \right) Z_\mu \quad (4.30)$$

It is a straightforward task to obtain the vector and axial vector terms of Eq. 2.18 for $Z \rightarrow e^+e^-$ from this expression. For the electron with $Q = -1$ and $T_3 = -\frac{1}{2}$ we have

$$\begin{aligned} & \frac{-ie}{\sin \theta_W \cos \theta_W} \left[\left(-\frac{1}{2} + \sin^2 \theta_W\right) \bar{e}_L \gamma_\mu e_L + \sin^2 \theta_W \bar{e}_R \gamma_\mu e_R \right] Z^\mu \\ &= \frac{-ie}{\sin \theta_W \cos \theta_W} \left[\left(-\frac{1}{4} + \sin^2 \theta_W\right) \bar{e} \gamma_\mu e + \frac{1}{4} \bar{e} \gamma_\mu \gamma^5 e \right] Z^\mu \\ &= \frac{-ie}{\sin \theta_W \cos \theta_W} \left[\frac{V_f \bar{e} \gamma_\mu e - a_f \bar{e} \gamma_\mu \gamma^5 e}{4} \right] Z^\mu. \end{aligned} \quad (4.31)$$

In the last line the substitutions $V_f = 2T_{3f} - 4Q_f \sin^2 \theta_W$ and $a_f = -2T_{3f}$ have been made, in agreement with the conventions adopted in Eq. 1.23.

By accounting for the e^+e^- annihilation formula of Eq. 1.18, we have come full circle, and are now back to the beginning of the lecture series. There are some missing pieces, however. We have not described what takes place in the theory to generate the heavy IVB masses, nor have we discussed the present experimental situation. These topics will be sketched in Lecture V.

HOMEWORK PROBLEM: Calculate the partial width for the decay of the Z into two neutrinos:

$$(Z \rightarrow \nu_e \bar{\nu}_e) = \frac{\alpha M_Z}{24 \sin^2 \theta_W \cos^2 \theta_W}$$

Hint: Replace Z_μ by its polarization vector ϵ_μ , which can be taken as the four vector $\epsilon_\mu = \{0, 0, 0, 1\}$. The decay rate is independent of the polarization direction, so the calculation for one choice equals the average value.

Lecture V

5.1 Symmetries and Lagrangians

In this lecture we will sketch the origin of the masses of the W and Z intermediate vector bosons according to the Electroweak Theory of Weinberg and Salam. We learned in the previous lecture how the weak current could be constructed with two charged fields ($W_\mu^{(+)}, W_\mu^{(-)}$), initially massless, which acquire equal masses via spontaneous symmetry breaking; and two neutral fields ($W_\mu^{(3)}, B_\mu$), also initially massless, which mix to form one massive vector field, Z_μ , and one massless field, the photon A_μ . There also is a neutral scalar field of unknown mass in the theory, called the Higgs field. The Higgs particle, so far unobserved experimentally, plays no direct role in the first order calculations of the theory of quarks and leptons, but is of critical importance for the spontaneous symmetry breaking mechanism to work. Other massless particles, called Goldstone bosons, appear and then conveniently disappear along the way.

Broken symmetries are a familiar theme in particle physics. Strong Isospin was proposed as an $SU(2)$ symmetry of the nuclear force (charge independence), which was assumed to be valid when the electromagnetic interaction was ignored, but explicitly broken by the electric charge. Invariance under space inversion (parity) was assumed to be valid for both the strong and electromagnetic interactions, but violated by the weak interactions. In this way one could imagine a hierarchy of interaction terms, the strongest being the most symmetric. As the strength of the force decreases, the number of valid symmetry operations decreases also.

There is another way to break symmetries, however, called spontaneous symmetry breaking. To understand this phenomenon it is desirable to use the language of Lagrangian field theory. One can do without it, but something is lost in translation, as with Pushkin's poetry. This may be unfortunate, because, although most readers are conversant with the Feynman rules and other apparatus used in these lectures so far, they may not be fully at ease in Lagrangian field theory land. Mindful of this possibility, and not too comfortable in this territory ourselves, we will tread carefully. The discussion presented here, which is adopted principally from Ref. 4, is essentially classical in nature. There are no field operators or commutation rules, nor, indeed, are any needed to describe how things work.

Lagrangians are useful in relativistic field theories because, unlike Hamiltonians, they are relativistically invariant. The symmetries— $SU(2)$, $SU(3)$, CP , etc.—are symmetries of the Lagrangian density, which is invariant under the appropriate transformations. For a scalar field $\phi(x)$ the Lagrangian is defined as the space integral of the Lagrangian density:

$$L = \int d^3x \mathcal{L}(\phi(x), \partial_\mu \phi(x)). \quad (5.1)$$

The Lagrangian density is a function of the field $\phi(x)$ (the potential energy), and its

spacetime derivatives $\partial_\mu\phi(x)$ (the kinetic energy). The differential equation for the field can be derived from the Lagrangian density via the Euler-Lagrange equations, which follow from a minimum principle:

$$\partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu\phi)} - \frac{\partial \mathcal{L}}{\partial\phi} = 0. \quad (5.2)$$

Because of the utility of this formalism for subsequent discussions, it is useful to review what the Lagrangians (it is convenient to drop the “density”) look like for more familiar particle/field wave equations. Some simple examples of free field Lagrangians are as follows:

$$1.) \text{ Scalar field } \mathcal{L} = \frac{1}{2} [(\partial^\mu\phi)(\partial_\mu\phi) - m^2\phi^2] \quad (5.3)$$

Substitution into Eq. (5.2) gives the Klein Gordon equation

$$\square\phi + m^2\phi = 0 \quad (5.4)$$

where

$$\square = \partial^\mu\partial_\mu = \frac{\partial^2}{\partial x_0^2} - \nabla^2. \quad (5.5)$$

This relativistic wave equation was first written down by Schroedinger, but he already had another equation in his name.

$$2.) \text{ Dirac field } \mathcal{L} = \bar{\psi}(x)(i\gamma^\mu\partial_\mu - m)\psi(x). \quad (5.6)$$

