DRAL
Daresbury Laboratory
Rutherford Appleton Laboratory

RAL Report
RAL-95-033

Proceedings of the School for Young High Energy Physicists

J Flynn I G Halliday R G Roberts and G M Shore

March 1995

Rutherford Appleton Laboratory Chilton DIDCOT Oxfordshire OX11 0QX

RAL-95-033

Central Laboratory of the Research Councils

Rutherford Appleton Laboratory
Chilton DIDCOT Oxon OX11 0QX

Proceedings of the School for Young High Energy Physicists

Editor: K J Peach

Rutherford Appleton Laboratory
September 5 - 16 1994

HEP SUMMER SCHOOL FOR YOUNG HIGH ENERGY PHYSICISTS

RUTHERFORD APPLETON LABORATORY/THE COSENER'S HOUSE, ABINGDON:

4 - 16 SEPTEMBER 1994

Director:	Dr K J Peach	University of Edinburgh
Lecturers:	Dr J Flynn	University of Southampton
	Prof I G Halliday	University College of Swansea
	Dr R G Roberts	Rutherford Appleton Laboratory
	Dr G Shore	University College Swansea
Tutors:	Dr S Cartwright	University of Sheffield
	Dr P D Dauncey	Rutherford Appleton Laboratory
	Dr P F Harrison	Queen Mary and Westfield College
	Dr W Scott	Rutherford Appleton Laboratory
Students:	Mr Nicholas Allen	Brunel University
	Mr Lawrence Angrave	University of Oxford
	Mr Richard Bates	University of Glasgow
	Mr Robin Boswell	University of Sheffield
	Mr Laurence Bryant	Royal Holloway & BNC
	Mr Steven Campbell-Robson	University of Bristol
	Mr Michael Carroll	University of Liverpool
	Mr Oliver Cooke	University of Manchester
	Mr Chris Cormack	University of Liverpool
	Mr Gordon Crawford	University of Lancaster
	Mr Paul Dixon	University of Lancaster
	Mr Simon Dorris	University of Glasgow
	Mr Jim Edwards	University of Manchester
	Mr Andrew Galloni	University of Liverpool
	Mr Holger Gunther	University of Lancaster
	Mr Alex Howard	Imperial College London
	Miss Victoria Hudgson	University of Birmingham
	Mr Bernard Hung	Imperial College London
	Mr Clive Lewis	University College London
	Mr M J Lyon	University of Oxford
	Mr Barry MacEvoy	Imperial College London
	Mr Spyros Manolopoulos	University of Sheffield
	Mr Jason D McFall	University of Oxford
	Mr Edward McKigney	University of Manchester
	Mr Andrew McNab	Queen Mary & Westfield
	Miss Chandrika Nath	University of Oxford
	Mr Dave Newbold	University of Bristol
	Mr Warrick Newton	University of Sheffield
	Mr Timothy Nicholls	University of Birmingham
	Mr Keith Norman	University of Birmingham
	Miss Ainsley M Normand	University of Oxford
	Mr Antonios Papanestis	University College London
	Mr Arnulf Quadt	University of Oxford
	Mr David Rigby	University of Birmingham
	Mr Eram Rizvi	Queen Mary & Westfield
	Mr Robert Saunders	University College London
	Mr Iain Scott	Queen Mary & Westfield
	Miss Anna Stacey	Imperial College
	Mr Argyris Stassinakis	University of Oxford
	Mr Ilhan Tapan	University of Bristol
	Miss Fiona K Thomson	University of Glasgow
	Mr Martin Utley	University of Glasgow
	Mr Ben Waugh	University of Manchester
	Miss Ursula Wielgosz	University of Oxford
	Mr Mark Williams	University of Lancaster
	Mr Kin Yip	University of Oxford

RAL Summer School for Young Experimental High Energy Physicists

Coseners House, 5 - 17 September 1994

PREFACE

Forty-six young experimental particle physicists students attended the 1994 Summer School, held as usual in Coseners House in Abingdon in early September. This year, the weather was mild and sunny for the most part, choosing to rain only for the Barbecue on the Friday evening.

The lectures reproduced here were given by Ian Halliday (Relativistic Quantum Field Theory), Jonathan Flynn (Relativistic Quantum Mechanics), Graham Shore (The Standard Model) and Dick Roberts (Phenomenology). The lectures were delivered with enthusiasm and good humour, and provided a solid foundation for deeper discussions in small groups over dinner or in the pub. We were very pleased that Dr Peter Williams, who is Chairman of both PPARC & Oxford Instruments, made time in his busy schedule to give us a very interesting and informative evening seminar. Bill David from DRAL gave a comprehensive and comprehensible overview of the range of physics performed at ISIS in another evening seminar, and Paul Harrison from QMW and Mike Whalley from the Durham HEP group gave afternoon seminars on permutation symmetry of the CKM matrix and information retrieval systems respectively.

All students gave a twelve minute seminar in one of the evening sessions; these were of an exceptionally high standard, delivered clearly and concisely and transmitting excitement and enthusiasm.

Chandy Nath and Lawrence Angrave provided an impromptu cabaret of gentle jazz for the formal dinner, followed by a witty and elegant after dinner speech by Sandy Donnachie, who has chaired every important committee in particle physics.

The work of the school was helped enormously by the hard work of the tutors - Susan Cartwright (Sheffield), Paul Dauncey (DRAL), Paul Harrison (QMW) and Bill Scott (DRAL).

Videos were shown on four or five evenings - the first organized by the director, the remainder by the students. The usual sports and other activities filled in those few gaps in the timetable. By the end of the two weeks of total immersion in particle physics, exhaustion tempered the sense of achievement at having survived the course.

The school benefited from the welcoming and calm atmosphere of Coseners House, and the efforts of all of the staff there were very much appreciated.

None of it would happen, of course, if it were not for the tireless enthusiasm and organizational skill of Ann Roberts (DRAL), whose gentle prodding ensures that everything that should happen does happen (and that those things which shouldn't happen don't).

This was undoubtedly one of the very best of the many summer schools with which I have been involved. To all who helped make it so - lecturers, tutors, staff at both Coseners and DRAL and above all the students - I extend my thanks and my good wishes.

Ken Peach (Director)
Department of Physics & Astronomy,
University of Edinburgh

CONTENTS

Pages

LECTURE COURSES

QUANTUM FIELD THEORY: CANONICAL AND PATH INTEGRAL APPROACHES

By Prof I G Halliday
University of Wales, Swansea

Lectures delivered at the School for Young High Energy Physicists
Rutherford Appleton Laboratory, September 1994

21/12/94

Quantum Field Theory:
Canonical and Path Integral Approaches.

I G Halliday,
Physics Department,
University of Wales, Swansea,
Singleton Park,
Swansea SA2 8PP.

Introduction.

This short lecture course is aimed at connecting the later courses to your undergraduate knowledge of Quantum Mechanics, Maxwell's equations, etc. I also have to provide a language and a series of results for the later lectures - a tall order. So, I would like to show you how to quantise an arbitrary theory; in particular show you why gauge theories are hard. On the other hand I would like to demonstrate Salam's remark that he has been surprised at how little he has had to change his ideas in over thirty years ! So the highlight of the course will be a "proof" of Feynman rules for the perturbative evaluation of field theories such as scalar fields, and Yang-Mills. I will make a point of starting far enough back to be comprehensible to everybody - I expect riots otherwise. A constant theme will be the way that different views make different aspects transparent. The prime example being the way Hamiltonian methods make Quantum Mechanics easy but hide symmetries. Lagrangian methods make symmetries easy but lose direct physical contact with your old view of Quantum Mechanics as a theory of operators, states, eigenvectors, etc..

Synopsis.

0) The examples done before coming to Coseners
   - (1) Harmonic Oscillator
   - (2) Gaussian Integrals
   - (3) Pictures in Quantum Mechanics
   - (4) Newton implies Lagrange
   - (5) Dirac δ function

1) Classical Mechanics - Lagrange vs Hamilton

2) Quantum Mechanics
   - (1) Schrödinger vs Heisenberg vs Dirac pictures
   - (2) Hamilton vs Lagrange, Dirac and Feynman
   - (3) Heisenberg Harmonic Oscillator

3) Free Boson
   - (1) Classical
   - (2) Quantum

4) Interacting Boson
   - (1) Feynman diagrams by operators
   - (2) Feynman diagrams by Functional integrals

5) Groups and Algebras
   - (1) Definitions and Examples

6) Gauge Theories
   - (1) Classical Maxwell Theory - Lorentz invariance
   - (2) Lagrangian formalism
   - (3) U(1) Covariant derivative
   - (4) Non-Abelian Gauge Theories

(5) Feynman Rules
(6) Gauge Fixing

7) Fermionic Integrals

8) Experimentally Virgin Territory

Acknowledgements

It is a great pleasure to thank Ken Peach for organising and running the school so well. His rumbustious humour contributed greatly to the atmosphere. The students, of course, made the school. This year I was impressed by their professional approach. The talks by the students were easily the best of the three years. Finally, I would thank Ann Roberts for her unobtrusive organisation in coping with lecturers changing schedules at very little notice.

Chapter 0

The prerequisites for the course

The purpose of this section of the notes is to provide you with a summary of what I expect you to know before we start in September. Everything here will be used in one form or another in the course. You must do all the examples before you arrive at Coseners. I stress that these examples and exercises are at the heart of quantum field theory either in operator form or the more trendy Functional or Path integral method. You must do all of these problems and two finger exercises. If not, you will end up having to understand both the mathematics and the field theory ideas simultaneously. With the exercises behind you the maths should be second nature. The mathematics at the heart of quantum field theory is constructed out of many oscillators. So the first aim is to be absolutely sure that you understand one oscillator by itself. Hence the way your undergraduate courses pounded away at this problem, probably without explaining why. To jump the gun slightly, the basic reason is that if we have a photon, pion, Higgs of energy $E(\underline{k}) = \hbar\omega(\underline{k})$ for momentum $\underline{k}$ then n such particles have energy $n\hbar\omega(\underline{k})$. The energy levels of a Harmonic Oscillator of frequency ω are $\hbar\omega(n+\frac{1}{2})$. Apart from the constant $\frac{1}{2}\hbar\omega$ these agree !

In the lectures I will use the Path integral formalism extensively to construct Feynman diagrams, fix gauges,... This trickery leans totally on a knowledge of Gaussian integrals. So again I include a revision section on these integrals and the two standard tricks of completing the square and introducing new parameters.

Many of you will have already have studied the concept of pictures in Quantum Mechanics; in particular the Schrödinger and Heisenberg pictures. We will use a new picture due to Dirac to construct perturbation theory. So here I revise the standard material.

Finally I will use extensively the ideas of Lagrangian and Hamiltonian mechanics. For completeness I give the standard derivation of Lagrange's equations from Newton's. In Quantum Field Theory Lagrangians play a major role because of the way they make the symmetries of the problem manifest. In Quantum Mechanics we usually start from the Hamiltonian which often hides the symmetries. We will thus develop a new formalism, the path integral formalism, which is the Lagrangian variant of Quantum Mechanics. This will make many calculations much easier and slicker. Here I remind you of the classical connection between Lagrangians and Hamiltonians.

We will use many properties of Dirac δ-functions. I remind you of the definitions, proofs and the results we will use.

0.1) Harmonic Oscillator

So let us redo the Harmonic Oscillator in Quantum Mechanics using an operator formalism. We use the most common picture in undergraduate texts - the Schrödinger picture - where operators are independent of time.

$$\hat{p} = -i\hbar\frac{\partial}{\partial q}$$

$$\hat{q} = q$$

The notation may look unusual, but here and in the lectures, I will try to avoid a standard confusion. The above equations are usually written

$$\hat{p} = -i\hbar\frac{\partial}{\partial x}$$

$$\hat{x} = x \tag{0.1.1}$$

where $\hat{x}$ is the dynamical variable corresponding to the position of a point particle. Later in Field theory we will have dynamical variables $\hat{\Phi}(\underline{x}, t)$ where $\underline{x}$ is not a dynamical variable but merely a label. i.e. $\hat{\Phi}(\underline{x}, t)$ is an operator at some fixed point $\underline{x}$ of space. The label $\underline{x}$ undergoes no dynamics, although $\hat{\Phi}(\underline{x}, t)$ certainly does. To avoid this confusion the dynamical variables for point particles are usually rewritten as momentum p, position q. The only property of (0.1.1) I will use is the commutation relation

$$[\hat{q}, \hat{p}] = i\hbar \tag{0.1.2}$$

and from now on we choose units such that $\hbar = 1$. In Quantum Mechanics the starting point is usually the Hamiltonian or energy operator

$$\hat{H}(\hat{p}, \hat{q}) = \frac{\hat{p}^2}{2m} + \frac{m\omega^2}{2}\hat{q}^2 \tag{0.1.3}$$

written as a function of the position and momentum. We will solve for the energy eigenstates and normalised wave functions of (0.1.3) **only using (0.1.2) and not (0.1.1).** This will be important later. Define

$$\begin{aligned} \hat{a}^\dagger &= \frac{1}{\sqrt{2}}(\hat{q}\sqrt{m\omega} - i\frac{\hat{p}}{\sqrt{m\omega}}) \\ \hat{a} &= \frac{1}{\sqrt{2}}(\hat{q}\sqrt{m\omega} + i\frac{\hat{p}}{\sqrt{m\omega}}) \end{aligned} \tag{0.1.4}$$

Since $\hat{q}^\dagger = \hat{q}$ and $\hat{p}^\dagger = \hat{p}$ it is clear that $\hat{a}, \hat{a}^\dagger$ are, in fact, Hermitian conjugates of one another so the notation makes sense. Now compute

$$[\hat{a}, \hat{a}^\dagger] = \frac{1}{2}[\hat{q}, -i\hat{p}] + \frac{1}{2}[i\hat{p}, \hat{q}] = 1 \tag{0.1.5}$$

by (0.1.2). Moreover

$$\begin{aligned} \hat{a}^\dagger\hat{a} &= \frac{1}{2}(\frac{\hat{p}^2}{m\omega} + \hat{q}^2 m\omega + i[\hat{q}, \hat{p}]) \\ &= \frac{1}{\omega}(\frac{\hat{p}^2}{2m} + \frac{1}{2}m\omega^2\hat{q}^2 - \frac{1}{2}\omega) \end{aligned} \tag{0.1.6}$$

where the $\hbar$ can be resurrected by dimensional analysis if required. Now let us compute the eigen values of $\hat{H}$. Suppose we have an eigenstate $|\alpha\rangle$ such that

$$\hat{a}^\dagger\hat{a}|\alpha\rangle = \alpha|\alpha\rangle$$

Then clearly

$$\hat{H}|\alpha\rangle = \hbar\omega(\alpha + \frac{1}{2})|\alpha\rangle$$

so an eigenstate of $\hat{a}^\dagger\hat{a}$ is an eigenstate of $\hat{H}$ and vice versa. Now I claim that $\hat{a}^\dagger|\alpha\rangle$ is also an eigenstate of $\hat{H}$:

$$\begin{aligned}
\hat{a}^\dagger\hat{a}\{\hat{a}^\dagger|\alpha\rangle\} &= \hat{a}^\dagger\{\hat{a}\hat{a}^\dagger\}|\alpha\rangle \\
&= \hat{a}^\dagger\{\hat{a}^\dagger\hat{a} + 1\}|\alpha\rangle \\
&= \hat{a}^\dagger\{\alpha + 1\}|\alpha\rangle \\
&= (\alpha + 1)\{\hat{a}^\dagger|\alpha\rangle\}
\end{aligned} \tag{0.1.7}$$

because $(\alpha + 1)$ is a number. Similarly

$$\hat{a}^\dagger\hat{a}\{\hat{a}|\alpha\rangle\} = (\alpha - 1)\{\hat{a}|\alpha\rangle\} \tag{0.1.8}$$

Thus, given any eigenstate with eigenvalue α, we can easily construct eigenstates with eigenvalues $\alpha, \alpha+1, \alpha+2, \cdots, \alpha-1, \alpha-2, \cdots$, by multiple applications of $\hat{a}^\dagger$ and $\hat{a}$. Do these sequences ever stop? To fix the limits, consider

$$\begin{aligned}
\langle\alpha|\hat{a}^\dagger\hat{a}|\alpha\rangle &= \int dq \Psi_\alpha^*(q)(\hat{a}^\dagger\hat{a})\Psi_\alpha(q) \\
&= \int dq \big(\hat{a}\Psi_\alpha(q)\big)^*\big(\hat{a}\Psi_\alpha(q)\big) \\
&= \int dq \Phi^*(q)\Phi(q) \geq 0
\end{aligned} \tag{0.1.9}$$

where we have used the fact that the Hamiltonian conjugate of $\hat{a}^\dagger$ is $\hat{a}$ and have written $\Phi(q) = \hat{a}\Psi_\alpha(q)$. Note that in (0.1.9) we get zero if and only if $\Phi(q) = 0$. On the other hand, if $|\alpha\rangle$ is an eigenstate of $\hat{a}^\dagger\hat{a}$, the left hand side is

$$\langle\alpha|\hat{a}^\dagger\hat{a}|\alpha\rangle = \alpha\langle\alpha|\alpha\rangle \tag{0.1.10}$$

Now above we constructed eigenstates $|\alpha\rangle, |\alpha-1\rangle, |\alpha-2\rangle\cdots$ and the above can be applied to any of them. So in (0.1.9) we have, using (0.1.10) that

$$(\alpha - n)\langle\alpha - n|\alpha - n\rangle \geq 0 \qquad ; n > 0 \tag{0.1.11}$$

Since $\langle\alpha - n|\alpha - n\rangle \geq 0$ we have $(\alpha - n) \geq 0$ for all n. Clearly an absurdity for sufficiently larger n. So somewhere above there is a mistake. Can you see it? Don't turn the page and cheat. The mistake is a standard mistake that lies at the heart of angular momentum theory, the theory of Lie algebras, Kac-Moody!

The way out is that we proved

$$\hat{a}^\dagger \hat{a}(\hat{a}^\dagger|\alpha\rangle) = (\alpha+1)(\hat{a}^\dagger|\alpha\rangle)$$

which seems to prove $\hat{a}^\dagger|\alpha\rangle$ is an eigenstate of $\hat{a}^\dagger\hat{a}$. **But** remember for any operator $\hat{O}$ the eigenvalue equation

$$\hat{O}\Phi(q) = \alpha\Phi(q)$$

always has the trivial solution $\Phi(q) = 0$. So we must check that $\hat{a}^\dagger|\alpha\rangle \neq 0$. Since we are getting a contradiction, in fact for some n we must always get, for the first time,

$$(\hat{a})^n|\alpha\rangle = 0$$

for some positive integer n. Then consider the state $|\beta\rangle = (\hat{a})^{n-1}|\alpha\rangle \neq 0$ such that $\hat{a}|\beta\rangle = 0$ and hence $\hat{a}^\dagger\hat{a}|\beta\rangle = 0|\beta\rangle$. Notice that the $(\hat{a}^\dagger)^n|\alpha\rangle$ eigenstates give us no problem; the analogue of (0.1.11) is

$$(\alpha+n)\langle\alpha+n|\alpha+n\rangle \geq 0 \qquad n \geq 0$$

which gives no sign problems. So we now construct everything on the basis of the lowest state $|\beta\rangle$ which must satisfy $\hat{a}|\beta\rangle = 0$. Then

$$\begin{aligned}\hat{H}(\hat{a}^\dagger)^n|\beta\rangle &= \omega(\hat{a}^\dagger\hat{a} + \frac{1}{2})(\hat{a}^\dagger)^n)|\beta\rangle \\ &= \omega(n+\frac{1}{2})(\hat{a}^\dagger)^n|\beta\rangle)\end{aligned}$$

So, as promised, the eigenstates of $\hat{H}$ are $\omega(n+\frac{1}{2})$. If you want to construct the wave functions, the $|\beta\rangle$ equation

$$\hat{a}|\beta\rangle = 0 \Rightarrow (\sqrt{m\omega}q + \frac{1}{\sqrt{m\omega}}\frac{\partial}{\partial q})\Psi(q) = 0$$

or

$$\Psi(q) = e^{\frac{-m\omega}{2}q^2}$$

Higher eigenstates can be computed by applying

$$\hat{a}^\dagger = (\sqrt{m\omega}q - \frac{1}{\sqrt{m\omega}}\frac{\partial}{\partial q})$$

to $\Psi(q)$.

An interesting exercise is to check the normalisation of the states. If we call the lowest states $|0\rangle, |1\rangle, |2\rangle$, of energy e-value $\frac{1}{2}\hbar\omega, \frac{3}{2}\hbar\omega, \frac{5}{2}\hbar\omega$ then the normalised state is

$$|n\rangle = \frac{1}{\sqrt{n!}}(\hat{a}^\dagger)^n|0\rangle$$

The proof of this is straightforward using the commutation relatives and induction. Take the normalised $|0\rangle$ state $\langle 0|0\rangle = 1$. Then

$$\begin{aligned}\langle 1|1\rangle &= \langle 0|\hat{a}\hat{a}^\dagger|0\rangle \\ &= \langle 0|(\hat{a}^\dagger\hat{a}+1)|0\rangle \\ &= \langle 0|0\rangle = 1\end{aligned}$$

because $\hat{a}|0\rangle = 0$. Then assume $\langle n|n\rangle = 1$ and compute

$$\begin{aligned}\langle n+1|n+1\rangle &= \frac{1}{(n+1)}\langle n|\hat{a}\hat{a}^\dagger|n\rangle \\ &= \frac{1}{(n+1)}\langle n|(\hat{a}^\dagger\hat{a}+1)|n\rangle \\ &= \frac{1}{(n+1)}\langle n|(n+1)|n\rangle \\ &= \langle n|n\rangle = 1\end{aligned}$$

This innocent looking $\sqrt{n!}$ is at the heart of all of laser physics. In that context $\hat{a}^\dagger, \hat{a}$ correspond to operators creating and annihilating a single photon of energy ω. Thus any atomic physics process where atoms decay and give off a photon must have an interaction proportional to $\hat{a}^\dagger$; a process where photons are absorbed must have an interaction proportional to $\hat{a}$. So consider two matrix elements

$$\begin{aligned}\langle n+1|\hat{a}^\dagger|n\rangle &= \sqrt{n+1}\langle n+1|n+1\rangle \\ &= \sqrt{n+1} \\ \langle n-1|\hat{a}|n\rangle &= \sqrt{n}\langle n|n\rangle = \sqrt{n}\end{aligned}$$

The first, when squared to give the probability of the transition, gives a factor of (n+1) i.e. the atom decays even if n = 0, increasing the number of photons from 0 to 1. **But**, if there already exist 10^{15} photons in that mode, the probability increases by 10^{15} i.e. stimulated emission and possibly lasing. So later we will see that a free scalar field corresponds to

$$\hat{H} = \int \frac{d^3k}{(2\pi)^3}\hat{a}^\dagger(\underline{k})\hat{a}(\underline{k})$$

i.e. the Hamiltonian is a sum of Harmonic Oscillators, one for each momentum $\underline{k}$. The odd normalisation is due to us making relativistic invariance explicit.

0.2) Gaussian Integrals

This is the mathematical trickery necessary to establish Feynman diagram expansions in field theory without driving yourself crazy commuting operators past each other. The method was one of Feynman's favourite ways of doing integrals. We build up slowly, from simple 1 dimensional integrals, to integrals over infinite numbers of variables.

a) Compute:

$$I(\alpha) = \int_{-\infty}^{+\infty} dx e^{-\alpha x^2}$$

Trick

$$I^2 = \int_{-\infty}^{+\infty} dx e^{-\alpha x^2} . \int_{-\infty}^{+\infty} dy e^{-\alpha y^2} = \int_{-\infty}^{+\infty} dx dy e^{-\alpha(x^2+y^2)}$$

Now change to polar coordinates r, θ.

$$I^2 = \int_0^{+\infty} r dr \int_0^{2\pi} d\theta e^{-\alpha r^2} = 2\pi \frac{1}{2} \int_0^{+\infty} dr^2 e^{-\alpha r^2} = \frac{\pi}{\alpha} e^{-\alpha r^2} \Big|_0^{\infty} = \frac{\pi}{\alpha}$$

so that

$$I(\alpha) = \sqrt{\frac{\pi}{\alpha}}$$

b) Compute:

$$I(\alpha, \beta) = \int_{-\infty}^{+\infty} dx e^{-\alpha x^2 + \beta x} = \int_{-\infty}^{+\infty} dx e^{-\alpha(x - \frac{\beta}{2\alpha})^2 + \alpha \frac{\beta^2}{4\alpha^2}}$$

Change variables to $y = x - \frac{\beta}{2\alpha}$, then

$$I(\alpha, \beta) = I(\alpha) e^{\frac{\beta^2}{4\alpha}} = \sqrt{\frac{\pi}{\alpha}} e^{\frac{\beta^2}{4\alpha}}$$

c) Compute:

$$I_m(\alpha) = \int_{-\infty}^{+\infty} x^m e^{-\alpha x^2} dx = \left\{ \left(\frac{\partial}{\partial \beta} \right)^m I(\alpha, \beta) \right\} \Big|_{\beta=0}$$

since differentiating with respect to β inside the integral just brings down extra factors of x.

$$I_1 = \frac{\partial}{\partial \beta} \sqrt{\frac{\pi}{\alpha}} e^{\frac{\beta^2}{4\alpha}} \Big|_{\beta=0} = 0$$

This is trivially correct as the integrand is odd.

$$\begin{aligned} I_2 &= \frac{\partial^2}{\partial \beta^2} \sqrt{\frac{\pi}{\alpha}} e^{\frac{\beta^2}{4\alpha}} \Big|_{\beta=0} \\ &= \sqrt{\frac{\pi}{\alpha}} \frac{\partial}{\partial \beta} \left\{ \frac{\beta}{2\alpha} e^{\frac{\beta^2}{4\alpha}} \right\} \Big|_{\beta=0} \\ &= \sqrt{\frac{\pi}{\alpha}} \left\{ \frac{1}{2\alpha} e^{\frac{\beta^2}{4\alpha}} + \frac{\beta^2}{4\alpha^2} e^{\frac{\beta^2}{4\alpha}} \right\} \Big|_{\beta=0} \\ &= \frac{1}{2\alpha} \sqrt{\frac{\pi}{\alpha}} \end{aligned} \tag{0.2.1}$$

This example deserves close scrutiny. What we are doing is swapping powers of x in the integrand for derivitives with respect to a paramater in the exponent. This is absolutely characteristic of Functional Integrals. Moreover in equation (0.2.1) above, we very characteristically get two terms. The first comes from differentiating the term brought down from the derivative of the exponent leaving no β term in the product. The second comes from differentiating the exponent twice. This leaves β factors which vanish. This trivial remark is at the heart of the construction of the Feynman diagram expansion.

d) The Most Important Integral in the World:

$$L = \int \prod_{i=1}^{n} d\phi_i e^{-\sum_{i,j} \phi_i K_{ij} \phi_j + \sum_k J_k \phi_k} \tag{0.2.2}$$

We have switched to integration variables called ϕ_i. Later these will be the values of scalar fields and i will be a label picking space-time points. If you were a lattice person the i's would label sites in the lattice. The K_{ij} form an $n \times n$ symmetric matrix and the J_k a column vector with n entries. Amazingly the above integral can be done for all K_{ij}. The answer is

$$L = \frac{\pi^{\frac{n}{2}}}{\sqrt{det(K)}} e^{+\frac{1}{4}\sum_{i,j} J_i (K^{-1})_{ij} J_j}$$

This result is the simplest and most elegant way of devoloping Feynman diagrams. The $(K^{-1})_{ij}$ will be the Feynman propagator from space point i to space point j. To prove this formula we reduce it to n copies of $I(\alpha, \beta)$. Since K_{ij} is a real symmetric matrix we can diagonalise it by an orthogonal matrix U such that

$$U^{-1} K U = K'$$

where

$$K' = \begin{pmatrix} \lambda_1 & & & & \\ & \lambda_2 & & & \\ & & \lambda_3 & & \\ & & & \ddots & \\ & & & & \lambda_n \end{pmatrix}$$

Here $U^T = U^{-1}$, $detU = 1$ Now define

$$\phi_i = \sum_j U_{ij} {\phi'}_j \qquad or \qquad \phi = U\phi'$$

Then

$$L = \int \prod d{\phi'}_i e^{-\sum_j {\phi'}_j \lambda_j {\phi'}_j + (\sum_{k,j} J_k U_{kj}){\phi'}_j}$$

since the Jacobian from ϕ to ϕ' is det U = 1.

Thus we have n copies of $I(\alpha, \beta)$. These give factors of $\sqrt{\pi}$ n-times and $\sqrt{\lambda_i}$. Now $detK = detK' = \prod \lambda_i$ so we get the correct $detK$ factor. Finally in the exponent we get

$$J^T.U.{K'}^{-1}.U^T.J = J^T.K^{-1}.J$$

as is easily checked. The announced answer.

0.3) Pictures in Quantum Mechanics

Normally in elementary Quantum Mechanics we are given the operators $\hat{p}$ and $\hat{q}$ as $-i\hbar\frac{\partial}{\partial x}$ and x operating on wave functions $\Phi(x)$. The first hint that this may not be the most general way of thinking about things is the realisation that we could equally well talk about wave functions $\Psi(p)$. This led Dirac to introduce the idea of pictures.

Schrödinger Picture

Normal Quantum Mechanics has operators that are independent of time, such as $\hat{p}$ and $\hat{q}$ as above. Wave functions however depend on time through the Schrödinger equation

$$i\hbar\frac{\partial}{\partial t}\Psi(x,t) = \hat{H}(\hat{p},\hat{q})\Psi(x,t) \tag{0.3.1}$$

The formal solution of this is easily written out setting $\hbar = 1$

$$\Psi(x,t) = e^{-i\hat{H}t}\Psi(x,0) \tag{0.3.2}$$

If this wave function happens to be an eigenstate of $\hat{H}$ then the exponent becomes a simple numerical phase e^{-iEt}. So usually the easiest way to compute the time dependence of any state is to expand in energy eigenstates and then give each term in the expansion its appropriate phase.

Heisenberg Picture

Here we switch all the time dependence to the operators leaving time independent wave functions. Remember Heisenberg's version of Quantum Mechanics was called matrix mechanics i.e. all the dynamics was in the time dependence of matrices (= operators) not wave functions.

Define, for any Schrödinger operator $\hat{O}_S$, the equivalent Heisenberg operator $\hat{O}_H$ by

$$\hat{O}_H(t) = e^{i\hat{H}t}\hat{O}_S e^{-i\hat{H}t} \tag{0.3.3}$$

and the Heisenberg wave function

$$\Psi_H = e^{i\hat{H}t}\Psi_S(t) \tag{0.3.4}$$

where I have suppressed the coordinate dependence of the wave functions and only explicitly shown the all important time dependence.

The crucial result is that all physical quantities such as probabilities, matrix elements etc. are unchanged by the switch of pictures. For example let us calculate the average value of an operator $\hat{O}$ in a state Ψ in each picture

$$\begin{aligned} Average &= \int \Psi_H^*\hat{O}_H(t)\Psi_H dq \\ &= \int (e^{i\hat{H}t}\Psi_S)^*(e^{i\hat{H}t}\hat{O}_S e^{-i\hat{H}t})(e^{i\hat{H}t}\Psi_S)dq \\ &= \int \Psi_S^*(t)\hat{O}_S\Psi_S(t)dq \end{aligned}$$

The Heisenberg wave functions are now time independent. The factor in eqn (0.3.4) was clearly chosen to cancel the explicit time dependence coming from the solution (0.3.2) of the Schrödinger equation.

The dynamics is now all hiding in the time dependence of the operators. So we need the equation of motion of the Heisenberg operators. Since the Schrödinger operators are time independent this is just a matter of differentiating the definition of the Heisenberg operators.

$$\begin{aligned} i\frac{\partial}{\partial t}\hat{O}_H(t) &= i\frac{\partial}{\partial t}\{e^{i\hat{H}t}\hat{O}_S e^{-i\hat{H}t}\} \\ &= -\hat{H}\hat{O}_H + \hat{O}_H\hat{H} \\ &= [\hat{O}_H, \hat{H}] \end{aligned}$$

It is also easy to check that the commutation relations of any two operators are usually unchanged in the two pictures. Thus

$$[\hat{q}, \hat{p}] = i$$

is true whether in Heisenberg or Schrödinger pictures. In particular in later lectures we will use the fact that the $\hat{a}^\dagger$ and $\hat{a}$ commutation relations are unchanged in the two pictures.

0.4) Classical Mechanics: Newton implies Lagrange

In order to gain an understanding of what the theoretical physicists are up to, we need to go back and quickly understand the developments of Newton's equations due to Lagrange. More particularly to see what the point was! Newton insists on inertial coordinates. Lagrange says any will do!

Suppose we have a system of masses with coordinates $\underline{x}_i \quad i = 1, \cdots N$. We parametrise these by coordinates $q_1, q_2, \cdots, q_n$, so that

$$\underline{x}_i = \underline{f}_i(q_i, \cdots q_n, t)$$

Now Newton's equations say

$$m_i\ddot{\underline{x}}_i = \underline{F}_i$$

where the $\underline{F}_i$ is the force on the i'th mass. Dot each side with $\frac{\partial \underline{f}_i}{\partial q_r}$ and sum over i.

$$\sum_i m_i \frac{\partial \underline{f}_i}{\partial q_r}.\ddot{\underline{x}}_i = \sum_i \frac{\partial \underline{f}_i}{\partial q_r}.\underline{F}_i$$

For ease we use the Einstein convention that repeated indices are summed $\sum_r q_r q_r \equiv q_r q_r$. But

$$\begin{aligned} \frac{\partial \dot{\underline{x}}_i}{\partial \dot{q}_r} &= \frac{\partial}{\partial \dot{q}_r}\left[\frac{\partial f_i}{\partial q_l}\dot{q}_l + \frac{\partial f_i}{\partial t}\right] \\ &= \frac{\partial f_i}{\partial q_r} \\ &= \frac{\partial x_i}{\partial q_r} \end{aligned}$$

So

$$
\begin{aligned}
\ddot{\underline{x}}_i.\frac{\partial \underline{f}_i}{\partial q_r} &= \ddot{\underline{x}}_i.\frac{\partial \dot{\underline{x}}_i}{\partial \dot{q}_r} \\
&= \frac{d}{dt}\left[\dot{\underline{x}}_i.\frac{\partial \dot{x}_i}{\partial \dot{q}_r}\right] - \dot{\underline{x}}_i.\frac{d}{dt}\left(\frac{\partial x_i}{\partial q_r}\right) \\
&= \frac{d}{dt}\left[\dot{\underline{x}}_i.\frac{\partial \dot{x}_i}{\partial \dot{q}_r}\right] - \dot{\underline{x}}_i.\left\{\frac{\partial^2 \underline{x}_i}{\partial q_l \partial q_r}\dot{q}_l + \frac{\partial^2 \underline{x}_i}{\partial q_r \partial t}\right\} \\
&= \frac{d}{dt}\left[\dot{\underline{x}}_i.\frac{\partial \dot{x}_i}{\partial \dot{q}_r}\right] - \dot{\underline{x}}_i.\frac{\partial}{\partial q_r}\left\{\frac{\partial \underline{x}_i}{\partial q_l}\dot{q}_l + \frac{\partial \underline{x}_i}{\partial t}\right\} \\
&= \frac{d}{dt}\left(\frac{\partial}{\partial \dot{q}_r}\frac{1}{2}\dot{\underline{x}}_i{}^2\right) - \frac{\partial}{\partial q_r}\left(\frac{1}{2}\dot{\underline{x}}_i{}^2\right)
\end{aligned}
$$

Write $T = \frac{1}{2}\sum_i m_i \dot{\underline{x}}_i^2$. Therefore

$$
\frac{d}{dt}\left(\frac{\partial T}{\partial \dot{q}_r}\right) - \frac{\partial}{\partial q_r}T = \sum_i \frac{\partial \underline{f}_i}{\partial q_r}.\underline{F}_i
$$

Now

$$
\sum_i \frac{\partial \underline{f}_i}{\partial q_r}.\underline{F}_i \delta q_r
$$

is the work done in a δq_r shift for the other q's fixed, since

$$
\delta \underline{x}_i = \frac{\partial \underline{f}_i}{\partial q_r}.\delta q_r
$$

Several remarks are now in order

a) Many forces never do any work e.g. the tension in the string of a pendulum, reactions at fixed points. They never show in Lagrange's equations, thus usually simplifying the problem greatly.

b) If the forces are conservative then the work done can be written directly for $V(q_i \cdots q_n)$

$$
-\frac{\partial V}{\partial q_r}.\delta q_r
$$

So define L = T - V and Newton reduces to

Lagrange:

$$
\frac{d}{dt}\left[\frac{\partial L}{\partial \dot{q}_r}\right] - \frac{\partial L}{\partial q_r} = 0.
$$

These equations are true for any coordinates you like, such that fixing $q_1 \cdots q_n$ fixes the $\underline{x}_i$. These coordinates can be accelerating, rotating, whatever. Lagrange takes care of it all.

Consider a simple pendulum for small oscillations. For θ the angle away from the vertical, $l =$ length,

$$\begin{aligned} L &= T - V \\ &= \frac{1}{2}ml^2\dot{\theta}^2 - mg(1-\cos\theta)l \\ &\approx \frac{1}{2}ml^2\dot{\theta}^2 - mg\theta^2\frac{l}{2} \qquad for \qquad \theta \ll 1 \end{aligned}$$

Lagrange:

$$\frac{\partial}{\partial t}\left[ml^2\dot{\theta}\right] + mgl\theta = 0$$

$$\ddot{\theta} + \frac{g}{l}\theta = 0$$

Simple harmonic motion - solutions

$$\theta = A\cos(\omega t + \delta); \qquad \omega = \sqrt{\frac{g}{l}}$$

Notice, unlike the Newtonian analysis, we don't need to introduce the tension in the string, but we also learn nothing about it.

Summary

Lagrange is superior to Newton in that any coordinates can be used. The Lagrangian L = T - V is a function $L(q_i, \dot{q}_i)$ of the coordinates and their derivatives.

0.5) Dirac δ-function

We need later many simple properties of the famous Dirac δ-function. Here I rehearse and make explicit the results we will use.

I remind you that the δ-function is defined by

$$\int_{-\infty}^{+\infty} f(x).\delta(x)dx = f(0)$$

for all functions f. More complicated integrals are always performed by using standard variable changes to rewrite them in this form.

For example, in

$$I = \int f(x).\delta(ax)dx$$

we set $y = ax$ and

$$I = \int_{-\infty}^{+\infty} f(\frac{y}{a})\delta(y)\frac{dy}{|a|} = \frac{1}{|a|}f(0)$$

and

$$\delta(ax) = \frac{1}{|a|}\delta(x)$$

The more complicated case of

$$\int f(x)\delta(g(x))dx$$

where $g(x)$ is zero, once only, at $x = x_0$ is solved by setting $y = g(x)$, $y_0 = g(x_0) = 0$.

$$\int f(g^{-1}(y))\delta(y)\frac{dy}{|\frac{dg}{dx}|_{x=x_0}} = f(x_0)\frac{1}{|\frac{dg}{dx}|_{x=x_0}}$$

and

$$\delta(g(x)) = \frac{1}{|\frac{dg}{dx}|_{g=0}}.\delta(x - x_0)$$

and $g(x_0) = 0$.

Multi-dimensional integrals are done the same way.

$$H = \int f(x_1, x_2, \cdots, x_n)\prod_{i=1}^{n}\delta(A_{ij}x_j - B_i)\prod_i dx_i$$

Define $y_i = A_{ij}x_j$ so that $x_j = \left(A^{-1}\right)_{jk}y_k$

$$H = \int f(x_1, x_2, \cdots, x_n)\prod \delta(y_i - B_I)\frac{\prod dy}{|detA_{ij}|}$$

$$= \frac{1}{|detA_{ij}|}.f((A^{-1})_{jk}B_k)$$

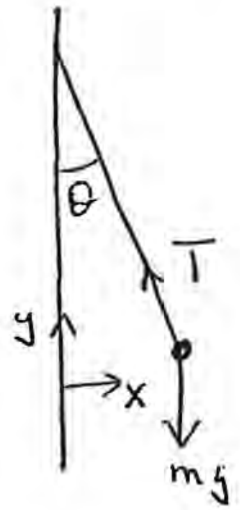
If f=1 notice this integral is independent of B ! This will be an important remark when we come to Faddeev-Popov.

Chapter 1

Classical Mechanics

In this brief chapter I would like to run through, without detailed proof, the status of Newton's equations versus Lagrange's methods and finally Hamilton's approach. These different methods already show effects that are of importance in Quantum Mechanics but there are important differences between the Classical and Quantum cases.

Since the Harmonic oscillator will play so large a rôle in our life for the next few days let me take it as the pedagogical example.



In Newton's method we need to take inertial coordinates i.e. coordinates where Newton's equations are correct. This excludes rotating coordinates which change the form of the equations from

$$m\ddot{x} = Force$$

Thus we are forced into coordinates such as (x, y) in the diagram. So the equations are, where t is the string tension,

$$m\ddot{x} = -t\sin\theta$$
$$m\ddot{y} = -mg + t\cos\theta$$

For small angle θ these reduce to

$$ml\ddot{\theta} = -t\theta$$
$$ml\frac{d^2}{dt^2}(1 - \frac{\theta^2}{2} + \cdots) = 0 = -mg + t \qquad ; ignoring \qquad \theta^2, etc.$$

Thus we get $\ddot{\theta} = -\frac{g}{l}\theta$, the standard equation for simple harmonic motion.

In Lagrangian form we can use any old coordinates and Lagrange's equations will sort out the mess. So for small oscillations in terms of the coordinate θ we have

$$T = \frac{1}{2}m(l\dot{\theta})^2$$
$$V = +\frac{1}{2}mgl(1-\cos\theta) \approx \frac{1}{2}mgl\theta^2$$

So Lagrange's equation, and notice we only have one at this stage, unlike Newton where we had two, becomes

$$\frac{d}{dt}\left\{\frac{\partial}{\partial\dot{\theta}}L\right\} - \frac{\partial}{\partial\theta}L = 0$$
$$\frac{d}{dt}\{ml^2\dot{\theta}\} - \{-mgl\theta\} = 0$$
$$ml^2\ddot{\theta} + mgl\theta = 0$$

The same as Newton !