Substitution into Eq. (5.2) gives the adjoint Dirac equation

$$i\partial_\mu\bar{\psi}(x)\gamma^\mu + m\bar{\psi}(x) = 0. \quad (5.7)$$

$$3.) \text{ Electromagnetic field } \mathcal{L} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} \quad (5.8)$$

where the antisymmetric field tensor which gives the electric and magnetic fields is defined in terms of the four vector potential $A^\mu = \{V, \vec{A}\}$ as

$$F^{\mu\nu} = \partial^\mu A^\nu - \partial^\nu A^\mu. \quad (5.9)$$

In this case the Euler-Lagrange equations give the wave equation

$$\partial^\nu\partial_\nu A_\mu - \partial_\mu\partial^\nu A_\nu = 0, \quad (5.10)$$

which has the familiar form $\square A_\mu = 0$, provided

$$\partial^\nu A_\nu = 0. \quad (5.11)$$

This is called a gauge condition. That the gradient of a scalar function can always be added to the vector potential in order to satisfy Eq. 5.11, and that the field tensor Eq. 5.9 is unchanged by such a transformation, is called gauge invariance.

$$4.) \text{ Massive vector field } \mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \frac{1}{2}M^2 A^\mu A_\mu. \quad (5.12)$$

where $F^{\mu\nu}$ is defined by Eq. 5.9. In this instance the Euler-Lagrange equations give the wave equation

$$\partial_\nu(\partial^\nu A^\mu - \partial^\mu A^\nu) + M^2 A^\mu = 0. \quad (5.13)$$

Here the presence of the M^2 term guarantees that $\partial_\mu A^\mu = 0$ (take the derivative of (5.13)), so there is no longer the freedom to choose the gauge condition. Equation 5.13 can then be written in the form

$$\square A_\mu + M^2 A_\mu = 0. \quad (5.14)$$

This is called Proca's equation, which is not gauge invariant.

HOMEWORK PROBLEM: Verify the forms given for the free particle spin zero, spin 1/2, massless spin 1, and massive spin 1 Lagrangian densities.

Gauge invariance plays a central role in the EWK theory. A local gauge transformation of the electromagnetic vector potential has the form

$$A^\mu \rightarrow A^{\mu'} = A^\mu + \partial^\mu \alpha(x), \quad (5.15)$$

where $\alpha(x)$ is an arbitrary, suitably behaved function of the spacetime coordinates. Since the field tensor $F_{\mu\nu}$ is antisymmetric in μ and ν , it is invariant under the transformation Eq. 5.15. If in addition $\square \alpha(x) = 0$, then if A^μ satisfies $\partial_\mu A^\mu = 0$, $\partial_\mu A^{\mu'} = 0$ also.

The interaction of an electron with an electromagnetic field is incorporated in the Dirac equation by substituting $p^\mu \rightarrow p^\mu - qA^\mu$, where q is the electron charge. This gives the equation

$$\gamma^\mu(i\partial_\mu - qA_\mu)\psi(x) - m\psi(x) = 0. \quad (5.16)$$

This equation is invariant under the gauge transformation Eq. 5.15 provided that the wave function is simultaneously altered by a spacetime dependent phase:

$$\psi(x) \rightarrow \psi'(x) = e^{-iq\alpha(x)}\psi(x). \quad (5.17)$$

This works because the two $\partial_\mu \alpha(x)$ terms which appear in the Dirac equation cancel, leaving $e^{-iq\alpha(x)}$ as an overall phase factor. Thus if $\gamma^\mu(i\partial_\mu - qA_\mu)\psi - m\psi = 0$, and $A'_\mu = A_\mu + \partial_\mu \alpha$, while $\psi' = e^{-iq\alpha}\psi$, then $\gamma^\mu(i\partial_\mu - qA'_\mu)\psi' - m\psi' = 0$. Note that the free particle Dirac equation, without the qA_μ term, is not invariant under the phase transformation Eq. 5.17. The interaction term has to be there. One can imagine using the argument backwards. That is to say, the requirement of spacetime

dependent phase invariance can generate an interaction with the Dirac particle. The interaction so generated is a vector field, because of the vector nature of the gauge transformation. It is also a massless field, since the M^2 term in Eq. (5.14) destroys its invariance under the transformation Eq. (5.15).

Gauge invariance and the Dirac equation can be summarized by the definition of the covariant derivative

$$D_\mu \equiv \partial_\mu + iqA_\mu. \quad (5.18)$$

The Dirac equation then reads $(i\gamma^\mu D_\mu - m)\psi = 0$, and $D'_\mu \psi' = e^{-iq\alpha} D_\mu \psi$ as we have shown above, so that the phase function factors out of the equation, leaving $i\gamma^\mu D'_\mu \psi' - m\psi' = 0$ also. The electric current $j^\mu = q\bar{\psi}\gamma^\mu\psi$ is obviously invariant under the gauge transformation.

HOMEWORK PROBLEM: Show that the Klein Gordon current $j^\mu = iq(\phi^*(\partial^\mu\phi - (\partial^\mu\phi)^*\phi)$ is invariant under $\phi \rightarrow \phi' = e^{-iq\alpha(x)}\phi$ provided that D^μ replaces ∂^μ , and A^μ is also transformed.

This group of gauge transformations is commutative, or Abelian, in the sense that the result of two phases $\alpha(x)$ and $\beta(x)$ is independent of the order. The group is called $U(1)$, and as we have seen guarantees that the photon is massless.

It is perhaps not immediately obvious how this helps us construct a theory of electroweak interactions, since at least one symmetry group is weak isospin, which is not Abelian, and since at least some of the vector fields are not massless. We will attack the massive field problem first. The non Abelian nature of the symmetry group is not really a fundamental problem.

5.2 Spontaneous Symmetry Breaking

“Spontaneous” is the name given to an alternate method of symmetry breaking. The first method is the hierarchical one described previously, where terms in the Lagrangian become less symmetric as their strength decreases. In spontaneous symmetry breaking the Lagrangian retains its symmetry. The symmetry instead is broken by the states of the system, beginning with the lowest energy state—called the vacuum.