This is a second order equation in time t. Hamilton was interested in obtaining two first order equations instead. Let me show you how this works in the simple case . I will then give a general proof.

The general idea is to define a new variable $p = \frac{\partial L}{\partial\dot{\theta}} = ml^2\dot{\theta}$. Then recognising this as the argument in the first term of Lagrange's equation we get, for free, that $\frac{d}{dt}p = \frac{d}{d\theta}L = -mgl\theta$. This then gives, as before, that $ml^2\ddot{\theta} = mgl\theta$. If we define $H = T + V = \frac{p^2}{2ml^2} + \frac{1}{2}mgl\theta^2$, then the equations of motion are, correctly,

$$\dot{p} = -\frac{\partial H}{\partial\theta} \qquad \dot{\theta} = \frac{\partial H}{\partial p}$$

This trick holds in general.

There is now a classic calculation which the thermodynamic whizzes among you should recognise. It is the equivalent of changing variables from V,S to V,T in going from the energy $dE = TdS - PdV$ to the free energy equation $dF = -SdT - PdV$. Here we go from $q, \dot{q}$ to q, p. So calculate, for r-coordinates q_r, and momenta $p_r = \frac{\partial L}{\partial\dot{q}_r}$,

$$\delta L = \sum_r\left\{\frac{\partial L}{\partial q_r}\delta q_r + \frac{\partial L}{\partial\dot{q}_r}\delta\dot{q}_r\right\}$$
$$= \sum_r\left\{\dot{p}_r\delta q_r + p_r\delta\dot{q}_r\right\}$$
$$= \delta\{\sum_r p_r\dot{q}_r\} + \sum_r\{\dot{p}_r\delta q_r - \dot{q}_r\delta p_r\}$$

So that, shuffling terms,

$$\delta\{-L + \sum p_r\dot{q}_r\} = \sum\dot{q}_r\delta p_r - \sum\dot{p}_r\delta q_r$$

This gives, as in thermodynamics,

$$\dot{q}_r = \frac{\partial H}{\partial p_r}$$
$$\dot{p}_r = -\frac{\partial H}{\partial q_r}$$

where $H = \sum_r p_r \dot{q}_r - L$. In many cases $H = T + V$ but this needs to be checked. In other words H is usually the total energy but not always.

These are the two first order equations which we promised. Notice H contains no $\dot{p}$ or $\dot{q}$ dependence. In particular the Hamiltonian $H(p_r, q_r)$ is now to be thought of as a function of p_r and q_r. The Lagrangian L is a function of $q_r, \dot{q}_r$. This is the classical reason for H appearing when we introduce p's and q's.

Now the big advantage of the Lagrangian method was it's ability to make use of any old coordinates. What is the equivalent statement for Hamiltonian systems ? Define for any two functions of the p's and q's, say $u(p,q)$ and $v(p,q)$, the Poisson brackets

$$\{u,v\} = \sum_r \left(\frac{\partial u}{\partial q_r}\frac{\partial v}{\partial p_r} - \frac{\partial u}{\partial p_r}\frac{\partial v}{\partial q_r} \right)$$

With this definition it is easy to see that

$$\{q_i, q_j\} = 0$$
$$\{q_i, p_j\} = \delta_{i,j}$$
$$\{p_i, p_j\} = 0$$

These should set bells off in your head as they look awfully like the q, p commutation relations. The statement of invariance of Hamilton's equations can now be simply stated. Given a change of variables from (q,p) to $Q(q,p), P(p,q)$ then Hamilton's equations remain invariant in form if and only if the Q, P have the same Poisson brackets as q, p. Such a transformation is called canonical. In his book Dirac claims that to Quantise any theory you just replace the Poisson brackets by $\frac{1}{i}[,]$. This is an enormously influential statement; which is unfortunately wrong. Using the above, any classical theory can be written using any canonically equivalent coordinates. But in general, if you use the Dirac prescription in one coordinate system, you do **not** get the result the Dirac prescription would predict for the other set of coordinates. In other words, although Classical Mechanics is invariant under Canonical Transformations, the equivalent quantised theories are not invariant under canonical changes due to operator ordering. This is a problem which has driven theorists mad for 70 years.

Chapter 2

Quantum Mechanics

The idea behind this section is to set up the two classic ways of computing perturbative expansions in Field Theory. Now Field Theory is just an example of Quantum Mechanics. So I will give the two formalisms in general Quantum Mechanics.

The physics behind these mathematical manipulations is actually rather complex and not well understood. As you gradually understand Quantum Field Theory you will slowly realise that it is an astonishingly complex mathematical structure. So either your understanding will go round and round in a convergent spiral or your head will spin.

The naïvest idea is that to first approximation, say in Quantum Electrodynamics, the electrons and photons can move about with little or no interaction. Thus it makes sense to split the Hamiltonian into two pieces. The first, soluble piece corresponds to free electrons and photons. We will see how to solve such a Hamiltonian in the next chapter. The interactions between them may then be treated as a small perturbation.

In the prerequisites I asked you to revise (or learn for the first time) the concept of a picture. In this chapter, in general Quantum Mechanics, I will introduce you to a third picture, the Dirac picture, which is explicitly defined to make perturbative calculations simple; well almost ! In the second section I will introduce you to the ideas of Feynman Path Integrals or Functional integrals; these, after you have tunnelled through a conceptual barrier, are the easy way of doing perturbative calculations. They are also the route which enables the Lattice people to simulate Quantum Field Theories on computers.

2.1) The Dirac Picture

We are in a Quantum Mechanical system where the Hamiltonian $\hat{H}$ is a sum of two parts; one soluble , one small.

$$\hat{H} = \hat{H}_0 + \hat{H}_I$$

Here $\hat{H}_I$ is usually called the interaction term. The Dirac picture is often called the interaction picture. The idea, starting from the Schrödinger picture, is to switch to the Heisenberg picture but only using the $\hat{H}_0$ term. Thus define

$$\begin{aligned}\hat{O}_I(t) &= e^{i\hat{H}_0 t}\hat{O}_S e^{-i\hat{H}_0 t}\\ &= e^{i\hat{H}_0 t}e^{-i\hat{H}t}\hat{O}_H(t)e^{i\hat{H}t}e^{-i\hat{H}_0 t}\\ &= U(t)\hat{O}_H(t)U^{-1}(t)\end{aligned}$$

I stress here that $\hat{O}_H$ is defined from the Schrödinger operator using the full Hamiltonian. The operator

$$\hat{U}(t) = e^{i\hat{H}_0 t}e^{-i\hat{H}t}$$

is crucial in what follows. Similarly we define for states

$$|a,t\rangle_I = e^{i\hat{H}_0 t}|a,t\rangle_S = U(t)|a\rangle_H$$

So we get the interaction picture from the Schrödinger operator by the free $\hat{H}_0$. Thus it satisfies

$$i\frac{\partial}{\partial t}\hat{O}_I(t) = [\hat{O}_I(t), \hat{H}_0]$$

Since $\hat{H}_0$ is soluble we can calculate this easily.

To calculate the Dirac picture operators we clearly need $U(t)$ so let us calculate its equation.

$$\begin{aligned} i\frac{\partial}{\partial t}U(t) &= -\hat{H}_0 e^{i\hat{H}_0 t}e^{-i\hat{H}t} + e^{i\hat{H}_0 t}e^{-i\hat{H}t}\hat{H} \\ &= e^{i\hat{H}_0 t}\hat{H}_I e^{-i\hat{H}t} \\ &= \hat{H}_I^I U(t) \end{aligned}$$

In the last eqution we easily see that $\hat{H}_I^I = \hat{H}_I(\hat{O}_I)$ i.e. the interaction Hamiltonian in the interaction picture is obtained by writing the interaction Hamiltonian in terms of interaction picture operators.

The crucial point is that this equation can easily be solved perturbatively. So we write

$$U(t) = 1 + U_1 + U_2 + U_3 + \cdots$$

where these terms are of order 0,1,2,3...in powers of the small $\hat{H}_I$. Substitute in the equation for U and compare equal powers of $\hat{H}_I$ on the two sides. The first term is clearly 1 since if $\hat{H}_I$ is 0 then U = 1.

$$i\frac{\partial}{\partial t}U_1 = \hat{H}_I(t)$$

Hence

$$U_1 = -i\int_0^t \hat{H}_I(t_1)dt_1$$

$$i\frac{\partial}{\partial t}U_2 = \hat{H}_I(t)U_1(t)$$

Hence

$$U_2(t) = (-i)^2\int_0^t dt_2\int_0^{t_2} dt_1 \hat{H}_I(t_2)\hat{H}_I(t_1)$$

You can guess the rest ?

Now let us massage this result into the standard form. Define the time ordered product of any two operators by

$$\begin{aligned} T(\hat{A}(t_1), \hat{B}(t_2)) &= \hat{A}(t_1)\hat{B}(t_2); t_1 > t_2 \\ &= \hat{B}(t_2)\hat{A}(t_1); t_2 > t_1 \end{aligned}$$

In general for many operators you move the earliest to the right, then the next earliest and so on. This has a beautiful effect, inside a time ordered expression we can permute operators in an arbitary way. The result is unchanged under such permutations. Notice in our expression for U that $t_1 < t_2$. So the integrand looks asymmetric. A much more symmetric way to write it is as follows

$$\begin{aligned} U_n &= (-i)^n \int_0^t dt_n \int_0^{t_n} dt_{n-1} \int_0^{t_{n-1}} dt_{n-2} \cdots \int_0^{t_2} dt_1 T(\hat{H}_I(t_n) \cdots \hat{H}_I(t_1)) \\ &= \frac{1}{n!} \int_0^t \prod_i dt_i T(\hat{H}_I(t_1)\hat{H}_I(t_2) \cdots \hat{H}_I(t_n)) \end{aligned} \tag{2.1.1}$$

These terms sum into an exponential

$$U = T\{exp(-i \int_0^t \hat{H}_I(t)dt)\} \tag{2.1.2}$$

To define this exponential we expand into the U_n terms. These are polynomials in $\hat{H}_I$ and we can apply the definition of the time-ordering T.

In the above we chose the Schrödinger and Heisenberg pictures to be equal at $t = 0$. This fixes the lower limit of the above integral (2.1.2) to be 0. It is often useful to be more general and fix them equal at $t = t_i$ the initial time and compute the $U(t_f, t_i)$ at the final time t_f. In (2.1.2) this merely changes the integral limits from 0,t to t_i, t_f.

In the next chapter we will use this formula extensively in the context of relativistic quantum field theory. I reiterate that in this case the physical model is of free physical particles which interact weakly. These interactions can then be treated as small perturbations. The theory itself will, of course, tell us whether this is internally consistent. Indeed many of the later lectures in QCD and the Salam Weinberg model will revolve around this problem.

2.2) Lagrangian Quantum Mechanics

In the preliminary reading for the course and the beginning of my lectures I stressed the importance of different views of dynamics. Above we have been very much concerned with the view of Quantum mechanics you were taught as undergraduates. Thus the equations are full of Hamiltonians and time dependence comes via Schrödinger equations. Unfortunately this is intrinsically non-relativistic in appearance. In special relativity space and time are supposed to be treated on an equal footing. This is impossible in a Hamiltonian approach. We need to switch back to Lagrangians. This means we need to address the problem of Lagrangian quantum mechanics. The pay-off will be a manifestly Lorentz symmetric formalism. In fact this is the chosen method for all problems with symmetries of any kind. Since Gauge symmetries dominate modern particle physics this is another reason for learning this method. The whole Faddeev Popov method comes from manipulation of Feynman integrals. I hasten to add that the Hamiltonian method, Schrödinger equation and all, is perfectly Lorentz symmetric but it is not **manifestly** symmetric. The Lagrangian methods solve the same equations and get the same answers. But the manifest imposition of symmetries often makes things easier to see.

I will develop the method first for a single dynamical degree of freedom q and its associated momentum p. Thus you should first understand these notes in this case. However if you now visualise q, p as column vectors for a finite number of degrees of freedom the proofs will be seen to be valid in this case also. Finally to reach field theory we need to make an intellectual leap and use the results for the infinite degrees of freedom implicit in Field theory. I will lead you up the garden path quite gently !

So, for the moment, consider a Quantum system with Hamiltonian

$$\hat{H}(\hat{p},\hat{q}) = \frac{\hat{p}^2}{2m} + V(\hat{q})$$

We would like to compute the amplitude for the particle to start at q_i at $t = t_i$ and move to q_f at time $t = t_f$. In the Schrödinger picture this is given by the amplitude $A = \langle q_f|e^{-i\hat{H}t}|q_i\rangle$ where $|q\rangle$ is the time independent eigenstate of position . Thus, in words, we start in the position eigenstate at $t = 0$, propagate in time for a time $t = t_f - t_i$ through the exponent and finally compute the overlap with the time independent final eigenstate q_f. The tricky bit is calculating the exponent. The Feynman trick is to split it into a lot of little steps. To each such step we can then apply perturbation theory. Thus write

$$e^{-i\hat{H}t} = e^{-i\hat{H}\Delta}.e^{-i\hat{H}\Delta}.\cdots.e^{-i\hat{H}\Delta}$$

with n terms in the product and $\Delta = \frac{(t_f - t_i)}{n}$. Then $\hat{H}\Delta$ is small if we take n large enough; and we will eventually let $n \to \infty$. We write

$$A = \langle q_f|e^{-i\hat{H}\Delta}.e^{-i\hat{H}\Delta}.\cdots e^{-i\hat{H}\Delta}|q_i\rangle$$

Now insert, many times, the Quantum Mechanical representations of 1

$$\int_q dq|q\rangle\langle q| = 1$$
$$\int_p dp|p\rangle\langle p| = 1$$

where $|q\rangle$, $|p\rangle$ are the complete sets of position and momentum eigenstates.

$$\int_{q_i,p_i} \langle q_f|p_n\rangle\langle p_n|e^{-i\hat{H}\Delta}|q_{n-1}\rangle\langle q_{n-1}|p_{n-1}\rangle\langle p_{n-1}|e^{-i\hat{H}\Delta}|q_{n-2}\rangle\cdots|q_1\rangle\langle q_1|p_1\rangle\langle p_1|e^{-i\hat{H}\Delta}|q_i\rangle$$

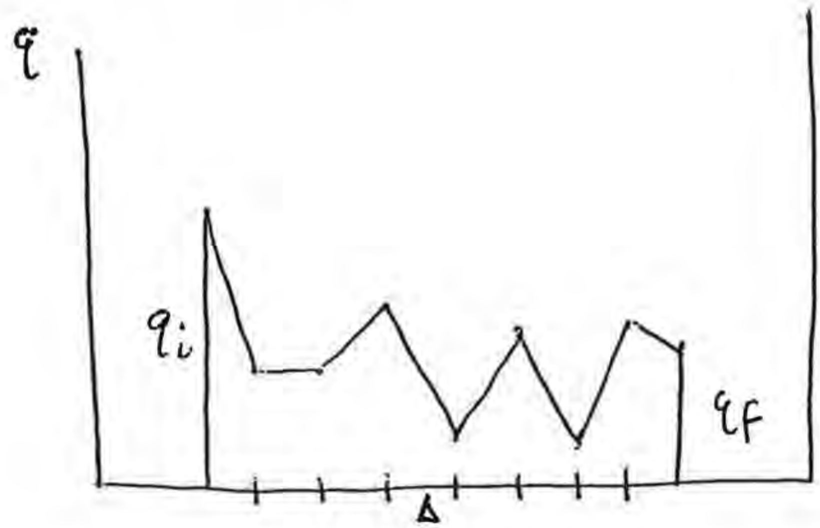
Here q_i, p_i should be thought of as the position and momentum after time $i \times \Delta$. Now $\langle q_n|p_n\rangle = e^{iq_np_n}$ (c.f. $\langle x|\psi\rangle = \Psi(x)$). Thus we can rewrite

$$\begin{aligned}\langle p_n|e^{-i\hat{H}\Delta}|q_{n-1}\rangle &= \langle p_n|(1 - i\hat{H}(\hat{p},\hat{q})\Delta)|q_{n-1}\rangle \\ &= \langle p_n|(1 - iH(p_n,q_{n-1})\Delta)|q_{n-1}\rangle \\ &= e^{-iH(p_n,q_{n-1})\Delta}.e^{-ip_n.q_{n-1}}\end{aligned}$$

Thus, substituting this in the expression for A, we get

$$
\begin{aligned}
A &= \int_{q_0=q_i, q_n=q_f} \prod_1^n dp_i \prod_1^{n-1} dq_i \left\{ e^{-i\sum H(p_j, q_{j-1})\Delta} . e^{-i\sum q_{j-1} . p_j} . e^{i \sum q_j . p_j} \right\} \\
&= \int \prod_1^n dp_i \prod_1^{n-1} dq_i \left[e^{i \sum \Delta \left[\frac{(q_j - q_{j-1})p_j}{\Delta} - H(p_j, q_{j-1}) \right]} \right] \qquad (2.2.1) \\
&= \int [dpdq] e^{i \int dt [p\dot{q} - H]}
\end{aligned}
$$

The last line comes from the approximation that $\frac{q_n - q_{n-1}}{\Delta}$ becomes $\dot{q}$ in the limit $\Delta \to 0$. The exponent is now the time integral of the Lagrangian. This is the Action. Thus the Lagrangian appears in Quantum Mechanics. The integrals correspond to integrating over all the paths connecting q_i with q_f.



Those of you wide awake should be saying "Hey, you said the Lagrangian was a function of $q, \dot{q}$ not q, p !". So we need one last trick. Rather than give you a general proof, let me consider a simple case. Assume $H = \frac{p^2}{2m} + V(q)$ i.e. a simple problem of a particle moving in a potential V. Then the p-integrals above are Gaussian.

$$
\int \prod dp_i e^{-i \frac{\Delta p_i{}^2}{2m}} . e^{ip_i(q_i - q_{i-1})} \approx e^{\frac{-i(q_i - q_{i-1})^2 m}{2\Delta}} \qquad by (0.2b)
$$

The p_i integral is trivially performed by replacing p_i by $\frac{(q_i - q_{i-1}).m}{\Delta} = \dot{q}_i.m$ in the Lagrangian. This is the Classical prescription in this simple case. Thus finally we get the Lagrangian expression for our amplitude

$$
A = \int [dq] e^{i \int dt L(q, \dot{q})} \qquad (2.2.2)
$$

We will use this formula extensively in Quantum Field Theory. It is the basic Quantum Mechanical input into the Functional method. In terms of Quantum Mechanics QED or QCD are just special choices of L. The lattice people spend all their lives trying to do these functional integrals numerically. Next we will show how this is a brilliant formalism for discussing Lorentz and Gauge symmetries. So this is the crucial modern starting point for all the discussions of Gauge fixing; the Faddeev Popov trickery. We will return to this.

2.3) Heisenberg Harmonic Oscillator's

In the pre-school exercises I asked you to make sure that you could solve the Harmonic Oscillator in what we now know as the Schrödinger representation. Here for pedagogical purposes and to ease your way into Field theory I solve it in the Heisenberg picture. So we have a Hamiltonian $\hat{H} = \frac{\hat{p}^2}{2m} + \frac{m\omega^2}{2}\hat{q}^2$. In the Heisenberg picture we have two time-dependent operators $\hat{q}(t)$ and $\hat{p}(t)$. These must satisfy the Heisenberg equations of motion

$$\begin{aligned} i\frac{\partial}{\partial t}\hat{q}(t) &= [\hat{q}(t), \hat{H}] \\ i\frac{\partial}{\partial t}\hat{p}(t) &= [\hat{p}(t), \hat{H}] \end{aligned} \tag{2.3.1}$$

Given the commutation relation $[\hat{q}, \hat{p}] = i$, which is unaffected by the switch to the Heisenberg picture, we easily see

$$\begin{aligned} i\frac{\partial}{\partial t}\hat{q}(t) &= \frac{1}{m}i\hat{p} \\ i\frac{\partial}{\partial t}\hat{p}(t) &= m\omega^2(-i\hat{q}) \end{aligned}$$

Differentiate one of these and use the other gives

$$\begin{aligned} \frac{\partial^2}{\partial t^2}\hat{q}(t) &= -\omega^2\hat{q}(t) \\ \frac{\partial^2}{\partial t^2}\hat{p}(t) &= -\omega^2\hat{p}(t) \end{aligned}$$

Although these are operator equations our true and tried methods work.

$$\begin{aligned} \hat{q}(t) &= \frac{1}{2m\omega}(e^{i\omega t}\hat{A}^\dagger + e^{-i\omega t}\hat{A}) \\ \hat{p}(t) &= \frac{1}{2m\omega}(im\omega e^{i\omega t}\hat{A}^\dagger - im\omega e^{-i\omega t}\hat{A}) \end{aligned}$$

are solutions for $\hat{A}, \hat{A}^\dagger$ independent of time. Solving for $\hat{A}^\dagger, \hat{A}$ we get

$$\begin{aligned} \hat{A} &= \sqrt{\frac{1}{2}}(\sqrt{m\omega}\hat{q} + i\frac{\hat{p}}{\sqrt{m\omega}}) \\ \hat{A}^\dagger &= \sqrt{\frac{1}{2}}(\sqrt{m\omega}\hat{q} - i\frac{\hat{p}}{\sqrt{m\omega}}) \end{aligned}$$

It is easy to check that these have the same commutation relations as the $\hat{a}, \hat{a}^\dagger$ of the prerequisites. Hence they have the same eigenvalues and eigenstates.

Chapter 3

Free Boson or Scalar Field Theory

The time has come to raise the stakes. We have been studying general Quantum Mechanics so far. Now we pick the special Lagrangians and Hamiltonians designed to give relativistically invariant field Theories. We will see that these provide exactly the language to describe particle production as you see at LEP. They will also have manifest symmetry properties. Altogether they provide the standard theoretical language to understand our world. I stress these results all come from standard Quantum Mechanics.

As always the first thing to get clear, even before writing down the Hamiltonian, is the set of independent degrees of freedom. In other words for which variables will we have to solve Heisenberg's equations ? In field theory the variables are the values of the fields at the different space points $\underline{x}$. You are used to this idea from Maxwell's equations where the values of the electric and magnetic field form the dynamical variables. In Quantum Mechanics we have to decide which variables turn into operators which then satisfy Heisenberg's equations of motion. Thus we have variables, for a single scalar field, $\phi(\underline{x}, t)$.

3.1) Classical treatment

First let us treat the problem classically. Then ϕ is a single real variable at each space-time point. Thus ϕ at any given point is the direct analogue of q for a single point particle. We expect a Kinetic energy of

$$T = \int d^3x \left(\frac{\partial\phi}{\partial t}\right)^2 \qquad c.f. \frac{\dot{q}^2}{2m}$$

The potential energy will have two terms

$$V = \frac{\mu^2}{2}\int d^3x \phi(\underline{x}, t)^2 \qquad c.f. \frac{1}{2}m\omega^2 q^2$$
$$+ \int d^3x \sum_i \frac{c^2}{2}\left(\frac{\partial\phi}{\partial x^i}\right)^2$$

with constants μ, c so far undefined physically. The gradient terms are necessary to enforce Lorentz invariance. Thus we have the Lagrangian

$$L = \int d^3x \left[\frac{1}{2}\left(\frac{\partial\phi}{\partial t}\right)^2 - \sum_i \frac{c^2}{2}\left(\frac{\partial\phi}{\partial x_i}\right)^2 - \frac{\mu^2}{2}\phi^2(\underline{x}, t)\right]$$

Then the action S can be written

$$S = \int dt L = \int d^3x dt \mathcal{L}(\underline{x}, t)$$

where the Lagrangian density $\mathcal{L}$ can easily be written down from above. Now classically we have a Lagrangian density that is a function of the dynamical variables ϕ, their time derivatives $\frac{\partial \phi}{\partial t}$ but also, unexpectedly, their spatial derivatives $\frac{\partial \phi}{\partial x_i}$ for each $i = 1, 2, 3$. Thus we must redo, say the Stationary action derivation of Lagrange, to obtain

$$\frac{\partial}{\partial t}\left[\frac{\partial \mathcal{L}}{\partial \dot{\phi}}\right] + \frac{\partial}{\partial x^i}\left[\frac{\partial \mathcal{L}}{\partial(\frac{\partial \phi}{dx^i})}\right] - \frac{\partial \mathcal{L}}{\partial \phi} = 0$$

Which in the case of our Lagrangian density gives

$$\frac{\partial^2 \phi}{\partial t^2} - c^2 \frac{\partial^2 \phi}{\partial {x_i}^2} + \mu^2 \phi = 0$$

Several comments are in order at this point. We use the Einstein summation convention so that the repeated i indices are summed. The above equation is Lorentz invariant if ϕ is a Lorentz scalar. Indeed if $\mu = 0$ this is the wave equation for light with c the velocity of light.

The general solution is not hard to write down. The expected result is that any solution will be a sum of plane waves. Notice the equation is linear in ϕ so that, given any two solutions, any linear combination is also a solution. Think of light.

So try a plane wave solution for ϕ.

$$\phi(\underline{x}, t) = A e^{i(\underline{k}.\underline{x} - \omega(\underline{k})t)}$$

Substituting in the equation of motion gives

$$A\left[-\omega^2(\underline{k}) + c^2 k^2 + \mu^2\right] = 0$$

In order for the solution to be non trivial we must have $\omega(\underline{k}) = \pm\sqrt{\mu^2 + c^2 \underline{k}^2}$. From now on ω will stand for the positive root of this equation. Then the general solution will be given, by superposition, as

$$\phi(\underline{x}, t) = \int \frac{d^3 k}{(2\pi)^3 . 2\omega(\underline{k})}\left[a(\underline{k}) e^{i(\underline{k}.\underline{x} - \omega t)} + a^*(\underline{k}) e^{-i(\underline{k}.\underline{x} - \omega t)}\right]$$

The factors of 2π and ω are conventional, but make Lorentz invariance manifest later. For the moment just think of them as factors extracted from a, a^*. The fact that ϕ is a real valued variable is guaranteed by a^* being the complex conjugate of a.

To switch to Hamiltonian methods we need to compute the momentum conjugate to $\phi(\underline{x}, t)$.

$$\Pi(\underline{x}, t) = \frac{\partial \mathcal{L}}{\partial \dot{\phi}(\underline{x}, t)} = \dot{\phi}(\underline{x}, t)$$

Writing this in terms of the a's we get

$$\Pi(\underline{x}, t) = \int \frac{d^3 k}{2\pi^3 . 2\omega(\underline{k})}\left[-i\omega a(\underline{k}) e^{i(\underline{k}.\underline{x} - \omega t)} + i\omega a^* e^{-i(\underline{k}.\underline{x} - \omega t)}\right]$$

Then we can easily write

$$\hat{H} = \int d^3y \left[\frac{\hat{\Pi}^2(\underline{y},t)}{2} + \frac{1}{2}\left(\frac{\partial \hat{\phi}}{\partial y}^2\right) + \frac{\mu^2}{2}\hat{\phi}^2 \right] \tag{3.1.1}$$

Now you can quantise in your sleep. We have the Hamiltonian in standard form written in terms of coordinates $\phi(\underline{x},t)$ and their conjugate momenta $\Pi(\underline{x},t)$.

3.2) Quantum Mechanics

Now we switch to Quantum Mechanics. As always the classical variables turn into operators. So corresponding to the classical single particle $\hat{q}$ we have an operator at each point of space time $\hat{\phi}(\underline{x})$. In the Schrödinger representation these will be time independent; in the Heisenberg representation they will be time dependent. Corresponding to the single particle $\hat{p}$ we will have a momentum corresponding to each $\phi(\underline{x})$. I stress again that the dynamical variables here are the values at the points $\underline{x}$ not the variables $\underline{x}$. Thus we have the momenta $\hat{\Pi}(\underline{x},t)$. Then directly copying your Undergraduate Quantum course we have the commutation relations.

$$\begin{aligned} [\hat{\phi}(\underline{x},t), \hat{\phi}(\underline{y},t)] &= 0 \\ [\hat{\Pi}(\underline{x},t), \hat{\phi}(\underline{y},t)] &= -i\delta^3(\underline{x}-\underline{y}) \\ [\hat{\Pi}(\underline{x},t), \hat{\Pi}(\underline{y},t)] &= 0 \end{aligned} \tag{3.2.1}$$

In other words the variables at different space points all commute, the only non-zero commutator is between a variable and its momentum at the same point. The normalisation is not obvious at this point, but we will see how natural it is later. Clearly the numerical factors can be changed by scaling ϕ. Here, because of the manifest time dependence, we have used Heisenberg operators. We showed before that commutation relations at equal times are unaffected by the switch of pictures.

Now we solve the Heisenberg equations of motion for the time dependence of these variables.

$$\begin{aligned} i\dot{\hat{\phi}}(\underline{x},t) &= [\hat{\phi}(\underline{x},t), \hat{H}] \\ &= \int d^3y [\hat{\phi}(\underline{x},t), \frac{\hat{\Pi}^2(\underline{y},t)}{2}] \\ &= \int d^3y [\hat{\phi}, \hat{\Pi}]\hat{\Pi}(\underline{y},t) \\ &= \int d^3y i\delta^3(\underline{x}-\underline{y})\hat{\Pi}(\underline{y},t) \\ &= i\hat{\Pi}(\underline{x},t) \end{aligned}$$

$$
\begin{aligned}
i\dot{\hat{\Pi}} &= \left[\hat{\Pi}(\underline{x},t),\hat{H}\right] \\
&= \int d^3y\left[\hat{\Pi}(\underline{x},t),\frac{\partial\hat{\phi}(\underline{y},t)}{\partial y}\right]\frac{\partial\hat{\phi}}{\partial y} + \mu^2\int d^3y\left[\hat{\Pi}(\underline{x},t),\hat{\phi}(\underline{y},t)\right]\hat{\phi}(\underline{y},t) \\
&= \int d^3y\left\{-i\frac{\partial}{\partial y}\delta^3(\underline{x}-\underline{y}).\frac{\partial\hat{\phi}}{\partial y} - i\mu^2\hat{\phi}(\underline{x},t)\right\} \\
&= i\frac{\partial^2\hat{\phi}}{\partial y^2} - i\mu^2\hat{\phi}(\underline{x},t)
\end{aligned}
$$

So if we put these together we get the equations for $\hat{\phi}$ and $\hat{\Pi}$ separately.

$$
\frac{\partial^2\hat{\phi}}{\partial t^2} = \nabla^2\phi - \mu^2\hat{\phi}^2
$$

and

$$
\hat{\Pi}(\underline{x},t) = \dot{\hat{\phi}}(\underline{x},t)
$$

You often hear nonsense, particularly in elementary field theory books, about the simple Schrödinger equation being changed into a relativistic wave equation. I stress here that this relativistic equation has been derived directly from the Heisenberg equation, in the Heisenberg picture. It is completely equivalent to the usual Schrödinger equation. What has changed is the Hamiltonian. Quantum mechanics is unchanged.

We can solve these operator equations exactly as in the classical case. Firstly, they are linear equations, so superpositions of solutions are solutions. We get, again with a funny choice of normalisation of the coefficients,

$$
\begin{aligned}
\hat{\phi}(\underline{x},t) &= \int\frac{d^3k}{(2\pi)^3.2\omega(\underline{k})}\left[\hat{a}(\underline{k})e^{i(\underline{k}.\underline{x}-\omega t)} + \hat{a}^\dagger(\underline{k})e^{-i(\underline{k}.\underline{x}-\omega t)}\right] \\
\hat{\Pi}(\underline{x},t) &= \int\frac{d^3k}{(2\pi)^3.2}\left[-i\hat{a}(\underline{k})e^{i(\underline{k}.\underline{x}-\omega t)} + i\hat{a}^\dagger(\underline{k})e^{-i(\underline{k}.\underline{x}-\omega t)}\right]
\end{aligned}
\tag{3.2.2}
$$

Given the commutation relations for the ϕ and Π we can compute those for the a and $a^\dagger$. Then we obtain

$$
\begin{aligned}
\left[\hat{a}(\underline{k}),\hat{a}(\underline{k}')\right] &= 0 \\
\left[\hat{a}^\dagger(\underline{k}),\hat{a}^\dagger(\underline{k}')\right] &= 0 \\
\left[\hat{a}(\underline{k}),\hat{a}^\dagger(\underline{k}')\right] &= (2\pi)^3.2\omega.\delta^3(\underline{k}-\underline{k}')
\end{aligned}
\tag{3.2.3}
$$

This is the result we are after. We have an infinite set of Harmonic Oscillators. For different $\underline{k}$ they commute. The next move is to compute the Hamiltonian. This is easily (well actually lengthily and tediously) proved, by substituting for $\hat{\phi}$ and $\hat{\Pi}$, to be given by

$$
\hat{H} = \int\frac{d^3k}{2(2\pi)^3}\hat{a}^\dagger(\underline{k})\hat{a}(\underline{k}) + Constant \tag{3.2.4}
$$

The Hamiltonian is a sum of independent commuting Harmonic Oscillators. The vacuum state is the state where all the oscillators are in their ground states. The excited states are obtained by applying the raising operators one or more at a time. For example $\hat{a}^\dagger(\underline{k})$ creates a particle of momentum $\underline{k}$ and energy $\hbar\omega(\underline{k})$. We can easily check that

$$\hat{H}\hat{a}^\dagger(\underline{p})|0\rangle = \omega(\underline{p})\hat{a}^\dagger(\underline{p})|0\rangle$$

The Hilbert space of the theory is given by the states $|0\rangle$, the vacuum; $\hat{a}^\dagger(\underline{k})|0\rangle$, the one particle states; $\hat{a}^\dagger(\underline{k})\hat{a}^\dagger(\underline{p})|0\rangle$, the two particle states; and so on. The commutation relations incidentally guarantee the two particle states are even under exchange i.e. the particles are guaranteed to be Bosons.

In the above Lorentz invariance was a bit hidden. We ended up with a nice covariant equation for ϕ but this felt rather accidental. The Heisenberg equations clearly treat $\underline{x}$ and t differently. The resolution for this lies in the Lagrangian method. So let us compute the Lagrangian or, in fact, the Action

$$Action = \int d^4x\left[\frac{1}{2}(\partial^\mu\phi)(\partial_\mu\phi) - \frac{\mu^2}{2}\phi^2\right]$$

This is manifestly Lorentz invariant. The measure d^4x is and so is the Lagrangian density. Thus the Functional method will start with a big advantage. Already, before doing anything, the formalism looks invariant. Contrast this with normal Quantum Mechanics where both the Schrödinger and Heisenberg equations treat time very differently to space coordinates. This would lead you to think they could not be invariant. They are invariant but not manifestly. This can lead to tedious, apparently non Lorentz invariant calculations, which mysteriously come right at the end. An example is the Heisenberg quation for ϕ which came out as the wave equation, eventually.

Before leaving scalars let me generalise slightly and introduce a pair of scalars. These will be necessary in the Salam-Weinberg theory. So we write down

$$\mathcal{L} = \sum_{r=1}^{2}\left[\frac{1}{2}(\partial^\mu\phi_r)(\partial_\mu\phi_r) - \frac{\mu^2}{2}\phi_r^2\right]$$

which involves two independent scalr fields ϕ_1 and ϕ_2. For the purposes of gauge invariance it is convenient to also have a formalism where we instead have two complex valued fields defined by

$$\chi = \frac{1}{\sqrt{2}}(\phi_1 - i\phi_2)$$
$$\chi^\dagger = \frac{1}{\sqrt{2}}(\phi_1 + i\phi_2)$$

An easy exercise is to check that

$$\mathcal{L} = \partial_\mu\chi^\dagger\partial_\mu\chi - \mu^2\chi^\dagger\chi$$

Then we can write down Heisenberg's equations as before and solve them. We find

$$\hat{\chi} = \int\frac{d^3x}{(2\pi)^3.2\omega}\left(\hat{b}(\underline{k})e^{-ik.x} + \hat{d}^\dagger(\underline{k})e^{ik.x}\right)$$

Since χ is not real, we have two independent operators $\hat{b}$ and $\hat{d}$. The Hermitian conjugate field is obtained by Hermitian conjugation.

The commutation relations can also be computed. The only non-zero terms come from

$$[\hat{b}(\underline{k}), \hat{b}^\dagger(\underline{q})] = 2\omega(2\pi)^3\delta^3(\underline{k}-\underline{q})$$
$$[\hat{d}(\underline{k}), \hat{d}^\dagger(\underline{q})] = 2\omega(2\pi)^3\delta^3(\underline{k}-\underline{q})$$

and, most importantly, the Hamiltonian is given by

$$\hat{H} = \int \frac{d^3k}{(2\pi)^3}\left(b^\dagger(\underline{k})b(\underline{k}) + d^\dagger(\underline{k})d(\underline{k})\right)$$

So again we have an independent set of oscillators. An interesting problem, is to determine the electric charge carried by each particle.

Chapter 4

Interacting Bosons

In the previous chapter we discussed the Quantum Mechanics of the Free Scalar Boson. This was a theory that was Lorentz invariant and corresponded to a set of free non-interacting particles. This means that it is a very boring theory. Particles never interact. No particles are created. In this chapter we introduce the interactions. We will do this first in the mathematically simplest theory of interacting scalar particles. You can think of this as a theory of Higgs particles, if you like. Later we will consider more realistic theories.

We will consider, for pedagogical reasons, a theory with a single scalar ϕ and a complex pair $\chi, \chi^\dagger$. So the Lagrangian density will be given by

$$\mathcal{L} = \mathcal{L}_\phi + \mathcal{L}_\chi + \mathcal{L}_{int}$$

where the interaction term is given by $\mathcal{L}_{int} = -g\hat{\chi}^\dagger\hat{\phi}\hat{\chi}$. This is Lorentz invariant since each field is a scalar. It is not hard to check that the Heisenberg equations of motion are

$$((\partial)^2 + \mu^2)\hat{\phi} + g\hat{\chi}^\dagger\chi = 0$$
$$((\partial)^2 + \mu^2)\hat{\chi} + g\hat{\phi}\hat{\chi} = 0$$

These are horrible non-linear operator equations. A Nobel prize for any solution. Only two ways to get information from this are known to man or woman. One is to assume the interaction term is small. The other is to put this on a lattice and do the functional integrals by brute force computing. My mission here is to explain perturbation theory. By being ingenious theorists have shown this is a correct move in a surprising number of cases. In fact it works much better than we have any right to expect.

So I will prove first, that the interaction term will lead to scattering, production and absorption of new particles. All the time the formalism will keep energy and momentum conservation correct.

I will first do a simple case using operators and an obvious Quantum Mechanical formalism. Then I will redo the calculation in the much slicker Functional formalism.

4.1) Feynman Diagrams from Operators

So, in Quantum Mechanics, in the Heisenberg picture, the crucial object to calculate is the $\hat{U}$ operator of Chapter 2. In lowest order of perturbation theory it is given by

$$\begin{aligned}\hat{U}(t_i, t_f) &= -i\int_{t_i}^{t_f} \hat{H}_I(t)dt \\ &= -ig\int d^4x\hat{\chi}^\dagger\hat{\phi}\hat{\chi}\end{aligned}$$

This, remember, is the operator which will propagate a t_i state to a t_f state. We apply it to the decay of a ϕ into a $\chi\chi^\dagger$ pair. For this to be kinematically possible, we must have the ϕ mass at least twice the χ mass. Then we take $t_i = -\infty$ and $t_f = +\infty$.

The states are therefore given by

$$|t = -\infty\rangle = \hat{a}^\dagger(\underline{k})|0\rangle$$
$$|t = +\infty\rangle = \hat{b}^\dagger(\underline{p})\hat{d}^\dagger(\underline{q})|0\rangle$$
$$\langle t = \infty| = \langle 0|\hat{d}(\underline{q})\hat{b}(\underline{p})$$

Being Heisenberg physical states they have no time dependence. So we need to calculate the matrix element

$$\langle t = \infty|\hat{U}(\infty, -\infty)|t = -\infty\rangle = -ig\langle|\hat{d}(\underline{q})\hat{b}(\underline{p})\left(\int d^4x\hat{\chi}^\dagger\hat{\phi}\hat{\chi}\right)\hat{a}(\underline{k})|0\rangle$$

substituting from above.

$$= -ig\langle 0|\hat{d}(\underline{q})\hat{b}(\underline{p})\int d^4x\int \not{d}p'\int \not{d}k'\int \not{d}q'\{\hat{d}(\underline{p}')e^{-ip'x} + \hat{b}^\dagger(\underline{p}')e^{-ip'.x}\}$$
$$\{\hat{a}(\underline{k}')e^{-ik'.x} + \hat{a}^\dagger(\underline{k}')e^{ik'.x}\}\{\hat{b}(\underline{q}')e^{-iq'.x} + \hat{d}^\dagger(\underline{q}')e^{iq'.x}\}\hat{a}^\dagger(\underline{k})|0\rangle$$

Now remember that an annihilation operator acting on $|0\rangle$ gives zero as does a creation operator acting on $\langle 0|$. Thus we can throw away the $\hat{a}^\dagger$, $\hat{b}$ and $\hat{d}$ terms. We have also introduced the Imperial notation $\not{d}^3k$ for $\frac{d^3k}{(2\pi)^3.2\omega}$ to save writing.

$$= -ig\langle 0|\hat{d}(q)\hat{b}(p)\int d^4x\int \not{d}p'\ \not{d}k'\ \not{d}q'\hat{b}^\dagger(\underline{p}')\hat{a}(\underline{k}')\hat{d}^\dagger(\underline{q}')\hat{a}^\dagger(\underline{k})|0\rangle e^{ix.(p'-k'-q')}$$

Now we have a creation and an annihilation operator for a,b,d. The only non-zero contribution comes from the right hand side of each commutator. Finally commuting all the terms past until they annihilate the vacuum we get

$$= -ig\delta^4(p - k - q)\langle 0|0\rangle.(2\pi)^4$$

The δ^4 contains 4-momentum conservation. This is our first real perturbative calculation. The method is general. Write down the initial and final states in terms of creation and annihilation operators. Write down the relevant term in $\hat{U}$. Start commuting all annihilation operators to the right. When they reach $|0\rangle$ they give zero. When a creation operator reaches the left and hits $\langle 0|$ it too gives zero. This is a finite procedure. It clearly needs systematic organisation. In operator formalism this goes under the name of Wick's theorem. We will duck this and use a much slicker Functional proof.

However, before we leave operators, I would like to present the operator formulation of the propagator. In the Interaction picture we saw how we obtained expressions of the form

$$\int dt_1dt_2\cdots dt_nT(H_I(t_1)H_I(t_2)H_I(t_3)\cdots)$$

Using the above trick we see that a crucial object will be the difference between two time ordered operators as appear in U and the so-called normal ordered form where we carry out the above procedure and commute all annihilation operators to the right of all creation operators.

Let us investigate this for two scalar fields. The time ordered form

$$T(\phi(x)\phi(y))$$

is quadratic in creation and annihilation operators. Moving all the annihilation operators to the right gives a normal ordered form plus non-operator terms which depend on x and y. Let us compute the non-operator term. Since it does not depend on any operator it is most easily extracted by taking the vacuum expectation value.

$$\langle 0|T\big(\phi(x)\phi(y)\big)|0\rangle$$

The normal ordered terms all give zero leaving the constant term.

$$\langle 0|T\big(\phi(x)\phi(y)\big)|0\rangle = i\Delta_F(x,y)$$

by definition of the Feynman propagator. Let us calculate it.

$$i\Delta_F = \langle 0| \int \not{d}k \int \not{d}q \big(\hat{a}(\underline{k})e^{i(\underline{k}.\underline{x}-\omega t)}\big)\big(\hat{a}^\dagger(\underline{q})e^{-i(\underline{q}.\underline{x}-\omega\tau)}\big)|0\rangle$$

The missing terms in the expansions of ϕ all give zero either on the initial or final vacuum. The two times are called t and τ. We assume $t \geq \tau$. Now commute the two terms past each other.

$$i\Delta_F(x,y) = \int \not{d}k \ \not{d}q (2\pi)^3 2\omega.\delta^3(\underline{k}-\underline{q}).e^{i\underline{k}.(\underline{x}-\underline{y})-i(t-\tau)\omega}$$

The delta function lets us do one of the momentum integrals. Notice that the above calculation assumed $t \geq \tau$ to put the operators in the above order. Thus we can write in general for any times

$$i\Delta_F(x,y) = \int \frac{d^3k}{(2\pi)^3.2\omega}\left[\theta(t-\tau)e^{i\underline{k}.(\underline{x}-\underline{y})-i(t-\tau)\omega} + \theta(\tau-t)e^{-i\underline{k}.(\underline{x}-\underline{y})-i(\tau-t)\omega}\right]$$

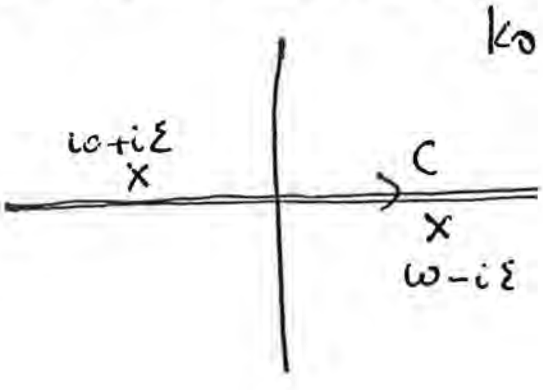
I now claim that

$$i\Delta_F(x,y) = \int \frac{d^4k e^{-ik.(x-y)}}{k^2-\mu^2+i\epsilon}$$

where we now use relativistic four vector notation and $a.b = a^0b^0 - \underline{a}.\underline{b}$. The proof of this statement is most easily given from Cauchy's theorem on complex integrals. In the above formula the $i\epsilon$ term fixes which side of the real k^0 contour integral contains the poles. The poles occur at solutions for k^0 of the equation

$$k^{0^2} - \underline{k}^2 - \mu^2 + i\epsilon$$

These occur at $k^0 = \pm\sqrt{\underline{k}^2 + \mu^2 - i\epsilon}$ or $k^0 = \pm\omega(\underline{k}) \mp i\epsilon$. So the k^0 plane looks like



The contour can now be lifted to ∞ up or down depending on the sign of $t - \tau$. If $t - \tau > 0$ then we close in the lower plane, if $t - \tau < 0$ then we close in the upper half plane. We pick up one pole in either case. To give the result

$$\theta(t-\tau).2\pi i \int \frac{-d^3 k e^{i\underline{k}.(\underline{x}-\underline{y})-i\omega(t-\tau)}}{2\omega} + \theta(\tau - t).2\pi i \int \frac{d^3 k e^{i\underline{k}.(\underline{x}-\underline{y})+i\omega(t-\tau)}}{-2\omega}$$

Again we see that an apparently non Lorentz invariant formalism gives invariant answers. The Feynman propagator is Lorentz invariant. Notice also the curious effect that suddenly we integrate not over d^3k but d^4k. In other words the Feynman propagator corresponds to particles off their mass-shell. Their energy, momentum does **not** satisfy $E^2 - \underline{p}^2 = m^2$.

Before leaving the operator methods let us very roughly outline how a more complicated process might go. This calculation also carries a health warning. We are going to calculate the fourth order contribution to the vacuum to vacuum transition matrix element. The operator expression for this is proportional to the integrals over the times x^0, y^0, z^0 and u^0.

$$\langle 0|T(H_I(x^0)H_I(y^0)H_I(z^0)H_I(u^0))|0\rangle$$

And each such Hamiltonian is an integral over the space components of the energy density $-g\chi^\dagger\phi\chi$. Thus we get the term

$$\int d^4x d^4y d^4z d^4u T\left\{\chi^\dagger(x)\phi(x)\chi(x).\chi^\dagger(y)\phi(y)\chi(y).\chi^\dagger(z)\phi(z)\chi(z).\chi^\dagger(u)\phi(u)\chi(u)\right\}$$

This is to be sandwiched between vacuum states. So when we turn the time ordering into the normal ordering no operators must be left. So as we commute terms past we must always pick up the Feynman terms, not the normal ordered terms. Crudely then we get the Feynman pairings of the points x, y, z, u in all possible ways. For example one possible term in the answer is

$$\int d^4x d^4y d^4z d^4u \Delta_\phi(x-y)\Delta_\phi(z-u)\Delta_\chi(x-y)\Delta_\chi(z-u)\Delta_\chi(x-u)\Delta_\chi(y-z)$$

There is a systematic procedure for writing down all possible Feynman diagrams. First draw all possible diagrams with vertices where $\phi, \chi, \chi^\dagger$ meet at each vertex. For each vertex a factor -ig, for each propagator a factor $\frac{i}{p^2-\mu^2+i\epsilon}$. At each vertex four momentum is conserved. This gives the overall conservation of four momentum.

4.2) The Functional Method of Deriving Feynman Diagrams

Now we turn to the standard modern approach to this problem of constructing Feynman diagrams. I assume you revised the Gaussian integrals section of the prerequisites. I claim that any free field theory reduces to doing Gaussian integrals. We can do any Gaussian integral using *the most important integral in the world.* I then would like to use the Gaussian trickery to calculate Feynman diagrams.

First let us write down the path integral for N particles of positions q_i. The canonical object to study is the function

$$W(J_i) = \int \prod_i [dq_i] e^{i\int_{t_i}^{t_f} L(q_i,\dot{q}_i)dt + \sum J_i q_i} \tag{4.2.1}$$

where $[dq_i]$ is the path integral over the i'th coordinate. The J_i are fake parameters put in to let us play Gaussian tricks. Thus

$$\frac{\partial W}{\partial J_l} = \int [dq_i] q_l e^{iS}$$

where S is the action. Thus J_i derivatives let us pull down factors of q_i in a systematic way. Note that we need to put $J_i = 0$ to get back to the original action S.

The generalisation to field theory is instantaneous if we remember what the dynamical degrees of freedom actually are. The analogues of the q_i are the field values $\phi(\underline{x})$. Just as i counts the individual degrees of freedom for Quantum Mechanics so $\underline{x}$ counts the degrees of freedom in field theory.

Thus the Functional integral in field theory is

$$W(J(x)) = \int \prod_x [d\phi] e^{-\int d^4x.\mathcal{L} + \int d^4x J(x)\phi(x)}$$

In other words, given a function J(x), we compute a number W. Thus we map from a function to a number...the old fashioned definition of a functional. As before we can take derivatives with respect to J(x) to pull down factors of $\phi(x)$. We have pulled a dirty theorists trick here. We continued t to it so the i in the exponent disappears. This makes the integral converge exponentially. We need to continue back at the end of the calculation. This is a big topic which I will duck.

Some care has to be taken with these derivatives. We need the concept of a Functional derivative rather than a normal derivative. Let us study a simple case. A functional is a map from a function to a number. The simplest example, with which you are familiar, is a normal integral. Given a function $J(x)$ the integral returns a number

$$W[J(x)] = \int \phi(x).J(x)dx$$

for an arbitrary function $J(x)$ and any fixed function $\phi(x)$. The Functional derivative is defined to be the limit as $\epsilon \to 0$ of

$$\frac{\delta W}{\delta J(y)} = lim \frac{W[J(x) + \epsilon\delta(x-y)] - W[J(x)]}{\epsilon} \tag{4.2.2}$$

In our integral case this gives

$$\begin{aligned}\frac{\delta W}{\delta J(y)} &= \int \delta(x-y)\phi(x)dx \\ &= \phi(y)\end{aligned}$$

The Dirac delta is necessary to give a non zero answer under the integral sign. The result of taking a Functional derivative of a constant Functional is a function of y, the point where the Functional derivative was evaluated.

So now let us turn to our free scalar first. Then we will show how to derive perturbation theory. The Lagrangian density is given by

$$\mathcal{L} = \frac{1}{2}\left(\frac{\partial\phi}{\partial t}\right)^2 + \frac{1}{2}(\partial\phi)^2 + \frac{1}{2}\mu^2\phi^2 + \frac{\lambda}{4!}\phi^4$$

The first three terms correspond to our free Boson, which we quantised previously by operators. The ϕ^4 term corresponds to the interaction term. In terms of Feynman diagrams we expect it to correspond to vertices where four particles meet.

First we solve the free part, then we add the perturbations. From the Gaussian prerequisites we know how to compute Gaussian integrals so we first rewrite the exponent in the form

$$\phi.\mathbf{Operator}.\phi$$

copying the Gaussian form

$$\sum \phi_i K_{ij} \phi_j$$

Since we now have an infinite number of degrees of freedom, labelled by $\underline{x}$ rather than i, we expect the $\sum$ to be replaced by $\int dx$. Thus

$$\int d^4x \left(\frac{\partial\phi}{\partial t}\right)^2 = -\int d^4x\phi\frac{\partial^2\phi}{\partial t^2} + Surface\ terms$$

Playing this game throughout and dropping surface terms we get, for the free theory,

$$W_0(J) = \int [d\phi] exp[-\frac{1}{2}\int d^4x d^4y \phi(x).K(x,y).\phi(y) + \int d^4z J(z)\phi(z)]$$

with

$$K(x,y) = \delta^4(x-y)(-\frac{\partial^2}{\partial t^2} - \nabla^2 + \mu^2) \tag{4.2.3}$$

The answer to the standard Gaussian is given in terms of the inverse matrix. So here we need the inverse operator to $K(x,y)$. In other words the solution to

$$\int d^4yK(x,y)\Delta(y,z) = \delta(x-z) \tag{4.2.4}$$

the analogue of

$$\sum_j K_{ij}K^{-1}_{jk} = \delta_{i,k}$$

in the discrete case.

It should come as no great surprise that the solution to this is our old friend the Feynman propagator.

$$\Delta_F = \int \frac{d^4k}{(2\pi)^4}\left\{\frac{e^{ik.(x-y)}}{k^2+\mu^2}\right\} \tag{4.2.5}$$

The only subtlety is that we take $k = (ik_0, \underline{k})$ i.e. Euclidean four vectors to guarantee exponential convergence of the integrals. To get physical answers we must continue back. So we see that the inverse operator is just the Fourier transform of the Feynman propagator. This is the basic result of the functional method. Later when we do gauge theories and when you add Dirac particles there will be additional indices for the gauge degrees of freedom, the spin indices, charge, etc., ... The propagator is the inverse, summing over all these degrees of freedom, as we will see.

So the exponent in the answer will be

$$\int dxdyJ(x)\Delta_F(x,y)J(y)$$

Now let us turn to interacting theory. Then we must calculate our perturbation series for U. To compute objects like

$$\int dt_1dt_2\langle 0|T(\phi(x)\phi(z)H_I(t_1)H_I(t_2))|0\rangle$$

we take Functional derivatives

$$\frac{\delta}{\delta J(x)}.\frac{\delta}{\delta J(z)}.\frac{\delta^4}{\delta J(u)^4}.\frac{\delta^4}{\delta J(w)^4} \tag{4.2.6}$$

to bring down all the operators, and then integrate $d^4u.d^4w$ to recover the two Hamiltonians integrated over time.

At the end of the calculation we must set all the $J=0$. Thus, after all the derivatives have been performed, no factors of J must remain in the numerator, such terms go away as $J \to 0$. Since the exponent is quadratic, this means that each term in the exponent must be differentiated twice. This must be done in all possible ways. Thus we get a sum of terms. In each term the derivatives are paired in that they both operate on the same term from the exponent. Such a paired derivative gives a factor

$$\frac{\delta}{\delta J(u)}.\frac{\delta}{\delta J(v)}.e^{\frac{1}{2}\int d^4xd^4yJ(x)\Delta_F(x,y)J(y)} = \Delta(u,v) \tag{4.2.7}$$

So each such pairing brings down a Feynman propagator, and the different pairings give all Feynman diagrams.

Chapter 5

Groups and Algebras

The big revolution in my lifetime has been the dominance of gauge theories. When I was a graduate student the philosophy was that the lightest, and therefore longest range hadron, was the pion. So the dynamics of this scalar particle was seen to be the most important. So everybody rushed around writing papers on the dynamics of scalar fields. Gauge theories were seen as **bad**. Seminars at CERN predicting that large p_T, which occurs automatically in gauge theories, might be interesting were treated with barely disguised derision.

Similarly a band of nutters went around writing down non-renormalisable, apparently non-predictive, gauge theories of the weak interactions. They were called Salam and Weinberg. This is all changed. The norm is gauge theories in all directions, as far as the eye can see. The nutters now do string theory. However being a nutter is not a sufficient reason to expect success.

The basic property of gauge theories is a vast symmetry called the Gauge Group. Except in a very few cases little is known about this vast symmetry group. The usual trickery of group theory is at a loss. Only the fearless theoretical physicists plunge into the unknown. The basic building blocks are the usual Lie groups. Since some of you are unfamiliar with these let me survey a couple of simple cases. These give a generic feel for the general case.

The language of symmetry in Quantum Mechanics is Group theory. So let us start with the definition of a group.

5.1) Definitions and Examples

A **Group** is a set of objects with a multiplication defined such that if a,b,c are arbitrary objects in the set then a.b is also in the set (closure), a^{-1} is also in the set (inverse), a unit e is in the set and we have the properties

$$a.(b.c) = (a.b).c$$

always,

$$a.a^{-1} = e = a^{-1}.a$$

$$e.a = a.e = a$$

e is often written 1.

Examples

a) The numbers $\{1, -1\}$ under multiplication.

b) The integers under addition.

c) More interesting. The set of 2×2 unitary matrices of determinant 1. Unitary means $a^\dagger = a^{-1}$. This is SU(2).

d) The set of 3×3 orthogonal matrices of determinant 1. Orthogonal means $a^T = a^{-1}$. This is SO(3). Since any rotation of three vectors can be written as a 3×3 orthogonal matrix this group is isomorphic to the rotation group.

The last two examples are known as Lie groups since the matrices depend smoothly on a finite number of parameters. Thus any rotation can be written as a product of rotations around the x,y, or z axes.

The rotation Group and Algebra

Such rotations can be parametrised

$$\begin{pmatrix} \cos\theta_3 & \sin\theta_3 & 0 \\ -\sin\theta_3 & \cos\theta_3 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} \cos\theta_2 & 0 & -\sin\theta_2 \\ 0 & 1 & 0 \\ \sin\theta_2 & 0 & \cos\theta_2 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta_1 & \sin\theta_1 \\ 0 & -\sin\theta_1 & \cos\theta_1 \end{pmatrix}$$

Now an interesting object appears if we consider the limit as the angles get small. This object is called the Lie algebra. This is important since it is the world inhabited by Gauge fields. Expanding the above we get $1+\theta_3\Sigma_3, 1+\theta_2\Sigma_2, 1+\theta_1\Sigma_1$ where

$$\Sigma_3 = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \Sigma_2 = \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \Sigma_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}$$

The interesting structure is not the product, but the commutator,

$$\begin{aligned}
[\Sigma_1, \Sigma_2] &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} - \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix} \\
&= \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} - \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \\
&= -\Sigma_3
\end{aligned}$$

Similarly $[\Sigma_2, \Sigma_3] = -\Sigma_1$ and $[\Sigma_3, \Sigma_1] = -\Sigma_2$. Thus these elements are closed under commutation. If we redefine $\Sigma \to i\Sigma$ then we retain the commutation relations of the angular momentum operators.

Thus the Lie algebra of the rotation group is the angular momentum algebra you have studied in great detail in your Quantum Mechanics courses. There is another way to see that angular momentum and rotations are closely linked. Consider a rotation of axes in the x,y plane such that

$$\begin{aligned}
x' &= x\cos\theta + y\sin\theta \approx x + \theta.y \\
y' &= -x\sin\theta + y\cos\theta \approx -x.\theta + y
\end{aligned}$$

Then a wave function transforms as

$$\Psi(x', y') = \Psi(x, y) + \theta(y\frac{\partial}{\partial x} - x\frac{\partial}{\partial y})\Psi(x, y) + O(\theta^2)$$

for small θ. The operator carrying out this transform is

$$(1 - i\theta\hat{L}_z \cdots)$$

In other words the infinitesimal rotations are generated by our old friend the angular momentum operator. This is the real reason that it appears everywhere in Quantum Mechanics. If a Quantum system is rotationally invariant then the Hamiltonian commutes with the Rotation operators i.e. the angular momentum operators. They can therefore be simultaneously measured. Hence eigenstates of energy can also be labelled by the total angular momentum.

SU(2) and su(2)

The most general SU(2) matrix can be written

$$g = \begin{pmatrix} a+ib & c+id \\ -c+id & a-ib \end{pmatrix} \tag{5.1.1}$$

where a,b,c,d are real numbers satisfying $a^2+b^2+c^2+d^2=1$. The unit matrix corresponds to $a \approx 1$ so, to get the algebra, we expand around $b, c, d \approx 0$. Thus we get

$$g \approx \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + ib\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} + id\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + ic\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}$$

Thus the elements of the algebra are the Pauli matrices. These also satisfy the angular momentum algebra. Thus the Lie Algebra of SU(2) is the same as the Lie Algebra of SO(3).

SU(3)

The group behind QCD is of course SU(3). So we need a quick survey of the group and its algebra. The group is the set of 3×3 unitary, determinant 1 matrices. The first obvious question is to compute the number of independent parameters or angles. Assume there are n $\theta_i; i = 1, \cdots, n$. Then expand an arbitrary SU(3) matrix in terms of the θ_i, assuming $\theta_i = 0$ corresponds to the unit matrix. Then $g = 1 + \sum \theta_i L_i \cdots$. Now it is easily seen that the condition that det g = 1 is equivalent to trace $L_i = 0$ (the trace of a matrix is the sum of the diagonal terms) while the condition that g is unitary implies that L_i are Hermitian i.e $L_i^\dagger = L_i$. There are only 8 independent traceless, Hermitian matrices. Gell-Mann wrote down a convenient set, which physicists have used ever since.

$$\lambda_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \lambda_2 = \begin{pmatrix} o & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \lambda_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \lambda_4 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}$$

$$\lambda_5 = \begin{pmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{pmatrix} \lambda_6 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \lambda_7 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix} \lambda_8 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{pmatrix} \tag{5.1.2}$$

These are closed under commutation ! The elements of the Lie algebra su(3) are linear combinations of these terms. This is the linear space in which the Gauge fields of SU(3) live. So be warned.

Lie's theorems

The heart of the theory is the realisation that, although the algebra can be computed in terms of infinitesimal group elements, from the algebra we can recover the Group. Let us see how this works for rotations. The matrix corresponding to a rotation by ψ about the unit vector $\underline{\hat{n}}$ is given by

$$\begin{aligned} R(\psi,\underline{\hat{n}}) &= \left[R(\frac{\psi}{N},\underline{\hat{n}})\right]^N \\ &= \left(1+i\sum \frac{\theta_i}{N}\Sigma_i\right)^N \\ &= e^{i\sum \theta_i\Sigma_i} \end{aligned} \tag{5.1.3}$$

where we have used the formula

$$\begin{aligned} \left(1+\frac{a}{N}\right)^N &= i + N.\frac{a}{N} + \frac{N(N-1)}{2}.\frac{a^2}{N^2} + \frac{N(N-1)(N-2)}{3!}.\frac{a^3}{n^3} \\ &\approx 1 + a + \frac{a^2}{2} + +\frac{a^3}{3!} + \cdots \\ &= e^a \end{aligned}$$

The basic idea is that a big rotation can be constructed from a lot of small ones. The small ones are fixed by the algebra. In the case of su(2) this gives, because $(\underline{\sigma}.\underline{\hat{n}})^2 = 1$, that

$$e^{i\frac{\psi}{2}\underline{\sigma}.\underline{\hat{n}}} = \cos\frac{\psi}{2} + i\sin\frac{\psi}{2}\underline{\sigma}.\underline{\hat{n}} \tag{5.1.4}$$

This is Hamilton's representation of finite rotations by quaternions.

Representations

In the above we defined a group as an abstract set with abstract properties. The Lie algebras were defined by expansions of explicit matrices about the identity. We could give a definition of the Lie Algebra as a linear vector space closed under commutation. Such things have been classified. The actual examples we had, in terms of matrices, are what mathematicians call representations. So in terms of Σ and the Pauli matrices we had two representations of su(2) in terms of 3×3 and 2×2 matrices. Lie lets us turn these into representations of SU(2). They correspond to the spin 1 and spin $\frac{1}{2}$ representations of SU(2).

Below we will see that, to define a gauge theory, we put the Gauge fields into the algebra and must prescribe which representations contain the other particles.

Chapter 6

Gauge theories

The particle physics interest is by now manifest. QCD is an SU(3) gauge theory. The Salam-Weinberg model is a gauge theory with gauge group $SU(2) \times U(1)$. Such theories are renormalisable, just like QED. They have remarkable properties which will be explored in the other lectures. Quark confinement, running coupling constants, chiral symmetry breaking, the Higgs effect are all properties of gauge theories. So we start at the beginning and follow our scalar route. First classical equations, then Lagrangians, then functional integrals, then compute perturbation theory.

6.1) Classical Maxwell Theory

Let us revise the Grand-daddy of all the gauge theories, due to Mr Maxwell. We will check that it is Lorentz invariant. Secondly we show how to derive it from a Lagrangian. Thirdly we discuss the U(1) gauge symmetry of this theory. Then we are all set to stick this classical Lagrangian into our Functional integral and derive the Feynman rules for the Maxwell U(1) Gauge theory.

The first thing to do is rewrite Maxwell in manifestly Lorentz invariant form. I choose units in which Maxwell's equations are

$$\begin{aligned} \nabla . \underline{B} &= 0 \\ \nabla \wedge \underline{E} &= -\frac{\partial \underline{B}}{\partial t} \\ \nabla . \underline{E} &= \rho \\ \nabla \wedge \underline{B} &= \underline{j} + \frac{\partial \underline{E}}{\partial t} \end{aligned}$$

Comprising eight equations two scalar, two vector for six unknowns $\underline{E}, \underline{B}$. Two of these equations can be solved by writing $\underline{B} = \nabla \wedge \underline{A}$ and $\underline{E} = -\nabla\phi - \frac{\partial \underline{A}}{\partial t}$. I remind you of the four vectors $x^\mu = (t, \underline{x})$, $x_\mu = (t, -\underline{x})$, $j^\mu = (\rho, \underline{j})$ and $j_\mu = (\rho, -\underline{j})$. Write $A^\mu = (\phi, \underline{A})$ and $A_\mu = (\phi, -\underline{A})$. Then we define the objects $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$. These naturally give the combinations $F_{01} = \partial_0 A_1 - \partial_1 A_0 = -\dot{A}_x - \frac{\partial}{\partial x} A_0 = -E_x$ and $F_{12} = \partial_1 A_2 - \partial_2 A_1 = -(\nabla \wedge \underline{A})_z = -\underline{B}_z$ so filling in all the terms we get

$$F_{\mu\nu} = \begin{pmatrix} 0 & E_x & E_y & E_z \\ -E_x & 0 & -B_z & B_y \\ -E_y & B_z & 0 & -B_x \\ -E_z & -B_y & B_x & 0 \end{pmatrix}$$

The covariant looking equation you might guess, would be $\partial_\mu F^{\mu\nu} = j^\nu$. We get for $\nu = 1$

$$\begin{aligned} \partial_\mu F^{\mu 1} &= \frac{\partial}{\partial t} F^{01} + \frac{\partial}{\partial x} F^{11} + \frac{\partial}{\partial y} F^{21} + \frac{\partial}{\partial z} F^{31} = -\dot{\underline{E}}_x + \frac{\partial}{\partial y} B_z + \frac{\partial}{\partial z}(-B_y) \\ &= (-\dot{\underline{E}} + \nabla \wedge \underline{B})_x = j_x = j^1 \end{aligned}$$

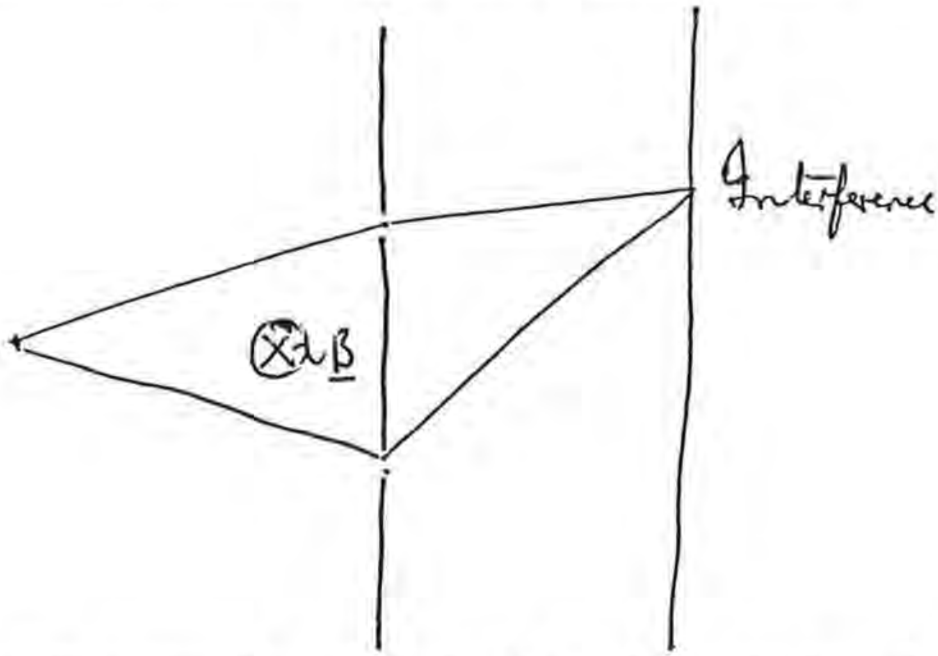
So this is one of Maxwell's equations. Now check the $\nu = 0$ term.

$$\frac{\partial}{\partial t}F^{00} + \frac{\partial}{\partial x}F^{10} + \frac{\partial}{\partial y}F^{20} + \frac{\partial}{\partial z}F^{30} = 0 + \frac{\partial}{\partial x}(E_x) + \frac{\partial}{\partial y}(E_y) + \frac{\partial}{\partial z}(E_z)$$
$$= \nabla.\underline{E} = j^0 = \rho$$

Thus two of Maxwell's equations are subsumed in $\partial_\mu F^{\mu\nu} = j^\nu$ The other two are subsumed in $\epsilon^{\mu\nu\rho\sigma}\partial_\nu F_{\rho\sigma} = 0$

We thus have our goal of a manifestly Lorentz invariant formalism . Note however that the equations have a funny little symmetry: $A_\mu \to A_\mu + \partial_\mu \phi$ leaves $F_{\mu\nu}$ unchanged. This was the first sign of a Gauge invariance. In classical physics it is a curiosity, since all of classical physics can be written safely in terms of the unaffected $\underline{E}, \underline{B}$ fields. In Quantum Mechanics it is a different matter as the $\underline{A}$ field is directly measurable.

The classic experiment to demonstrate this is due to Bohm and Aharanov.



An electron two-slit experiment is carried out with the addition of a long solenoid between the slits. The solenoid carries current and so there is a magnetic field $\underline{B}$ inside the coil. Outside the coil there is no magnetic field but $\underline{A}$ is not zero. The integral of $\underline{A}$ around a loop encircling the coil is given by Stokes

$$\int \underline{A}.d\underline{l} = \int d\underline{S}.(\nabla \wedge \underline{A}) = \int d\underline{S}.\underline{B} = B flux$$

So the A field is non zero outside the coil.

The interference pattern changes when the field is switched on. Hence, in Quantum physics, the dynamical variables are not merely the $\underline{E}, \underline{B}$ fields.

6.2) The Lagrangian Formalism

Now we return to rewrite the classical Maxwell theory in Lagrangian formalism. From the adverts above, we expect a Lorentz invariant Lagrangian density.

$$\mathcal{L} = -\frac{1}{4} F_{\mu\nu} F^{\mu\nu}$$

for the simple case of no fields except the Maxwell fields. The independent degrees of freedom are, at first sight, the four A_μ fields. So we get only four equations from Lagrange but remember that the A_μ already solve two of Maxwell's equations.

Thus the A_0 equation gives

$$\begin{aligned} \frac{\partial}{\partial t}\left[\frac{\partial \mathcal{L}}{\partial \dot{A}_0}\right] + \frac{\partial}{\partial x}\left[\frac{\partial \mathcal{L}}{\partial(\partial_x A_0)}\right] + \cdots - \frac{\partial \mathcal{L}}{\partial A_0} &= 0 \\ 0 + \frac{\partial}{\partial x}\left[-F_{01}\right] + \frac{\partial}{\partial y}\left[-F_{02}\right] \cdots - 0 &= 0 \\ \nabla . \underline{E} &= 0 \end{aligned}$$

As promised, one of Maxwell's equations. There is however a problem, $\Pi_{A^0} = 0$ since the Lagrangian has no dependence on $\frac{\partial A^0}{\partial t}$. So what on earth can the commutation relations be ?

Now turn to the $\underline{A}_x$ equation.

$$\frac{\partial}{\partial t}\left[\frac{\partial \mathcal{L}}{\partial \dot{\underline{A}}_x}\right] + \frac{\partial}{\partial x}\left[\frac{\partial \mathcal{L}}{\partial(\partial_x A_x)}\right] + \frac{\partial}{\partial y}\left[\frac{\partial \mathcal{L}}{\partial(\partial_y A_x)}\right] + \frac{\partial}{\partial z}\left[\frac{\partial \mathcal{L}}{\partial(\partial_z A_x)}\right] - \frac{\partial \mathcal{L}}{\partial A_x} = 0$$

$$\frac{\partial}{\partial t}\left[-F_{01}\right] + 0 + \frac{\partial}{\partial y}\left[-F_{12}\right] + \frac{\partial}{\partial z}\left[F_{31}\right] = 0$$

$$\frac{\partial}{\partial t}\left[-\underline{E}_x\right] + \frac{\partial}{\partial y}\underline{B}_z - \frac{\partial}{\partial z}\underline{B}_y = 0$$

$$\frac{\partial}{\partial t}\underline{E} = \nabla \wedge \underline{B}$$

The other Maxwell equation. Thus this Lagrangian gives the expected answers. So we have connected Maxwell's equations to our world of Lagrangians and, if necessary, Hamiltonians. The only fly in the ointment is the bizarre non-existence of the momentum conjugate to the scalar electric potential A^0. If we treated Gauge theories in Hamiltonian formulation this would cause us serious aggravation. The jargon, again due to Dirac, is the language of constrained systems and Dirac Brackets as opposed to Poisson brackets. We will avoid this problem by sticking to our Lagrangian, functional treatment. These problems will not go away but will reappear in a different disguise. The no free lunch theorem.

6.3) The U(1) Covariant Derivative

In a short while we will launch into the complexities of Non-Abelian gauge theories. Before we do this I introduce a crucial construction, the covariant derivative, in the simple U(1) case. We saw above that Maxwell's equations were invariant under the transformation

$$A^\mu(x) \to A'^\mu = A^\mu(x) + \partial^\mu \chi(x).$$

Now consider some field $\psi(x)$ which transforms as

$$\psi(x) \to \psi'(x) = e^{-ig\chi(x)}\psi(x)$$

for an arbitrary real function $\chi(x)$. Notice that unlike in the case of Lorentz transformations or rotations no changes happen to $\underline{x}$.

We would like to construct Lagrangians, with interaction terms involving both A_μ and ψ. So first turn to the transformation of the $\partial_\mu\psi(x)$ term.

$$\begin{aligned}
\partial_\mu\psi(x) \to \partial_\mu\psi'(x) &= \partial_\mu\Big\{e^{-ig\chi(x)}\psi(x)\Big\} \\
&= \frac{\partial\psi}{\partial x^\mu}.e^{-ig\chi(x)} - ig\frac{\partial\chi}{\partial x^\mu}.e^{-ig\chi}.\psi \\
&= \left\{\frac{\partial\psi}{\partial x^\mu} - ig\frac{\partial\chi}{\partial x^\mu}.\psi\right\}.e^{-ig\chi(x)} \\
&\neq e^{-ig\chi(x)}.\frac{\partial\psi}{\partial x^\mu}
\end{aligned}$$

So the derivative does not transform in the same way as the original function. But try

$$\begin{aligned}
\Big(\partial^\mu + igA^\mu\Big)\psi \to \Big(\partial^\mu + ig\big(A^\mu + \partial^\mu\chi(x)\big)\Big).e^{-ig\chi(x)}.\psi& \\
= e^{-ig\chi(x)}\Big[\partial^\mu\psi - ig\partial^\mu\chi.\psi + iga^\mu.\psi + ig\partial^\mu\chi.\psi\Big]& \\
= e^{-ig\chi}\Big[\partial^\mu + igA^\mu\Big]\psi&
\end{aligned}$$

We see that the combination $D^\mu\psi(x) = (\partial^\mu + igA^\mu)\psi(x)$ transforms in exactly the same way as ψ. This makes it easy to construct gauge invariant Lagrangian densities. This is the covariant derivative. Those of you who have studied general relativity will recognise the name, the style, but not the details.

For example, the standard kinetic term for a free spin $\frac{1}{2}$ Dirac particle is $\bar{\psi}\gamma_\mu\partial^\mu\psi$. Now ψ transforms as $e^{-ig\chi}\psi(x)$ and the kinetic term is not gauge invariant. However the combination $\bar{\psi}\gamma_\mu\Big[\partial^\mu + igA^\mu\Big]\psi$ is gauge invariant. The imposition of gauge invariance then ties the free quadratic term to the interacting term containing three fields. In this sense gauge invariance fixes the interactions, given the free terms.

6.4) Non-abelian Gauge Theories

We take the case of SU(2). More general theories are easy once you understand this case. The trick, as above, is to construct the covariant derivative. Suppose we have a field ψ transforming as

$$\psi(x) \rightarrow \psi'(x) = U(x)\psi(x) \tag{6.4.1}$$

where $U(x)$ is an element of SU(2). This element can be different at every point of space time. In other words this gauge symmetry will turn out to be a vast group. The elements of the group are fields with group values. In other words at every point in space time we attach an SU(2) matrix, possibly different at every point. Group multiplication is defined by multiplying the elements at each space time point

$$U(x).V(x) = U.V(x)$$

giving another set of SU(2) matrices at each space time point. We can choose independent SU(2) rotations at every point in space and have them depend arbitrarily on time. Compare this with the angular momentum /rotation operator which rotates all points of space by the same amount. This is why the gauge symmetry is often called a local symmetry. We can choose to have a gauge transformation which is 1 (i.e. no transformation) everywhere except a finite local region. We can carry out symmetry transformations independently on the moon and on earth.

So we introduce the non-abelian gauge field for SU(2). As promised it lives in the su(2) algebra. Thus the Gauge field can be thought of as

$$W^{\mu} = \frac{\tau_1}{2}W_1^{\mu} + \frac{\tau_2}{2}W_2^{\mu} + \frac{\tau_3}{2}W_3^{\mu} = \frac{1}{2}\underline{\tau}.\underline{W}^{\mu} \tag{6.4.2}$$

Here the three τ's are the three Pauli matrices of the su(2) algebra. They are given a different name just to avoid confusion with any angular momentum Pauli matrices that might be around. There are 12 W fields. One for each Pauli matrix and each such term is a four vector, hence the Lorentz indices μ. You will therefore often see the gauge field given as 2×2 matrix W^{μ} which can be rewritten as a linear combination of Pauli matrices with coefficients the actual fields W_i^{μ}.

Now put these together to construct the covariant derivative.

$$D^{\mu}.\psi = \left(\partial^{\mu} + iW^{\mu}\right)\psi \tag{6.4.3}$$

Unlike Maxwell we do not know how the gauge field must transform under gauge transformations. We let the covariant derivative tell us. Thus we assume that after the gauge transformation

$$D^{\mu}\psi \rightarrow D'^{\mu}\psi' = U(x).D^{\mu}\psi$$

From this we see that

$$D' = UDU^{-1}$$

Rewriting this in terms of the fields W^{μ} and W'^{μ}, we arrive at

$$\partial^{\mu} + i\frac{\tau}{2}.\underline{W}'^{\mu} = U.\Big(\partial^{\mu} + i\frac{\tau}{2}.\underline{W}^{\mu}\Big)U^{-1}$$

or in terms of the 2×2 matrices W^{μ}

$$W'^{\mu} = W^{\mu U} = U.W^{\mu}U^{-1} - iU(x).\partial^{\mu}U^{-1}(x) \tag{6.4.4}$$

This somewhat strange formula is defined entirely to make the covariant derivative of a gauge field transform in the same way as the field.

Now we construct the analogue of the $F_{\mu\nu} = \partial_{\mu}A_{\nu} - \partial_{\nu}A_{\mu}$ field. This is defined to be

$$F^{\mu\nu} = \partial^{\mu}W^{\nu} - \partial^{\nu}W^{\mu} - i[W^{\mu}, W^{\nu}] \tag{6.4.5}$$

This is done with only one aim in mind; to construct gauge invariant Lagrangians. So how does $F^{\mu\nu}$ transform ? The result is

$$F'^{\mu\nu} = U.F^{\mu\nu}.U^{-1} \tag{6.4.6}$$

Notice that all these terms are 2×2 matrices which do not necessarily commute. In the Abelian, commuting case the U's would cancel. This happens in the Maxwell case where the U's are just phases.

To check this by brutal calculation we just substitute the transformation properties of the W's into the definition of the covariant derivative and grind.

$$\begin{aligned}
F'^{\mu\nu} &= \partial^{\mu}W'^{\nu} - \partial^{\nu}W'^{\mu} + i[W'^{\mu}, W'^{\nu}] \\
&= \partial^{\mu}\Big(U.W^{\nu}.U^{-1} - iU.\partial^{\nu}U^{-1}\Big) \\
&- \partial^{\nu}\Big(U.W^{\mu}.U^{-1} - iU.\partial^{\mu}U^{-1}\Big) \\
&+ i\Big[U.W^{\mu}.U^{-1} - iU\partial^{\mu}U^{-1}, U.W^{\nu}.U^{-1} - iU.\partial^{\nu}U^{-1}\Big] \\
&= U\Big[\partial^{\mu}W^{\nu} - \partial^{\nu}W^{\mu} + i[W^{\mu}, W^{\nu}]\Big]U^{-1} \\
&+ \partial^{\mu}U.W^{\nu}.U^{-1} + U.W^{\nu}.\partial^{\mu}U^{-1} - i\partial^{\mu}U.\partial^{\nu}U^{1} - iU.\partial^{\mu}\partial^{\nu}U^{-1} \\
&- \partial^{\nu}U.W^{\mu}.U^{-1} - U.W^{\mu}.\partial^{\nu}U^{-1} + i\partial^{\nu}U.\partial^{\mu}U^{1} + iU.\partial^{\mu}\partial^{\nu}U^{-1} \\
&+ i\Big[U.W^{\mu}.U^{-1}, -iU.\partial^{\nu}U^{-1}\Big] + \Big[U.\partial^{\mu}U^{-1}, U.W^{\nu}.U^{-1}\Big] - i\Big[U.\partial^{\mu}U^{-1}, U.\partial^{\nu}U^{-1}\Big]
\end{aligned}$$

The first line is the expected answer. We have to show the rest cancels. The trick here is to take the derivative of the identity relation $\partial^{\mu}(U(x).U^{-1}(x)) = \partial^{\mu}1 = 0$. Which gives $\partial^{\mu}U.U^{-1} + U.\partial^{\mu}U^{-1} = 0$ so that $\partial^{\mu}U^{-1} = -U.\partial^{\mu}U.U^{-1}$. This lets us get rid of all derivatives of U^{-1} in the above, to give for the right hand side

$$= U.F^{\mu\nu}.U^{-1} + \partial^{\mu}U.W^{\nu}.U^{-1} - U.W^{\nu}.U^{-1}.\partial^{\mu}U.U^{-1} + i\partial^{\mu}U.U^{-1}.\partial^{\nu}U.U^{-1}$$

$$-iU.\partial^\mu\partial^\nu U^{-1} - \partial^\nu.W^\mu.U^{-1} + U.W^\mu.U^{-1}.\partial^\nu U.U^{-1} - i\partial^\nu U.U^{-1}.\partial^\mu U.U^{-1}$$
$$+iU.\partial^\mu\partial^\nu U^{-1} - U.W^\mu.U^{-1}.\partial^\nu U.U^{-1} + \partial^\nu U.U^{-1}.U.W^\mu.U^{-1} - U.U^{\iota}.\partial^\mu U.U^{-1}.U.W^\nu U^{-1}$$
$$+U.W^\nu.U^{-1}\partial^\mu U.U^{-1} - i\partial^\mu U.U^{-1}.\partial^\nu U.U^{-1} + i\partial^\nu U.U^{-1}\partial^\mu U.U^{-1}$$
$$= U.F^{\mu\nu}.U^{-1}$$

It is now an easy exercise to check that the following is a gauge invariant, Lorentz invariant Lagrangian density.

$$\mathcal{L} = -\frac{1}{4}trace\big[F_{\mu\nu}F^{\mu\nu}\big] \tag{6.4.7}$$

This is the **Yang-Mills** Lagrangian density.