The most oft cited example of this phenomenon in classical physics is ferromagnetism, where at the domain level a favored direction in space is chosen, even though the Lagrangian is rotationally invariant. Applying pressure to a vertical knitting needle also works. The needle, which is cylindrically symmetrical and subjected to forces which also possess this symmetry, pokes out in one direction under the pressure. The ammonia molecule is an example of two states of opposite parity—one with the nitrogen on top of the three hydrogens in the tetrahedron, and one with it on the bottom. Transitions from one configuration to the other by tunneling are strongly suppressed. The ammonia molecule effective one dimensional potential resembles the Higgs potentials discussed below.

Following Chapter 5 of Ref. 4, we consider the application of this idea to a Lagrangian first studied by Goldstone:

$$\mathcal{L} = (D^\mu \phi)^* D_\mu \phi - \mu^2 \phi^* \phi - |\lambda|(\phi^* \phi)^2 - \frac{1}{4} F_{\mu\nu} F^{\mu\nu} \quad (5.19)$$

The charged scalar ϕ is a two component complex field

$$\phi(x) = \frac{\phi_1 + i\phi_2}{\sqrt{2}}, \quad \phi^* = \frac{\phi_1 - i\phi_2}{\sqrt{2}}. \quad (5.20)$$

The covariant derivative includes the interaction term between the charged field ϕ and the electromagnetic field

$$D_\mu = \partial_\mu + iqA_\mu \quad (5.21)$$

The second and third terms, quadratic and quartic in $|\phi|$, form $-V$, the potential energy, while the last term is the free photon field. The Lagrangian is invariant under rotations about the axis perpendicular to $(\phi_1, i\phi_2)$, or to transformations of the form $\phi \rightarrow e^{i\omega} \phi$, for constant ω .

HOMEWORK PROBLEM: Use the Euler-Lagrange equations to show that Eq. 5.19 gives the correct equation for a scalar particle in an electromagnetic field if $|\lambda| = 0$.

$$\left[(\partial_\mu + iqA_\mu) (\partial^\mu + iqA^\mu) + \mu^2 \right] \phi = 0.$$

For a certain range of choices of the real parameters μ^2 and $|\lambda|$, it is possible to pick a ground state which has a favored orientation in $(\phi_1, i\phi_2)$ space. Small oscillations about this vacuum state will be described by the same Lagrangian, but it will be transformed in such a way that the “electromagnetic” vector field acquires a mass. The mass term will be proportional to the ground state offset—i.e., to the symmetry breaking. Of the two components of the motion, the radial term remains massive—the Higgs; the azimuthal term is massless, and is made to vanish by a gauge transformation—the Goldstone boson. This toy model therefore has many of the important ingredients of the EWK theory. It is not quite complicated enough to handle four vector fields and give masses to three of them, and it has no fermions, but it is interesting nonetheless.

In the spirit of a classical Lagrangian, we identify the potential $V(\phi) = \mu^2 \phi^* \phi + |\lambda|(\phi^* \phi)^2$. If $\mu^2 > 0$, the minimum of V is at $\phi = 0$. The potential is a smooth bowl, and nothing interesting happens, since the lowest state is at the origin. On the other hand, if $\mu^2 < 0$, then $\frac{\partial V}{\partial \phi} = 0$ at $\phi = 0$ and at $|\phi|^2 = -\mu^2/2|\lambda|$. In this case the potential is shaped like the bottom of a wine bottle, as shown in Fig. 5.1. The Lagrangian is cylindrically symmetrical. The lowest value of $V = -\mu^4/4|\lambda|$ at $|\phi|^2 = -\mu^2/2|\lambda|$. The complex field $\langle \phi \rangle$ has the same ground state energy for any location on this circle, and is therefore degenerate. If ϕ_1 and ϕ_2 were space coordinates, and this were a quantum mechanics problem, then the ground state would indeed be cylindrically symmetrical; the uncertainty principle would forbid

localization of the wave function in some region of the circle. But the ϕ 's are fields, so perhaps we can select a state which breaks the symmetry.

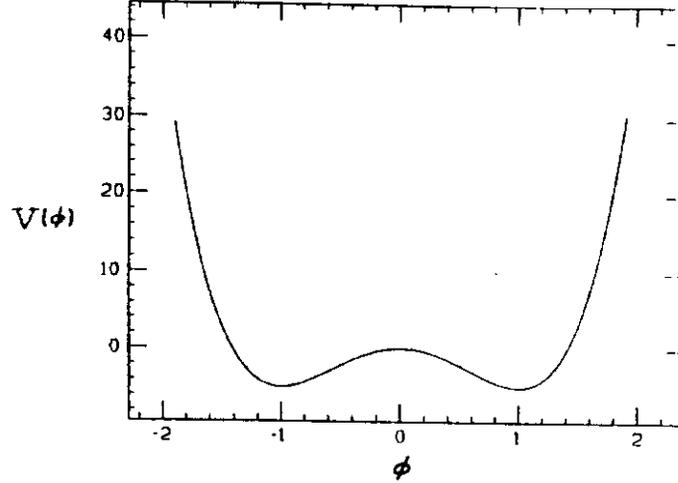


Figure 5.1. $V(\phi) = -10|\phi|^2 + 5|\phi|^4$.