Here trace is the sum of the diagonal terms in the matrices. It is easy to prove the crucial property that tr(A.B.C)=tr(C.A.B). This term is also Lorentz invariant due to the way we contracted the Lorentz $\mu\nu$ indices. So we are in good shape with both symmetries manifest.

If we wanted to do QCD with its SU(3) symmetry we would have taken W^μ to be a linear superposition of the 8 Gell-Mann matrices. There would have been 4(Lorentz)$\times$8(su(3) generators) gauge fields. Similar arguments apply to any Gauge group.

6.5) Feynman Rules

We are now ready to compute the Feynman rules for the gauge field. Before plunging into details let us look crudely at the Lagrangian. We see there are terms quadratic in the fields W^μ of the form $(\partial^\mu W^\nu - \partial^\nu W^\mu)^2$ plus terms cubic and quartic in the W's. The quadratic terms will describe free fields. The others, the unavoidable self-interaction terms of a non-abelian Yang-Mills theory. They are the reason why QCD, even without fermions, is a highly non-trivial field theory.

We follow our Gaussian tricks to the end. So we need again to write the quadratic part of the action as **Field.Operator.Field**. Consider the Abelian U(1) case. The problem here lies in the Lorentz indices not the gauge group indices.

$$\int (\partial_\mu A_\nu - \partial_\nu A_\mu)(\partial_\mu A_\nu - \partial_\nu A_\mu) d^4x$$

$$= \int \Big\{ A_\nu(-\partial_\mu\partial_\mu A_\nu + A_\nu\partial_\mu\partial_\nu A_{\mu\nu} + A_\mu\partial_\nu\partial_\mu A_\nu - A_\mu\partial_\nu\partial_\nu A_\mu \Big\} d^4x$$

So in terms of operators we need the inverse of the operator

$$\int d^4x d^4y A_\mu(x).\Big[-2(\partial_\mu\partial_\nu)g_{\mu\nu} + 2\partial_\mu\partial_\nu\Big]\delta^4(x-y).A_\nu(y)$$

Remember we solved the free spin zero particle by Fourier transforms. Here we have the added complication of the Lorentz indices. The operator is a two index tensor in

Lorentz indices. So we need the inverse of this matrix/tensor in momentum space. After Fourier we get, in the numerator,

$$P_{\mu\nu} = (k_\mu k_\nu - k^2.g_{\mu\nu})$$

This has an unfortunate property. Compute the product of two of them

$$\begin{aligned} P_{\mu\nu}P_{\nu\rho} &= (k_\mu k_\nu - k^2 g_{\mu\nu})(k_\nu k_\rho - k^3 g_{\nu\rho}) \\ &= k_\mu k_\rho.k^2 - k^2.k_\mu k_\rho - k^2 k_\mu k_\rho + k^4.g_{\mu\rho} \\ &= -k^2(k_\mu k_\rho - k^2 g_{\mu\rho}) \end{aligned}$$

Now any matrix whose square is proportional to itself does not have an inverse. Try multiplying the equation by the inverse. So the propagator does not exist !

What's up ? There are several views on this.

1). Canonical quantisation is in trouble. Remember we got $\Pi_{A_0} = 0$. Which is inconsistent with the standard commutation relations. We need to start again using Dirac's theory of constraints.

2). Another way of saying the same thing is the fact that there are only actually two photon states c.f. right and left polarised; whereas we have a vector A^μ describing the photon i.e. four degrees of freedom.

3) In path integral formalism we should not integrate over the gauge equivalent field configurations. Thus if W^μ is gauge equivalent to W'^μ then to count both in the path integral is a bit strange. The Faddeev Popov trick is a systematic way of removing this double counting.

6.6) Gauge Fixing or Faddeev Popov

Let me start with a trivial, rotationally invariant, integral and use a large sledgehammer to crack it. The sledgehammer will however also crack the gauge problem.

So consider the integral

$$I = \int f(x,y)dxdy$$

where f is invariant under rotations. In terms of polar coordinates $f(x,y) = F(r,\theta) = F(r)$. Trivially

$$I = \int F(r)rdrd\theta = 2\pi \int F(r)rdr$$

Here we think of 2π as the volume of the rotation group.

I would like to rewrite this familiar calculation in the language of invariance and group transformations. In this case trivial rotations, in the general case these will be full gauge transformations. The rotational invariance of the function f will be replaced by the gauge invariance of the Yang-Mills Action.

First I redo the trivial calculation. Then I will use a general coordinate change. These will correspond to different choices of **gauges** in the Yang-Mills case. Given a point $\underline{r} = (r, \theta)$ we define a rotated point $\underline{r}_\phi = (r, \theta + \phi) = R(\phi)\underline{r}$. Thus R corresponds to an operator which rotates the coordinates. The function f is clearly invariant under such an operation; in fact this is what rotational invariance means. We only want to count one point from each circle. All other points on a given circle have the same value of f. The jargon is that the circles form the orbits of the rotation group.

So define

$$W_\phi = \int d^2r f(x,y)\delta(\theta - \phi) = \int r dr d\theta F(r,\theta)\delta(\theta - \phi) = \int r dr F(r,\phi)$$

It is clear immediately that

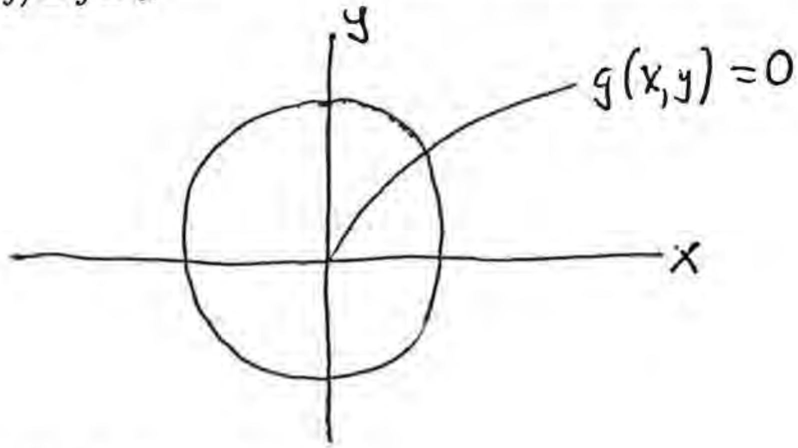
$$W = \int d^2r f(x,y) = \int d\phi W_\phi$$

W_ϕ corresponds, obviously, to integrating straight out along a radial path at angle ϕ. Because of the invariance of F it is immediate that $W_\phi = W_{\phi'}$ for all ϕ, ϕ'. Thus the integral

$$W = \int d\phi W_\phi = 2\pi . W_\phi$$

since W_ϕ is independent of ϕ. This trick, of using the invariance to extract the group volume, is what we need to use in general.

Consider the same example but integrated along a arbitrary curve which cuts each orbit once and once only. Suppose the equation of this curve is $g(\underline{r}) = 0$. For example the x-axis is $g(x,y) = y = 0$.



First consider the integral

$$\Delta_g^{-1}(\underline{r}) = \int d\phi \delta(g(\underline{r}_\phi))$$

The reason for the name will become clear. At each radius there exists a ϕ which rotates an arbitrary $\underline{r}$ onto our curve. Thus this δ function has a zero in ϕ. The answer is

$$\frac{1}{|\frac{\partial g}{\partial \theta}|_{g=0}}$$

Hence the name, it is the inverse Jacobian.

But now we prove that Δ is invariant i.e. unchanged by rotations of $\underline{r}$.

$$\Delta_g^{-1}(\underline{r}_{\phi'}) = \int d\phi\delta(g(\underline{r}_{\phi+\phi'})) = \int d\psi\delta(g(\underline{r}_{\psi})) = \Delta_g^{-1}(\underline{r})$$

where we changed integration variable to $\psi = \phi + \phi'$. Thus, inserting a factor of 1,

$$W = \int d\phi W_\phi$$

where

$$W_\phi = \int d^2 r f(x,y).\Delta_g(\underline{r})\delta(g(\underline{r}_\phi)) = W_{\phi'}$$

The proof of this latter statement goes as follows. First calculate the rotation R that takes $\underline{r}_\phi \to \underline{r}_{\phi'}$. Then change variables from $\underline{r}$ to $R.\underline{r} = \underline{r}'$. The three terms d^2r,$f(\underline{r})$ and Δ are all invariant while the δ term changes as required. In the general case we will need to prove the invariance of the measure, the function and the Δ term. Thus finally we have

$$W = 2\pi.W_\phi$$

The 2π is, once again, the group volume and W_ϕ integrates once over each orbit as required.

Now turn to the gauge case. We want to integrate over the W^μ. But we only want to count the gauge equivalent fields once. Thus we need the equivalent of g above. Equations such that they only have one solution under gauge transformations.

Before plunging into the details let me give two examples.

a). **Axial Gauge**: For each W_a^μ field we set $g^a = W_a^3 = 0$. In other words if we call the field W_a^μ, transformed by the gauge transformation U, $W_a^{\mu U}$ we substitute this in g and solve for U. At each space-time point we can rotate the W field so that it's third space component is zero for each su(2) index. At this point we will lose manifest Lorentz invariance. This is possible because the gauge transformations of (6.4.4) are x dependent and so show up differently in the different Lorentz components even although they originally operate on the SU(2) labels.

b). **Covariant Gauge**: This is defined by taking $g^a = \partial_\mu A_a^\mu$.

Now with a wild generalisation we write

$$\Delta_g^{-1}(W^\mu) = \int [\prod_x dU(x)] \prod_x \delta(g^a(W_a^{\mu U})) = \Delta_g^{-1}(W^{\mu V})$$

for all gauge transformations V. Moreover we can show, as above, that

$$\Delta_g^{-1}(W^\mu) = \int [\prod dU] \prod \delta(g_a(W_a^{\mu U}) - B_a) = \Delta_g^{-1}(W^{\mu V})$$

independent of V and B. This follows from standard properties of the δ-function as summarised in the prerequisites. The $\prod$ over all space time points makes the notation cumbersome so we drop it, but keep in mind that we now have a functional integral over separate degrees of freedom at each space time point. An essential part of our simple rotation model was that $d\theta = d(\theta + \phi)$ or in other words that the measure is invariant under rotations. Similarly above we integrate dU over the gauge group. It is a known result that all Lie groups like SU(2), SU(3) have invariant integration measures. These are called Haar measures. So into our gauge field functional integral we insert the factors of

$$1 = \Delta_f(W_a^\mu). \int [dU]\delta(g_a(W_a^{\mu VU}) - B_a)$$

$$Const = \int [dB] e^{-\frac{1}{2\xi}\int d^4x B^2(x)}$$

giving

$$\int [dW_a^\mu]\,[dU]\, e^{-Action}.\Delta_g(W_a^\mu)\prod \delta(g_a(W_b^{\mu U}) - B_a). \int [dB]\, e^{-\frac{1}{2\xi}\int d^4x B^2(x)}$$
$$= \int [dW_a^\mu]\,[dU]\, e^{-Action}\Delta_g(W_a^\mu)\delta(g_a(W_a^\mu) - B_a). \int [dB]\, e^{-\frac{1}{2\xi} d^4x B^2(x)}$$
$$= \int [dW_a^\mu]\,[dU]\, e^{-Action}\Delta_g(W_a^\mu) e^{-\frac{1}{2\xi}\int d^4x (g_a)^2}$$
$$= \left\{Volume\ of\ gauge\ Group\right\} \times \int [dW_a^\mu] e^{-Action}\Delta_g(W_a^\mu) e^{-\frac{1}{2\xi}\int d^4x g_a^2}$$

The \$64,000 question at this point is whether we have solved the propagator problem? In other words, throwing away the gauge group measure above, does the quadratic term operator now have an inverse? The real change is the additional term g_a^2. In covariant gauge this alters the Fourier transformed operator numerator into

$$k_\mu k_\nu \{1 - \frac{1}{\xi}\} - k^2.g_{\mu\nu}$$

This now has an inverse and the propagator in momentum space is

$$\frac{\left[g^{\mu\nu} - \frac{(1-\xi)k^\mu k^\nu}{k^2}\right]}{k^2 + i\epsilon}$$

The other term in our new integral Δ gives rise to ghosts although in some gauges there is no such contribution. I do not have time for this. There is clearly a large scale industry in trying funny gauges. Some people develop this to an art. Occasionally you can eliminate almost all Feynman graphs by clever choices.

Chapter 7

Fermionic Integrals

The previous two years I gave this course it was decided that I finish with the Higg's model. This year rather than rushing it at the end it was put rightfully at the start of the Standard Model course. I was given the task of explaining how the above functional trickery could be expanded to include fermions. This was more in the lines of whetting the appetite rather than a full meal.

A crucial ingredient in the functional technique is the ability to do Gaussian integrals. Here I would like to show simply how this works for integrals over anti-commuting variables.

Consider a standard real integral

$$\int_{-\infty}^{+\infty} f(x)dx = \int_{-\infty}^{\infty} f(x+c)dx$$

for any real constant c. In other words the integration is translation invariant.

We take this as a necessary condition for a fermionic integral. So first consider a general function of one anti-commuting variable a. In other words $\{a, a\} = a^2 + a^2 = 0$. Thus on Taylor expanding any such function

$$f(a) = f(0) + \frac{df}{da}(0)a + \frac{d^2f}{2da^2}(0)a^2 + \dots$$

we see only the first two terms survive. The coefficients $f(0)$, $\frac{df}{dx}$ are real or complex numbers independent of any anti-commuting variable. The remainder are zero. Thus the most general function of one anti-commuting variable is a linear function ! Analysis is easy.

Looking at the most general, one dimensional integral we get

$$\int f(a)da = \int [f(0) + \frac{df}{da}(0)a]da = \int f(a+c)da = \int [f(0) + \frac{df}{da}(0)(a+c)]da$$

This must be true for all c. Thus we easily see that $\int da = 0$. By definition $\int daa = 1$.

Then $\int f(a)da = \frac{df}{da}(0)$. So, for fermions, integration equals differentiation not it's inverse.

Now turn to multi-dimensional anti-commuting integrals.

$$\int f(a, b)dadb$$

Here $\{a, b\} = ab + ba = 0$. So we need to agree an ordering.

$$f(a, b) = f_0 + f_a a + f_b b + f_{ab} ab$$

where the f_i are real or complex numbers, is the most general function of two anti-commuting variables.

So define

$$\int db.daf(a,b) = f_{ab}$$

Notice we get a sign change on interchanging da and db. The formula also agrees with the answer we get by first doing da then db.

This methodology obviously extends to functions of more variables.

Now we turn to the fermionic analogue of Gaussian integrals. Consider a quadratic function of 4 variables a_1, a_2 and b_1, b_2

$$A = a^T.A.b = a_1 A_{11} b_1 + a_1 A_{12} b_2 + a_2 A_{21} b_1 + a_2 A_{22} b_2$$

Think of this an action with inependent a and b fields. So consider

$$I = \int da_1 da_2 db_1 db_2 e^{-A}$$

We expand the exponential. Clearly only terms with four factors of a or b will give a non-zero contribution. Thus only the third term in the exponential expansion counts. We must also have no squares.

$$I = \int da_1 da_2 db_1 db_2 \frac{A.A}{2} = \int da_1 da_2 db_1 db_2$$

$$\frac{\left[a_1 A_{11} b_1 . a_2 A_{22} b_2 + a_2 A_{22} b_2 . a_1 A_{11} b_1 + a_1 A_{12} b_2 . a_2 A_{21} b_1 + a_2 A_{21} b_1 . a_1 A_{12} b_2\right]}{2}$$

$$= -A_{11}.A_{22} + A_{12}.A_{21} = -detA$$

This should be compared with the equivalent real Gaussian integral. There we get $detA^{-\frac{1}{2}}$ factors.

Thus in functional integral language we need to generalise the above to the Dirac action. Then we ahve an anti-commuting quadratic $\bar{\psi}(x)(i\gamma.\frac{\partial}{\partial x} - m)\psi(x)$

Here $\bar{\psi}$ and ψ anti-commute being momentum and coordinate respectively.

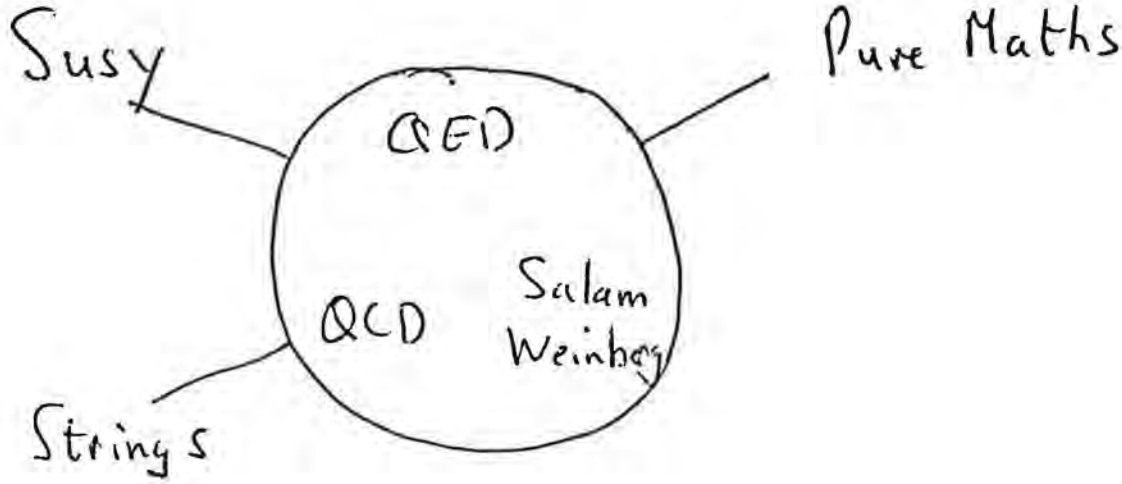
Apart from signs all the machinery of our Gaussian trickery can be developed to give fermion propagators. An easy exercise.

Another unexpected place these integrals turn up is in our gauge fixed Yang-Mills. The Faddeev Popov trick leaves factors of a determinant $\Delta_g(W_a^\mu)$ in the numerator. These determinants can be written as integrals over some new fields. These are the Ghosts.

Chapter 8

Virgin Experimental Territory

In 1994 we had a request for a very fast overview of things theoretical that had not led to positive experimental results. All in half an hour.



In the diagram I show a plan of Theorist territory, unconquered by experimentalists. Such territory may be empty or not exist of course.

In our lectures we have stressed the rôle of symmetry. As you may have noticed our symmetries never mix fields of different spin. Thus the Gauge transformations mix gauge bosons with gauge bosons, fermions with fermions. Theorists have however found a symmetry whivh mixes particles of different spins. This is super-symmetry. Here a representation of a super-symmetry typically joins scalars and fermions or scalars, fermions and spin 1 particles. Complicated versions join gravity with gauge bosons.

These theories have many beautiful properties. Often the divergences of Feynman diagrams cancel. Discovery of such super partners to standard particles would be an impressive breakthrough. This was a motivation for HERA.

Another way to pull together lots of particles in one structure is string theory. Here instead of the fields $\phi(x)$ dependent on one space time point x, which we have grown used to, we introduce variable $X^{\mu}(\sigma, \tau)$. Here σ, τ parametrise the points on the surface X^{μ} swept out in space time as σ, τ vary. The individual modes of this object correspond to an infinite number of particles of different spins and masses.

Again such theories have beautiful properties but nobody has seen a clear way to reconstruct our actual world.

Many Gauge Theories have Soliton like solutions. These can for example be magnetic monopoles or strings or other defects. An area of great current interest is to study whether such objects can seed galaxy formation. This connects us up to the world of Astronomy and COBE.

A surprising development over the past 15 years or so has been the influence of Quantum Field Theory on Pure Mathematics. In the 19'th century Riemann showed how two-dimensional comlex surfaces could be exhaustively studied by considering electric fields or fluid flow on the surface. To take a trivial example the Riemann sphere, effectively the complex plane with one point at infinity, is different from the torus. This is most easily seen by seeing that the sphere cannot have fluid flow without a singulaity whereas on the torus flow is possible without singularity.

Since the early 80's Atiyah, Donaldson and Witten have shown how detailed Geometrical information can be extracted from considering Quantum Field Theories on surfaces. Thus our Gauge theory work has profound geometrical consequences.

More recently claims have even been made that these geometrical approaches may lead to a soluble 4 dimensional field theory. That would set the world alight.

Questions and Exercises

0.1.1) Prove (0.1.2).

0.1.2) Compute the normalised wave functions $\Psi_1(q), \Psi_2(q)$ for the first and second excited states of the Harmonic Oscillator.

0.1.3) Check explicitly that $(\hat{a})^2\Psi_1(q) = 0$.

0.1.4) Prove that the Hermitian conjugate of $\hat{a}$ is $\hat{a}^\dagger$.

0.2.1) Compute

$$\int_{-\infty}^{+\infty} x^4 e^{-\alpha x^2} dx$$

0.3.1) Consider a free quantum mechanical particle, of mass m, moving along a 1 dimensional line. Given a wave function $\Psi = e^{-x^2}$ at t = 0, calculate, in both Heisenberg and Schrödinger pictures, the probability density of finding it at x = 0 at time t.

0.3.2) Check $[\hat{q}, \hat{p}] = i$ in both Schrödinger and Heisenberg pictures.

0.3.3) Check that the $\hat{a}, \hat{a}^\dagger$ commutation relations are the same in the two pictures.

0.4.1) Derive the Lagrangian for a particle falling, under gravity, near the surface of the earth. Derive the equations of motion, and solutions, in both Lagrangian and Hamiltonian form.

0.5.1) Compute

$$\int_{-\infty}^{+\infty} dx e^{-x^2} \delta(2x-1)$$

$$\int_{-\infty}^{+\infty} dx e^{-x^2} \delta(2x^2-1)$$

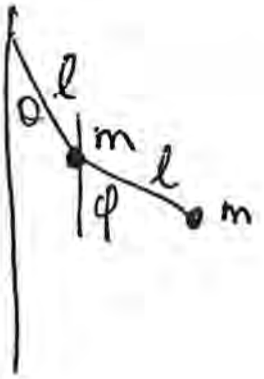
0.5.2) Prove that

$$lim_{\lambda\to\infty}\left\{\sqrt{\frac{\lambda}{\pi}}e^{\frac{-x^2}{2\lambda}}\right\} = \delta(x)$$

0.5.3) Compute

$$\int_{-\infty}^{+\infty} dx dy e^{-(x^2+y^2)} \delta(2x-y-1)\delta(x+y-2)$$

1.1.1) Using Lagrange's equation, solve for the motion of a double pendulum undergoing small oscillations.



1.2.1) What are the momenta conjugate to θ, ϕ?

1.3) Compute the Poisson brackets

$$\{q^2, p\}, \{q, p^2\}$$

How do they compare with

$$[\hat{q}^2, \hat{p}], [\hat{q}, \hat{p}^2]$$

1.4) Prove that Hamilton's Classical equations can be rewritten

$$\dot{q}_i = \{q_i, H\}$$

$$\dot{p}_i = \{p_i, H\}$$

Does this satisfy Dirac's Classical to Quantum prescription?

2.1.1) Derive an expression for $U_3(t)$. Prove it is given by the time ordered prescription.

2.1.2) Prove that

$$\sum_{q_i} |q_i\rangle\langle q_i| = 1$$

2.2.2) For a free particle with Hamiltonian

$$\hat{H} = \frac{\hat{p}^2}{2m}$$

write down the Δ - time slice approximation explicitly for (2.2.1) and (2.2.2)

2.2.3) Prove that if $A(t_i, t_f, q_i, q_f)$ is the amplitude to get from position q_i at t_i to position q_f at t_f then this function satisfies

$$\int dq' A(t_i, t', q_i, q') A(t', t_f, q', q_f) = A(t_i, t_f, q_i, q_f)$$

for $t_i < t' < t_f$.

2.3.1) Check that equations (2.3.1) are obtained by applying Dirac's quantisation prescription to Hamilton's Classical equations.

3.1.1) Derive equation 3.1.1.

3.2.2) Derive equation 3.2.1 from 3.2.3.

3.2.3) Derive (3.2.4)

3.2.4) Check the ϕ field describes a boson.

4.1.1) Using the free Lagrangian for ϕ but a new interaction $\lambda\phi^4$ compute, by commutation, the amplitude for 2 $\phi's$ of momenta q_1, q_2 to scatter into two ϕ's of momenta p_1, p_2 in lowest order.

4.2.1) Compute $\frac{\delta}{\delta J(y)}$ and $\frac{\delta^2}{\delta J(y)\delta J(z)}$ of

$$a) \int \phi(x)J(x)dx$$

$$b) \left[\int \phi(x)J(x)dx\right]^2$$

$$c) \int \phi^2(x)J(x)dx$$

4.2.2) Prove (4.2.7)

5.1.1) Check the product of 2 matrices of the form 5.1.1 gives another matrix of the same form.

5.1.2) Check that the commutators of the Gell-Mann λ_i matrices give linear combinations of the λ_i.

5.1.3) Prove 5.1.4.

5.1.4) Prove that the operators ($F_i = \frac{\lambda_i}{2}$)

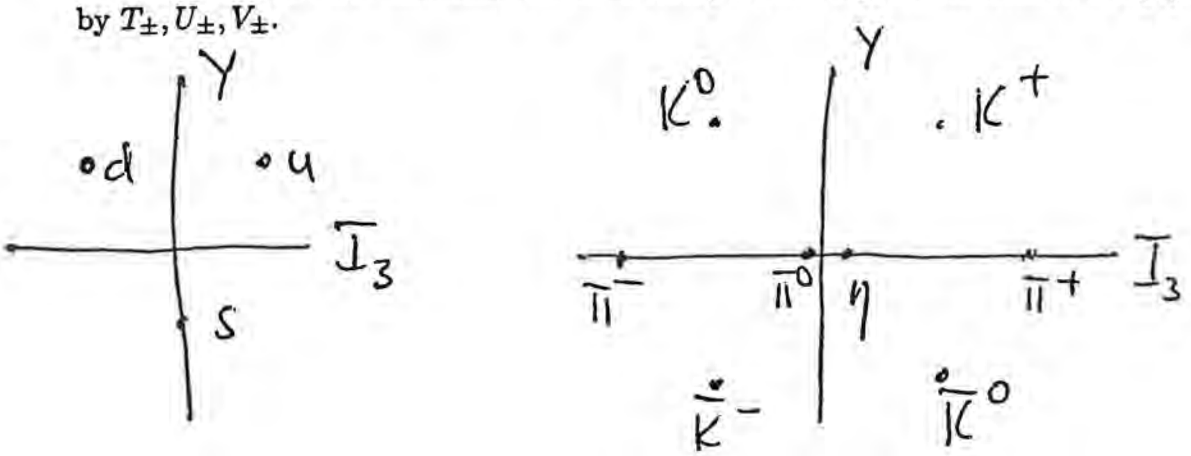
$$T_\pm = F_1 \pm iF_2, U_\pm = F_6 \pm iF_7, V_\pm = F_4 \pm iF_5$$

satisfy ($T_3 = F_3, Y = \frac{2}{\sqrt{3}}F_8$)

$$\begin{aligned} [T_3, T_\pm] &= \pm T_\pm & [Y, T_\pm] &= 0 \\ [T_3, U_\pm] &= \mp\frac{1}{2}U_\pm & [Y, U_\pm] &= \pm U_\pm \\ [T_3, V_\pm] &= \pm\frac{1}{2}V_\pm & [Y, V_\pm] &= \pm V_\pm \end{aligned}$$

Hence show that they operate as raising and lowering operators on eigenstates of T_3, Y.

Hence show how the states in the 3 dim., 8 dim. representations of $SU(3)$ are related by $T_\pm, U_\pm, V_\pm$.



6.1.1) You are used to the electric field due to a static charge q being given by a scalar potential

$$\phi(\underline{r}, t) = \frac{1}{4\pi\epsilon_0} . \frac{q}{r}$$

$$\underline{A}(\underline{r}, t) = 0$$

Prove it is equally well described by

$$\phi(\underline{x}, t) = 0$$

$$\underline{A}(\underline{r}, t) = \frac{tq\underline{r}}{4\pi\epsilon_0 . r^3}$$

So a voltmeter had better not measure ϕ ! What does it measure?

6.4.1) Another way of defining $F_{\mu\nu} = [D_\mu, D_\nu]$. Prove that this is equivalent to 6.4.5. Use this to prove 6.4.6.

INTRODUCTION TO QUANTUM ELECTRODYNAMICS AND QUANTUM CHROMODYNAMICS

By Dr J Flynn
University of Southampton

Lectures delivered at the School for Young High Energy Physicists
Rutherford Appleton Laboratory, September 1994

Introduction to Quantum Electrodynamics and Quantum Chromodynamics

Jonathan Flynn
Department of Physics
University of Southampton
Highfield, Southampton SO17 1BJ

Lectures presented at the School for Young High Energy Physicists, Rutherford Appleton Laboratory, September 1994.

Contents

1 Introduction

The aim of this course is to teach you how to calculate amplitudes, cross-sections and decay rates, particularly for quantum electrodynamics, QED, but in principle also for quantum chromodynamics, QCD. By the end of the course you should be able to go from a Feynman diagram such as the one for $e^+e^- \to \mu^+\mu^-$ in Figure 1.1(a), to a number for the cross-section, for example.

We will restrict ourselves to calculations at *tree level* but will also look qualitatively at higher order *loop* effects which amongst other things are responsible for the running of the QCD coupling constant. This running underlies the useful application of perturbative QCD calculations to high-energy processes. As you can guess, the sort of diagrams which are important here have closed loops of particle lines in them: in Figure 1.1(b) is one example contributing to the running of the strong coupling (the curly lines denote gluons).

In order to do our calculations we will need a certain amount of technology. In particular, we will need to describe particles with spin, especially the spin-1/2 leptons and quarks. We will therefore spend some time looking at the Dirac equation and its free particle solutions. After this will come revision of Fermi's golden rule to find probability amplitudes for transitions, followed by some general results on normalisation, flux factors and phase space, which will allow us to obtain formulas for cross sections and decay rates.

With these tools in hand, we will look at some examples of tree level QED processes. Here you will get hands-on experience of calculating transition amplitudes and getting from them to cross sections. We then move on to QCD. This will entail a brief introduction to renormalisation in both QED and QCD. We will introduce the idea of the running coupling constant and look at asymptotic freedom in QCD.

In reference [1] you will find a list of textbooks which may be useful.

1.1 Units and Conventions

I will use natural units, $c = 1$, $\hbar = 1$, so mass, energy, inverse length and inverse time all have the same dimensions.

$$\begin{array}{lll} \text{4-vector} & a^\mu \quad \mu = 0,1,2,3 & a = (a^0, \mathbf{a}) \\ \text{scalar product} & a{\cdot}b = a^0b^0 - \mathbf{a}{\cdot}\mathbf{b} = g_{\mu\nu}a^\mu b^\nu & \end{array} \tag{1.1}$$

From the scalar product you see that the metric is:

$$g = \mathrm{diag}(1,-1,-1,-1), \qquad g^{\mu\lambda}g_{\lambda\nu} = \delta^\mu_\nu = \begin{cases} 1 & \text{if } \mu = \nu \\ 0 & \text{if } \mu \neq \nu \end{cases} \tag{1.2}$$

For $c = 1$, $g^{\mu\nu}$ and $g_{\mu\nu}$ are numerically the same.

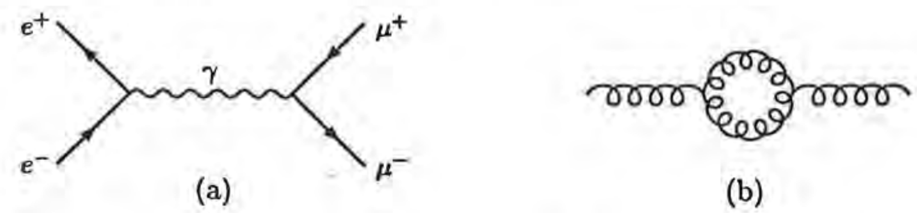


Figure 1.1 Examples of Feynman diagrams contributing to (a) $e^+e^- \to \mu^+\mu^-$ and (b) the running of the strong coupling constant.

From the above, you would think it natural to write the space components of a 4-vector as a^i for $i = 1, 2, 3$. However, for 3-vectors I will normally write the components as a_i. This is confusing only when you convert between ordinary vector equations and their covariant forms, when you have to remember the sign difference between a^i and a_i.

Note that ∂_μ is a covector,

$$\partial_\mu = \frac{\partial}{\partial x^\mu}, \quad \partial_\mu x^\nu = \delta^\nu_\mu, \tag{1.3}$$

so $\nabla^i = -\partial^i$ and $\partial^\mu = (\partial^0, -\nabla)$

My convention for the totally antisymmetric Levi-Civita tensor is:

$$\epsilon^{\mu\nu\lambda\sigma} = \begin{cases} +1 & \text{if } \{\mu,\nu,\lambda,\sigma\} \text{ an even permutation of } \{0,1,2,3\} \\ -1 & \text{if an odd permutation} \\ 0 & \text{otherwise} \end{cases} \tag{1.4}$$

Note that $\epsilon^{\mu\nu\lambda\sigma} = -\epsilon_{\mu\nu\lambda\sigma}$, and $\epsilon^{\mu\nu\lambda\sigma} p_\mu q_\nu r_\lambda s_\sigma \to (\det\Lambda)\, \epsilon^{\mu\nu\lambda\sigma} p_\mu q_\nu r_\lambda s_\sigma$ for Λ in the Lorentz group.

1.2 Relativistic Wave Equations?

Imagine you are working in the 1920's. You already know quantum mechanics based on Schrödinger's equation and you know relativity. You might ask if you can come up with some relativistic version of a quantum mechanical wave equation. If you do this, you encounter difficulties arising from the one-particle viewpoint, thinking of the equations describing a wave function. These difficulties are solved by ditching the wave function in favour of a *quantum field*, the subject of your quantum field theory course.

What is the problem with the one particle interpretation? Trouble arises from combining the uncertainty principle with the relativistic equivalence of mass and energy-momentum. If you try to localise a particle in a region with dimensions of order L, the particle's momentum and (in the relativistic regime) energy are uncertain by $\sim 1/L$. As the dimension L becomes smaller than the particle's inverse mass, $1/m$, states with more than one particle become energetically accessible. The more you try to localise a particle, the more you become uncertain whether you have one or any number of particles. Relativistic causality is inconsistent with a single particle theory and the real world evades the conflict through pair creation.

What happens in quantum field theory is that field *operators*, which can create or destroy multiparticle states, satisfy Heisenberg equations of motion. If there are no interactions, then the relevant equations are the Klein-Gordon equation for scalar fields or the Dirac equation for spin-half fields (such as the electron). The free quantum fields are expanded as linear combinations of plane wave solutions of these equations, but with operator valued coefficients which can create and destroy single particles. Thus we need to know the properties of the plane wave solutions. This is trivial for the scalar field, but is more interesting for the Dirac field. All the problems with "negative energy solutions" in the wave function approach are non-problems in quantum field theory: the negative energy parts multiply operators which destroy particles.

In fairness I should mention that you can get quite far with the one particle interpretation if you consider external forces which vary slowly on scales of order $1/m$, and thereby don't have enough energy to create new particle pairs. Notably, you can use the Dirac equation, which we'll meet below, in the presence of an electromagnetic field, to calculate fine structure in the spectra of hydrogen-like atoms (see textbooks such as Itzykson and Zuber [1] section 2.3 for example).

1.3 The Klein-Gordon Equation

In your quantum field theory course, you will show that the Heisenberg equations of motion for a free scalar field and its canonical conjugate give the *Klein-Gordon* equation

$$(\Box + m^2)\,\phi(x) = 0 \tag{1.5}$$

where

$$\Box = \partial_\mu \partial^\mu = \partial^2/\partial t^2 - \nabla^2 \tag{1.6}$$

and x is the 4-vector $(t, \mathbf{x})$. Using the substitutions,

$$E \to i\frac{\partial}{\partial t}, \qquad \mathbf{p} \to -i\boldsymbol{\nabla}, \tag{1.7}$$

you can see that the objects created or destroyed by ϕ satisfy the relativistic energy-momentum relation

$$E^2 = \mathbf{p}^2 + m^2. \tag{1.8}$$

The operator $\Box$ is Lorentz invariant, so the Klein-Gordon equation is relativistically covariant (that is, transforms into an equation of the same form) if ϕ is a scalar function. That is to say, under a Lorentz transformation $(t, \mathbf{x}) \to (t', \mathbf{x}')$,

$$\phi(t, \mathbf{x}) \to \phi'(t', \mathbf{x}') = \phi(t, \mathbf{x})$$

so ϕ is invariant. In particular ϕ is then invariant under spatial rotations so it represents a spin-zero particle (more on spin when we come to the Dirac equation), there being no preferred direction which could carry information on a spin orientation.

The Klein-Gordon equation has plane wave solutions

$$\phi(x) = Ne^{-i(Et - \mathbf{p}\cdot\mathbf{x})} \tag{1.9}$$

where N is a normalisation constant and $E = \pm\sqrt{\mathbf{p}^2 + m^2}$. Thus, there are both positive and negative energy solutions. In the quantum field ϕ, these are just associated with operators which create or destroy particles. However, they are a severe problem if you try to interpret ϕ as a wavefunction. The spectrum is no longer bounded below, and you can extract arbitrarily large amounts of energy from the system by driving it into ever more negative energy states. Any external perturbation capable of pushing a particle across the energy gap of $2m$ between the positive and negative energy continuum of states can uncover this difficulty.

A second problem with the wavefunction interpretation arises when you try to find a probability density. Since ϕ is Lorentz invariant, $|\phi|^2$ doesn't transform like a density. To search for a candidate we derive a continuity equation, rather as you did for the Schrödinger equation in the pre-school problems. Defining ρ and $\mathbf{J}$ by

$$\begin{aligned} \rho &\equiv i\left(\phi^* \frac{\partial \phi}{\partial t} - \phi \frac{\partial \phi^*}{\partial t}\right) \\ \mathbf{J} &\equiv -i\left(\phi^* \boldsymbol{\nabla}\phi - \phi \boldsymbol{\nabla}\phi^*\right) \end{aligned} \tag{1.10}$$

you obtain (see problem) a covariant conservation equation

$$\partial_\mu J^\mu = 0 \tag{1.11}$$

where J is the 4-vector $(\rho, \mathbf{J})$. It is natural to interpret ρ as a probability density and $\mathbf{J}$ as a probability current. However, for a plane wave solution (1.9), $\rho = 2|N|^2 E$, so ρ is not positive definite since we've already found E can be negative.

▷**Exercise 1.1**
Derive the continuity equation (1.11). Start with the Klein-Gordon equation multiplied by ϕ^* and subtract the complex conjugate of the K-G equation multiplied by ϕ.

Thus, ρ may well be considered as the density of a conserved quantity (such as electric charge), but we cannot use it for a probability density. To Dirac, this and the existence of negative energy solutions seemed so overwhelming that he was led to introduce another equation, first order in time derivatives but still Lorentz covariant, hoping that the similarity to Schrödinger's equation would allow a probability interpretation. In fact, with the interpretation of ϕ as a quantum field, these problems are not problems at all: the negative energy solutions will find an explanation in terms of antiparticles and ρ will indeed be a charge density as hinted above. Moreover, Dirac's hopes were unfounded because his new equation also turns out to admit negative energy solutions. Fortunately it is just what we need to describe particles with half a unit of spin angular momentum, so we will now turn to it.

2 The Dirac Equation

Dirac wanted an equation first order in time derivatives and Lorentz covariant, so it had to be first order in spatial derivatives too. His starting point was

$$i\frac{\partial\psi}{\partial t} = -i\,\boldsymbol{\alpha}\cdot\boldsymbol{\nabla}\psi + \beta m\psi \tag{2.1}$$

Remember that in field theory, the Dirac equation is the equation of motion for the field operator describing spin-1/2 fermions. In order for this equation to be Lorentz covariant, it will turn out that ψ cannot be a scalar under Lorentz transformations. In fact this will be precisely how the equation turns out to describe spin-1/2 particles. We will return to this below.

If ψ is to describe a free particle it is natural that it should satisfy the Klein-Gordon equation so that it has the correct energy-momentum relation. This requirement imposes relationships among the $\boldsymbol{\alpha}$ and β. To see these, apply the operator on each side of equation (2.1) twice,

$$-\frac{\partial^2\psi}{\partial t^2} = -\alpha^i\alpha^j\nabla^i\nabla^j - i\,(\beta\alpha^i + \alpha^i\beta)m\nabla^i\psi + \beta^2 m^2\psi$$

The Klein-Gordon equation will be satisfied if

$$\begin{aligned} \alpha_i\alpha_j + \alpha_j\alpha_i &= 2\delta_{ij} \\ \beta\alpha_i + \alpha_i\beta &= 0 \\ \beta^2 &= 1 \end{aligned} \tag{2.2}$$

for $i, j = 1, 2, 3$. It is clear that the α_i and β cannot be ordinary numbers, but it is natural to give them a realisation as matrices. In this case, ψ must be a multi-component *spinor* on which these matrices act.

▷**Exercise 2.1**
Prove that any matrices $\boldsymbol{\alpha}$ and β satisfying equation (2.2) are traceless with eigenvalues ± 1. Hence argue that they must be even dimensional.

In two dimensions a natural set of matrices for the $\boldsymbol{\alpha}$ would be the Pauli matrices

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \qquad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{2.3}$$

However, there is no other independent 2×2 matrix with the right properties for β, so the smallest dimension for which the Dirac matrices can be realised is four. One choice is the *Dirac representation*

$$\boldsymbol{\alpha} = \begin{pmatrix} 0 & \boldsymbol{\sigma} \\ \boldsymbol{\sigma} & 0 \end{pmatrix}, \qquad \beta = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{2.4}$$

Note that each entry above denotes a two-by-two block and that the 1 denotes the 2×2 identity matrix.

There is a theorem due to Pauli which states that all sets of matrices obeying the relations in (2.2) are equivalent. Since the Hermitian conjugates $\boldsymbol{\alpha}^\dagger$ and $\beta^\dagger$ clearly obey the relations, you can, by a change of basis if necessary, assume that $\boldsymbol{\alpha}$ and β are Hermitian. All the common choices of basis have this property. Furthermore, we would like α_i and β to be Hermitian so that the Dirac Hamiltonian (2.14) is Hermitian.

▷**Exercise 2.2**
Derive the continuity equation $\partial_\mu J^\mu = 0$ for the Dirac equation with

$$\rho = J^0 = \psi^\dagger(x)\psi(x), \qquad \mathbf{J} = \psi^\dagger(x)\boldsymbol{\alpha}\psi(x). \tag{2.5}$$

We will see in section 2.6 that $(\rho, \mathbf{J})$ does indeed transform as a four-vector.

2.1 Free Particle Solutions I: Interpretation

We look for plane wave solutions of the form

$$\psi = \begin{pmatrix} \phi \\ \chi \end{pmatrix} e^{-i(Et-\mathbf{p}\cdot\mathbf{x})}$$

where ϕ and χ are two-component spinors, independent of x. Using the Dirac representation, the Dirac equation gives

$$E\begin{pmatrix} \phi \\ \chi \end{pmatrix} = \begin{pmatrix} m & \boldsymbol{\sigma}\cdot\mathbf{p} \\ \boldsymbol{\sigma}\cdot\mathbf{p} & -m \end{pmatrix}\begin{pmatrix} \phi \\ \chi \end{pmatrix},$$

so that

$$\chi = \frac{\boldsymbol{\sigma}\cdot\mathbf{p}}{E+m}\phi, \qquad \phi = \frac{\boldsymbol{\sigma}\cdot\mathbf{p}}{E-m}\chi.$$

For $E \neq -m$ there are solutions,

$$\psi(x) = \begin{pmatrix} \phi \\ \frac{\boldsymbol{\sigma}\cdot\mathbf{p}}{E+m}\phi \end{pmatrix} e^{-i(Et-\mathbf{p}\cdot\mathbf{x})}, \tag{2.6}$$

while for $E \neq m$ there are solutions,

$$\psi(x) = \begin{pmatrix} \frac{\boldsymbol{\sigma}\cdot\mathbf{p}}{E-m}\chi \\ \chi \end{pmatrix} e^{-i(Et-\mathbf{p}\cdot\mathbf{x})}, \tag{2.7}$$

for arbitrary constant ϕ and χ. Now, since $E^2 = \mathbf{p}^2 + m^2$ by construction, we find, just as we did for the Klein-Gordon equation (1.5), that there exist positive and negative energy solutions given by equations (2.6) and (2.7) respectively. Once again, the existence of negative energy solutions vitiates the interpretation of ψ as a wavefunction.

Dirac interpreted the negative energy solutions by postulating the existence of a "sea" of negative energy states. The vacuum or ground state has all the negative energy states full. An additional electron must now occupy a positive energy state since the Pauli exclusion principle forbids it from falling into one of the filled negative energy states. By promoting one of these negative energy states to a positive energy one, by supplying energy, you create a pair: a positive energy electron and a hole in the negative energy sea corresponding to a positive energy positron. This was a radical new idea, and brought pair creation and antiparticles into physics. Positrons were discovered in cosmic rays by Carl Anderson in 1932.

The problem with Dirac's hole theory is that it doesn't work for bosons, such as particles governed by the Klein Gordon equation, for example. Such particles have no exclusion principle to stop them falling into the negative energy states, releasing their energy. We need a new interpretation and turn to Feynman for our answer.

According to Feynman and quantum field theory, we should interpret the emission (absorption) of a negative energy particle with momentum p^μ as the absorption (emission)

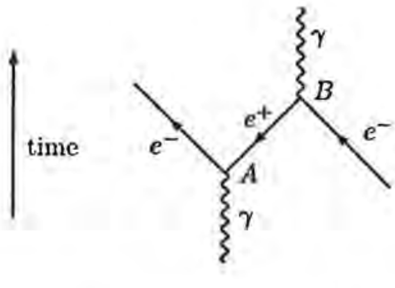


Figure 2.1 Feynman interpretation of a process in which a negative energy electron is absorbed. Time increases moving upwards.

of a positive energy antiparticle with momentum $-p^\mu$. So, in Figure 2.1, for example, an electron–positron pair is created at point A. The positron propagates to point B where it is annihilated by another electron.

Thus Feynman tells us to keep both types of free particle solution. One is to be used for particles and the other for the accompanying antiparticles. Let's return to our spinor solutions and write them in a conventional form. Take the positive energy solution of equation (2.6) and write,

$$\sqrt{E+m}\begin{pmatrix} \chi_r \\ \frac{\sigma\cdot\mathbf{p}}{E+m}\chi_r \end{pmatrix} e^{-ip\cdot x} \equiv u_p^r e^{-ip\cdot x}. \tag{2.8}$$

For the former negative energy solution of equation (2.7), change the sign of the energy, $E \to -E$, and the three-momentum, $\mathbf{p} \to -\mathbf{p}$, to obtain,

$$\sqrt{E+m}\begin{pmatrix} \frac{\sigma\cdot\mathbf{p}}{E+m}\chi_r \\ \chi_r \end{pmatrix} e^{ip\cdot x} \equiv v_p^r e^{ip\cdot x}. \tag{2.9}$$

In these two solutions E is now (and for the rest of the course) always positive and given by $E = (\mathbf{p}^2 + m^2)^{1/2}$. The subscript r takes the values $1, 2$, with

$$\chi_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \qquad \chi_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}. \tag{2.10}$$

At this point I would like to introduce another notation, and define

$$\omega_p \equiv \sqrt{\mathbf{p}^2 + m^2}. \tag{2.11}$$

so that, ω_p is the energy (positive) of a particle or anti-particle with three-momentum $\mathbf{p}$ (I write the subscript p instead of $\mathbf{p}$, but you should remember it really means the three-momentum). I will tend to use E or ω_p interchangeably.

The u-spinor solutions will correspond to particles and the v-spinor solutions to antiparticles. The role of the two χ's will become clear in the following section, where it will be shown that the two choices of r are spin labels. Note that each spinor solution depends on the three-momentum $\mathbf{p}$, so it is implicit that $p^0 = \omega_p$. In the expansion of the Dirac quantum field operator in terms of plane waves,

$$\psi(x) = \int \frac{d^3p}{(2\pi)^3\, 2\omega_p} \sum_{r=1,2} \left[b(p,r) u_p^r e^{-ip\cdot x} + d^\dagger(p,r) v_p^r e^{ip\cdot x} \right] \tag{2.12}$$

the operator b annihilates a fermion of momentum $(\omega_p, \mathbf{p})$ and spin r, whilst $d^\dagger$ creates an antifermion of momentum $(\omega_p, \mathbf{p})$ and spin r. The Hermitian conjugate Dirac field contains operators which do the opposite. This discussion should be clearer after your quantum field theory lectures.

The vacuum state $|0\rangle$ is defined by,

$$b(p,r)\,|0\rangle = d(p,r)\,|0\rangle = 0, \tag{2.13}$$

for every momentum $p = (\omega_p, \mathbf{p})$ and spin label r. This ensures the interpretation above: particles are created by the "daggered" operators and destroyed by the undaggered ones.

2.2 Free Particle Solutions II: Spin

Now it's time to justify the statements we have been making that the Dirac equation describes spin-1/2 particles. The Dirac Hamiltonian in momentum space is

$$H = \boldsymbol{\alpha}\cdot\mathbf{p} + \beta m \tag{2.14}$$

and the orbital angular momentum operator is

$$\mathbf{L} = \mathbf{r} \times \mathbf{p}.$$

Normally you have to worry about operator ordering ambiguities when going from classical objects to quantum mechanical ones. For $\mathbf{L}$ the problem does not arise — why not?

Evaluating the commutator of $\mathbf{L}$ with H,

$$\begin{aligned} [\mathbf{L}, H] &= [\mathbf{r}\times\mathbf{p}, \boldsymbol{\alpha}\cdot\mathbf{p}] \\ &= [\mathbf{r}, \boldsymbol{\alpha}\cdot\mathbf{p}] \times \mathbf{p} \\ &= i\boldsymbol{\alpha}\times\mathbf{p}, \end{aligned} \tag{2.15}$$

we see that the orbital angular momentum is not conserved. We'd like to find a *total* angular momentum $\mathbf{J}$ which *is* conserved, by adding an additional operator $\mathbf{S}$ to $\mathbf{L}$,

$$\mathbf{J} = \mathbf{L} + \mathbf{S}.$$

To this end, consider the three matrices,

$$\boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\sigma} & 0 \\ 0 & \boldsymbol{\sigma} \end{pmatrix} = -i\alpha_1\alpha_2\alpha_3\boldsymbol{\alpha}. \tag{2.16}$$

The $\boldsymbol{\Sigma}/2$ have the correct commutation relations to represent angular momentum, since the Pauli matrices do, and their commutators with $\boldsymbol{\alpha}$ and β are,

$$[\boldsymbol{\Sigma}, \beta] = 0, \qquad [\Sigma_i, \alpha_j] = 2i\epsilon_{ijk}\alpha_k. \tag{2.17}$$

▷**Exercise 2.3**
Verify the commutation relations in equation (2.17).

From the relations in (2.17) we find that

$$[\boldsymbol{\Sigma}, H] = -2i\boldsymbol{\alpha}\times\mathbf{p}.$$

Comparing this with the commutator of $\mathbf{L}$ with H in equation (2.15), you readily see that

$$\left[\mathbf{L} + \tfrac{1}{2}\boldsymbol{\Sigma}, H\right] = 0,$$

and we can set

$$\mathbf{S} = \frac{1}{2}\mathbf{\Sigma}.$$

We interpret $\mathbf{S}$ as an angular momentum *intrinsic* to the particle. Now

$$\mathbf{S}^2 = \frac{1}{4}\begin{pmatrix} \boldsymbol{\sigma}\cdot\boldsymbol{\sigma} & 0 \\ 0 & \boldsymbol{\sigma}\cdot\boldsymbol{\sigma} \end{pmatrix} = \frac{3}{4}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

and recalling that the eigenvalue of $\mathbf{J}^2$ for spin j is $j(j+1)$, we conclude that $\mathbf{S}$ represents spin-1/2 and the solutions of the Dirac equation have spin-1/2 as promised.

We worked in the Dirac representation for convenience, but the result is of course independent of the representation.

Now consider the u-spinor solutions u_p^r of equation (2.8). Choose $\mathbf{p} = (0, 0, p_z)$ and write

$$u_\uparrow = u_{p_z}^1 = \begin{pmatrix} \sqrt{E+m} \\ 0 \\ \sqrt{E-m} \\ 0 \end{pmatrix}, \qquad u_\downarrow = u_{p_z}^2 = \begin{pmatrix} 0 \\ \sqrt{E+m} \\ 0 \\ -\sqrt{E-m} \end{pmatrix}. \tag{2.18}$$

It is easy to see that,

$$S_z u_\uparrow = \frac{1}{2}u_\uparrow, \qquad S_z u_\downarrow = -\frac{1}{2}u_\downarrow.$$

So, these two spinors represent spin up and spin down along the z-axis respectively. For the v-spinors, with the same choice for $\mathbf{p}$, write,

$$v_\downarrow = v_{p_z}^1 = \begin{pmatrix} \sqrt{E-m} \\ 0 \\ \sqrt{E+m} \\ 0 \end{pmatrix}, \qquad v_\uparrow = v_{p_z}^2 = \begin{pmatrix} 0 \\ -\sqrt{E-m} \\ 0 \\ \sqrt{E+m} \end{pmatrix}, \tag{2.19}$$

where now,

$$S_z v_\downarrow = \frac{1}{2}v_\downarrow, \qquad S_z v_\uparrow = -\frac{1}{2}v_\uparrow.$$

This apparently perverse choice of up and down for the v's is because, as you see in equation (2.12) for the quantum Dirac field, $u_\uparrow$ multiplies an annihilation operator which *destroys* a particle with momentum p_z and spin up, whereas $v_\downarrow$ multiplies an operator which *creates* an antiparticle with momentum p_z and spin up.

2.3 Normalisation, Gamma Matrices

We have included a normalisation factor $\sqrt{E+m}$ in our spinors. With this factor,

$$u_p^{r\dagger} u_p^s = v_p^{r\dagger} v_p^s = 2\omega_p \delta^{rs}. \tag{2.20}$$

This corresponds to the standard relativistic normalisation of $2\omega_p$ particles per unit volume. It also means that $u^\dagger u$ transforms like the time component of a 4-vector under Lorentz transformations as we will see in section 2.6.

▷ **Exercise 2.4**
Check the normalisation condition for the spinors in equation (2.20).

I will now introduce (yet) more standard notation. Define the *gamma matrices*,

$$\gamma^0 = \beta, \qquad \boldsymbol{\gamma} = \beta\boldsymbol{\alpha}. \tag{2.21}$$

In the Dirac representation,

$$\gamma^0 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \qquad \boldsymbol{\gamma} = \begin{pmatrix} 0 & \boldsymbol{\sigma} \\ -\boldsymbol{\sigma} & 0 \end{pmatrix}. \tag{2.22}$$

In terms of these, the relations between the $\boldsymbol{\alpha}$ and β in equation (2.2) can be written compactly as,

$$\{\gamma^\mu, \gamma^\nu\} = 2g^{\mu\nu}. \tag{2.23}$$

Combinations like $a_\mu\gamma^\mu$ occur frequently and are conventionally written as,

$$\not{a} = a_\mu\gamma^\mu = a^\mu\gamma_\mu,$$

pronounced "a slash." Note that γ^μ is not, despite appearances, a 4-vector — it just denotes a set of four matrices. However, the notation is deliberately suggestive, for when combined with Dirac fields you can construct quantities which transform like vectors and other Lorentz tensors (see the next section).

Let's close this section by observing that using the gamma matrices the Dirac equation (2.1) becomes

$$(i\not{\partial} - m)\psi = 0, \tag{2.24}$$

or in momentum space,

$$(\not{p} - m)\psi = 0. \tag{2.25}$$

The spinors u and v satisfy

$$\begin{aligned} (\not{p} - m)u_p^r &= 0 \\ (\not{p} + m)v_p^r &= 0 \end{aligned} \tag{2.26}$$

▷**Exercise 2.5**
Derive the momentum space equations satisfied by u_p^r and v_p^r.

2.4 Lorentz Covariance

We want the Dirac equation (2.24) to preserve its form under Lorentz transformations (LT's). Let $\Lambda^\mu{}_\nu$ represent an LT,

$$x^\mu \to x'^\mu = \Lambda^\mu{}_\nu x^\nu \tag{2.27}$$

The requirement is,

$$(i\gamma^\mu\partial_\mu - m)\psi(x) = 0 \quad \longrightarrow \quad (i\gamma^\mu\partial'_\mu - m)\psi'(x') = 0,$$

where $\partial_\mu = \Lambda^\sigma{}_\mu\partial'_\sigma$. We know that 4-vectors get their components mixed up by LT's, so we expect that the components of ψ might get mixed up also,

$$\psi(x) \to \psi'(x') \equiv S(\Lambda)\psi(\Lambda^{-1}x') \tag{2.28}$$

where $S(\Lambda)$ is a 4×4 matrix acting on the spinor index of ψ. Note that the argument $\Lambda^{-1}x'$ is just a fancy way of writing x.

To determine S we rewrite the Dirac equation in terms of the primed variables,

$$(i\gamma^\mu \Lambda^\sigma{}_\mu \partial'_\sigma - m)\psi(\Lambda^{-1}x') = 0. \tag{2.29}$$

The matrices $\gamma'^\sigma \equiv \gamma^\mu \Lambda^\sigma{}_\mu$ satisfy the same anticommutation relations as the γ^μ's in equation (2.23),

$$\{\gamma'^\mu, \gamma'^\nu\} = 2g^{\mu\nu}. \tag{2.30}$$

▷**Exercise 2.6**
Check relation (2.30).

Now we invoke the theorem (Pauli's theorem) which states that any two representations of the gamma matrices are equivalent. This means that there is a matrix $S(\Lambda)$ such that

$$\gamma'^\mu = S^{-1}(\Lambda)\gamma^\mu S(\Lambda). \tag{2.31}$$

This allows us to rewrite equation (2.29) as

$$(i\gamma^\mu \partial'_\mu - m)S(\Lambda)\psi(\Lambda^{-1}x') = 0,$$

so that the Dirac equation does indeed preserve its form. To construct S explicitly for an infinitesimal LT, let,

$$\Lambda^\mu{}_\nu = \delta^\mu{}_\nu - \epsilon(g^{\rho\mu}\delta^\sigma{}_\nu - g^{\sigma\mu}\delta^\rho{}_\nu) \tag{2.32}$$

where ϵ is an infinitesimal parameter and ρ and σ are fixed. Since this expression is antisymmetric in ρ and σ there are six choices for the pair (ρ, σ) corresponding to three rotations and three boosts. Writing,

$$S(\Lambda) = 1 + i\epsilon s^{\rho\sigma} \tag{2.33}$$

where $s^{\rho\sigma}$ is a matrix to be determined, we find that equation (2.31) for γ' is satisfied by,

$$s^{\rho\sigma} = \frac{i}{4}[\gamma^\rho, \gamma^\sigma] \equiv \frac{1}{2}\sigma^{\rho\sigma}. \tag{2.34}$$

▷**Exercise 2.7**
Verify that equation (2.31) relating γ' and γ is satisfied by $s^{\rho\sigma}$ defined through equations (2.33) and (2.34).

We have thus determined how ψ transforms under LT's. To find quantities which are Lorentz invariant, or transform as vectors or tensors, we need to introduce the Pauli and Dirac adjoints. The Pauli adjoint $\overline{\psi}$ of a spinor ψ is defined by

$$\overline{\psi} \equiv \psi^\dagger \gamma^0 = \psi^\dagger \beta. \tag{2.35}$$

The Dirac adjoint is defined by

$$(\overline{\psi} A \phi)^* = \overline{\phi}\,\overline{A}\psi. \tag{2.36}$$

For Hermitian γ^0 it is easy to show that

$$\overline{A} = \gamma^0 A^\dagger \gamma^0. \tag{2.37}$$

Some properties of the Pauli and Dirac adjoints are:

$$\begin{aligned} \overline{(\lambda A + \mu B)} &= \lambda^* \overline{A} + \mu^* \overline{B}, \\ \overline{AB} &= \overline{B}\,\overline{A}, \\ \overline{A\psi} &= \overline{\psi}\,\overline{A}. \end{aligned}$$

With these definitions, $\overline{\psi}$ transforms as follows under LT's:

$$\overline{\psi} \to \overline{\psi}' = \overline{\psi} S^{-1}(\Lambda) \tag{2.38}$$

▷**Exercise 2.8**

(1) Verify that $\gamma^{\mu\dagger} = \gamma^0\gamma^\mu\gamma^0$. This says that $\overline{\gamma}^\mu = \gamma^\mu$.

(2) Using (2.33) and (2.34) verify that $\gamma^0 S^\dagger(\Lambda)\gamma^0 = S^{-1}(\Lambda)$, i.e. $\overline{S} = S^{-1}$. So S is not unitary in general, although it *is* unitary for rotations (when ρ and σ are spatial indices). This is because the rotations are in the unitary $O(3)$ subgroup of the nonunitary Lorentz group. Here you show the result for an infinitesimal LT, but it is true for finite LT's.

(3) Show that $\overline{\psi}$ satisfies the equation

$$\overline{\psi}(-i\overleftarrow{\not\partial} - m) = 0$$

where the arrow over $\not\partial$ implies the derivative acts on $\overline{\psi}$.

(4) Hence prove that $\overline{\psi}$ transforms as in equation (2.38).

Note that result (2) of the problem above can be rewritten as $\overline{S}(\Lambda) = S^{-1}(\Lambda)$, and equation (2.31) for the similarity transformation of γ^μ to γ'^μ takes the form,

$$\overline{S}\gamma^\mu S = \Lambda^\mu{}_\nu\gamma^\nu. \tag{2.39}$$

Combining the transformation properties of ψ and $\overline{\psi}$ from equations (2.28) and (2.38) we see that the bilinear $\overline{\psi}\psi$ is Lorentz invariant. In section 2.6 we'll consider the transformation properties of general bilinears.

Let me close this section by recasting the spinor normalisation equations (2.20) in terms of "Dirac inner products." The conditions become,

$$\begin{array}{rcll} \overline{u}^r_p u^s_p & = & 2m\delta^{rs} & \\ \overline{u}^r_p v^s_p & = & 0 & = \overline{v}^r_p u^s_p \\ \overline{v}^r_p v^s_p & = & -2m\delta^{rs} & \end{array} \tag{2.40}$$

▷**Exercise 2.9**

Verify the normalisation properties in the above equations (2.40).

2.5 Parity

In the next section we are going to construct quantities bilinear in ψ and $\overline{\psi}$, and classify them according to their transformation properties under LT's. We normally use LT's which are in the connected Lorentz Group, $SO(3,1)$, meaning they can be obtained by a continuous deformation of the identity transformation. Indeed in the last section we considered LT's very close to the identity in equation (2.32). The full Lorentz group has four components generated by combining the $SO(3,1)$ transformations with the discrete operations of parity or space inversion, P, and time reversal, T,

$$\Lambda_P = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}, \qquad \Lambda_T = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

LT's satisfy $\Lambda^T g\Lambda = g$ (see the preschool problems), so taking determinants shows that $\det\Lambda = \pm 1$. LT's in $SO(3,1)$ have determinant 1, since the identity does, but the P and T operations have determinant -1.

Let's now find the action of parity on the Dirac wavefunction and determine the wavefunction ψ_P in the parity-reversed system. According to the discussion of the previous section, and using the result of equation (2.39), we need to find a matrix S satisfying

$$\overline{S}\gamma^0 S = \gamma^0, \qquad \overline{S}\gamma^i S = -\gamma^i.$$

It's not hard to see that $S = \overline{S} = \gamma^0$ is an acceptable solution, from which it follows that the wavefunction ψ_P is

$$\psi_P(t,\mathbf{x}) = \gamma^0\psi(t,-\mathbf{x}). \tag{2.41}$$

In fact you could multiply γ^0 by a phase and still have an acceptable definition for the parity transformation.

In the nonrelativistic limit, the wavefunction ψ approaches an eigenstate of parity. Since

$$\gamma^0 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix},$$

the u-spinors and v-spinors at rest have opposite eigenvalues, corresponding to particle and antiparticle having opposite *intrinsic* parities.

2.6 Bilinear Covariants

Now, as promised, we will construct and classify the bilinears. To begin, observe that by forming products of the gamma matrices it is possible to construct 16 linearly independent quantities. In equation (2.34) we have defined

$$\sigma^{\mu\nu} \equiv \frac{i}{2}[\gamma^\mu, \gamma^\nu],$$

and now it is convenient to define

$$\gamma_5 \equiv \gamma^5 \equiv i\gamma^0\gamma^1\gamma^2\gamma^3, \tag{2.42}$$

with the properties,

$$\gamma_5^\dagger = \gamma_5, \qquad \{\gamma_5, \gamma^\mu\} = 0.$$

Then the set of 16 matrices

$$\Gamma : \{1, \gamma_5, \gamma^\mu, \gamma^\mu\gamma_5, \sigma^{\mu\nu}\}$$

form a basis for gamma matrix products.

Using the transformations of ψ and $\overline{\psi}$ from equations (2.28) and (2.38), together with the similarity transformation of γ^μ in equation (2.39), construct the 16 fermion bilinears and their transformation properties as follows:

$$\begin{array}{rcll} \overline{\psi}\psi & \to & \overline{\psi}\psi & \text{S scalar} \\ \overline{\psi}\gamma_5\psi & \to & \det(\Lambda)\,\overline{\psi}\gamma_5\psi & \text{P pseudoscalar} \\ \overline{\psi}\gamma^\mu\psi & \to & \Lambda^\mu{}_\nu\overline{\psi}\gamma^\nu\psi & \text{V vector} \\ \overline{\psi}\gamma^\mu\gamma_5\psi & \to & \det(\Lambda)\,\Lambda^\mu{}_\nu\overline{\psi}\gamma^\nu\gamma_5\psi & \text{A axial vector} \\ \overline{\psi}\sigma^{\mu\nu}\psi & \to & \Lambda^\mu{}_\lambda\Lambda^\nu{}_\sigma\overline{\psi}\sigma^{\lambda\sigma}\psi & \text{T tensor} \end{array} \tag{2.43}$$

▷ **Exercise 2.10**
Verify the transformation properties of the bilinears in equation (2.43).

Observe that $\overline{\psi}\gamma^\mu\psi = (\rho, \mathbf{J})$ is just the current we found earlier in equation (2.5). Classically ρ is positive definite, but for the quantum Dirac field you find that the space integral of ρ is the charge operator, which counts the number of electrons minus the number of positrons,

$$Q \sim \int d^3x\, \psi^\dagger\psi \sim \int d^3p\, [b^\dagger b - d^\dagger d].$$

The continuity equation $\partial_\mu J^\mu = 0$ expresses conservation of electric charge.

2.7 Charge Conjugation

There is one more discrete invariance of the Dirac equation in addition to parity. It is charge conjugation, which takes you from particle to antiparticle and vice versa. For scalar fields the symmetry is just complex conjugation, but in order for the charge conjugate Dirac field to remain a solution of the Dirac equation, you have to mix its components as well:

$$\psi \to \psi_C = C\overline{\psi}^T.$$

Here $\overline{\psi}^T = \gamma^{0T}\psi^*$ and C is a matrix satisfying the condition

$$C\gamma_\mu^T C^{-1} = -\gamma_\mu.$$

In the Dirac representation,

$$C = i\gamma^2\gamma^0 = \begin{pmatrix} 0 & -i\sigma^2 \\ -i\sigma^2 & 0 \end{pmatrix}.$$

I refer you to textbooks such as [1] for details.

When Dirac wrote down his equation everybody thought parity and charge conjugation were exact symmetries of nature, so invariance under these transformations was essential. Now we know that neither of them, nor the combination *CP*, are respected by the standard electroweak model.

2.8 Neutrinos

In the particle data book [2] you will find only upper limits for the masses of the three neutrinos, and in the standard model they are massless. Let's look therefore at solutions of the Dirac equation with $m = 0$. Specialising from equation (2.1), we have, in momentum space,

$$|\mathbf{p}|\psi = \boldsymbol{\alpha}\cdot\mathbf{p}\,\psi.$$

For such a solution,

$$\gamma_5\psi = \gamma_5\frac{\boldsymbol{\alpha}\cdot\mathbf{p}}{|\mathbf{p}|}\psi = 2\frac{\mathbf{S}\cdot\mathbf{p}}{|\mathbf{p}|}\psi,$$

using the spin operator $\mathbf{S} = \frac{1}{2}\boldsymbol{\Sigma} = \frac{1}{2}\gamma_5\boldsymbol{\alpha}$, with $\boldsymbol{\Sigma}$ defined in equation (2.16). But $\mathbf{S}\cdot\mathbf{p}/|\mathbf{p}|$ is the projection of spin onto the direction of motion, known as the *helicity*, and is equal to $\pm 1/2$. Thus $(1+\gamma_5)/2$ projects out the neutrino with helicity $1/2$ (right handed) and $(1-\gamma_5)/2$ projects out the neutrino with helicity $-1/2$ (left handed). To date, only left handed neutrinos have been observed, and only left handed neutrinos appear in the standard model. Since

$$\gamma^0\frac{1}{2}(1-\gamma_5)\psi = \frac{1}{2}(1+\gamma_5)\gamma^0\psi,$$

any theory involving only left handed neutrinos necessarily violates parity.

The standard model contains only left handed massless neutrinos. It is really the electroweak symmetry which prevents them having masses, not the fact that they are left handed only. It would be possible to doctor the standard model to contain so-called Majorana neutrinos which can be massive. However, this would entail relinquishing lepton number conservation and break the electroweak symmetry (or involve the introduction of new particles).

2.9 Dirac Lagrangian

In the spirit of the field theory course, we could have started out by looking for objects, transforming in the right way under Lorentz transformations and rotations, to represent spin-1/2 particles. This would have led us to Dirac spinors, for which we would have shown that

$$\mathcal{L} = \overline{\psi}(i\partial\!\!\!/ - m)\psi$$

is a Lorentz invariant Lagrangian.

Then Lagrange's equations immediately give the Dirac equation, as you can see simply from $\partial\mathcal{L}/\partial\overline{\psi} = 0$ (observing that $\mathcal{L}$ is independent of $\partial\overline{\psi}/\partial t$). Now you could quantise by Hamiltonian or path integral methods. A new feature that appears is that, for consistency, you must impose canonical *anti*commutation relations in the Hamiltonian form, or use *anti*commuting (Grassman) variables in the path integral. Thus, the connection between spin and statistics appears.

3 Cross Sections and Decay Rates

In section 4 we will learn how to calculate quantum mechanical amplitudes for electromagnetic scattering and decay processes. These amplitudes are obtained from the Lagrangian of QED, and contain information about the dynamics underlying the scattering or decay process. This section is a brief review of how to get from the quantum mechanical amplitude to a cross section or decay rate which can be measured. We will commence by recalling Fermi's golden rule for transition probabilities.

3.1 Fermi's Golden Rule

Consider a system with Hamiltonian H which can be written

$$H = H_0 + V \tag{3.1}$$

We assume that the eigenstates and eigenvalues of H_0 are known and that V is a small, possibly time-dependent, perturbation. The equation of motion of the system is,

$$i\frac{\partial}{\partial t}\left|\psi(t)\right\rangle = (H_0 + V)\left|\psi(t)\right\rangle . \tag{3.2}$$

If V vanished, we could calculate the time evolution of $|\psi(t)\rangle$ by expanding it as a linear combination of energy eigenstates. To develop a perturbation theory in V we will change our basis of states from the Schrödinger picture to the *interaction* or Dirac picture, where we hide the time evolution due to H_0 and concentrate on that due to V. Thus we define the interaction picture states and operators by,

$$\left|\psi_I(t)\right\rangle \equiv e^{iH_0t}\left|\psi(t)\right\rangle , \qquad \mathcal{O}_I(t) \equiv e^{iH_0t}\mathcal{O}(t)e^{-iH_0t}, \tag{3.3}$$

so that the interaction picture and Schrödinger picture states agree at time $t = 0$, $|\psi_I(0)\rangle = |\psi(0)\rangle$, with a similar relation for the operators. In the new basis, the equation of motion becomes,

$$i\frac{\partial}{\partial t}\left|\psi_I(t)\right\rangle = V_I(t)\left|\psi_I(t)\right\rangle , \tag{3.4}$$

which can be integrated formally as an infinite series in V,

$$\left|\psi_I(t)\right\rangle = \left[1 + \sum_{n=1}^{\infty}\frac{1}{i^n}\int_{-T/2}^{t}dt_1\int_{-T/2}^{t_1}dt_2\cdots\int_{-T/2}^{t_{n-1}}dt_n\, V_I(t_1)V_I(t_2)\cdots V_I(t_n)\right]\left|\psi(-T/2)\right\rangle . \tag{3.5}$$

Here, we have chosen to start with some (known) state $|\psi_I(-T/2)\rangle$, at time $-T/2$, and have evolved it to $|\psi_I(t)\rangle$ at time t. The evolution is done by the operator, U, that you've seen in the field theory course:

$$\left|\psi_I(t)\right\rangle = U(t,-T/2)\left|\psi_I(-T/2)\right\rangle .$$

Now consider the calculation of the probability of a transition to an eigenstate $|b\rangle$ at time t. The amplitude is,

$$\begin{aligned}\langle b|\psi(t)\rangle &= \langle b_I|\psi_I(t)\rangle \\ &= \langle b|\, e^{-iH_0t}\left|\psi_I(t)\right\rangle \\ &= e^{-iE_bt}\langle b|\psi_I(t)\rangle,\end{aligned}$$

so $\left|\langle b|\psi(t)\rangle\right|^2 = \left|\langle b|\psi_I(t)\rangle\right|^2$. We let V be time independent and consider the amplitude for a transition from an eigenstate $|a\rangle$ of H_0 at $t = -T/2$ to an orthogonal eigenstate $|b\rangle$ at $t = T/2$. The idea is that at very early or very late times H_0 describes some set of free particles. We allow some of these particles to approach each other and scatter under the influence of V, then look again a long time later when the outgoing particles are propagating freely under H_0 again. To first order in V,

$$\langle b\,|\psi_I(T/2)\rangle = -i\int_{-T/2}^{T/2} \langle b|\, V_I(t)\,|a\rangle\, dt = -i\langle b|V|a\rangle \int_{-T/2}^{T/2} e^{i\omega_{ba}t}dt,$$

where $\omega_{ba} = E_b - E_a$.

▷Exercise 3.1

Show that for $T \to \infty$ the first order transition amplitude for general V can be written in the covariant form

$$\langle b\,|\psi_I(\infty)\rangle = -i\int d^4x\, \phi_b^*(x) V \phi_a(x),$$

where $\phi_i(x) \equiv \phi_i(\mathbf{x})e^{-E_i t}$ and $\phi_i(\mathbf{x})$ is the usual Schrödinger wavefunction for a stationary state of H_0, with energy E_i.

The transition rate for time independent V is,

$$\frac{\left|\langle b\,|\psi_I(T/2)\rangle\,\right|^2}{T} = \left|\langle b|V|a\rangle\right|^2 \frac{4\sin^2(\omega_{ba}T/2)}{\omega_{ba}^2 T}.$$

If $E_b \neq E_a$, this probability tends to zero as $T \to \infty$. However, for $E_b = E_a$ we use the result,

$$\frac{1}{2\pi T}\,\frac{\sin^2(\omega_{ba}T/2)}{(\omega_{ba}/2)^2} \xrightarrow{T\to\infty} \delta(\omega_{ba}). \tag{3.6}$$

For long times the transition rate becomes,

$$R_{ba} = 2\pi\left|\langle b|V|a\rangle\right|^2 \delta(E_b - E_a). \tag{3.7}$$

We need V small for the first order result to be useful and T large so that the delta-function approximation is good. However, T cannot be too large since the transition probability grows with time and we don't want probabilities larger than one.

If we allow for a number of final states $|b\rangle$, with density $\rho(E_b)$ around energy E_b, the transition rate becomes,

$$\int 2\pi\left|\langle b|V|a\rangle\right|^2 \delta(E_b - E_a)\rho(E_b)dE_b = 2\pi\rho(E_a)\left|\langle b|V|a\rangle\right|^2. \tag{3.8}$$

This is *Fermi's golden rule.*

▷Exercise 3.2

Justify the result of equation (3.6) and hence verify Fermi's golden rule in equation (3.8).

I'll stop at first order in V. The answer you get from the formal solution in equation (3.5) depends on the form of V and the initial conditions. Your field theory course gives you a systematic way to perform perturbative calculations of transition amplitudes in field theories by the use of Feynman diagrams. In particular, you've seen the operator method of generating these diagrams, which I've mirrored in deriving the Golden Rule. Let's now move on to see how to get from these amplitudes to cross-sections and decay rates.

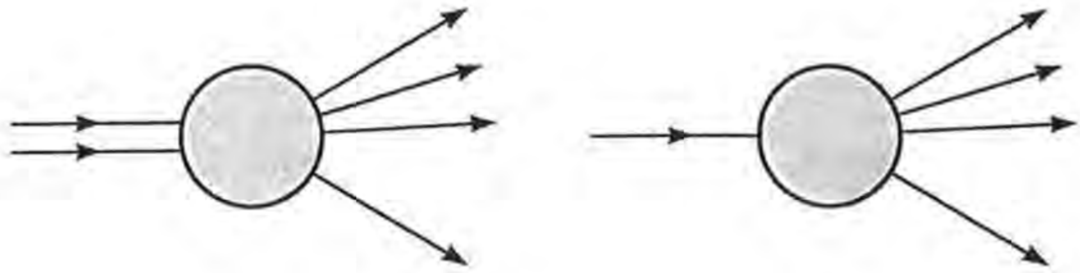


Figure 3.1 Scattering (a) and decay (b) processes.

3.2 Phase Space

We saw in the previous section that $\langle b|\psi_I(\infty)\rangle$ gives the probability amplitude to go from state $|a\rangle$ in the far past to state $|b\rangle$ in the far future. In quantum field theory you calculate the amplitude to go from state $|i\rangle$ to state $|f\rangle$ to be,

$$i\mathcal{M}_{fi}(2\pi)^4\delta^4(P_f - P_i), \tag{3.9}$$

where $i\mathcal{M}_{fi}$ is the result obtained from a Feynman diagram calculation, and the overall energy-momentum delta function has been factored out (so when you draw your Feynman diagrams you conserve energy-momentum at every vertex). We have in mind processes where two particles scatter, or one particle decays, as shown in Figure 3.1.

Attempting to take the squared modulus of this amplitude produces a meaningless square of a delta function. This is a technical problem because our amplitude is expressed between non-normalisable plane wave states. These states extend throughout space-time so the scattering process occurs everywhere all the time. To deal with this properly you can construct normalised wavepacket states which do become well separated in the far past and the far future. We will be low-budget and put our system in a box of volume V. We also imagine that the interaction is restricted to act only over a time of order T. The final answers come out independent of V and T, reproducing the luxury wavepacket ones.

Relativistically normalised one particle states satisfy,

$$\langle k|k'\rangle = (2\pi)^3 2\omega_k \delta^3(\mathbf{k} - \mathbf{k}'), \tag{3.10}$$

but the discrete nonrelativistically normalised box states satisfy,

$$\langle \mathbf{k}|\mathbf{k}'\rangle = \delta_{\mathbf{k}\mathbf{k}'}. \tag{3.11}$$

We want to know the transition probability from an initial state of one or two particles to a set of final states occupying some region of $\mathbf{k}$-space, where the density of states in the box normalisation is,

$$\text{box state density} = \frac{d^3\mathbf{k}}{(2\pi)^3} V, \tag{3.12}$$

recalling that the spacing of allowed momenta is $2\pi/L$. A particular final state is labelled, $|f\rangle = |\mathbf{k}_1, \ldots, \mathbf{k}_n\rangle$, and the initial state is,

$$|i\rangle = \begin{cases} |\mathbf{k}\rangle & \text{one particle} \\ |\mathbf{k}_1, \mathbf{k}_2\rangle \sqrt{V} & \text{two particles} \end{cases} \tag{3.13}$$

Note the factor of $\sqrt{V}$ in the two particle case. Without this, as V becomes large the probability that the two particles are anywhere near each other goes to zero. From the viewpoint of one particle hitting another, the one particle state is normalised to one (probability 1 of being somewhere in the box), and the two particle state is normalised as a density (think of one particle having probability 1 of being in any unit volume and the second having probability 1 of being somewhere in the box).

The transition probability from i to f is given by (3.9). We want to convert this to the box normalisation. One ingredient of the conversion is the delta function of momentum conservation, arising from,

$$\int_{VT} d^4x \, e^{i(P_f - P_i)\cdot x} = (2\pi)^4 \delta^4_{VT}(P_f - P_i),$$

using the box normalisation. Now,

$$\int \frac{d^4p}{(2\pi)^4} \left|(2\pi)^4 \delta^4_{VT}(p)\right|^2 = \int_{VT} d^4x = VT,$$

so we will say,

$$\left|(2\pi)^4 \delta^4_{VT}(p)\right|^2 \simeq VT \, (2\pi)^4 \delta^4(p).$$

The second ingredient is a factor of $1/(2E_iV)^{1/2}$ for every particle in the initial or final state (here I am using E_i synonymously with ω_{k_i}). This comes from converting between relativistic and box normalisations for the states.

To see where this arises from we write here the expression for a free field expanded in terms of annihilation and creation operators using three different normalisations: nonrelativistic, $\langle \mathbf{k}|\mathbf{k}'\rangle = \delta^3(\mathbf{k}-\mathbf{k}')$; relativistic, $\langle k|k'\rangle = (2\pi)^3 2\omega_k \delta^3(\mathbf{k}-\mathbf{k}')$; box, $\langle \mathbf{k}|\mathbf{k}'\rangle = \delta_{\mathbf{k}\mathbf{k}'}$.

$$\begin{aligned}
\phi(x) &= \int \frac{d^3k}{(2\pi)^{3/2}\sqrt{2\omega_k}} \left[a_{\mathbf{k}} e^{-ik\cdot x} + a^\dagger_{\mathbf{k}} e^{ik\cdot x}\right] && \text{nonrelativistic} \\
&= \int \frac{d^3k}{(2\pi)^3 \, 2\omega_k} \left[a(k) e^{-ik\cdot x} + a^\dagger(k) e^{ik\cdot x}\right] && \text{relativistic} \\
&= \sum_{\mathbf{k}} \frac{1}{\sqrt{2\omega_k}\sqrt{V}} \left[a_{\mathbf{k}} e^{-ik\cdot x} + a^\dagger_{\mathbf{k}} e^{ik\cdot x}\right] && \text{box}
\end{aligned}$$

Since the discrete sum on $\mathbf{k}$ in the box case corresponds to $\int d^3\mathbf{k}\, V/(2\pi)^3$, we see that,

$$|k\rangle_{\text{rel}} \longleftrightarrow \sqrt{2\omega_k}\sqrt{V}|\mathbf{k}\rangle_{\text{box}}.$$

The box states are normalised to one particle in volume V and the relativistic states have $2\omega_k$ particles per unit volume.