Thus suppose that the ground state (the vacuum) picks a particular direction in $(\phi_1, i\phi_2)$ space—the real axis for example: $\langle\phi\rangle_0 = v/\sqrt{2}$, where $v^2/2 = \frac{-\mu^2}{2|\lambda|}$. To consider small oscillations about this equilibrium point, we introduce the shifted field

$$\phi' = \phi - \langle\phi_0\rangle = \frac{\eta + i\zeta}{\sqrt{2}} \quad (5.22)$$

or

$$\phi = \frac{\eta + i\zeta + v}{\sqrt{2}} \quad (5.23)$$

The radial amplitude is η , while the azimuthal one is ζ . Substitution of (5.23) into (5.19) gives many terms. In the small oscillation approximation, η, ζ and A_μ are all small, so we keep only terms up to second order in these fields. Constant terms can be dropped. Some key cancellations occur because of the relation $v^2 = -\mu^2/|\lambda|$. The Lagrangian in this approximation is

$$\begin{aligned} \mathcal{L} = & \frac{1}{2}(\partial^\mu\eta\partial_\mu\eta + 2\mu^2\eta^2) + \frac{1}{2}(\partial^\mu\zeta\partial_\mu\zeta) \\ & + qA^\mu\partial_\mu\zeta v + \frac{q^2}{2}A^\mu A_\mu v^2 - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} \end{aligned} \quad (5.24)$$

We see that the mass term $2\mu^2$ is retained by the radial amplitude η . This is called the Higgs field. The vector field A_μ has acquired a mass $M^2 = \frac{q^2 v^2}{2}$, while the ζ field is the massless Goldstone boson, and couples to the vector potential via $qA^\mu\partial_\mu\zeta v$. The A^μ parts of the covariant derivatives do not appear because of the small oscillations approximation, so the gauge invariance of the complete Lagrangian is no longer manifest.

The terms in Eq. (5.24) involving ζ and A_μ can be factored as follows:

$$\frac{q^2 v^2}{2} \left(\frac{\partial^\mu \zeta \partial_\mu \zeta}{q^2 v^2} + \frac{2(\partial_\mu \zeta) A^\mu}{qv} + A_\mu A^\mu \right) = \frac{q^2 v^2}{2} \left(\frac{\partial^\mu \zeta}{qv} + A^\mu \right) \left(\frac{\partial_\mu \zeta}{qv} + A_\mu \right). \quad (5.25)$$

Therefore, if a gauge transformation is made

$$A'_\mu = \frac{1}{qv} \partial_\mu \zeta + A_\mu, \quad (5.26)$$

$F_{\mu\nu}$ remains the same, and the Lagrangian becomes

$$\mathcal{L} = \frac{1}{2} (\partial^\mu \eta \partial_\mu \eta + 2\mu^2 \eta^2) + \frac{q^2 v^2}{2} A'_\mu A^{\mu'} - \frac{1}{4} F_{\mu\nu} F^{\mu\nu}. \quad (5.27)$$

This is Eq. (5.3.15) in Ref. 4. The Higgs field η has a mass; the A_μ field has a mass proportional to v , and the ζ field has disappeared. The degree of freedom represented by the ζ field was used to form the extra polarization state of the massive vector field. The number of polarization states for a vector field changes discontinuously from two to three as the mass becomes nonzero.

HOMEWORK PROBLEM: Substitute Eq. (5.23) into Eq. (5.19) to obtain the effective Lagrangian Eq. (5.27).

So what happened? Perhaps this all went by too fast. The hand is quicker than the eye. Certainly this discussion should be considered suggestive rather than a proof of anything. We started out with a perfectly respectable Lagrangian, although the $|\lambda|(\phi^* \phi)^2$ term is something we have not seen before. Then we were free to choose a vacuum state which broke the rotational symmetry because of the degeneracy. Clever definitions of the constants led to nice looking formulas for the Higgs mass and the mass of the vector field. Although this procedure does not have to be carried out in every gauge theory—indeed QCD is a gauge theory of the SU(3) color group where the gluons remain massless and this machinery is not used—there are no inconsistencies encountered in doing so.

5.3 The Standard Model Again

The standard model, while more complicated than our toy, is nevertheless the simplest one which does the job. Since we are not dealing with a problem uniquely specified from first principles, whatever that means, the possibilities are endless. We have seen that there must be an SU(2) gauge symmetry group, weak isospin, because of the charge changing weak interaction. Models have been constructed using this group alone, and a Higgs field with three real components (isospin one). This produces two massive charged IVB's, and one massless neutral one, which could be $W_\mu^{(+)}$, $W_\mu^{(-)}$,

neutral currents exist, such a model can be discarded as too simple. Something else must be added.

A U(1) symmetry, weak hypercharge, is introduced as well, to give SU(2)×U(1), and four vector gauge fields as described in Lecture IV. The Higgs field is a complex doublet in weak isospin with four real components

$$\phi = \begin{pmatrix} \phi^{(+)} \\ \phi^{(0)} \end{pmatrix} \quad \begin{aligned} \phi^{(+)} &= (\phi_1 + i\phi_2)/\sqrt{2} \\ \phi^{(0)} &= (\phi_3 + i\phi_4)/\sqrt{2} \end{aligned} \quad (5.28)$$

This structure gives two charged fields and two neutral fields. The (e, ν) EWK Lagrangian before the spontaneous symmetry breaking is:

$$\begin{aligned} \mathcal{L}_1 &= \bar{\chi}_L \gamma^\mu [i\partial_\mu - g\frac{1}{2}\vec{\tau} \cdot \vec{W}_\mu - g'(-\frac{1}{2})B_\mu] \chi_L \\ &\quad + \bar{e}_R \gamma^\mu [i\partial_\mu - g'(-1)B_\mu] e_R - \frac{1}{4} \vec{W}_{\mu\nu} \cdot \vec{W}^{\mu\nu} - \frac{1}{4} B_{\mu\nu} B^{\mu\nu}. \end{aligned} \quad (5.29)$$

Here $\vec{\tau} \cdot \vec{W}_\mu = \sum_{i=1}^3 \tau_i \cdot W_\mu^{(i)}$ is a sum over SU(2) indices. The extra terms in the square brackets are the covariant derivatives, which for the non Abelian SU(2) gauge group involves an operator. The field tensor for the U(1) field $B_{\mu\nu}$ looks like the electromagnetic tensor:

$$B_{\mu\nu} = \partial_\mu B_\nu - \partial_\nu B_\mu, \quad (5.30)$$

while the SU(2) $W_{\mu\nu}$ field has a more complicated form, involving a commutator in SU(2) space:

$$W_{\mu\nu} = \partial_\mu W_\nu - \partial_\nu W_\mu + ig [W_\mu, W_\nu]. \quad (5.31)$$