So in the box normalisation, with one or two particles in the initial state and any number in the final state,

$$\text{box amp} = i\mathcal{M}_{fi}(2\pi)^4\delta^4(P_f - P_i) \prod_{\text{out}} \left[\frac{1}{\sqrt{2E_fV}}\right] \prod_{\text{in}} \left[\frac{1}{\sqrt{2E_i}}\right] \frac{1}{\sqrt{V}},$$

where the initial state energy product depends on the choice of normalisation in equation 3.13 above. The squared matrix element is thus:

$$|\text{box amp}|^2 = |\mathcal{M}_{fi}|^2 T (2\pi)^4 \delta^4(P_f - P_i) \prod_{\text{out}} \left[\frac{1}{2E_fV}\right] \prod_{\text{in}} \left[\frac{1}{2E_i}\right],$$

and the differential transition probability into a region of phase space becomes,

$$\frac{\text{differential prob}}{\text{unit time}} = S\,|\mathcal{M}_{fi}|^2 \prod_{\text{in}} \left[\frac{1}{2E_i}\right] \times \begin{pmatrix}\text{relativistic density}\\ \text{of final states}\end{pmatrix}, \tag{3.14}$$

where the *relativistic density of final states*, or rdfs, is,

$$\text{rdfs} = D \equiv (2\pi)^4 \delta^4(P_f - P_i) \prod_{\text{out}} \frac{d^3\mathbf{k}_f}{(2\pi)^3\, 2E_f}. \tag{3.15}$$

You also sometimes hear the name LIPS, standing for Lorentz invariant phase space. Observe that everything in the transition probability is Lorentz invariant save for the initial energy factor (using $d^3k\,/2E = d^4k\,\delta^4(k^2 - m^2)\theta(k^0)$, which is manifestly Lorentz invariant, where $E = (\mathbf{k}^2 + m^2)^{1/2}$). I have smuggled in one extra factor, S, in equation (3.14) for the transition probability. If there are some identical particles in the final state, we will overcount them when integrating over all momentum configurations. The symmetry factor S takes care of this. If there n_i identical particles of type i in the final state, then

$$S = \prod_i \frac{1}{n_i!}. \tag{3.16}$$

▷**Exercise 3.3**
Show that the expression for two-body phase space in the centre of mass frame is given by

$$\frac{d^3k_1}{(2\pi)^3\, 2\omega_{k_1}} \frac{d^3k_2}{(2\pi)^3\, 2\omega_{k_2}} (2\pi)^4 \delta^4(P - k_1 - k_2) = \frac{1}{32\pi^2 s} \lambda^{1/2}(s, m_1^2, m_2^2) d\Omega^*, \tag{3.17}$$

where $s = P^2$ is the centre of mass energy squared, $d\Omega^*$ is the solid angle element for the angle of one of the outgoing particles with respect to some fixed direction, and

$$\lambda(a, b, c) = a^2 + b^2 + c^2 - 2ab - 2bc - 2ca. \tag{3.18}$$

3.3 Cross Sections

The cross section for two particles to scatter is a sum of the differential cross sections for scattering into distinct final states:

$$d\sigma = \frac{\text{transition prob}}{\text{unit time} \times \text{unit flux}} = \frac{1}{|\vec{v}_1 - \vec{v}_2|} \frac{1}{4E_1E_2} S|\mathcal{M}_{fi}|^2 D, \tag{3.19}$$

where the velocities in the flux factor, $1/|\vec{v}_1 - \vec{v}_2|$, are subtracted *nonrelativistically*. I denote them with arrows to remind you that they are ordinary velocities, not the spatial parts of 4-velocities. The amplitude-squared and phase space factors are manifestly Lorentz invariant. What about the initial velocity and energy factors? Observe that

$$E_1E_2(\vec{v}_1 - \vec{v}_2) = E_2\mathbf{p}_1 - E_1\mathbf{p}_2.$$

In a frame where $\mathbf{p}_1$ and $\mathbf{p}_2$ are collinear,

$$|E_2\mathbf{p}_1 - E_1\mathbf{p}_2|^2 = (p_1 \cdot p_2)^2 - m_1^2 m_2^2,$$

and the last expression is manifestly Lorentz invariant. Hence the differential cross section is Lorentz invariant, as is the total cross section,

$$\sigma = \frac{1}{|\vec{v}_1 - \vec{v}_2|} \frac{1}{4E_1E_2} S \sum \int_{\text{final states}} |\mathcal{M}_{fi}|^2 D. \tag{3.20}$$

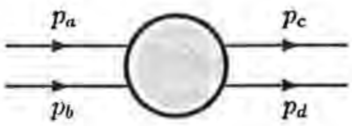


Figure 3.2 $2 \to 2$ scattering.

3.3.1 Two-body Scattering

An important special case is $2 \to 2$ scattering (see Figure 3.2),

$$a(p_a) + b(p_b) \to c(p_c) + d(p_d).$$

▷ **Exercise 3.4**
Show that in the centre of mass frame the differential cross section is,

$$\frac{d\sigma}{d\Omega^*} = \frac{S\,\lambda^{1/2}(s, m_c^2, m_d^2)}{64\pi^2 s\,\lambda^{1/2}(s, m_a^2, m_b^2)}|\mathcal{M}_{fi}|^2. \tag{3.21}$$

The result of equation (3.21) is valid for any $|\mathcal{M}_{fi}|^2$, but if $|\mathcal{M}_{fi}|^2$ is a constant you can trivially get the total cross section.

Invariant $2 \to 2$ scattering amplitudes are frequently expressed in terms of the *Mandelstam variables*, defined by,

$$\begin{aligned} s &\equiv (p_a + p_b)^2 &= (p_c + p_d)^2 \\ t &\equiv (p_a - p_c)^2 &= (p_b - p_d)^2 \\ u &\equiv (p_a - p_d)^2 &= (p_b - p_c)^2 \end{aligned} \tag{3.22}$$

In fact there are only two independent Lorentz invariant combinations of the available momenta in this case, so there must be some relation between s, t and u.

▷ **Exercise 3.5**
Show that

$$s + t + u = m_a^2 + m_b^2 + m_c^2 + m_d^2.$$

▷ **Exercise 3.6**
Show that for two body scattering of particles of equal mass m,

$$s \geq 4m^2, \qquad t \leq 0, \qquad u \leq 0.$$

3.4 Decay Rates

With one particle in the initial state,

$$\frac{\text{total decay prob}}{\text{unit time}} = \frac{1}{2E} S \sum \int_{\text{final states}} |\mathcal{M}_{fi}|^2 D.$$

Only the factor $1/2E$ is not manifestly Lorentz invariant. In the rest frame, for a particle of mass m,

$$\left.\frac{\text{tdp}}{\text{ut}}\right|_{\text{rest frame}} \equiv \Gamma \equiv \frac{1}{2m} \sum \int_{\text{final states}} |\mathcal{M}_{fi}|^2 D. \tag{3.23}$$

This is the "decay rate." In an arbitrary frame we find, $(\text{tdp/ut}) = (m/E)\Gamma$, which has the expected Lorentz dilatation factor. In the master formula (equation 3.14) this is what the product of $1/2E_i$ factors for the initial particles does.

3.5 Optical Theorem

When discussing the Golden Rule, we encountered the evolution operator $U(t',t)$, which you also met in the field theory course. This takes a state at time t and evolves it to time t'. The scattering amplitudes we calculate in field theory are between states in the far past and the far future: hence they are matrix elements of $U(\infty,-\infty)$, which is known as the S-matrix,

$$S \equiv U(\infty,-\infty) = \mathrm{T}\,\exp -i\int_{-\infty}^{\infty} dt\, H_I(t).$$

Since the S-matrix is unitary, we can write,

$$(S-I)(S^\dagger - I) = -\Big((S-I)+(S-I)^\dagger\Big). \tag{3.24}$$

Note that $S-I$ is the quantity of interest, since we generally ignore cases where there is no interaction (the "I" piece of S). In terms of the invariant amplitude,

$$\begin{aligned}\langle f|\,S-I\,|i\rangle &= i\mathcal{M}_{fi}(2\pi)^4\delta^4(P_f-P_i)\\ \langle f|\,(S-I)^\dagger\,|i\rangle &= -i\mathcal{M}^*_{if}(2\pi)^4\delta^4(P_f-P_i)\end{aligned}$$

Sandwiching the above unitarity relation (equation 3.24) between states $|i\rangle$ and $|f\rangle$, and inserting a complete set of states between the factors on the left hand side,

$$\begin{aligned}&\sum_m \langle f|\,S-I\,|m\rangle\,\langle m|\,S^\dagger - I\,|i\rangle\\ &= \sum_m \mathcal{M}_{fm}\mathcal{M}^*_{im}(2\pi)^8\delta^4(P_f-P_m)\delta^4(P_i-P_m)\prod_{j=1}^{r_m}\frac{d^3\mathbf{k}_j}{(2\pi)^3\,2E_j}\\ &= \sum_m \int \mathcal{M}_{fm}\mathcal{M}^*_{im}(2\pi)^4\delta^4(P_f-P_i)D_m\end{aligned}$$

where D_m is the phase space factor for the state labelled by m, containing r_m particles, $D_m \equiv D_{r_m}(P_i;k_1,\ldots,k_{r_m})$. Hence,

$$\sum_m \int \mathcal{M}_{fm}\mathcal{M}^*_{im}D_m = i(\mathcal{M}^*_{if}-\mathcal{M}_{fi}).$$

If the intermediate state m contains n_i identical particles of type i, there is an extra symmetry factor S, with,

$$S = \prod_i \frac{1}{n_i!}$$

on the left hand side of the above equation to avoid overcounting. The same factor (see equation 3.16) appears in the cross section formula (equation 3.19) when some of the final state particles are identical.

If $|i\rangle$ and $|f\rangle$ are the same two particle state,

$$4E_T p_i\,\sigma = 2\,\mathrm{Im}\,\mathcal{M}_{ii}. \tag{3.25}$$

this is the *optical theorem*, relating the forward part of the scattering amplitude to the total cross section. If particles of masses m_1 and m_2 scatter, then $E_T = s^{1/2}$ and $4sp_i^2 = \lambda(s,m_1^2,m_2^2)$, where λ is the function defined in eqaution (3.18). Then the optical theorem reads, $\mathrm{Im}\,\mathcal{M}_{ii} = \lambda^{\frac{1}{2}}(s,m_1^2,m_2^2)\,\sigma$.

4 Quantum Electrodynamics

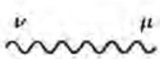
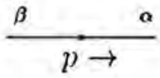
In this section we are going to get some practice calculating cross sections and decay rates in QED. The starting point is the set of Feynman rules derived from the QED Lagrangian,

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} - \frac{1}{2}(\partial_\mu A^\mu)^2 + \overline{\psi}(i\not{D} - m)\psi. \tag{4.1}$$

Here, $D_\mu = \partial_\mu + ieA_\mu$ is the electromagnetic covariant derivative, $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ and $(\partial{\cdot}A)^2/2$ is the gauge fixing term for Feynman gauge. This gives the rules in Table 4.1.

The fermion propagator is (up to factors of i) the inverse of the operator, $\not{p} - m$, which appears in the quadratic term in the fermion fields, as you expect from your field theory course. The derivation of the gluon propagator, along with the need for gauge fixing, is also discussed in the field theory course. The external line factors are easily derived by considering simple matrix elements in the operator formalism, where they are left behind from the expansions of fields in terms of annihilation and creation operators, after the operators have all been (anti-)commuted until they annihilate the vacuum. In path integral language the natural objects to compute are Green functions, vacuum expectation values of time ordered products of fields: it takes a little more work to convert them to transition amplitudes and see the external line factors appear.

The spinor indices in the Feynman rules are such that matrix multiplication is performed in the opposite order to that defining the flow of fermion number. The arrow on the fermion line itself denotes the fermion number flow, *not* the direction of the momentum associated with the line: I will try always to indicate the momentum flow separately

For every ...	draw ...	write ...
Internal photon line		$\dfrac{-ig^{\mu\nu}}{q^2 + i\epsilon}$
Internal fermion line		$\dfrac{i(\not{p} + m)_{\alpha\beta}}{p^2 - m^2 + i\epsilon}$
Vertex		$-ie\gamma^\mu_{\alpha\beta}$
Outgoing electron		$\overline{u}^s_p$
Incoming electron		u^s_p
Outgoing positron		v^s_p
Incoming positron		$\overline{v}^s_p$
Outgoing photon		$\epsilon^{*\mu}$
Incoming photon		ϵ^μ

- Attach a directed momentum to every internal line
- Conserve momentum at every vertex

Table 4.1 Feynman rules for QED. μ, ν are Lorentz indices and α, β are spinor indices.

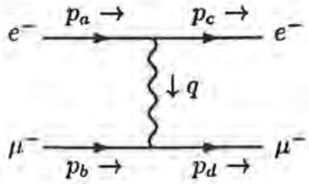


Figure 4.1 Lowest order Feynman diagram for electron–muon scattering.

as in Table 4.1. This will become clear in the examples which follow. We have already met the Dirac spinors u and v. I will say more about the photon polarisation vector ϵ when we need to use it.

4.1 Electron–Muon Scattering

To lowest order in the electromagnetic coupling, just one diagram contributes to this process. It is shown in Figure 4.1. The amplitude obtained from this diagram is

$$i\mathcal{M}_{fi} = (-ie)\,\overline{u}(p_c)\gamma^\mu u(p_a)\left(\frac{-ig_{\mu\nu}}{q^2}\right)(-ie)\,\overline{u}(p_d)\gamma_\nu u(p_b). \tag{4.2}$$

Note that I have changed my notation for the spinors: now I label their momentum as an argument instead of as a subscript, and I drop the spin label unless I need to use it. In constructing this amplitude we have followed the fermion lines backwards with respect to fermion flow when working out the order of matrix multiplication.

The cross-section involves the squared modulus of the amplitude, which is

$$|\mathcal{M}_{fi}|^2 = \frac{e^4}{q^4}\,L^{\mu\nu}_{(e)}L_{(\mu)\,\mu\nu},$$

where the subscripts e and μ refer to the electron and muon respectively and,

$$L^{\mu\nu}_{(e)} = \overline{u}(p_c)\gamma^\mu u(p_a)\overline{u}(p_a)\gamma^\nu u(p_c),$$

with a similar expression for $L^{\mu\nu}_{(\mu)}$.

▷**Exercise 4.1**
Verify the expression for $|\mathcal{M}_{fi}|^2$.

Usually we have an unpolarised beam and target and do not measure the polarisation of the outgoing particles. Thus we calculate the squared amplitudes for each possible spin combination, then average over initial spin states and sum over final spin states. Note that we square and then sum since the different possibilities are in principle distinguishable. In contrast, if several Feynman diagrams contribute to the same process, you have to sum the amplitudes first. We will see examples of this below.

The spin sums are made easy by the following results (I temporarily restore spin labels on spinors):

$$\begin{aligned} \sum_r u^r(p)\,\overline{u}^r(p) &= \not{p} + m \\ \sum_r v^r(p)\,\overline{v}^r(p) &= \not{p} - m \end{aligned} \tag{4.3}$$

▷**Exercise 4.2**
Derive the spin sum relations in equation (4.3).

Using the spin sums we find,

$$\frac{1}{4}\sum_{\text{spins}}|\mathcal{M}_{fi}|^2 = \frac{e^4}{4q^4}\,\mathrm{tr}\left(\gamma^\mu(\not{p}_a+m_e)\gamma^\nu(\not{p}_c+m_e)\right)\mathrm{tr}\left(\gamma_\mu(\not{p}_b+m_\mu)\gamma_\nu(\not{p}_d+m_\mu)\right).$$

Since all calculations of cross sections or decay rates in QED require the evaluation of traces of products of gamma matrices, you will generally find a table of "trace theorems" in any quantum field theory textbook [1]. All these theorems can be derived from the fundamental anticommutation relations of the gamma matrices in equation (2.23) together with the invariance of the trace under a cyclic change of its arguments. For now it suffices to use,

$$\begin{aligned}\mathrm{tr}(\not{a}\not{b}) &= 4\,a{\cdot}b\\ \mathrm{tr}(\not{a}\not{b}\not{c}\not{d}) &= 4(a{\cdot}b\,c{\cdot}d - a{\cdot}c\,b{\cdot}d + a{\cdot}d\,b{\cdot}c)\\ \mathrm{tr}(\gamma^{\mu_1}\cdots\gamma^{\mu_n}) &= 0 \quad \text{for } n \text{ odd}\end{aligned} \tag{4.4}$$

▷ **Exercise 4.3**
Derive the trace results in equation (4.4)

Using these results, and expressing the answer in terms of the Mandelstam variables of equation (3.22), we find,

$$\frac{1}{4}\sum_{\text{spins}}|\mathcal{M}_{fi}|^2 = \frac{2e^4}{t^2}\left(s^2+u^2-4(m_e^2+m_\mu^2)(s+u)+6(m_e^2+m_\mu^2)^2\right).$$

This can now be used in the $2 \to 2$ cross section formula (3.21) to give, in the high energy limit, $s, u \gg m_e^2, m_\mu^2$,

$$\frac{d\sigma}{d\Omega^*} = \frac{e^4}{32\pi^2 s}\,\frac{s^2+u^2}{t^2}. \tag{4.5}$$

for the differential cross section in the centre of mass frame.

▷ **Exercise 4.4**
Derive the result for the electron–muon scattering cross section in equation (4.5).

Other calculations of cross sections or decay rates will follow the same steps we have used above. You draw the diagrams, write down the amplitude, square it and evaluate the traces (if you are using spin sum/averages). There are one or two more wrinkles to be aware of, which we will meet below.

4.2 Electron–Electron Scattering

Since the two scattered particles are now identical, you can't just replace m_μ by m_e in the calculation we did above. If you look at the diagram of Figure 4.1 (with the muons replaced by electrons) you will see that the outgoing legs can be labelled in two ways. Hence we get the two diagrams of Figure 4.2.

The two diagrams give the amplitudes,

$$\begin{aligned} i\mathcal{M}_1 &= \frac{ie^2}{t}\overline{u}(p_c)\gamma^\mu u(p_a)\overline{u}(p_d)\gamma_\mu u(p_b),\\ i\mathcal{M}_2 &= -\frac{ie^2}{u}\overline{u}(p_d)\gamma^\mu u(p_a)\overline{u}(p_c)\gamma_\mu u(p_b).\end{aligned}$$

Figure 4.2 Lowest order Feynman diagrams for electron–electron scattering.

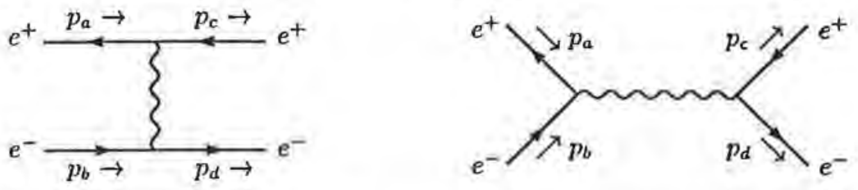


Figure 4.3 Lowest order Feynman diagrams for electron–positron scattering in QED.

Notice the additional minus sign in the second amplitude, which comes from the anti-commuting nature of fermion fields. You should accept as part of the Feynman rules for QED that when diagrams differ by an interchange of two fermion lines, a relative minus sign must be included. This is important because

$$|\mathcal{M}_{fi}|^2 = |\mathcal{M}_1 + \mathcal{M}_2|^2,$$

so the interference term will have the wrong sign if you don't include the extra sign difference between the two diagrams.

4.3 Electron–Positron Annihilation

4.3.1 $e^+e^- \rightarrow e^+e^-$

For this process the two diagrams are shown in Figure 4.3, with the one on the right known as the annihilation diagram. They are just what you get from the diagrams for electron-electron scattering in Figure 4.2 if you twist round the fermion lines. The fact that the diagrams are related this way implies a relation between the amplitudes. The interchange of incoming particles/antiparticles with outgoing antiparticles/particles is called *crossing*. This is a case where the general results of crossing symmetry can be applied, and our diagrammatic calculations give an explicit realisation. Theorists spent a great deal of time studying such general properties of amplitudes in the 1960's when quantum field theory was unfashionable.

4.3.2 $e^+e^- \rightarrow \mu^+\mu^-$ and $e^+e^- \rightarrow$ hadrons

If electrons and positrons collide and produce muon–antimuon or quark-antiquark pairs, then the annihilation diagram is the only one which contributes. At sufficiently high energies that the quark masses can be neglected, this immediately gives the lowest order QED prediction for the ratio of the annihilation cross section into hadrons to that into $\mu^+\mu^-$,

$$R \equiv \frac{\sigma(e^+e^- \rightarrow \text{hadrons})}{\sigma(e^+e^- \rightarrow \mu^+\mu^-)} = 3\sum_f Q_f^2, \tag{4.6}$$

where the sum is over quark flavours f and Q_f is the quark's charge in units of e. The 3 comes from the existence of three colours for each flavour of quark. Historically this

Figure 4.4 Feynman diagrams for Compton scattering.

was important: you could look for a step in the value of R as your e^+e^- collider's CM energy rose through a threshold for producing a new quark flavour. If you didn't know about colour, the height of the step would seem too large. Incidentally, another place the number of colours enters is in the decay of a π^0 to two photons. There is a factor of **3** in the amplitude from summing over colours, without which the predicted decay rate would be one ninth of its real size.

At the energies used today at LEP, of course, you have to remember the diagram with a Z replacing the photon. We will say some more about this later.

▷**Exercise 4.5**
Show that the cross-section for $e^+e^- \to \mu^+\mu^-$ is equal to $4\pi\alpha^2/(3s)$, neglecting the lepton masses.

4.4 Compton Scattering

The diagrams which need to be evaluated to compute the Compton cross section for $\gamma e \to \gamma e$ are shown in Figure 4.4. For unpolarised initial and/or final states, the cross section calculation involves terms of the form

$$\sum_\lambda \epsilon_\lambda^{*\mu}(p)\, \epsilon_\lambda^\nu(p), \tag{4.7}$$

where λ represents the polarisation of the photon of momentum p. Since the photon is massless, the sum is over the two transverse polarisation states, and must vanish when contracted with p_μ or p_ν. In addition, however, since the photon is coupled to the electromagnetic current $J^\mu = \overline{\psi}\gamma^\mu\psi$ of equation (2.5), any term in the polarisation sum (4.7) proportional to p^μ or p^ν does not contribute to the cross section. This is because the current is conserved, $\partial_\mu J^\mu = 0$, so in momentum space $p_\mu J^\mu = 0$. The upshot is that in calculations you can use,

$$\sum_\lambda \epsilon_\lambda^{*\mu}(p)\, \epsilon_\lambda^\nu(p) = -g^{\mu\nu}, \tag{4.8}$$

since the remaining terms on the right hand side do not contribute.

4.5 Form Factors

So far we have considered processes where the strong interactions were absent, or ignored. There are many electroweak processes where a complete computation would require a better understanding of QCD, especially its non-perturbative aspects, than we currently have. However, by using Lorentz and gauge invariance, and any other known symmetries of a process, we can parcel up the strong interaction effects in a small number of invariant functions. Let's see how this goes in an example, the electromagnetic form factors of pions and nucleons.

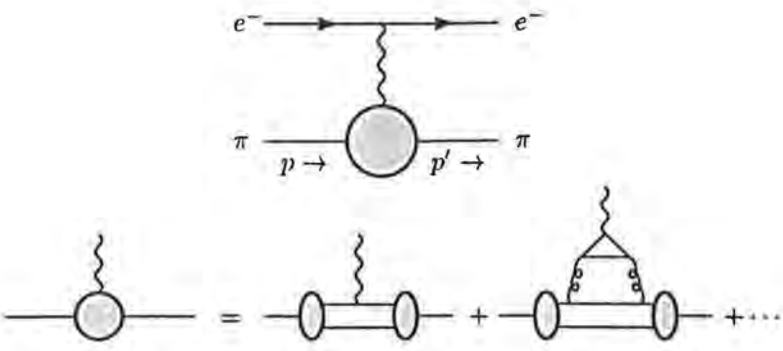


Figure 4.5 Electron–pion scattering (top diagram) and some contributions to the pion electromagnetic form factor (lower diagrams). Wavy lines denote photons and curly lines are gluons. Ordinary lines between the shaded ellipses denote quarks.

4.5.1 Pion Form Factor

Consider electron–pion scattering as depicted in the top diagram in Figure 4.5. The shaded blob represents all the strong interaction effects in the pion electromagnetic form factor. In the lower part of the figure are represented some contributions to the shaded blob. Note that the blob itself contains more blobs (the shaded ellipses) indicating the unknown wavefunction of the pion in terms of quarks. The electron's coupling to the photon is understood in QED and has been discussed above. Let's see how much we can say about the pion's coupling to photons. This coupling is given by the matrix element $\langle\pi(p')|\, J^\mu(0)\,|\pi(p)\rangle$, where $J^\mu(0)$ is the electromagnetic current at the origin (we put in *all* strong interaction corrections, but work to *first* order in the electromagnetic interaction). Using Lorentz covariance we can write,

$$\langle\pi(p')|\, J^\mu(0)\,|\pi(p)\rangle = e\left[F(q^2)(p+p')^\mu + G(q^2)q^\mu\right],$$

where $q = p - p'$. Electromagnetic gauge invariance implies that $q_\mu J^\mu = 0$ so that $G(q^2) = 0$. Hence all the strong interaction effects are contained in $F(q^2)$ and

$$\langle\pi(p')|\, J^\mu(0)\,|\pi(p)\rangle = e\, F(q^2)(p+p')^\mu. \tag{4.9}$$

▷**Exercise 4.6**
Starting from the kinetic term in the Lagrangian for a free charged scalar field, $\partial_\mu\phi^*\partial^\mu\phi$, and introducing the electromagnetic field by minimal substitution, $\partial_\mu \to \partial_\mu - ieA_\mu$, show that, to lowest order in perturbation theory $F(q^2) = 1$ for all q^2. Note that the change of sign in the coupling compared to QED is because QED involves the negatively charged electron, whilst here ϕ is taken as the field which destroys positively charged objects and creates negatively charged ones. You may need to normal order the current.

An additional general piece of information is that $F(0) = 1$ since at $q^2 = 0$ the photon cannot resolve the structure of the pion. This result is a consequence of the conservation of the electromagnetic current, since the space integral of J^0 gives the charge operator. For $q^2 \neq 0$ we expect $F(q^2)$ to fall with q^2 owing to the pion's composite nature.

▷**Exercise 4.7**
Given that the electric charge operator is defined by

$$eQ = \int d^3x\, J^0(x),$$

show that current conservation implies Q is time independent, and that $F(0) = 1$ for a positively charged pion.

4.5.2 Nucleon Form Factor

For nucleons there are two form factors consistent with Lorentz covariance, current conservation and parity conservation (which holds for electromagnetic and strong interactions). They are defined as follows (again we are working to first order in electromagnetism):

$$\langle N(p', s')|\, J^\mu \,|N(p, s)\rangle = e\,\overline{u}^{s'}(p') \left[\gamma^\mu F_1(q^2) + \frac{i\kappa}{2M} F_2(q^2)\sigma^{\mu\nu}q_\nu\right] u^s(p), \tag{4.10}$$

where u and $\overline{u}$ are the nucleon spinors, and M the nucleon mass. At zero momentum transfer only the first term contributes and $F_1(0) = 1[0]$ for the proton[neutron]. The factor κ is chosen so that $F_2(0) = 1$: κ is 1.79 for the proton and -1.91 for the neutron. In writing the expression (4.10), use is made of the *Gordon identity*,

$$\overline{u}(p')\gamma^\mu u(p) = \frac{1}{2m}\overline{u}(p')\left[(p + p')^\mu + i\sigma^{\mu\nu}(p' - p)_\nu\right]u(p),$$

to replace a term in $(p + p')^\mu$ with terms of the form given. Given the form factor expression you can compute the angular distribution of electrons in electron–nucleon scattering in terms of F_1 and F_2.

▷ **Exercise 4.8**
Use Lorentz covariance, current conservation and parity invariance to show that there are two electromagnetic form factors for the nucleon in (4.10).

5 Quantum Chromodynamics

In the 1960's most theorists lost interest in quantum field theory. They were discouraged by the apparent non renormalisability of massive vector boson theories which precluded a field theory description of weak interactions. For the strong interactions, their strength and the menagerie of hadrons seemed also to preclude a field theory description. The renaissance of field theory came with the realisation that spontaneous symmetry breaking, the Higgs mechanism and the property of asymptotic freedom made renormalisable gauge theories viable candidates to describe the electroweak and strong interactions.

Our discussion in this section will lead to the property of *asymptotic freedom* which enables us to make phenomenological predictions using perturbation theory for QCD. Since perturbative calculations beyond tree level are not in the scope of this course, the discussion will necessarily be somewhat qualitative. We'll proceed by going back to QED to introduce the idea of renormalisation then work up to the running coupling in QCD and thence to asymptotic freedom.

QCD is a theory of interactions between spin-1/2 quarks and spin-1 gluons. It is a nonabelian gauge theory based on the group $SU(3)$, with Lagrangian,

$$\mathcal{L} = -\frac{1}{4} G^a_{\mu\nu} G^{a\,\mu\nu} + \sum_f \overline{\psi}_f (i\not{D} - m_f)\psi_f + \begin{matrix}\text{gauge fixing and}\\ \text{ghost terms}\end{matrix} \tag{5.1}$$

Here, a is a colour label, taking values from 1 to 8 for $SU(3)$, and f runs over the quark flavours. The covariant derivative and field strength tensor are given by,

$$\begin{aligned} D_\mu &= \partial_\mu - igA^a_\mu T^a, \\ G^a_{\mu\nu} &= \partial_\mu A^a_\nu - \partial_\nu A^a_\mu + gf^{abc} A^b_\mu A^c_\nu, \end{aligned} \tag{5.2}$$

where the f^{abc} are the structure constants of $SU(3)$ and the T^a are a set of eight independent Hermitian traceless 3×3 matrix generators in the fundamental or defining representation (see the pre school problems and the quantum field theory course).

As in QED gauge fixing terms are needed to define the propagator and ensure that only physical degrees of freedom propagate. The gauge fixing procedure is more complicated in the nonabelian case and necessitates, for certain gauge choices, the appearance of Faddeev–Popov ghosts to cancel the contributions from unphysical polarisation states in gluon propagators. However, the ghosts first appear in loop diagrams, which we will not compute in this course.

There are no Higgs bosons in pure QCD. The only relic of them is in the masses for the fermions which are generated via the Higgs mechanism, but in the electroweak sector of the standard model.

A fundamental difference between QCD and QED is the appearance in the nonabelian case of interaction terms (vertices) containing gluons alone. These arise from the nonvanishing commutator term in the field strength of the nonabelian theory in equation (5.2). The photon is electrically neutral, but the gluons carry the colour charge of QCD (specifically, they transform in the adjoint representation). Since the force carriers couple to the corresponding charge, there are no multi photon vertices in QED but there are multi gluon couplings in QCD. This difference is crucial: it is what underlies the decreasing strength of the strong coupling with increasing energy scale.

In QCD, hadrons are made from quarks. Colour interactions bind the quarks, producing states with no net colour: three quarks combine to make baryons and quark–antiquark

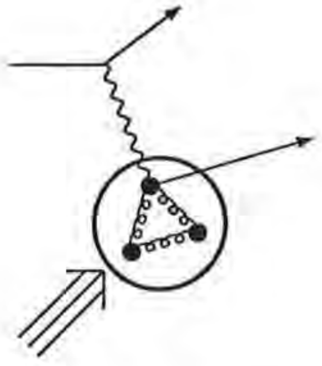


Figure 5.1 Schematic depiction of deep inelastic scattering. An incident lepton radiates a photon which knocks a quark out of a proton. The struck quark is detected indirectly only after hadronisation into observable particles.

pairs give mesons. It is generally believed that the binding energy of a quark in a hadron is infinite. This property, called *confinement*, means that there is no such thing as a free quark. Because of asymptotic freedom, however, if you hit a quark with a high energy projectile it will behave in many ways as a free (almost) particle. For example, in deep inelastic scattering, or DIS, a photon strikes a quark in a proton, say, imparting a large momentum to it. Some strong interaction corrections to this part of the process can be calculated perturbatively. As the quark heads off out of the proton, however, the brown muck of myriad low energy strong interactions cuts in again and "hadronises" the quark into the particles you actually detect. This is illustrated schematically in Figure 5.1.

5.1 Renormalisation: An Introduction

5.1.1 Renormalisation in Quantum Electrodynamics

Let's start by going back to QED and considering how the electric charge is defined and measured. This will bring up the question of what happens when you try to compute higher loop corrections. In fact, the expansion in the number of loops is an expansion in Planck's constant $\hbar$, as you can show if you put back the factors of $\hbar$ for once.

The electric charge $\hat{e}$ is usually defined as the coupling between an on-shell electron and an on-shell photon: that is, as the vertex on the left hand side of Figure 5.2 with $p_1^2 = p_2^2 = m^2$, where m is the electron mass, and $q^2 = 0$. It is $\hat{e}$ and not the Lagrangian parameter e which we measure. That is,

$$\frac{\hat{e}^2}{4\pi} = \frac{1}{137}.$$

We call $\hat{e}$ the renormalised coupling constant of QED. We can calculate $\hat{e}$ in terms of e in perturbation theory. To one loop, the relevant diagrams are shown on the right hand side of Figure 5.2, and the result takes the form,

$$\hat{e} = e + e^3\Big[a_1 \ln\frac{M^2}{m^2} + b_1\Big] + \cdots \tag{5.3}$$

where a_1 and b_1 are constants obtained from the calculation. The e^3 term is divergent, so we have introduced a cutoff M to regulate it. This is called an ultraviolet divergence since it arises from the propagation of high momentum modes in the loops. The cutoff

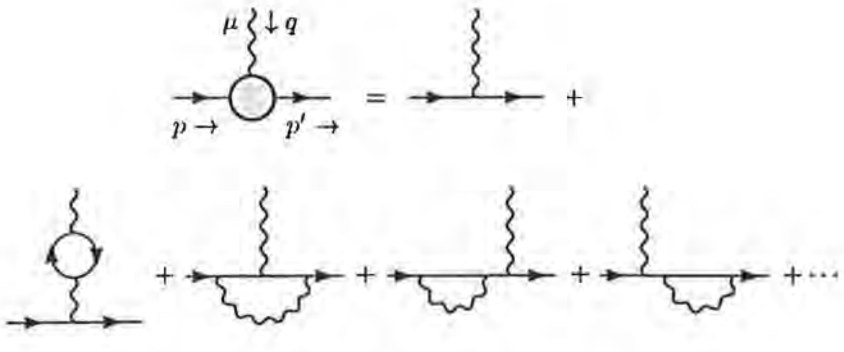


Figure 5.2 Diagrams for vertex renormalisation in QED up to one loop.

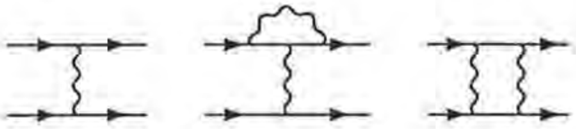


Figure 5.3 Some diagrams for electron–electron scattering in QED up to one loop.

amounts to selecting only those modes where each component of momentum is less than M in magnitude. Despite the divergence in (5.3), it still relates the measurable quantity $\hat{e}$ to the coupling e we introduced in our theory. This implies that e itself must be divergent. The property of renormalisability ensures that in any relation between physical quantities the ultraviolet divergences cancel: the relation is actually independent of the method used to regulate divergences.

As an example, consider the amplitude for electron–electron scattering, which we considered at tree level in section 4.2. Some of the contributing diagrams are shown in Figure 5.3, where the crossed diagrams are understood (we showed the crossed tree level diagram explicitly in Figure 4.2). Ultraviolet divergences are again encountered when the diagrams are evaluated, and the result is of the form,

$$i\mathcal{M}_{fi} = c_0 e^2 + e^4 \Big[c_1 \ln \frac{M^2}{m^2} + d_1\Big] + \cdots \tag{5.4}$$

where c_0, c_1 and d_1 are constants, determined by the calculation. In order to evaluate $\mathcal{M}_{fi}$ numerically, however, we must express it in terms of the known parameter $\hat{e}$. Combining (5.3) and (5.4) yields,

$$i\mathcal{M}_{fi} = c_0 \hat{e}^2 + \hat{e}^4 \Big[(c_1 - 2a_1 c_0) \ln \frac{M^2}{m^2} + d_1 - 2b_1 c_0\Big] + \cdots \tag{5.5}$$

where the ellipsis denotes terms of order $\hat{e}^6$ and above. Since $|\mathcal{M}_{fi}|^2$ is measurable, consistency (renormalisability) requires,

$$c_1 = 2a_1 c_0.$$

This result is indeed borne out by the actual calculations, and the relation between $\mathcal{M}_{fi}$ and $\hat{e}$ contains no divergences:

$$i\mathcal{M}_{fi} = c_0 \hat{e}^2 + \hat{e}^4 (d_1 - 2b_1 c_0) + \mathcal{O}(\hat{e}^6). \tag{5.6}$$

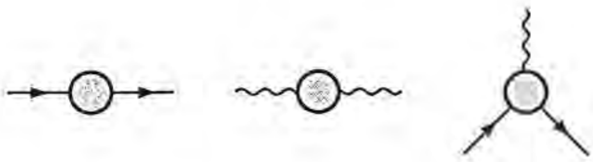


Figure 5.4 Primitive divergences of QED.

To understand how this cancellation of divergences happened we can study the convergence properties of loop diagrams (although we shall not evaluate them). Consider the third diagram on the right hand side in Figure 5.2 and the middle diagram in Figure 5.3. These both contain a loop with one photon propagator, behaving like $1/k^2$ at large momentum k, and two electron propagators, each behaving like $1/k$. To evaluate the diagram we have to integrate over all momenta, leading to an integral,

$$I \sim \int_{\text{large } k} \frac{d^4k}{k^4}, \tag{5.7}$$

which diverges logarithmically, leading to the $\ln M^2$ terms in (5.3) and (5.4). Notice, however, that the divergent terms in these two diagrams must be the same, since the divergence is by its nature independent of the finite external momenta (the factor of two in equation (5.5) arises because there is a divergence associated with the coupling of each electron in the scattering process). In this way we can understand that at least some of the divergences are common in both (5.3) and (5.4). What about diagrams such as the third box-like one in Figure 5.3? Now we have two photon and two electron propagators, leading to,

$$I \sim \int_{\text{large } k} \frac{d^4k}{k^6}.$$

This time the integral is convergent.

Detailed study like this reveals that ultraviolet divergences always disappear in relations between physically measurable quantities. We discussed above the definition of the physical electric charge $\hat{e}$. A similar argument applies for the electron mass: the Lagrangian bare mass parameter m is divergent, but we can define a finite physical mass $\hat{m}$.

In fact you find that all ultraviolet divergences in QED stem from graphs of the type shown in Figure 5.4 and known as the *primitive divergences*. Any divergent graph will be found on inspection to contain a divergent subgraph of one of these basic types. For example, Figure 5.5 shows a graph where the divergence comes from the primitive divergent subgraph inside the dashed box. Furthermore, the primitive divergences are always of a type that would be generated by a term in the initial Lagrangian with a divergent coefficient. Hence by rescaling the fields, masses and couplings in the original Lagrangian we can make all physical quantities finite (and independent of the exact details of the adjustment such as how we regulate the divergent integrals). This is what we mean by renormalisability.

This should be made clearer by an example. Consider calculating the vertex correc-

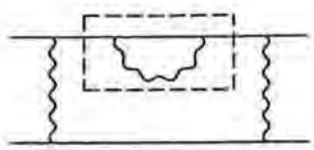
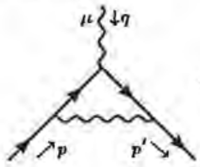


Figure 5.5 Diagram containing a primitive divergence.

tion in QED to one loop,



$$= \overline{u}(p')\Big[A\gamma^{\mu} + B\sigma^{\mu\nu}q_{\nu} + Cq^2\gamma^{\mu} + \cdots\Big]u(p).$$

The calculation shows that A is divergent. However, we can absorb this by adding a cancelling divergent coefficient to the $\overline{\psi}\not{A}\psi$ term in the QED Lagrangian (4.1). The B and C terms are finite and unambiguous. This is just as well, since an infinite part of B, for example, would need to be cancelled by an infinite coefficient of a term of the form,

$$\overline{\psi}\sigma^{\mu\nu}F_{\mu\nu}\psi,$$

which is not available in (4.1).

In fact, the B term gives the QED correction to the magnetic dipole moment, g, of the electron or muon (see page 160 of the textbook by Itzykson and Zuber [1]). These are predicted to be 2 at tree level. You can do the one-loop calculation (it was first done by Schwinger between September and November 1947) with a few pages of algebra to find,

$$g = 2\left(1 + \frac{\alpha}{2\pi}\right).$$

This gives $g/2 = 1.001161$, which is already impressive compared to the experimental values [2]:

$$\begin{aligned}(g/2)_{\text{electron}} &= 1.001159652193(10),\\ (g/2)_{\text{muon}} &= 1.001165923(8).\end{aligned}$$

Higher order calculations show that the electron and muon magnetic moments differ at two loops and above. Kinoshita and collaborators have devoted their careers to these calculations and are currently at the four loop level. Theory and experiment agree for the electron up to the 11th decimal place.

The C term gives the splitting between the $2s_{1/2}$ and $2p_{1/2}$ levels of the hydrogen atom, known as the Lamb shift. Bethe's calculation of the Lamb shift was an early triumph for quantum field theory. Here too, the current agreement between theory and experiment is impressive.

5.1.2 Bare Versus Renormalised

In discussing the vertex correction in QED, we said that the divergent part of the A term could be absorbed by adding a cancelling divergent coefficient to the $\overline{\psi}\not{A}\psi$ term in the QED Lagrangian (4.1). When a theory is renormalisable, *all* divergences can be removed in this way. Thus, for QED, if the original Lagrangian is (ignoring the gauge-fixing term),

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + i\overline{\psi}\not{\partial}\psi - e\overline{\psi}\not{A}\psi - m\overline{\psi}\psi,$$

then redefine everything by:

$$\psi = Z_2^{1/2}\psi_R, \qquad A^\mu = Z_3^{1/2} A_R^\mu,$$
$$e = Z_e \hat{e} = \frac{Z_1}{Z_2 Z_3^{1/2}}\hat{e}, \qquad m = Z_m \hat{m},$$

where the subscript R stands for "renormalised." In terms of the renormalised fields,

$$\mathcal{L} = -\frac{1}{4} Z_3 F_{R\mu\nu} F_R^{\mu\nu} + i Z_2 \overline{\psi}_R \partial\!\!\!/ \psi_R - Z_1 \hat{e} \overline{\psi}_R A\!\!\!/_R \psi_R - Z_m Z_2 \hat{m} \overline{\psi}_R \psi_R.$$

Writing each Z as $Z = 1 + \delta Z$, reexpress the Lagrangian one more time as,

$$\mathcal{L} = -\frac{1}{4} F_{R\mu\nu} F_R^{\mu\nu} + i \overline{\psi}_R \partial\!\!\!/ \psi_R - \hat{e} \overline{\psi}_R A\!\!\!/_R \psi_R - \hat{m} \overline{\psi}_R \psi_R + (\delta Z \text{ terms}).$$

Now it looks like the old lagrangian, but written in terms of the renormalised fields, with the addition of the δZ *counterterms*. Now when you calculate, the counterterms give you new vertices to include in your diagrams. The divergences contained in the counterterms cancel the infinities produced by the loop integrations, leaving a finite answer.