A complete discussion of how these terms arise is given in Refs. 3 and 4. The Higgs field Lagrangian is

$$\mathcal{L}_2 = \left| \left(i\partial_\mu - \frac{g\vec{\tau}}{2} \cdot \vec{W}_\mu - g'\frac{Y}{2}B_\mu \right) \phi \right|^2 - V(\phi) \quad (5.32)$$

where $V(\phi)$ has the familiar form:

$$V(\phi) = \mu^2 \phi^\dagger \phi + |\lambda| (\phi^\dagger \phi)^2. \quad (5.33)$$

The vacuum expectation value of the Higgs field is chosen real and neutral:

$$\langle \phi \rangle_0 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \end{pmatrix} \quad (5.34)$$

There are three massless Goldstone bosons, two charged and one neutral, which get absorbed by gauge transformations to become the third components of the W and Z polarization vectors. There is one massive neutral field remaining, which is the Higgs particle. The masses of the W^\pm are related to Eq. 5.34:

$$M_{W^\pm} = gv/2. \quad (5.35)$$

Similarly, the mass of the Z is

$$\begin{aligned}
 M_Z &= v(g^2 + g'^2)^{1/2}/2 \\
 &= M_W(1 + g'^2/g^2)^{1/2} \\
 &= M_W/\cos\theta_W.
 \end{aligned}
 \tag{5.36}$$

Here we have used Eq. 4.28 and Eq. 5.35. As in our toy model above, the Higgs mass becomes

$$M_\phi = -2\mu^2 > 0, \tag{5.37}$$

which is an undetermined quantity in the theory.

This outline completes our discussion of the standard model. All of the important relations have either been derived, or made plausible by simpler derivations.

5.4 Experimental Situation

The standard EWK model with heavy W^\pm and Z intermediate bosons has been well established by a decade of experimental activity, beginning with the discovery experiments at the CERN SPS collider, and continuing at the Fermilab Tevatron, LEP at CERN, and the Linear Collider at SLAC. It is beyond the scope of these notes to review every aspect of these studies, but it is safe to say that so far, after intense experimental scrutiny, there are no data inconsistent with the model.

The cross section formula for $e^+e^- \rightarrow \bar{f}f$ through the Z resonance (Eq. 1.18), has been checked, and the number of neutrino families—three—has been obtained from the peak cross section. The neutral current coupling assignments for quarks and leptons given in Eq. 1.23 have been verified. The energy dependence of the Z decay asymmetry, Fig. 1.4, has been observed.

Hadronic production of Z 's and W 's, described in Lectures II and III, is in accord with theoretical expectations. At 1.8 TeV the contribution from sea quark structure functions cannot be neglected, so the formulas have more terms than those displayed in Eqs. 2.19 and 3.1, but the concept is unchanged. Data for the W transverse mass distribution were shown in Fig. 3.6. In order to obtain a value for the W mass from these data points, Monte Carlo programs were written which included all experimental resolution effects. Expected curves were then generated for various values of M_W with the width fixed at the standard model value: $\Gamma_W = 2.1$ GeV. The best fit W masses extracted in this way by the UA2 and CDF groups were averaged to obtain the value for M_W quoted below.

The Standard Model was constructed to be consistent with experimental results in weak decays and neutrino reactions, discussed in Lectures III and IV, and although searches are ongoing for lepton number violation, charged right handed currents, etc., there are so far no contradictions.

There are two missing particles: the top quark, and the scalar Higgs. Experimentally the b quark appears to have $T_3 = -\frac{1}{2}$, which implies a doublet partner with

The Standard Model was constructed to be consistent with experimental results in weak decays and neutrino reactions, discussed in Lectures III and IV, and although searches are ongoing for lepton number violation, charged right handed currents, etc., there are so far no contradictions.

There are two missing particles: the top quark, and the scalar Higgs. Experimentally the b quark appears to have $T_3 = -\frac{1}{2}$, which implies a doublet partner with $T_3 = +\frac{1}{2}$. The mass difference $m_t - m_b > M_W$, making the top quark very heavy, which leads to a small production cross section expected for $\bar{p}p \rightarrow \bar{t}t + X$ at the Tevatron. The discovery of the top quark is anticipated in 1992 or 1993.

Finding the Higgs will probably be more difficult. This particle, which is required by the dynamical symmetry breaking, does not affect the weak interactions of quarks and leptons in lowest order. If $m_\phi < M_Z$, then $Z \rightarrow \phi + e^+e^-$ is possible; the signal at LEP would be a peak in the e^+e^- final state mass spectrum below the Z mass when the storage ring is set at the Z resonance. The absence of such a signal leads to the limit $m_\phi > 57$ GeV. LEP 200 could extend this search to the 85–90 GeV mass range by searching for $e^+e^- \rightarrow Z^* \rightarrow Z + \phi$.

The Higgs contributes to higher order loop diagrams, and hence does affect the standard model formulas through radiative corrections. Equation 5.36 can be used to eliminate the Weinberg angle term $\sin^2 \theta_W$ from Eq. 4.29. The result is a quadratic equation for M_W^2 , which has the root

$$M_W^2 = \frac{M_Z^2}{2} \left(1 + \left(1 - \frac{4A}{M_Z^2} \right)^{1/2} \right), \quad (5.38)$$

where $A = \frac{\pi\alpha}{\sqrt{2}G_F} = 1389.83 \text{ GeV}^2$. This formula gives M_W in terms of three very well known quantities: α , G_F and M_Z . Substituting $M_Z = 91.161 \pm 0.031 \text{ GeV}$ gives $M_W = (80.91 \pm 0.03) \text{ GeV}$, within 1% of the experimental number $M_W = (80.15 \pm 0.31) \text{ GeV}$! Although small, this 1% difference is statistically significant, and is attributed to a radiative correction term to the constant A . The constants α and G_F are defined by very low energy processes, much smaller than the masses of the W and the Z . A correction term can be inserted into the formula for M_W^2 to allow for changes in α and G_F as a result of the large increase in energy scale. This term, Δr , modifies Eq. (4.29):

$$M_W^2 = \frac{A}{\sin^2 \theta_W (1 - \Delta r)} \quad (5.39)$$

This can be solved for Δr to give

$$\Delta r = 1 - \frac{A}{M_W^2 (1 - M_W^2/M_Z^2)} \quad (5.40)$$

The radiative corrections depend upon all of the masses in the theory, two of which, the t quark and the Higgs, are unknown. The dependence on the t quark mass is quadratic, whereas the Higgs mass enters logarithmically. Figure 5.2 shows the result

of a calculation of Δr as a function of m_t for two different choices of m_ϕ . Setting M_W to the experimental number quoted above in Eq. 5.40 gives $\Delta r = 0.048 \pm 0.018$. With this number the top quark mass can be read from Fig. 5.2: $m_t = 130^{+40}_{-60}$ GeV. As the error in M_W improves, so will this prediction, although eventually the unknown m_ϕ will dominate the error on the high mass side, unless the Higgs is discovered in the meantime. When the top quark is found, it will be interesting to see if the result is consistent with expectations from the Standard Model.

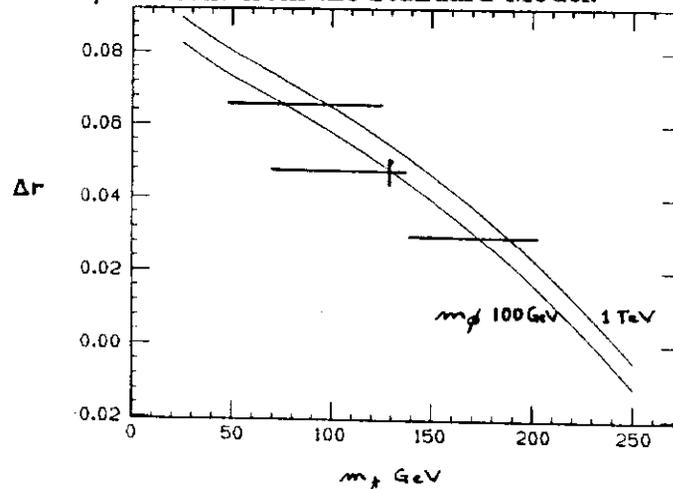


Figure 5.2. Δr vs. top quark mass.

Acknowledgements

I want to thank my hosts in Mexico City, Professors Rodrigo Huerta and Miguel Perez Angon, for the opportunity to lecture at their institution. I especially thank Dr. Aurelio Espiritu for aid in setting up the Monte Carlo programs on the local VAX cluster. I appreciate the help of Vincent Fuh in the preparation of this manuscript.

Appendix

This is a brief description of the Monte Carlo technique applied to some simple calculations in collider physics. Many effects encountered in the comparison of experimental data with theoretical expectations either become very cumbersome when expressed analytically in closed form, or cannot be written down at all. Typical examples are: the folding of two different angular distributions, like particle production and decay; the effects of quark and gluon structure functions on particle kinematics in the laboratory; fragmentation of partons into observed particles, and the clustering into jets; and distortions of the theoretical curves due to experimental resolutions, detection efficiencies, etc.

The programs described here are simplified toys, as emphasized in the lectures. The real stuff is considerably more complicated, but the principles are the same.

The Monte Carlo technique means in this instance the generation of “events” using random numbers.²⁵ The computer produces a random floating point number between zero and one. A histogram of 1000 such numbers binned in 0.1 unit bins would have 100 ± 10 events per bin. Thus continuous curves calculated this way exhibit statistical fluctuations. If the angular distribution of a reaction is isotropic, and if r is the random number, then choosing $\cos(\theta) = -1 + 2 * r$ produces the desired events. There are two ways commonly used to generate events according to a distribution $f(x)$ which is not flat. We will call them A and B. Method A works as follows. If the indefinite integral

$$F(x) = \int_0^x f(x') dx' \quad (\text{A.1})$$

exists, and if $F(1) = 1$, then the inversion

$$x = F^{-1}(r) \quad (\text{A.2})$$

where r is a random number, generates events according to $f(x)$ in the range $0 < x < 1$. This works because the distribution in r is flat, and $\Delta r = f(x)\Delta x$. Since quadratic functions are quite common, one often has to solve a cubic equation for the inversion Eq. A.2.

Method B is less efficient in computer time, but is often used when the inversion of the integral in Eq. A.1 is not convenient. Suppose that $f(x)$ is a smooth function, and that the maximum value $f_{\text{MAX}} = 1$, in the range $0 < x < 1$. Then pick a random number r_1 , evaluate $f(r_1)$, and compare $f(r_1)$ to a second random number r_2 . If $r_2 < f(r_1)$, the event is retained; otherwise it is rejected. Looping through this procedure to generate 1000 events, say, produces the desired curve $f(x)$.

Another useful property of random numbers is that the sum of 12 of them is Gaussian distributed with mean six and variance one:

$$\sum_{i=1}^{12} r_i - 6 = x \quad \text{generates} \quad e^{-x^2/2}. \quad (\text{A.3})$$

The probability of being outside $\pm 6\sigma$ is of course zero, but normally this is not a serious defect. The number 12 has the same origin as the relation between a flat distribution of width a and the Gaussian equivalent σ : $\sigma = a/\sqrt{12}$.

What follows are descriptions of the Monte Carlo programs which generate W 's and allow them to decay into leptons, and fragment quarks into jets of particles. These programs use the appropriate functions and the random number techniques outlined above. Rather than actually presenting the FORTRAN code, which is terribly boring, we will simply go through the steps necessary to get the job done, and the reader can write her own code.

function $G(x)$ in the form:

$$xG(x) = g(x) \sum_{k=0}^3 (a_k + \sqrt{x}b_k)C_k(x) \quad (\text{A.4})$$

where $g(x) = (1-x)^\alpha$. The exponent is an integer which depends on quark flavor: $\alpha_u = 3$, $\alpha_d = 4$, and $\alpha_s = 8$. $C_k(x)$ is the Chebyshev polynomial

$$C_k(x) = \cos(k \cos^{-1}(2x-1)). \quad (\text{A.5})$$

As discussed in Chapter 28 of Ref. 25, Chebyshev polynomials are really cosine functions— $\cos(k\phi)$ —masquerading as something more complicated. They satisfy a differential equation very similar to Legendre polynomials. It is straightforward to evaluate the expressions for $xu(x)$ and $xd(x)$ given the coefficients of the fits in Table IV of Ref. 10. They are plotted in Fig. 2.2.