The old A and ψ are called the *bare* fields, and e and m are the bare coupling and mass.

Note that to maintain the original form of $\mathcal{L}$, you want $Z_1 = Z_2$, so that the $\partial\!\!\!/$ and $\hat{e}A\!\!\!/$ terms combine into a covariant derivative term. This relation does hold, and is a consequence of the electromagnetic gauge symmetry: it is known as the *Ward identity*.

5.2 Renormalisation in Quantum Chromodynamics

We now try to repeat the procedure we used for the coupling in QED, but this time in QCD, which is also a renormalisable theory. If we define the renormalised coupling $\hat{g}$ as the strength of the quark–gluon coupling, then in addition to the diagrams of Figure 5.2, with the photons replaced by gluons, there are more diagrams at one loop, shown in Figure 5.6. Looking at the second of these new diagrams, it is ultraviolet divergent (containing a $\ln M^2$ term), but also infrared divergent, since there is no mass to regulate the low momentum modes. In QED all the loop diagrams contain at least one electron propagator and the electron mass provides an infrared cutoff (you still have to worry when the electron is on-shell, but this is not our concern here). In the second diagram of Figure 5.6 there is no quark in the loop. Now suppose we choose to define the renormalised coupling off-shell at some non-zero q^2. The finite value of q^2 provides the infrared regulator and the diagram has a term proportional to $\ln(M^2/q^2)$.

Thus in QCD we can't define a physical coupling constant from an on-shell vertex. This is not really a serious restriction since the QCD coupling is not directly measurable anyway. Now the renormalised coupling depends on how we define it and therefore on at least one momentum scale (in almost all practical cases, only one momentum scale). The renormalised strong coupling is thus written,

$$\hat{g}(q^2).$$

When physical quantities are expressed in terms of $\hat{g}(q^2)$ the coefficients of the perturbation series are finite.

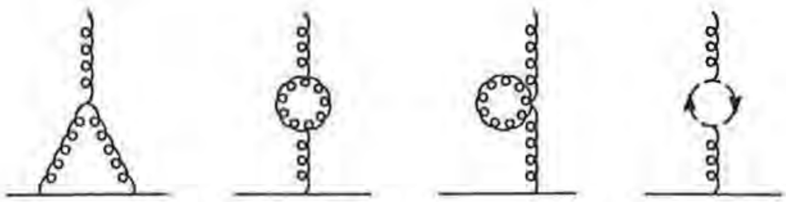


Figure 5.6 Additional diagrams for vertex renormalisation in QCD up to one loop. The dashed line denotes a ghost. For some gauge choices and some regularisation methods not all of these are required.

It would of course be possible to define the renormalised QED coupling to depend on some momentum scale. However, the on-shell definition used above is a natural one to pick.

You can define counterterms for QCD in the same way as was demonstrated for QED. Now the gauge coupling g enters in many terms where it could get renormalised in different ways. In fact, the gauge symmetry imposes a set of relations between the renormalisation constants, known as the *Slavnov–Taylor* identities, which generalise the Ward identity of QED.

5.3 Asymptotic Freedom

We have just seen that the renormalised coupling in QCD, $\hat{g}(q^2)$, depends on the momentum at which it is defined. We say it depends on the *renormalisation scale*, and commonly refer to $\hat{g}$ as the "running coupling constant." We would clearly like to know just how $\hat{g}$ depends on q^2, so we calculate the diagrams in Figures 5.2 and 5.6, to get the first terms in a perturbation theory expansion:

$$\hat{g}(\mu) = g + g^3\left[a_1 \ln \frac{M^2}{\mu^2} + b_1\right] + \cdots \tag{5.8}$$

where a_1 and b_1 are constants and g is the "bare" coupling from the Lagrangian (5.1). I have switched to using μ^2 in place of q^2, and have written $\hat{g}$ as a function of μ for convenience. From this equation it follows that,

$$\mu \frac{\partial \hat{g}}{\partial \mu} \equiv \beta(\hat{g}) = -2a_1\,\hat{g}^3 + \cdots \tag{5.9}$$

The discovery by Politzer and by Gross and Wilczek, in 1973, that $a_1 > 0$ led to the possibility of using perturbation theory for strong interaction processes, since it implies that the strong interactions get weaker at high momentum scales — $\hat{g}(\infty) = 0$ is a stable solution of the differential equation (5.9). Keeping just the $\hat{g}^3$ term, we can solve (5.9) to find,

$$\alpha_s(\mu) \equiv \frac{\hat{g}^2(\mu)}{4\pi} = \frac{4\pi}{\beta_0 \ln(\mu^2/\Lambda^2)}, \tag{5.10}$$

where Λ is a constant of integration and $\beta_0 = 32\pi^2 a_1$. Thus $\alpha_s(\mu)$ decreases logarithmically with the scale at which it is renormalised, as shown in Figure 5.7. If for some process the natural renormalisation scale is large, there is a chance that perturbation theory will be applicable. The value of β_0 is,

$$\beta_0 = 11 - \frac{2}{3}n_f, \tag{5.11}$$

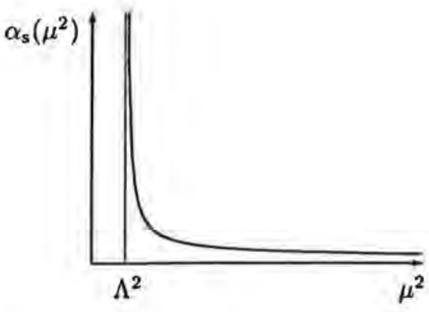


Figure 5.7 Running of the strong coupling constant with renormalisation scale.

where n_f is the number of quark flavours. The crucial discovery when this was first calculated was the appearance of the "11" coming from the self-interactions of the gluons via the extra diagrams of Figure 5.6. Quarks, and other non-gauge particles, always contribute negatively to β_0. Nonabelian gauge theories are the only ones we know where you can have asymptotic freedom (providing you don't have too much "matter" — providing the number of flavours is less than or equal to 16 for QCD).

What is the significance of the integration constant Λ? The original QCD Lagrangian (5.1) contained only a dimensionless bare coupling g (the quark masses don't matter here, since the phenomenon occurs for a pure glue theory), but now we have a dimensionful parameter. The real answer is that the radiative corrections (in all field theories except finite ones) break the scale invariance of the original Lagrangian. In QED there was an implicit choice of scale in the on-shell definition of $\hat{e}$. Lacking such a canonical choice for QCD, you have to say "measure α_s at $\mu = M_Z$" or "find the scale where $\alpha_s = 0.2$," so that a scale is necessarily involved. The phenomenon was called *dimensional transmutation* by Coleman. Λ is given by,

$$\Lambda = \mu \exp\left(- \int^{\hat{g}(\mu)} \frac{dg}{\beta(g)} \right), \tag{5.12}$$

and is μ-independent. The explicit μ dependence is cancelled by the implicit μ dependence of the coupling constant. Today it has become popular to specify the coupling by giving the value of Λ itself.

We've seen that the coupling depends on the scale at which it is renormalised. Moreover, there are many ways of defining the renormalised coupling at a given scale, depending on just how you have regulated the infinities in your calculations and which momentum scales you set equal to μ. The value of $\hat{g}(\mu)$ thus depends on the *renormalisation scheme* you pick, and with it, Λ. In practice, the most popular scheme today is called modified minimal subtraction, $\overline{\text{MS}}$, in which integrals are evaluated in $4 - \epsilon$ dimensions and divergences show up as poles of the form ϵ^{-n} for positive integer n. In the particle data book [2] you will find values quoted for $\Lambda_{\overline{\text{MS}}}$ around 260 MeV (it also depends on the number of quark flavours). Don't buy a value of Λ unless you know which renormalisation scheme was used to define it.

In Figure 5.7 you see that the coupling blows up at $\mu = \Lambda$. This is an artifact of using perturbation theory. We can't trust our calculations if $\alpha_s(\mu) > 1$. In practice, you can perhaps use scales for μ down to about 1 GeV, but not much lower, and 2 GeV is probably safer. This region is a murky area where people try to match perturbative calculations onto results obtained from a variety of more or less kosher techniques.

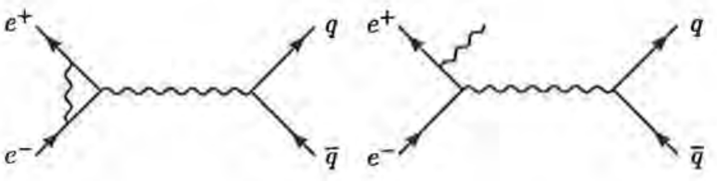


Figure 5.8 QED radiative corrections in e^+e^- annihilation.

▷**Exercise 5.1**

Extending the expansion of $\hat{g}$ in terms of g in (5.8) to two loops gives

$$\hat{g}(\mu) = g + g^3\Big[a_1 \ln \frac{M^2}{\mu^2} + b_1\Big] + g^5\Big[a_2 \ln^2 \frac{M^2}{\mu^2} + b_2 \ln \frac{M^2}{\mu^2} + c_2\Big],$$

with a similar equation for $\hat{g}(\mu_0)$ in terms of g. Renormalisability implies that $\hat{g}(\mu)$ can be expanded in terms of $\hat{g}(\mu_0)$,

$$\hat{g}(\mu) = \sum_{n=0}^{\infty} \hat{g}^{2n+1}(\mu_0) X_n,$$

where the X_n are finite coefficients. Show that this implies that a_2 is determined once the one loop coefficient a_1 is known. In fact a_1 determines all the terms $(\alpha_s \ln \mu)^n$, called the leading logarithms: from a one loop calculation, you can sum up all the leading logarithms.

For QED there is no positive contribution to the beta function, so the electromagnetic coupling has a logarithmic increase with renormalisation scale. However the effect is small even going up to LEP energies: α goes from $1/137$ to about $1/128$. The so called Landau pole, where α blows up, is safely hidden at an enormous energy scale.

5.4 Applications

In this section we will briefly consider some places where perturbative QCD can be applied.

5.4.1 $e^+e^- \to$ hadrons

In section 4.3.2 we considered the ratio R of the annihilation cross section for e^+e^- into hadrons to that into $\mu^+\mu^-$. The result we found from the lowest order annihilation diagram proceeding via an intermediate virtual photon was,

$$R \equiv \frac{\sigma(e^+e^- \to \text{hadrons})}{\sigma(e^+e^- \to \mu^+\mu^-)} = 3 \sum_f Q_f^2, \tag{5.13}$$

where I remind you that sum is over quarks f with Q_f the quark's charge in units of e. Now I would like to extend the discussion in two ways: QED and QCD corrections, and contributions of intermediate Z bosons.

Turning first to QED corrections, consider the two diagrams in Figure 5.8 illustrating two possibilities. The graph on the left contributes to the order α correction to the amplitude. It is ultraviolet divergent, but we have discussed above how to deal with this. However, it is also infrared divergent when the momentum of the photon in the loop goes to zero. The treatment of this problem involves a cancellation of divergences between

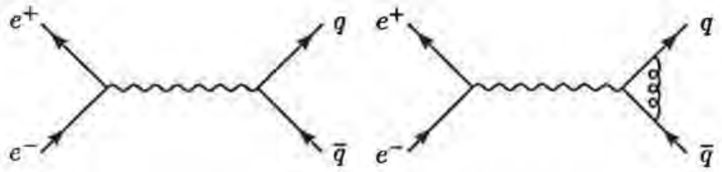


Figure 5.9 QCD radiative corrections in e^+e^- annihilation.

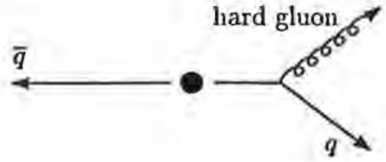


Figure 5.10 QCD bremsstahlung producing a three jet event.

this graph and the bremsstrahlung diagram on the right of Figure 5.8. Physically, limited detector resolution means you can't tell if the final state you detect is accompanied by one (or infinitely many) very soft photons. So, the rate you calculate should also include these undetected photons, and in summing all the terms, the infrared divergences disappear. Since quarks have electric charge, we can also, of course, have QED corrections where the photon lines connect to the quark legs of the annihilation diagram

For the strong interactions, if α_s is not too large and we aren't near a hadronic resonance, then we expect that calculating the diagrams in Figure 5.9 will give the leading QCD corrections. The gluon is exchanged only between the quarks since the incoming e^+e^- don't feel the strong force. The result of the computation is

$$R = 3\sum_f Q_f^2\Big(1 + \frac{\alpha_s(\mu)}{\pi} + \cdots\Big).$$

What value should we choose for μ in this expression? To answer this you need to know that higher order terms in the perturbation series contain powers of $\ln(s/\mu^2)$, where s is the square of the centre of mass energy. So, to avoid large coefficients in the higher order terms, the preferred choice is $\mu^2 \sim s$. Observe that the leading order graph predicts a back-to-back $q\bar{q}$ pair. Owing to hadronisation, what we actually see is a pair of back-to-back jets. Experimentally, the jets follow the angular distribution predicted for the underlying $q\bar{q}$ process, that is, a $(1 + \cos^2\theta)$ distribution, where θ is the scattering angle in the centre of mass frame. Three jet events can arise from QCD bremsstrahlung where a "hard" (high momentum) gluon radiates from one of the quark legs (see Figure 5.10). The observation of such three jet events at DESY in the 1980's was hailed as the "discovery of the gluon."

At present day e^+e^- colliders, the most important contributions to e^+e^- annihilation come from other diagrams in the standard model. In Figure 5.11 we show two diagrams where the e^+e^- can annihilate into a neutral Z boson or a neutral Higgs scalar, H^0. The Z and Higgs propagators contain factors $1/(q^2 - m^2)$ where $q^2 = s$ and m refers to the Z or Higgs mass respectively. For the Z graph, the ratio of its amplitude to the QED amplitude is,

$$\frac{\mathcal{M}_Z}{\mathcal{M}_{\text{QED}}} \sim \frac{q^2}{q^2 - m_Z^2}.$$

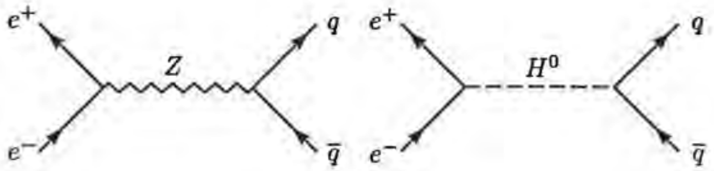


Figure 5.11 Z bosons and Higgs particles in e^+e^- annihilation.

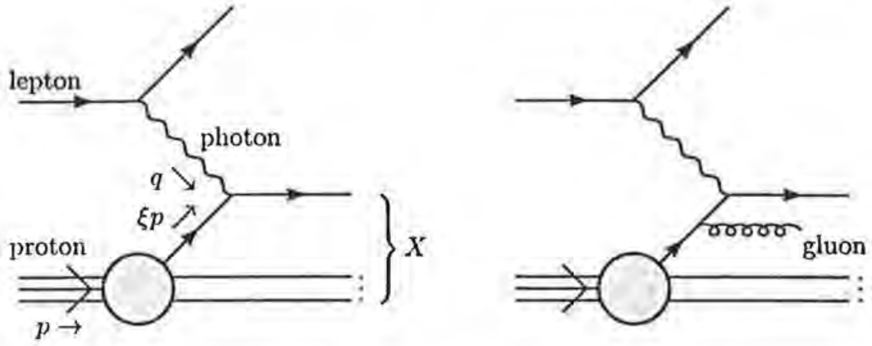


Figure 5.12 DIS process and a QCD correction.

In the Higgs case the ratio is,

$$\frac{\mathcal{M}_{H^0}}{\mathcal{M}_{\text{QED}}} \sim \frac{q^2}{q^2 - m_Z^2} \frac{m_e m_q}{m_W^2}.$$

The extra factors of the electron and quark masses for the Higgs contribution arise because of the standard model mass generation mechanism (see your standard model lectures), and the factor of m_e means that the Z contribution is most important. These amplitude ratios make it clear that as the centre of mass energy approaches m_Z, the Z process will dominate the pure QED one. This, of course, is exactly the situation at LEP.

I will not go further with this subject, but in closing I note that the agreement between the LEP results and the standard model depends on the inclusion of radiative corrections. This agreement provides compelling evidence for the quantum field theoretic aspects of the standard model.

5.4.2 Deep Inelastic Lepton Hadron Scattering

The process of interest is

$$\text{lepton} + \text{hadron} \rightarrow \text{lepton} + X,$$

where X denotes "anything" and the momentum transfer q between the initial and final leptons is large. The initial state lepton may be an electron, muon or neutrino, and the interaction can proceed via the exchange of a photon, W or Z. In Figure 5.12 we illustrate this for electron-proton deep inelastic scattering (DIS), mediated by a photon. The photon couples to one of the qurks in the proton, and since the interaction of the photon and lepton is understood, the strong interaction physics resides in the virtual-photon–proton scattering amplitude.

Choose a Lorentz frame in which the proton is highly relativistic and let the struck quark carry a fraction ξ of the proton's momentum p. Neglecting the struck quark's transverse momentum, since the transverse momentum of secondary particles in hadronic experiments is generally small, we have,

$$(\xi p + q)^2 \approx 0,$$

where we assume the struck quark is nearly on shell and has negligible mass. This leads to,

$$\xi = \frac{-q^2}{2p\cdot q} \equiv x. \tag{5.14}$$

The fraction of the proton's momentum carried by the struck quark is given by the kinematic variable x (known as Bjorken's x variable). Measurements of the differential DIS cross section thus provide information about the momentum distribution of quarks inside hadrons.

What can we say about this process in perturbation theory? In calculating higher order contributions such as that from gluon radiation in the right hand diagram in Figure 5.12, there is an important difference from the calculation of the R ratio for e^+e^- annihilation in (5.13). The region of phase space where the struck quark is nearly on shell is important, as was anticipated above in the identification of ξ with x in (5.14). This manifests itself in the appearance of terms of the form $\alpha_s^n \ln^n(q^2/\lambda^2)$, where λ is some lower cutoff on the quark's momentum. The choice of λ depends on details of the proton wavefunction and hence these terms can't be calculated in perturbation theory. In other words, the relevant momenta are small, and DIS cross sections are not calculable in perturbation theory. However, for large q^2, it *is* possible to to compute the evolution of these cross sections with q^2, since these effects depend on the region of phase space where the quark is far off shell ($q^2 \gg \Lambda^2$). So, in summary, although DIS cross sections are not themselves calculable, their dependence on q^2 is. This is sufficient for a considerable amount of phenomenology — see your course on the Physics of Structure.

DIS cross sections, and hence the momentum distribution of quarks in a proton, depend on q^2. As q^2 increases, theory predicts that there should be fewer quarks at large x and more at small x. This result has a physical interpretation. Imagine probing a proton with a virtual photon and seeing a quark carrying fraction y of the proton's momentum. If you increase the photon energy, you may see that what you thought was a quark with momentum yp is actually a quark with momentum xp together with a gluon of momentum $(y-x)p$. Thus the total momentum of the quark and gluon is yp and the quantum numbers of the pair are those of a single quark. In the first case, the pair was not resolved, but in the second case we see that since $x < y$ the quark's momentum is less now than when we looked with a lower energy photon.

There is currently great interest in DIS processes at HERA, which is allowing us to explore smaller values of x, giving a new testing ground for theoretical ideas.

5.4.3 Drell–Yan and Related Processes

Now consider a process with two initial state hadrons. For illustration, consider the Drell–Yan process,

$$\text{hadron} + \text{hadron} \to e^+(\mu^+) + e^-(\mu^-) + X,$$

where the centre of mass energy of the hadrons and the invariant mass of the lepton pair are large and comparable. A parton model for this process, proposed by Drell and

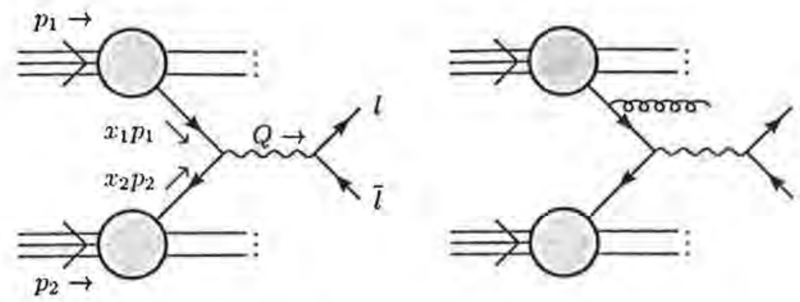


Figure 5.13 Drell–Yan process and a QCD correction.

Yan is illustrated in Figure 5.13. A quark from one of the initial hadrons, labelled with subscript 1 in the figure, annihilates an antiquark from the other hadron, producing a virtual photon which in turn decays into a lepton–antilepton pair.

The momentum distribution of the quarks in the initial state hadrons can be determined from DIS experiments, so the process is calculable in terms of those distributions:

$$\frac{d\sigma}{dQ^2} = \frac{4\pi\alpha^2}{9Q^4}\sum_f Q_f^2 \int_0^1 dx_1\, dx_2$$
$$\times\delta(x_1 x_2 - Q^2/s)x_1 x_2\big[q_{1f}(x_1)\overline{q}_{2f}(x_2) + q_{2f}(x_2)\overline{q}_{1f}(x_1)\big], \qquad (5.15)$$

where $s = (p_1 + p_2)^2$ and $q_{if}(x)$ is the probability density for finding a quark of flavour f in hadron i carrying a fraction x of its momentum (similarly for $\overline{q}_{if}$). Now consider some higher order correction such as the gluon radiation graph on the right of Figure 5.13. Just as for DIS there are important contributions from the "long-distance" region of phase space, where the quark and antiquark are almost on-shell. However, close study reveals that these long-distance contributions are precisely the same as in DIS, so can be absorbed into the quark distribution functions. Thus the Drell–Yan and DIS cross sections can be related in perturbation theory. The relation is just equation (5.15) with the $q_{if}(x_i)$ replaced by $q_{if}(x_i, M^2)$, which is the probability density determined from DIS experiments with $q^2 = M^2$. There are further perturbative corrections to (5.15), but the large logarithms coming from long-distance physics can always be absorbed into the distribution functions.

The factorisation of long distance effects into the distribution functions is a common feature of hard inclusive processes, including, besides Drell–Yan production, the production of particles or jets with large transverse momenta. In each case the cross section is a convolution of the partonic distribution functions with the cross section for the quark or gluon hard scattering process. Thus hadrons can be viewed as broad band beams of quarks and gluons, with a known (experimentally determined) momentum distribution. These beams are what we use to search for the Higgs scalar, or signals of new physics such as technicolour or supersymmetry — but that is all material for another course.

Acknowledgements

I would like to thank Tim Jones and Chris Sachrajda for lending me copies of their notes from previous YHEP schools. These notes are based heavily on those, and some sections are copied almost verbatim.

It is also a pleasure to thank Ken Peach for organising and cheerleading the school, Ann Roberts for keeping everything running smoothly, and the students for listening.

References

[1] I J R Aitchison and A J G Hey *Gauge theories in particle physics, 2nd ed* Adam Hilger 1989
B F Hatfield *Quantum Field Theory of Point Particles and Strings* Addison Wesley 1992
C Itzykson and J-B Zuber *Quantum Field Theory* McGraw-Hill 1980
M Le Bellac *Quantum and Statistical Field Theory* Oxford 1991
F Mandl and G Shaw *Quantum Field Theory* Wiley 1984
P Ramond *Field Theory: a Modern Primer, 2nd ed* Addison Wesley 1989
L H Ryder *Quantum Field Theory* Cambridge 1991 (reprint)

[2] Particle Data Group, Phys. Rev. D **45** (1992) S1

THE STANDARD MODEL

By Dr G M Shore
University of Wales, Swansea

Lectures delivered at the School for Young High Energy Physicists
Rutherford Appleton Laboratory, September 1994

The Standard Model

Graham M. Shore

Department of Physics
University of Wales Swansea
Singleton Park
Swansea SA2 8PP, U.K.

Lectures presented at the School for Young High Energy Physicists,
Rutherford Appleton Laboratory, September 1994

Contents

1 Introduction

The standard model is a beautifully crafted and brilliantly predictive theory of all known phenomena in elementary particle physics. It was conceived in the decade from the mid 1960s to the mid 1970s when quantum field theory made a spectacular revival and non-abelian gauge theories were shown to provide a quantitative understanding of particle physics. Those were (I am told) heady times for theorists. Writing in 1984, Sidney Coleman remembers them nostalgically:

"This was a great time to be a high-energy theorist, the period of the famous triumph of quantum field theory. And what a triumph it was, in the old sense of the word: a glorious victory parade, full of wonderful things brought back from far places to make the spectator gasp with awe and laugh with joy."

Since then, the $SU(3)_C \times SU(2)_L \times U(1)_Y$ standard model, the fusion of quantum chromodynamics with the electroweak theory of Glashow, Salam and Weinberg, has successfully described (or at least not contradicted) all experimental data.

These lectures describe the construction of the standard model, with particular reference to the symmetry structure and tree-level dynamics of the electroweak interactions. They are complementary to the other lecture courses in this volume, which describe in more depth the quantum dynamics of gauge theories.

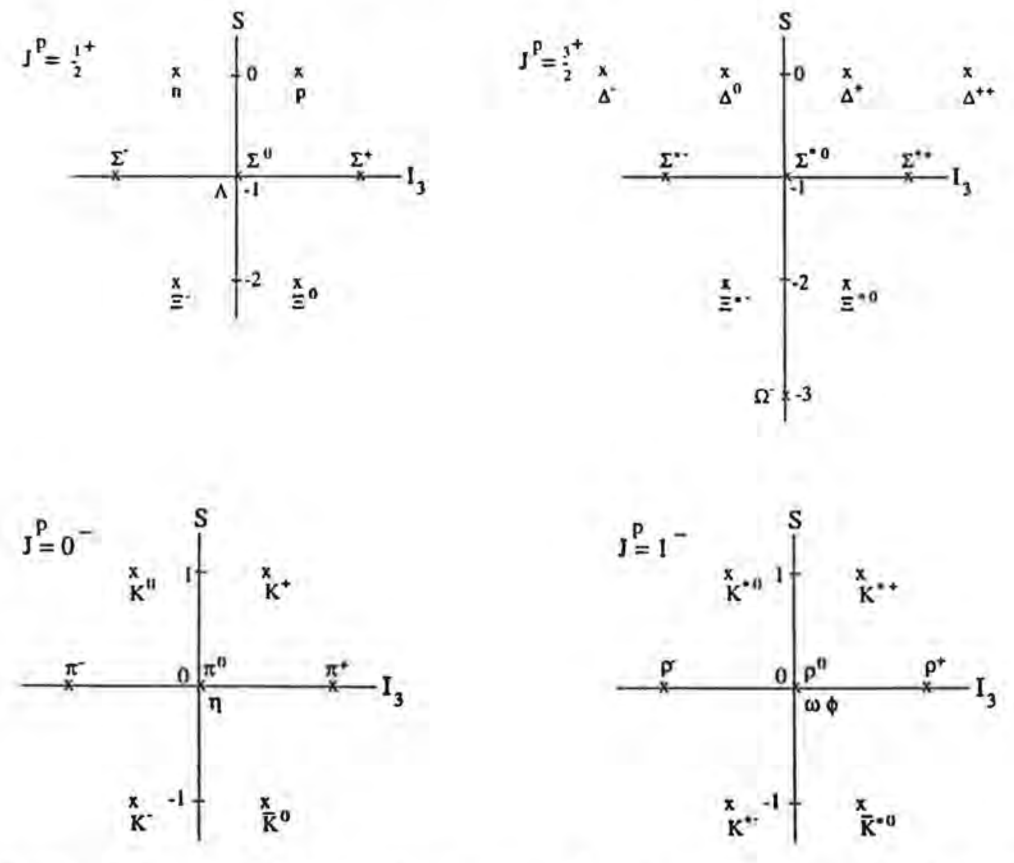
There are many excellent books on gauge theories and the standard model. The description given in these lectures follows quite closely the presentation in the book by Halzen and Martin, 'Quarks and Leptons'. This would provide a good source of supplementary reading and further examples.

2 Elementary particles, QED and QCD

We begin by listing the elementary particles which are currently known to exist in nature. These are the leptons ℓ, quarks q and the gauge bosons which mediate the fundamental forces.

Leptons	$\ell=$	e	μ	τ	ν_e	ν_μ	ν_τ
	mass(MeV)	0.51	105.6	1784	$\leq$46eV	$\leq$0.25	$\leq$70
Quarks	q=	u	d	s	c	b	t
	mass	7MeV	15MeV	200MeV	1.3GeV	4.8GeV	?175 GeV
	charge	$\frac{2}{3}$	$\frac{-1}{3}$	$\frac{-1}{3}$	$\frac{2}{3}$	$\frac{-1}{3}$	$\frac{2}{3}$

The quarks do not exist as free particles, but are permanently bound into hadrons. This is confinement. If we consider just the three quarks u, d, s, we form the baryon and meson octets and decuplets of 'flavour' $SU(3)$:



With the discovery of charm, bottom, ... the picture can be extended. New hadrons exist and fit into multiplets of higher flavour symmetries $SU(4)$, ... For example, there are the charmed mesons such as $D^+ = c\bar{d}$ with $m = 1.86GeV$ which decays by $D^+ \to K^-\pi^+\pi^+$. Of course, because of the mass differences between the quarks, these flavour symmetries are only approximate. All this phenomenology establishes the quarks as the elementary particles; mesons and baryons are bound states.

The next category of elementary particles are the gauge bosons:

$$\gamma \qquad \mathrm{g} \qquad W^{\pm} \quad \mathrm{Z}$$

The photon γ mediates the electromagnetic interaction, described by quantum electrodynamics (QED). It is massless. The strong (inter-quark, not inter-nuclear) force is mediated by a 'colour' octet of massless gluons and described by another gauge theory, quantum chromodynamics (QCD). Finally, the gauge bosons corresponding to the weak interactions are the charged $W^{\pm}$ and neutral Z, with masses of 80 and $91GeV$ respectively. These were discovered in 1983 by the UA1 and UA2 collaborations at CERN.

Finally, as we shall see, a further ingredient is required to make the picture work. The minimal standard model also predicts the existence of a scalar particle H^0, the famous Higgs boson.

In the standard quantum field theory model, all these elementary particles are considered to be the quanta of elementary fields.

The simplest example of a gauge theory of this type is QED, describing the interaction of electrons and photons. The action is

$$S = \int dx \left[\bar{\psi}\gamma^{\mu}(\partial_{\mu} - ieA_{\mu})\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} + m^2\bar{\psi}\psi \right] \tag{1}$$

where ψ is the electron field and A_{μ} is the photon field. Green functions (and hence S-matrix elements, etc.) are constructed from the path integral,

$$Z = e^{W[J,K,\bar{K}]} = \int \mathcal{D}\psi\mathcal{D}\bar{\psi}\mathcal{D}A \; e^{i\int dx \; \mathcal{L} + J^{\mu}A_{\mu} + \bar{K}\psi + \bar{\psi}K} \tag{2}$$

usually using perturbation theory, Feynman diagrams, etc.

In the early 1970s, it was realised that the strong interaction could be described by a non-abelian gauge theory, quantum chromodynamics. Each quark is assigned a colour quantum number, corresponding to the gauge group $SU(3)_C$. QCD is 'flavour blind', i.e. independent of the type of quark.The action is

$$S = \int dx\mathcal{L} = \sum_{flavour} \int dx \left[\bar{\psi}\gamma^{\mu}(\partial_{\mu} + igT^aA^a_{\mu})\psi - \frac{1}{4}G^a_{\mu\nu}G^{a\mu\nu} + m^2\bar{\psi}\psi \right] \tag{3}$$

where ψ is the colour triplet quark-field, A^a_{μ} is the colour octet gluon field and T^a is the matrix specifying the quark representation (for quarks, the fundamental representation of $SU(3)_C$).

The physics of non-abelian gauge theories is very different from QED. In particular QCD exhibits asymptotic freedom – the effective coupling $\to 0$ at small distances. This implies simple (quasi-free) behaviour of quarks in deep inelastic scattering experiments probing the structure of the proton. The inverse effect (infrared slavery), viz. the increase in the effective coupling at long distances, is related to confinement.

At this point, with QED and QCD, we have a theory of the strong and the electromagnetic interactions:

Gauge group	$SU(3)_c \times U(1)_{em}$				
Elementary fields	g	γ			
	e	μ	τ		
	u	d	s	c	$\dots$

$SU(3)_C$ acts only on the colour degree of freedom of the quarks. $U(1)_{em}$ acts on all charged particles. The theory is parametrised by two coupling constants e and g, the latter being traded for Λ_{QCD} according to dimensional transmutation, plus masses. There are no constraints on the masses, mass terms in $\mathcal{L}$ being gauge invariant.

This leaves the weak interactions to be incorporated. These are much more complicated – they act on the flavour degrees of freedom of the quarks and between ν, e, etc. The following are examples of weak interaction processes:

$$\begin{aligned} n &\to p\, e^- \bar{\nu}_e \\ (d &\to u\, e^- \bar{\nu}_e) \\ \pi^- &\to \mu^- \bar{\nu}_\mu \\ \mu^- &\to e^- \bar{\nu}_e \nu_\mu \\ \nu_\mu e^- &\to \mu^- \nu_e \\ \nu_\mu N &\to \mu^- X \end{aligned}$$

If the weak interactions were really distinct from the other two, we would simply have to enlarge the gauge group to include a new 'quantum flavourdynamics' group G_W acting on the quark flavours and lepton types. However, the picture which will emerge from the following discussion is more subtle. The weak interactions mix with electromagnetism and weave together the intricate tapestry that is the standard model.

3 Weak Interactions

3.1 Effective current-current interaction

The weak interactions were originally described by a phenomenological current-current interaction. From a modern viewpoint, we understand this interaction in terms of the effective low-energy Lagrangian implied by a gauge theory of massive vector bosons.

To motivate this, consider integrating out the gauge field in the QED Lagrangian

$$S = \int dx \left[\frac{1}{2} A^\mu D_{\mu\nu} A^\nu \ - \ eA^\mu J_\mu^{\rm em}\right] \tag{4}$$

where $D_{\mu\nu} = \partial^2 g_{\mu\nu} - (1-\alpha)\partial_\mu \partial_\nu$, where α is the gauge-fixing parameter. Completing the square, we find

$$\int \mathcal{D}A e^{iS} \ = \ \int \mathcal{D}A e^{i\int \frac{1}{2}(A-eJ/D)D(A-eJ/D)-e^2J^2/2D} \tag{5}$$

$$= \ e^{i\int \mathcal{L}_{eff}} \tag{6}$$

with

$$\mathcal{L}_{eff} = -\frac{i}{2} e^2 J_\mu^{\rm em} \Delta^{\mu\nu} J_\nu^{\rm em} \tag{7}$$

where Δ is the photon propagator, $D\Delta = i\delta$.

The QED interaction is therefore of current-current type, but mediated by a propagator $\sim \frac{1}{q^2}$. It is therefore a long-range interaction. The electromagnetic current is

$$J_\mu^{\rm em} = -\bar{e}\gamma_\mu e - \bar{\mu}\gamma_\mu \mu + \ldots \tag{8}$$

where e, μ, ... are Dirac fields for the electron, muon, etc.

At low energies ($q \ll m_W$), the weak interactions can be well described by an effective theory comprising a current-current interaction. Since the weak interactions are short range, a good approximation is to replace the propagator by a constant, which is equivalent to a point interaction, i.e.

$$\mathcal{L}_{eff}^{\rm weak} = G^{\rm AB} J_\mu^{\rm A} J^{\mu \rm B} \tag{9}$$

Notice that G has dimensions of $mass^{-2}$ which implies that this is a non-renormalisable interaction. It violates unitarity (cross sections $\sigma \sim s$ for large energy). This means that the current-current interaction cannot be fundamental. Nevertheless, it gives an excellent description of weak interaction processes for momenta below m_W.

Our aim is to build a renormalisable gauge theory of the weak interactions. The next step, therefore, is to extract the form of the weak currents $J_\mu^{\rm A}$ from the phenomenology of weak interactions.

3.2 Lorentz structure of currents

The general Lorentz structure for a bilinear fermion current is

$$J = \bar{\psi}\Gamma\psi \tag{10}$$

where $\Gamma = 1,\ \gamma_5,\ \gamma_\mu,\ \gamma_\mu\gamma_5,\ \sigma_{\mu\nu}$ (total=16)

Now, if the current-current interaction is derived from a gauge theory with vector bosons, we will have either $\Gamma = \gamma_\mu$ or $\gamma_\mu\gamma_5$ (the so-called V or A currents). Extensive studies of weak interaction phenomenology in the 1950s showed that this is indeed true – the other forms (S, P and T) are excluded by experiment.

The original assumption was that Γ must be γ_μ, based on the analogy with the electromagnetic current. This was the basis of the 1932 Fermi theory of β decay. The $\gamma_\mu\gamma_5$, or A, interaction would violate parity.

However, in 1956, Lee and Yang surveyed weak interaction data and concluded that parity may not be conserved (e.g. $K^+ \to \pi\pi$ and $\pi\pi\pi$ both occur). The experimental confirmation of parity violation by Wu ($^{60}C \to {}^{60}Ni\, e^-\, \bar{\nu}_e$, polarised beta decay), Ledermann ($\pi^- \to \mu^-\bar{\nu}_\mu$ followed by $\mu^- \to e^-\, \bar{\nu}_e\, \nu_\mu$) and others followed shortly after.

The cumulative experimental evidence led to the identification (by Feynman and Gell-Mann and Marshak et al.) of the Lorentz structure of the charged weak current as V−A, i.e. $\Gamma = (1-\gamma_5)\gamma_\mu$. Also, only the left-handed (helicity $-\frac{1}{2}$, $\nu_L = \frac{1}{2}(1-\gamma_5)\nu$) neutrino seems to occur in nature, together with the right-handed antineutrino. There is no ν_R state.

Because left and right handed states enter differently in weak interaction theory, it is convenient to use the left and right handed projections for all particles. So, e.g.

$$e_L = \frac{1}{2}(1-\gamma_5)e \tag{11}$$

$$e_R = \frac{1}{2}(1+\gamma_5)e \tag{12}$$

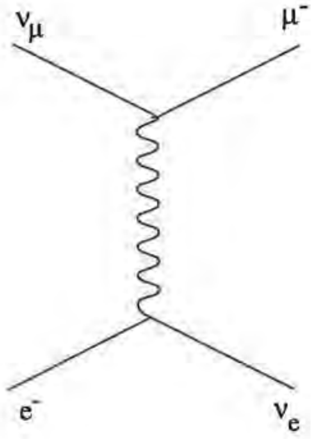
are the helicity $-\frac{1}{2}$ and $+\frac{1}{2}$ components of the electron.

Under parity ($P^{-1}\gamma_5 P = -\gamma_5$)

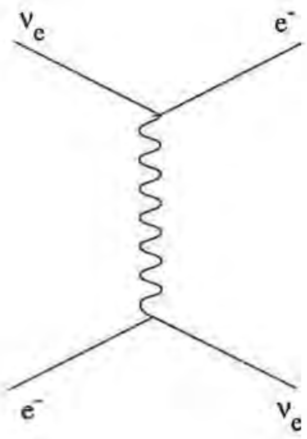
$$e_L \overset{P}{\longleftrightarrow} e_R \tag{13}$$

Under charge conjugation ($\psi_C = C\gamma^0\psi^* = C\bar{\psi}^T$)

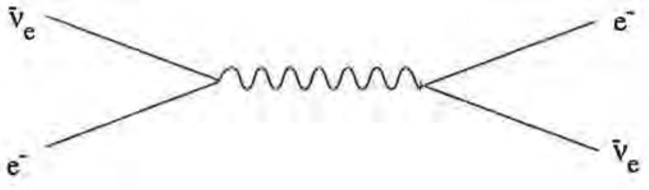
$$e \overset{C}{\longleftrightarrow} \bar{e}^T \tag{14}$$



3. $\nu_e\, e^- \to \nu_e\, e^-$



4. $\bar{\nu}_e\, e^- \to \bar{\nu}_e\, e^-$



The crossed diagrams for (3) and (4) do not occur. These would require gauge bosons carrying lepton number.

Elastic $\nu_e\, e^-$ and $\bar{\nu}_e\, e^-$ scattering also have neutral current contributions. However, only the charged current contributes to (2).

All these interactions are of the current-current form:

$$\mathcal{L}_{eff} = \frac{4\,G}{\sqrt{2}}\, J^{CC}_\mu\, J^{CC\,\mu\dagger} \tag{19}$$

where the lepton charged current is

$$\begin{aligned} J^{CC}_\mu &= \bar{\nu}_e\,\gamma_\mu\,\tfrac{1}{2}\,(1-\gamma_5)\,e + \bar{\nu}_\mu\,\gamma_\mu\,\tfrac{1}{2}\,(1-\gamma_5)\,\mu + \bar{\nu}_\tau\,\gamma_\mu\,\tfrac{1}{2}\,(1-\gamma_5)\,\tau & (20)\\ &= \bar{\nu}_{eL}\,\gamma_\mu\,e_L + \bar{\nu}_{\mu L}\,\gamma_\mu\,\mu_L + \bar{\nu}_{\tau L}\,\gamma_\mu\,\tau_L & (21)\end{aligned}$$

and

$$J^{CC\,\dagger}_\mu = \bar{e}_L\,\gamma_\mu\,\nu_{eL} + \bar{\mu}_L\,\gamma_\mu\,\nu_{\mu L} + \bar{\tau}_L\,\gamma_\mu\,\nu_{\tau L} \tag{22}$$

Notice that cross terms linking, for example, electron and muon type currents are possible – the gauge bosons are independent of the generation. This is known as "universality" of the weak interactions.