The program first evaluates $xu(x)$ and $xd(x)$ in 100 bins from $0 < x < 1$. It finds the maximum values of the functions, and plots the functions by MC B described above. Then it picks the rapidity of the W at random: $y_W = -y_{\max} + 2 * y_{\max} * r$, where y_{\max} is given by Eq. 1.10; and calculates x_1 and x_2 from Eq. 2.6. With these numbers it evaluates the product $x_1x_2u(x_1)d(x_2)$, and declares the chosen value of y_W successful if it satisfies the criterion of MC B. Looping through this procedure generates the distribution of Fig. 3.1.

While in the loop calculation of y_W , the program generates the leptonic decay $W \rightarrow \ell + \nu$. The angular distribution is given in Eq. 3.11. Here it is convenient to integrate $(1 - \cos\theta)^2$ and use MC A. The rapidity of the charged lepton in the W rest frame is calculated from Eq. 1.9 with $E_\ell = M_W/2$, and the rapidity in the laboratory is the sum $y_\ell = y_\ell^* + y_W$. The results are shown in Fig. 3.2. A lepton p_ℓ cut can be imposed at this stage.

W_TRANS_MASS

A separate bit of code was used to generate the transverse mass distribution of Fig. 3.5. Since M_\perp is independent of the longitudinal motion of the W , we let the W be at rest in the laboratory: $x_1 = x_2 = x$. Then the parton level variable $\hat{s} = x^2s$, where s is a constant. Let $\zeta = M_\perp/\sqrt{\hat{s}}$. Then Eqs. 3.7 and 3.10 may be rewritten apart from constant factors in the form

$$\frac{dN}{d\zeta} = \frac{\zeta(2-\zeta^2)}{(1-\zeta^2)^{1/2}} \frac{xu(x)xd(x)}{(x^2s - M_W^2)^2 + M_W^2\Gamma^2} \quad (\text{A.6})$$

As defined, ζ depends on x . Now assume that the structure functions can be considered constant over the width of the Breit-Wigner. Then we can integrate the B - W , and pick x according to the MC A prescription. The function $f(\zeta) = \frac{\zeta(2-\zeta^2)}{(1-\zeta^2)^{1/2}}$ can

As defined, ζ depends on x . Now assume that the structure functions can be considered constant over the width of the Breit-Wigner. Then we can integrate the B - W , and pick x according to the MC A prescription. The function $f(\zeta) = \frac{\zeta(2 - \zeta^2)}{(1 - \zeta^2)^{1/2}}$ can also be integrated, resulting in a cubic equation for ζ^2 . Setting the equation equal to a random number, and solving it for ζ^2 (it has one real root) picks ζ^2 according to MC A.

The transverse mass is given by

$$M_{\perp} = x\sqrt{s}\zeta \quad (\text{A.7})$$

This is the histogrammed variable.

FRAGGER

This program is based on the algorithm of Ref. 14. The fragmentation function used by these authors is

$$f(z) = 1 - a + 3 * a * ((1 - z)^2), \quad \text{with } a = .77. \quad (\text{A.8})$$

Starting with an initial quark of momentum p , a (q, \bar{q}) meson is formed with momentum $z * p$, leaving behind a quark with momentum $(1 - z) * p$. The function $f(z)$ is then used to make a second meson with the left over quark. This process is repeated until the produced meson momentum is below some threshold value—set to the pion mass in the program. An upper limit of 50 is placed on the number of particles from a single quark, although this cutoff is rarely reached in the calculation with $p = 100$ GeV. The transverse momentum of the meson so created is chosen at random in azimuth relative to the parent quark, and with a Gaussian distributed p_t of variance $\sigma = .3$ GeV. The p_t of the meson and the remaining quark are balanced. Their polar angles differ, since they have different longitudinal momenta, while their azimuthal angles are 180° apart.

Equation A.8 can be integrated to give a cubic equation, which then is set equal to a random number and solved for z according to the method of MC A. The p_t is picked using Eq. A.3. The only complicated part of the calculation is the rotation operation. The polar and azimuthal angles of the meson are given relative to the parent quark direction, which moves from one fragmentation process to the next. The components of the momentum of the meson relative to the primary quark direction—the z axis—are the desired quantities. The situation is sketched in Fig. A.1. Sequential rotations through the Euler angles ϕ and θ are required. If the (x', y', z') frame is oriented at angles (ϕ, θ) relative to the (x, y, z) frame—rotate first about z through ϕ , and then about y' through θ ; and if the components of a vector are (px', py', pz') in the rotated frame, then (px, py, pz) are given by the operation

$$\begin{pmatrix} px \\ py \\ pz \end{pmatrix} = \begin{pmatrix} \cos \phi \cos \theta & -\sin \phi & \sin \theta \cos \phi \\ \sin \phi \cos \theta & \cos \phi & \sin \theta \sin \phi \\ -\sin \theta & 0 & \cos \theta \end{pmatrix} \begin{pmatrix} px' \\ py' \\ pz' \end{pmatrix} \quad (\text{A.9})$$

The first meson requires no rotations. Subsequent meson momentum components must be rotated according to Eq. A.9, where (ϕ, θ) are the angles of the parent quark relative to the primary quark.

Results for this program are shown in Figs. A.2 and A.3 for an initial quark momentum of 100 GeV. Although very simplified, this model reproduces the characteristics of 100 GeV jets—multiplicity, invariant mass, z distribution—reasonably well.

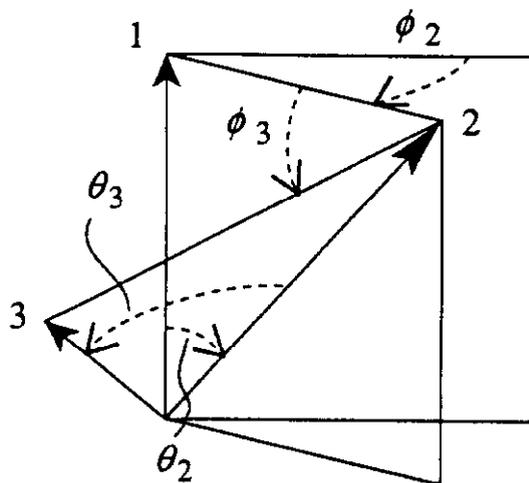


Figure A.1. Two vectors for subsequent parent quarks, showing the angles.

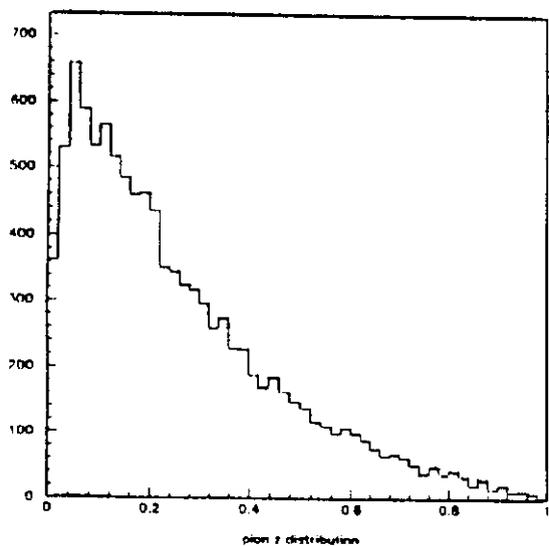


Figure A.2. Meson z distribution.

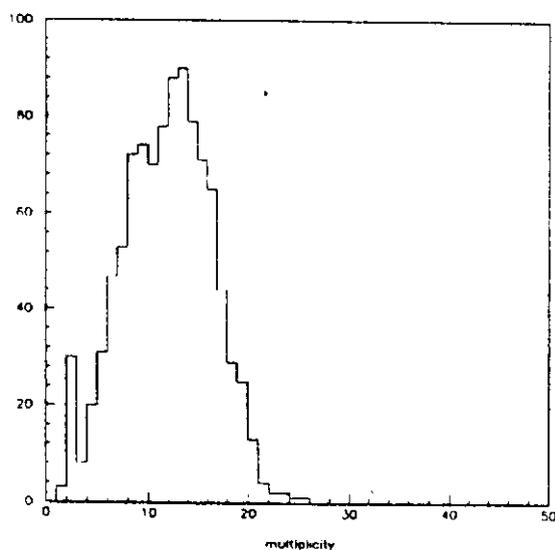


Figure A.3. Particle multiplicity distribution.

References

1. L.G. Pondrom, in *Proceedings of the XXV International Conference on High Energy Physics* (World Scientific, Singapore, 1991).
2. V.D. Barger and R.J.N. Phillips, *Collider Physics* (Addison Wesley, New York, 1987).
3. F. Halzen and A.D. Martin, *Quarks & Leptons* (John Wiley, New York, 1984).
4. C. Quigg, *Gauge Theories of the Strong, Weak, and Electromagnetic Interactions* (Benjamin-Cummings, Reading, Mass., 1983).
5. I.J.R. Aitchison and A.J.G. Hey, *Gauge Theories in Particle Physics*, 2nd ed. (Adam Hilger, Philadelphia, 1989).
6. K. Gottfried and V. Weisskopf, *Concepts of Particle Physics Vol II* (Oxford University Press, New York, 1986).
7. Review of Particle Properties, *Phys. Lett. B* **239**, (1990) 1.
8. J.D. Bjorken & S.D. Drell, *Relativistic Quantum Fields* (McGraw Hill, New York, 1965).
9. R.E. Taylor, H.W. Kendall, & J.I. Friedman, *Rev. Mod. Phys.* **63** (1991) 573 ff. These papers summarize the presentations delivered at the Nobel Prize ceremony in 1990.
10. R.P. Feynman and R.D. Field, *Phys. Rev. D* **15**, (1977) 2590.
11. The closest version to this treatment of structure functions that is known to the author is in Ref. 6. This must all have been intuitively obvious to Feynman (*High Energy Collisions, Stony Brook 1969* (Gordon and Breach, New York, 1969)), but I have never been able to follow his reasoning. It is like trying to improve one's tennis by watching Ivan Lendl. World class athletes do things the rest of us can't do.
12. F. Abe *et al.* (CDF Collaboration), *Phys. Rev. D* **43** (1991) 2070.
13. F. Abe *et al.* (CDF Collaboration), *Phys. Rev. Lett.* **64** (1990) 152; see also C. Albajar *et al.*, (UA1 Collaboration) *Phys. Lett. B* **253** (1991) 503.
14. R.P. Feynman and R.D. Field, *Nucl. Phys. B* **136** (1978) 1.
15. B. Andersson *et al.*, *Phys. Rep.* **97** (1983) 31.

16. J. Alitti *et al.*, (UA2 Collaboration), *Z. Phys. C* **49** (1991) 17.
17. R.E. Marshak, Riazuddin, and C.P. Ryan, *Theory of Weak Interactions in Particle Physics* (Wiley-Interscience, New York, 1969) p. 228.
18. M. Schwartz, *Rev. Mod. Phys.* **61** (1989) 527.
19. S. Weinberg, *Rev. Mod. Phys.* **52** (1980) 515; A. Salam, *ibid.*, p. 525; S. Glashow, *ibid.*, p. 539.
20. P. Galison, *Rev. Mod. Phys.* **55** (1983) 477.
21. J. Steinberger, *Rev. Mod. Phys.* **61** (1989) 533.
22. This is a procedure first introduced by Glashow, Iliopoulos, and Maiani to cancel the experimentally unobserved flavor changing neutral currents. It required the introduction of a fourth quark with charge $+2/3$, dubbed the charmed quark, which had not been observed either.
23. U. Amaldi, *et al.*, *Phys. Rev. D* **36** (1987) 1385.
24. G. Arnison *et al.*, (UA1 Collaboration), *Phys. Lett.* **122B** (1983) 103.
25. R.W. Hamming, *Numerical Methods for Scientists and Engineers*, 2nd edition, McGraw Hill, New York (1973), Chap. 8.