The coupling strength G (the Fermi constant) is the same for all these processes. This indicates a single underlying explanation. If we postulate that the interaction is due to the exchange of a massive vector boson $W^\pm$ with propagator

$$\Delta_{\mu\nu} = i\,\frac{1}{q^2 - m_W^2}\left(-g_{\mu\nu} + \frac{q_\mu\,q_\nu}{m_W^2}\right) \overset{q\ small}{\sim} i\left(\frac{g_{\mu\nu}}{m_W^2}\right) \tag{23}$$

then the effective Lagrangian becomes

$$\mathcal{L}_{eff} = -i\left(\frac{g}{\sqrt{2}}\right)^2 J^{CC}_\mu\,\Delta^{\mu\nu}\,J^{CC\,\dagger}_\nu \sim \frac{g^2}{2\,m_W^2} J^{CC}_\mu\,J^{CC\,\mu\dagger} \tag{24}$$

So we can identify

$$\frac{G}{\sqrt{2}} = \frac{g^2}{8\,m_W^2} \tag{25}$$

The V$-$A structure can be verified from $\nu\,e$ scattering. If we assume a Lorentz structure $J^{CC}_\mu = \bar{\nu}_\ell\,\gamma_\mu(a + b\gamma_5)\ell$ for the weak current, then process (2) gives

$$\frac{d\sigma}{d\Omega}(\nu_\mu\,e^- \to \nu_e\mu^-) = \frac{G^2\,s}{32\,\pi^2}\left(A^+ + A^-\,\cos^4\frac{\theta}{2}\right) \tag{26}$$

where $A^\pm = (a^2 + b^2)^2 \pm 4a^2b^2$.
(See Halzen and Martin, sect. 12.7 for cross-sections for charged current $\nu\,e$ and $\bar{\nu}\,e$ scattering.)

from a non-abelian gauge theory with interaction $J_\mu^A A^{A\mu}$, then the currents must form a representation of the gauge group.

For one lepton generation, these currents are

$$\begin{aligned} J_\mu^{CC} = J_\mu^- &= \bar{\nu}_L \gamma_\mu e_L \\ J_\mu^{CC\dagger} = J_\mu^+ &= \bar{e}_L \gamma_\mu \nu_L \\ J_\mu^{NC} &= \frac{1}{2}\Big(\bar{\nu}_L \gamma_\mu \nu_L \ + \ c_L^e \bar{e}_L \gamma_\mu e_L \ + \ c_R^e \bar{e}_R \gamma_\mu e_R\Big) \end{aligned} \tag{32}$$

The charged currents $J_\mu^\pm$ can form two of the three components of the adjoint representation of $SU(2)$. In the fundamental (2-dimensional) representation of $SU(2)$, the generators are $T^A = \frac{1}{2}\tau^A$ and satisfy the commutation relations

$$[T^A, T^B] = i\epsilon^{ABC} T^C, \qquad A = 1,2,3 \tag{33}$$

For the charged components,

$$\tau^\pm = \frac{1}{2}(\tau^1 \pm i\tau^2) \tag{34}$$

$$\Rightarrow \ \tau^+ = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \tau^- = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$$

Now construct lepton doublets

$$\chi_L = \begin{pmatrix} \nu_{eL} \\ e_L \end{pmatrix}, \qquad \text{2-dim } SU(2)_L \text{ representation} \tag{35}$$

The currents $J_\mu^\pm$ can then be written as

$$J_\mu^\pm = \bar{\chi}_L \gamma_\mu \tau^\pm \chi_L = J_\mu^1 \pm i J_\mu^2 \tag{36}$$

i.e. as components of the $SU(2)_L$ current

$$J_\mu^A = \bar{\chi}_L \gamma_\mu T^A \chi_L \tag{37}$$

The remaining component is

$$\begin{aligned} J_\mu^3 &= \frac{1}{2}\bar{\chi}_L \gamma_\mu \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \chi_L \\ &= \frac{1}{2}\left(\bar{\nu}_L \gamma_\mu \nu_L - \bar{e}_L \gamma_\mu e_L\right) \end{aligned} \tag{38}$$

However, this has no right-handed part and so it obviously cannot be identified with the remaining current J_μ^{NC}.

The solution is to introduce a new current, corresponding to a new $U(1)_Y$ interaction. The proposal is to define

$$J_\mu^Y = -\bar{\nu}_L \gamma \nu_L - \bar{e}_L \gamma_\mu e_L - 2\bar{e}_R \gamma_\mu e_R \tag{39}$$

so that $U(1)_Y$ commutes with $SU(2)_L$.

Now try and express J^{NC}_μ and J^{em}_μ as linear combinations of J^3_μ and J^Y_μ :

$$\begin{aligned} J^{em}_\mu &= -\bar{e}_L\gamma_\mu e_L - \bar{e}_R\gamma_\mu e_R = J^3_\mu + \frac{1}{2}J^Y_\mu \\ J^{NC}_\mu &= aJ^3_\mu + bJ^Y_\mu \end{aligned} \tag{40}$$

$$\Rightarrow \left.\begin{aligned} a - 2b &= 1 \\ a + 2b &= -c_L \\ 4b &= -c_R \end{aligned}\right\} \qquad \text{solution only if } 1 + c_L - c_R = 0 \tag{41}$$

Then, $a = \frac{1}{2}(1 - c_L)$ and $b = -\frac{1}{4}c_R$. So, provided we have $1 + c_L - c_R = 0$, we can express

$$J^{NC}_\mu = \frac{1}{2}(1 - c_L)J^3_\mu - \frac{1}{4}c_R J^Y_\mu \tag{42}$$

$$= J^3_\mu - \frac{1}{2}c_R J^{em}_\mu \tag{43}$$

where recall $c_R = c_V - c_A$

The condition $1 + c_L - c_R = 0$ requires $c_A = -\frac{1}{2}$. This is necessary for this mixing scheme to work.

To incorporate this structure into a gauge theory, choose a gauge group $SU(2)_L \times U(1)_Y$, with gauge bosons W^A_μ and B_μ. The interaction term in the Lagrangian is

$$\mathcal{L}_{int} = -g\, J^A_\mu W^{A\,\mu} - \frac{g'}{2} J^Y_\mu B^\mu \tag{44}$$

In terms of $J^\pm_\mu$, J^{em}_μ and J^{NC}_μ we have

$$\begin{aligned} \mathcal{L}_{int} = &- \frac{g}{\sqrt{2}}\left(J^+_\mu W^{+\,\mu} + J^-_\mu W^{-\,\mu}\right) \\ &- J^{em}_\mu\left(g\,\tfrac{1}{2}\,c_R W^{3\,\mu} + g'(1 - \frac{1}{2}c_R)B^\mu\right) \\ &- J^{NC}_\mu\left(gW^{3\,\mu} - g'B^\mu\right) \end{aligned} \tag{45}$$

where $W^\pm_\mu = \frac{1}{\sqrt{2}}\left(W^1_\mu \mp iW^2_\mu\right)$.

In the Weinberg-Salam model, the mixing between W^3_μ and B_μ to give A_μ and Z_μ is of the following form:-

$$\begin{aligned} Z_\mu &= W^3_\mu \cos\theta_W - B_\mu \sin\theta_W \\ A_\mu &= W^3_\mu \sin\theta_W + B_\mu \cos\theta_W \end{aligned} \tag{46}$$

Then the Lagrangian becomes

$$\mathcal{L} = - \frac{1}{\sqrt{2}} \frac{e}{\sin\theta_W} \left(J_\mu^+ W^{+\mu} + J_\mu^- W^{-\mu} \right) \tag{47}$$

$$- e J_\mu^{\mathrm{em}} A_\mu - \frac{e}{\sin\theta_W \cos\theta_W} J_\mu^{NC} Z^\mu \tag{48}$$

where we identify

$$e = g \sin\theta_W = g' \cos\theta_W \tag{49}$$

and

$$\frac{1}{2} c_R = \sin^2\theta_W \tag{50}$$

This last result implies that

$$J_\mu^{NC} = J_\mu^3 - \sin^2\theta_W J_\mu^{\mathrm{em}} \tag{51}$$

The resulting current-current effective interaction is

$$\mathcal{L} = \left(\frac{g}{\sqrt{2}}\right)^2 J_\mu^{CC} \frac{1}{m_W^2} J^{CC\,\mu\dagger} + \left(\frac{g}{\cos\theta_W}\right)^2 J_\mu^{NC} \frac{1}{m_Z^2} J^{NC\,\mu} \tag{52}$$

Comparing with

$$\mathcal{L} = \frac{4G}{\sqrt{2}} \left(J_\mu^{CC} J^{CC\,\mu\dagger} + 2\rho J_\mu^{NC} J^{NC\,\mu} \right) \tag{53}$$

we identify

$$\frac{G}{\sqrt{2}} = \frac{g^2}{8\, m_W^2} = \frac{e^2}{8\, m_W^2 \sin^2\theta_W} \tag{54}$$

and

$$\rho = \frac{m_W^2}{m_Z^2 \cos^2\theta_W} \tag{55}$$

4.2 Weinberg-Salam Lagrangian (leptons)

The $SU(2)_L \times U(1)_Y$ Lagrangian is therefore

$$\begin{aligned} \mathcal{L} = & - \frac{1}{4} F_{\mu\nu}^A F^{\mu\nu\,A} - \frac{1}{4} F_{\mu\nu}^a F^{\mu\nu\,a} \\ & + \sum_{i=e,\mu,\tau} i \bar{\chi}_L^i \left(\partial_\mu + i g T^A W_\mu^A + i \frac{g'}{2} Y B_\mu \right) \chi_L^i \\ & + \sum_{i=e,\mu,\tau} i \bar{\psi}_R^i \left(\partial_\mu + i \frac{g'}{2} Y B_\mu \right) \psi_R^i \end{aligned} \tag{56}$$

where $\quad T^A = \frac{1}{2} \tau^A \quad$ determines the $SU(2)_L$ representation of χ_L

$$\text{and} \quad Y = -1 \quad \text{for } \chi_L = \begin{pmatrix} \nu_{eL} \\ e_L \end{pmatrix}, \begin{pmatrix} \nu_{\mu L} \\ \mu_L \end{pmatrix}, \begin{pmatrix} \nu_{\tau L} \\ \tau_L \end{pmatrix} \tag{57}$$

$$Y = -2 \quad \text{for } \psi_R = e_R, \mu_R, \tau_R \tag{58}$$

The parameters are

$$g, g', m_W, m_Z$$

or equivalently

$$e, \sin\theta_W, m_W, m_Z$$

The interaction terms are

$$\mathcal{L} = -e\, J_\mu^{\rm em} A^\mu - \frac{e}{\sin\theta_W \cos\theta_W} J_\mu^{NC} Z^\mu \tag{59}$$

where

$$\begin{aligned} Z_\mu &= W_\mu^3 \cos\theta_W - B_\mu \sin\theta_W \\ A_\mu &= W_\mu^3 \sin\theta_W + B_\mu \cos\theta_W \end{aligned} \tag{60}$$

and

$$\begin{aligned} J_\mu^{\rm em} &= J_\mu^3 + \frac{1}{2} J_\mu^Y \\ J_\mu^{NC} &= J_\mu^3 - \sin^2\theta_W J_\mu^{\rm em} \end{aligned} \tag{61}$$

The effective interaction is

$$\mathcal{L}_{int} = \frac{4G}{\sqrt{2}} \left(J_\mu^{CC} J^{CC\,\mu\dagger} \; + \; 2\rho J_\mu^{NC} J^{NC\,\mu} \right) \tag{62}$$

together with electromagnetism. This phenomenological description has the parameters e, G, ρ, c_V^e, c_A^e.

The Weinberg-Salam model requires $c_A^e = -\frac{1}{2}$. The other equivalences are

$$\frac{G}{\sqrt{2}} = \frac{e^2}{8 m_W^2 \sin^2\theta_W} \tag{63}$$

$$\rho = \frac{m_W^2}{m_Z^2 \cos^2\theta_W} \tag{64}$$

$$c_V^e = -\tfrac{1}{2} + 2\sin^2\theta_w \tag{65}$$

Universality implies that $c_{V,A}^e = c_{V,A}^\mu = c_{V,A}^\tau$ as well as a single G, ρ.

In fact, the full Weinberg-Salam model including the Higgs mechanism also implies $\rho = 1$ because of an additional (custodial $SU(2)$) symmetry which is built into the model. (See section 9.)

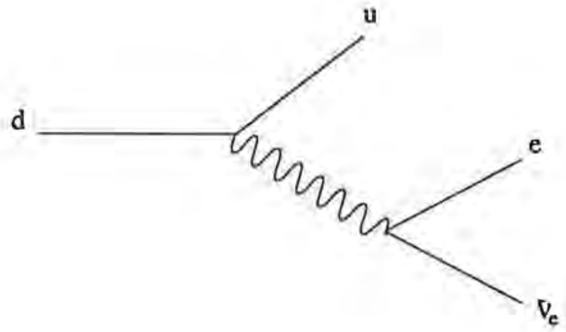
5 Quarks in the Electroweak Model

An analysis of weak interactions involving hadrons leads to a very similar structure for the quark sector of the electroweak model.

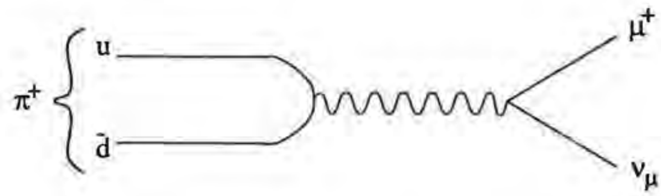
5.1 Charged weak current (quarks)

A selection of key processes includes the following:

1. β decay $\quad n \rightarrow p\, e^- \, \bar{\nu}_e \qquad$ i.e. $d \rightarrow u\, e^- \, \bar{\nu}_e$

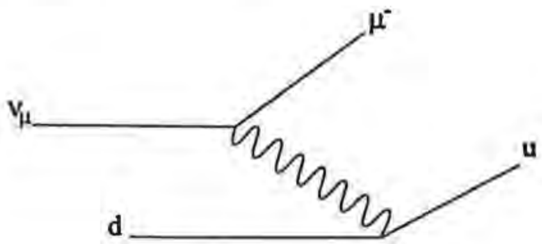


2. π decay $\quad \pi^+ \rightarrow \mu^+ \, \nu_\mu$

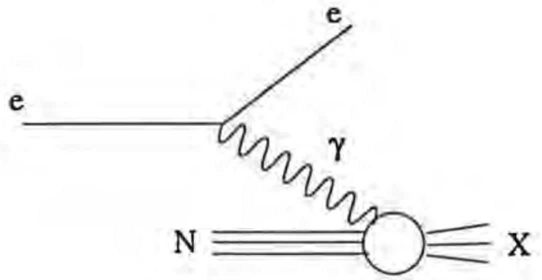


(see Halzen and Martin, sect. 12.6 for a discussion of the "hadronisation" of $u\bar{d}$ into π^+.)

3. $\nu_\mu \, N \rightarrow \mu^- \, X \qquad$ e.g. $\nu_\mu \, d \rightarrow \mu^- \, u$



This is realised in deep inelastic scattering. It is the weak interaction analogue of $e^- N \to e^- X$



All these processes can be described at low energies, $q \ll m_W$, by an effective current-current interaction, also of Lorentz structure $V - A$:

$$\mathcal{L}_{int} = \frac{4G}{\sqrt{2}} J_\mu^{CC} J^{CC\,\mu\dagger} \tag{66}$$

with $J_\mu^{CC} \simeq \bar{u_L}\,\gamma_\mu\, d_L$. $\mathcal{L}_{int}$ uses the *same* G as before. This extends electron-muon universality to lepton-quark universality.

In fact, this is too simple. Consider the next generation, with strange and charm quarks. These almost obey

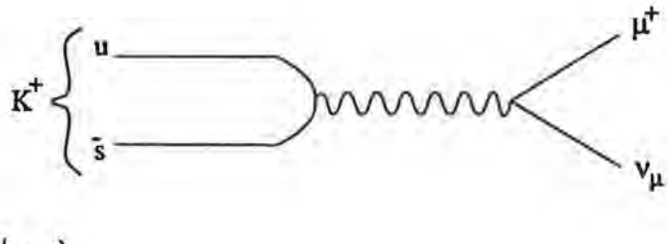
$$J_\mu^{CC} \simeq \bar{u}_L\,\gamma_\mu\, d_L + \bar{c}_L\,\gamma_\mu\, s_L \tag{67}$$

This corresponds to a structure like the leptons, with families of $SU(2)_L$ doublets

$$\begin{pmatrix} u_L \\ d_L \end{pmatrix}, \quad \begin{pmatrix} c_L \\ s_L \end{pmatrix}.$$

However, processes such as $K^+ \to \mu^+ \nu$ also occur, involving a $u - s$ transition:

$K^+ \to \mu^+\, \nu_\mu$



(cf. $\pi^+ \to \mu^+ \nu_\mu$)

To incorporate this flavour mixing, we add an ad-hoc quark mixing angle (the Cabibbo angle) and define

$$\begin{aligned} d' &= d\cos\theta_c + s\sin\theta_c \\ s' &= -d\sin\theta_c + s\cos\theta_c \end{aligned} \qquad (68)$$

so that the $SU(2)_L$ eigenstates are

$$\begin{pmatrix} u_L \\ d'_L \end{pmatrix}, \qquad \begin{pmatrix} c_L \\ s'_L \end{pmatrix}.$$

Then

$$\frac{\Gamma(K^+ \to \mu^+\nu_\mu)}{\Gamma(\pi^+ \to \mu^+\nu_\mu)} \sim \sin^2\theta_c \qquad \text{(up to kinematic factors)}$$

The Cabibbo angle is small: $\theta_c = 13^o$, $\sin\theta_c = 0.23$ (see also the discussion of the CKM matrix in sect. 5.4).

5.2 Neutral current (quarks)

Processes such as $\nu_\mu N \to \nu_\mu X$ were observed at CERN (Gargamelle) in 1973 with strength

$$\frac{\sigma(\nu N \to \nu X)}{\sigma(\nu N \to \mu X)} \sim 0.3 \qquad (69)$$

They can be described by

$$\mathcal{L}_{int} = \frac{4G}{\sqrt{2}} 2\,\rho J_\mu^{NC} J^{NC\mu} \qquad (70)$$

with

$$J_\mu^{NC} = \frac{1}{2}\bar{q}\,\gamma_\mu\,(c_V^q - c_A^q\gamma_5)\;q \qquad (71)$$

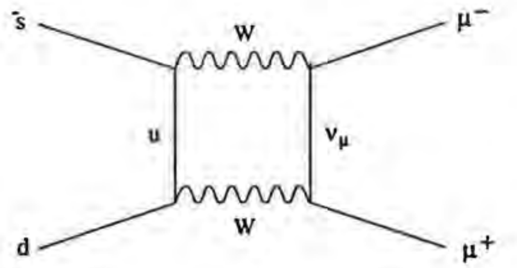
This is the same as for the leptons, except for the different c_V, c_A parameters (see sect. 5.6).

5.3 Charm and flavour changing neutral currents

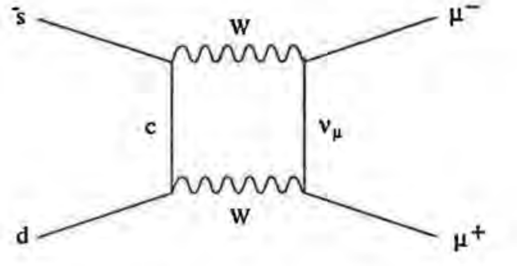
Suppose there was no c quark. With any $u \leftrightarrow d'$ transitions, we would have the flavour changing neutral current ($\Delta S = 1$) decay $K^0 \to \mu^+\mu^-$ from the top diagram overleaf.

The Cabibbo factors from the (du) (us) vertices give $\cos\theta_c\sin\theta_c$. However, experimentally, flavour changing neutral current (FCNC) decays are found to be strongly suppressed, e.g.

$$\frac{\Gamma(K_L^0 \to \mu^+\mu^-)}{\Gamma(K_L^0 \to anything)} \sim 10^{-8} \qquad (72)$$



$\sim \cos\theta_c \, \sin\theta_c$



$\sim -\cos\,\theta_c \, \sin\theta_c$

With charm, there is another diagram, shown above. The Cabibbo factors have the opposite sign, giving a cancellation. So FCNCs are strongly suppressed. This is the famous GIM (Glashow, Iliopoulos, Maiani) mechanism.

This was one of the main motivations for the proposal of charm by GIM in 1970. Another was anomalies.

Theoretical Interlude – Anomalies

It can happen that a symmetry which holds in the classical theory is no longer a good symmetry in the corresponding quantum theory. This is known as an anomaly. This phenomenon is particularly associated with chiral symmetries (i.e. involving γ_5) such as occur in the electroweak model.

As the simplest example, consider massless QED with the action

$$S = \int dx \mathcal{L} = \int dx \ \left[i\bar{\psi}(\partial_\mu - ieA_\mu)\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} \right] \tag{73}$$

This is invariant under the (global) chiral transformation,

$$\psi \to e^{i\alpha\gamma_5}\,\psi \tag{74}$$

By Noether's theorem, there is a conserved current $J_{\mu 5} = \bar{\psi}\gamma_\mu\gamma_5\psi$ corresponding to this symmetry. It satisfies the equation of motion (conservation law)

$$\partial^\mu J_{\mu 5} = 0 \tag{75}$$

Does this remain true in the quantum theory? The equivalent statement would be the chiral Ward identity for, e.g. the two-point Green function

$$< 0|T^*\partial^\mu \ J_{\mu 5}\Phi|0 > \stackrel{?}{=} \langle \delta\Phi \rangle \tag{76}$$

where $\langle\delta\Phi\rangle$ is the vacuum expectation value of the chiral variation of some arbitrary (elementary or composite) field Φ.

To compute Green functions in the quantum theory, we need the generating functional,

$$W = \int \mathcal{D}\psi\mathcal{D}\bar{\psi}\mathcal{D}A\, e^{i\int dx\mathcal{L}} \tag{77}$$

Now consider the behaviour of W under a change of integration variable, $\psi \rightarrow e^{i\alpha\gamma_5}\psi$, $\bar{\psi} \rightarrow e^{-i\alpha\gamma_5}\bar{\psi}$. Since this is only a change of variable, W does not change. So (taking $\alpha = \alpha(x)$ as a technical device), we get

$$0 = \frac{\delta W}{\delta\alpha(x)} \stackrel{?}{=} -i\int \mathcal{D}\psi\mathcal{D}\bar{\psi}\mathcal{D}A\, \partial^\mu J_{\mu5}\, e^{i\int dx\mathcal{L}} \tag{78}$$

since $\frac{\delta S}{\delta\alpha(x)} = -\partial^\mu J_{\mu5}$. Since the variation is a total derivative, the global transformation is a symmetry. This gives the naive Ward identity.

However, the integration measure $\mathcal{D}\psi\mathcal{D}\bar{\psi}$, which is the key ingredient in taking us from the classical to the quantum theory, is *not* invariant under chiral transformations. In fact,

$$\mathcal{D}\psi\mathcal{D}\bar{\psi} \rightarrow \mathcal{D}\psi\mathcal{D}\bar{\psi}e^{i\int dx\alpha(x)\frac{e^2}{16\pi^2}F_{\mu\nu}\tilde{F}^{\mu\nu}} \tag{79}$$

where $\tilde{F}^{\mu\nu} = \epsilon^{\mu\nu\alpha\beta}F_{\alpha\beta}$. The derivation of this is subtle and difficult. However, the final result for the Ward identity is simple:

$$0 = -i\int \mathcal{D}\psi\mathcal{D}\bar{\psi}\mathcal{D}A\, (\partial^\mu J_{\mu5} - \frac{e^2}{16\pi^2}F_{\mu\nu}\tilde{F}^{\mu\nu})e^{i\int dx\mathcal{L}} \tag{80}$$

that is,

$$< 0|T^*\partial^\mu J_{\mu5}\ \Phi|0> - < 0|T^*\frac{e^2}{16\pi^2}F_{\mu\nu}\tilde{F}^{\mu\nu}\ \Phi|0> = \langle\delta\Phi\rangle \tag{81}$$

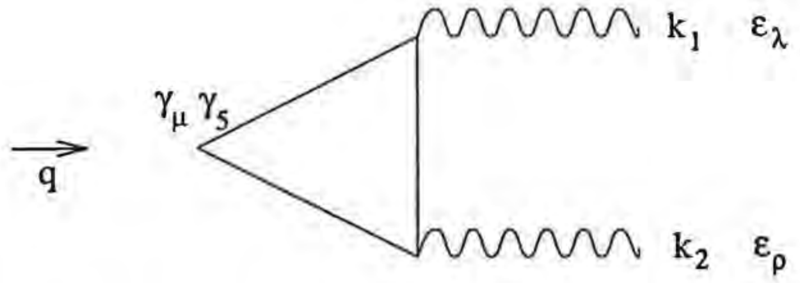
This result is exact and non-perturbative. In fact, with an appropriate choice of renormalisation for the composite operators $J_{\mu5}$ and $F_{\mu\nu}\tilde{F}^{\mu\nu}$, it holds in the same form to all orders (Adler-Bardeen theorem). This is the anomalous chiral Ward identity.

In perturbation theory, the anomaly is manifested in the 1-loop triangle diagram shown overleaf.

Naively, we expect this amplitude to satisfy $q^\mu M_{\mu\lambda\rho} \stackrel{?}{=} 0$ because of the classical current conservation. However a careful treatment of the divergent integrals involved in its calculation actually gives

$$q^\mu M_{\mu\lambda\rho} = \frac{ie^2}{4\pi^2}\epsilon_{\mu\nu\lambda\rho}k_1^\mu k_2^\nu \neq 0 \tag{82}$$

in accordance with the anomalous Ward identity.



The result for non-abelian currents is similar. In this case, the currents at the vertices of the triangle diagram (or equivalently the external gauge fields) include group generators T^a , T^b and T^c. The anomaly is then proportional to

$$\mathcal{A} = Tr\{T^a, T^b\}T^c \tag{83}$$

We have described the 'AVV' anomaly. In theories such as the electroweak model which also has axial gauge bosons there are also 'AAA' and higher-point anomalies.

The physical significance of anomalies depends entirely on whether or not the axial current is or is not coupled to gauge fields.

Global currents:

This is the case where the current is not coupled to a gauge field. Here, there is no problem. The quantum theory (anomalous Ward identity) does not look like the classical theory (conserved current), but this does not damage the consistency of the theory. In fact, the existence of these anomalies is an essential and experimentally verified part of the standard model.

For example, the anomaly is essential for the neutral pion decay $\pi^0 \to \gamma\gamma$. The pion couples to the axial current $J_{\mu 5}$ according to $< 0|J_{\mu 5}|\pi >= ik_\mu f_\pi$ where f_π is the pion decay constant, $93MeV$ (see section 7). This allows us to calculate the $\pi^0 \to \gamma\gamma$ decay amplitude from the matrix element $< 0|J_{\mu 5}|\gamma\gamma >$. The divergence of this would vanish if the naive Ward identity was true, predicting $\pi^0 \not\to \gamma\gamma$. In fact, because of the anomaly,

$$\begin{aligned} < 0|\partial^\mu J_{\mu 5}|\gamma\gamma > &= \frac{e^2}{16\pi^2}\sum_f Q_f^2 < 0|F_{\mu\nu}\tilde{F}^{\mu\nu}|\gamma\gamma > \\ &\neq 0 \end{aligned} \tag{84}$$

and this permits a non-zero decay amplitude $\pi^0 \to \gamma\gamma$ in QED and QCD.

The constant multiplying the anomaly, $\sum_f Q_f^2$, measures the sum of the squares of the charges for all the fermions which make up $J_{\mu 5}$ (i.e. which go round the loop in the triangle diagram). Initial calculations with quarks gave a result for the decay amplitude 3 times smaller than experiment. This is resolved if we take the number of colours into account. So, experiment and the anomaly explanation of $\pi^0 \to \gamma\gamma$ implies that QCD must have $N_C = 3$.

Gauged currents:

The situation is quite different if we couple a dynamical gauge field to the anomalous current (i.e. promote the anomalous symmetry to a local transformation). Here, the anomaly completely destroys the consistency of the quantum theory. The gauge symmetry is broken (since the current is not conserved) and the quantum theory is non-unitary (the unphysical ghost degrees of freedom do not decouple).

This second case is dangerous for a chiral gauge theory such as the electroweak model, since we have gauge fields coupled to the axial current. The theory will only be unitary if all the potential anomalies vanish.

Rewriting in terms of left and right-handed fields, the anomaly coefficient is proportional to

$$\mathcal{A} = \sum_{reps} Tr\Big[\{T_L^a, T_L^b\}T_L^c - \{T_R^a, T_R^b\}T_R^c\Big] \tag{85}$$

There are four possible anomalies to check in the electroweak sector:

(1) a, b, c all $SU(2)_L$ currents:-
All fermions are in doublets, so

$$\begin{aligned}\mathcal{A} &= \sum_{\chi_L^q, \chi_L^\ell} Tr\Big\{\frac{\tau^a}{2}, \frac{\tau^b}{2}\Big\}\frac{\tau^c}{2} \\ &\sim \delta^{ab}\sum TrT^c = 0\end{aligned} \tag{86}$$

since the trace of an $SU(2)$ matrix vanishes.

(2) $a = SU(2)_L$ and $b, c = U(1)_Y$:-
In this case,

$$\mathcal{A} = \sum 2Tr\frac{\tau^a}{2} Y_L^2 = 0 \tag{87}$$

for the same reasons.

(3) $a, b = SU(2)_L$ and $c = U(1)_Y$:-
Here,

$$\begin{aligned}\mathcal{A} &= \sum Tr\Big\{\frac{\tau^a}{2}, \frac{\tau^b}{2}\Big\}Y_L \\ &\sim \delta^{ab}\sum TrY_L \sim \sum_{Lreps} TrQ\end{aligned} \tag{88}$$

since for the left-handed representations $\frac{1}{2}Y_L = -T_L^3 + Q$

(4) a, b, c all $U(1)_Y$:-
Here,

$$\begin{aligned}\mathcal{A} &= 2\Big\{\sum_{Lreps} TrY_L^3 - \sum_{Rreps} TrY_R^3\Big\} \\ &\sim \sum_{Lreps} TrQ\end{aligned} \tag{89}$$

since for the right-handed representations $\frac{1}{2}Y_R = Q$.

So, the anomalies of type (1) and (2) necessarily vanish. But the anomalies for type (3) and (4) vanish if and only if

$$\sum_f Q_f = 0 \tag{90}$$

i.e. anomaly cancellation requires the sum of the electric charges of the fermions to vanish.

In the standard model, this is true individually for each generation:-

$$\begin{aligned}\sum_{f=\nu_e,e,u,d} Q_f &= 0 - 1 + N_C(\frac{2}{3} - \frac{1}{3}) = -1 + \frac{1}{3}N_C \\ &= 0 \quad \text{for} \ \ N_C = 3\end{aligned} \tag{91}$$

This theoretical analysis tells us several important things about the standard model

1. The $SU(N_C) \times SU(2)_L \times U(1)_Y$ gauge theory (QCD plus electroweak) with the known quark and lepton spectrum *must* have $N_C = 3$
2. Anomaly cancellation within each generation means that a model with two lepton generations and the quarks u, d, s does not exist. Anomaly freedom implies that charm exists!
3. 3 lepton generations implies that top exists.
4. The condition $\sum_f Q_f = 0$ as described above looks contrived. This suggests that the quarks and leptons may have originally been in some single larger representation. This hints at some form of grand unification.

5.4 Third generation and the CKM matrix

To keep the success of the GIM mechanism when we extend the standard model to three generations, we assign the quarks to $SU(2)_L$ eigenstates $\begin{pmatrix} u_{iL} \\ d'_{iL} \end{pmatrix}$, where $u_{iL} = u_L,\, c_L,\, t_L$ and $d_{iL} = d_L,\, s_L,\, b_L$, with

$$d'_{iL} = V_{ij}\, d_{jL} \tag{92}$$

If V is unitary ($V^\dagger V = 1$), then

$$\bar{d}'_i\, d'_i \;=\; \bar{d}\, V^\dagger V\, d \;=\; \bar{d}_i\, d_i \tag{93}$$

This property suppresses the FCNC diagrams discussed in sect. 5.3 and ensures the neutral current is flavour diagonal. V enters the vertices with the $W^\pm$ but not with the Z.

V is the CKM (Cabibbo-Kobayashi-Maskawa) matrix. First, note the parameter counting for an arbitrary number N of generations:

$$\begin{aligned} \text{Unitary } N \times N \text{ matrix} \quad &\rightarrow \quad N^2 \text{ parameters.} \\ \text{Orthogonal } N \times N \text{ matrix} \quad &\rightarrow \quad \frac{1}{2}N(N-1) \text{ parameters.} \end{aligned}$$

But $(2N-1)$ relative phases for the quarks are irrelevant. So V_{CKM} has $\frac{1}{2}N(N-1)$ real parameters and $N^2 - (2N-1) - \frac{1}{2}N(N-1) = \frac{1}{2}(N-1)(N-2)$ phases.

In the standard model, $N = 3$, so V_{CKM} has 3 angles and 1 phase (important for CP violation). The Kobayashi-Maskawa parametrisation is

$$\begin{aligned} V_{CKM} &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_2 & s_2 \\ 0 & -s_2 & c_2 \end{pmatrix} \begin{pmatrix} c_1 & s_1 & 0 \\ -s_1 & c_1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & e^{i\delta} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & c_3 & s_3 \\ 0 & -s_3 & c_3 \end{pmatrix} \\ &= \begin{pmatrix} c_1 & s_1 c_3 & s_1 s_3 \\ -s_1 c_2 & c_1 c_2 c_3 - s_2 s_3 e^{i\delta} & c_1 c_2 s_3 + s_2 c_3 e^{i\delta} \\ s_1 s_2 & -c_1 s_2 c_3 - c_2 s_3 e^{i\delta} & -c_1 s_2 s_3 + c_2 c_3 e^{i\delta} \end{pmatrix} \end{aligned} \tag{94}$$

where we let $c_1 = \cos\theta_1$ etc.

The current experimental values are

$$V_{CKM} = \begin{pmatrix} |V_{ud}| = 0.974 & |V_{us}| = 0.221 & |V_{ub}| = 0.003 \\ |V_{cd}| = 0.20 & |V_{cs}| = 1.1 & |V_{cb}| = 0.04 \\ |V_{td}| = 0.008 & |V_{ts}| \simeq 0 & |V_{tb}| \simeq 1 \end{pmatrix} \tag{95}$$

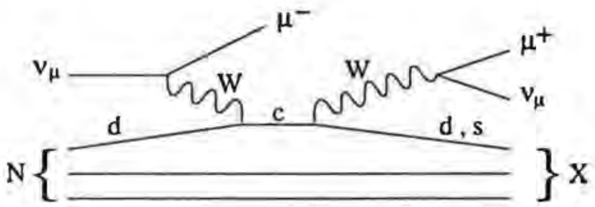
These arise from:-

$$V_{ud} \;:\; \beta\text{decay} \quad n \rightarrow p e^- \bar{\nu}_e \quad , \quad \pi^+ \rightarrow \pi^0 e^+ \nu_e$$

V_{us} : $K^+ \to \pi^0 e^+ \nu_e$, $K^0 \to \pi^- e^+ \nu_e$
semileptonic hyperon decays $\Lambda \to p e^- \bar{\nu}_e$

V_{ub} : $b \to u e^- \bar{\nu}_e$, need B decays with no K in final state

V_{cd} : $\nu_\mu d \to \mu^- e$, as in the diagram

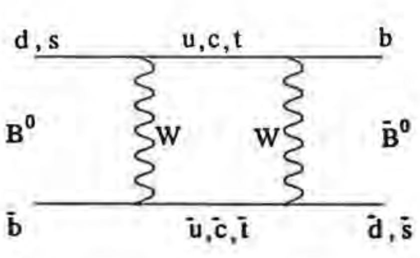


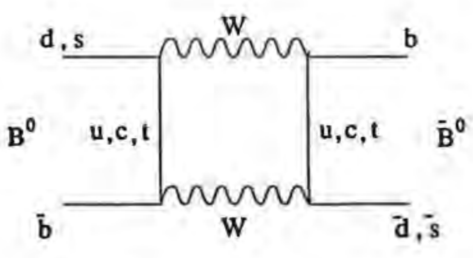
V_{cs} : $\nu_\mu s \to \mu^- c$, needs estimate of s-content of nucleon
$D^+ \to \bar{K}^0 e^+ \nu_e$

V_{cb} : $B \to D^* \ell \bar{\nu}_\ell$, plus heavy quark effective theory

δ : The phase is determined by the ϵ parameter in $K^0 - \bar{K}^0$

V_{td} : $B^0 - \bar{B}^0$ mixing, from diagrams like





Clearly a great deal of experimental work (and theoretical analysis – much depending on models and approximations for heavy quark states) is being done to determine the quark mixing parameters and verify the assumption that V_{CKM} is a unitary matrix. There is at present no accepted theory of these mixing angles – they are all free parameters in the standard model.

5.5 CP violation and the CKM matrix

Much of the interest in V_{CKM} is because it is the only source of CP violation in the standard model. We show here why the appearance of a phase in V_{CKM} leads to CP violation. Let

$$\begin{aligned} M &= (\bar{u}_k\gamma^\mu(1-\gamma_5)V_{ki}u_i)(\bar{u}_j\gamma_\mu(1-\gamma_5)V_{jl}u_l)^\dagger \\ &= V_{ki}\,V_{jl}^*(\bar{u}_k\gamma^\mu(1-\gamma_5)u_i)(\bar{u}_l\gamma_\mu(1-\gamma_5)u_j) \end{aligned} \tag{96}$$

be the charged-current induced matrix element for $q_iq_j \rightarrow q_kq_l$. u are the appropriate Dirac spinors. If we can show that the CP transformed matrix element satisfies $M_{CP} = M^\dagger$, then the theory conserves CP. Otherwise, CP is violated.

Under C,

$$\begin{aligned} u &\rightarrow u_C = C\bar{u}^T \\ \bar{u} &\rightarrow \bar{u}_C = -u^TC^{-1} \end{aligned} \tag{97}$$

where $C^{-1}\gamma_\mu C = -\gamma_\mu^T$, $C^{-1}\gamma_\mu\gamma_5 C = (\gamma_\mu\gamma_5)^T$.

Under P, $P^{-1}\gamma_\mu(1+\gamma_5)P = \gamma_\mu^\dagger(1-\gamma_5)$ where $\gamma_0^\dagger = \gamma_0$, $\gamma_i^\dagger = -\gamma_i$.
So,

$$\bar{u}_k\gamma^\mu(1-\gamma_5)V_{ki}u_i \xrightarrow{CP} -V_{ki}\bar{u}_i\gamma_\mu^\dagger(1-\gamma_5)u_k \tag{98}$$

and then,

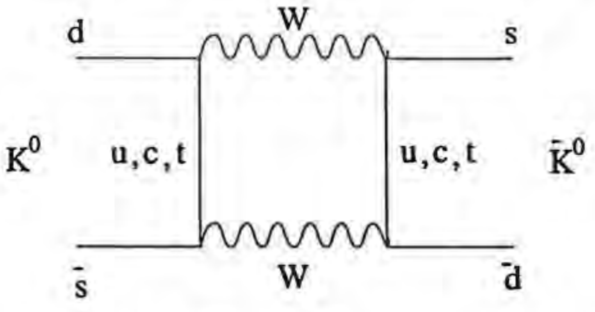
$$M_{CP} = V_{ki}\,V_{jl}^*(\bar{u}_i\gamma^\mu(1-\gamma_5)u_k)(\bar{u}_j\gamma_\mu(1-\gamma_5)u_l) \tag{99}$$

compared with

$$M^\dagger = V_{ki}^*\,V_{jl}(\bar{u}_i\gamma^\mu(1-\gamma_5)u_k)(\bar{u}_j\gamma_\mu(1-\gamma_5)u_l) \tag{100}$$

We find that $M_{CP} = M^\dagger$ provided V_{ij} are *real.*

It follows that in the three generation model where V_{CKM} has a complex parameter, CP is violated. This will show up in $K^0-\bar{K}^0$ or $B^0-\bar{B}^0$ mixing:-



In the two generation model, K^0 and $\bar{K}^0$ are linear combinations of the $CP = +1, -1$ eigenstates, $K_S = \frac{1}{\sqrt{2}}(K^0+\bar{K}^0)$ and $K_L = \frac{1}{\sqrt{2}}(K^0-\bar{K}^0)$, which decay by $K_S \rightarrow 2\pi$ and $K_L \rightarrow 3\pi$. However $K_L \rightarrow 2\pi$ does occur with a small branching ratio of $\sim 10^{-3}$.

5.6 Gauge boson-current interaction (quarks)

The same construction as for the lepton sector now goes through essentially unchanged. In the electroweak Lagrangian, the interaction of the A, W and Z bosons with the quarks is

$$\mathcal{L} = -\frac{g}{\sqrt{2}}\left(J_\mu^+ W^{+\mu} + J_\mu^- W^{-\mu}\right) - eJ_\mu^{em} A^\mu - \frac{g}{\cos\theta_W} J_\mu^{NC} Z^\mu \qquad (101)$$

where

$$\begin{aligned} J_\mu^+ &= \bar{\chi}_{iL}\gamma_\mu \tau^+ \chi_{iL} = \bar{u}_{iL}\gamma_\mu V_{ij} d_{jL} \\ J_\mu^- &= \bar{\chi}_{iL}\gamma_\mu \tau^- \chi_{iL} = \bar{d}_{iL}\gamma_\mu V_{ij}^\dagger u_{jL} \end{aligned} \qquad (102)$$

since

$$\chi_{iL} = \begin{pmatrix} u_{iL} \\ d'_{iL} \end{pmatrix} \qquad \text{with } d'_{iL} = V_{ij}\, d_{jL} \qquad (103)$$

and

$$J_\mu^3 = \bar{\chi}_{iL}\gamma_\mu \tau^3 \chi_{iL} = \bar{u}_{iL}\gamma_\mu u_{iL} - \bar{d}_{iL}\gamma_\mu d_{iL} \qquad (104)$$

the CKM matrix V dropping out due to its assumed unitarity. The electromagnetic and weak neutral currents are

$$J_\mu^{em} = \bar{q}_f \gamma^\mu Q\, q_f, \qquad f = u, c, t, d, s, b. \qquad (105)$$

and

$$\begin{aligned} J_\mu^{NC} &= J_\mu^3 - \sin^2\theta_W J_\mu^{em} \\ &= \bar{q}_f \gamma^\mu \left(\tfrac{1}{2}(1-\gamma_5)\, t^3 - \sin^2\theta_W Q\right) q_f \\ &= \bar{q}_f \gamma^\mu \tfrac{1}{2}\left(c_V^{q_f} - c_A^{q_f}\gamma_5\right) q_f \end{aligned} \qquad (106)$$

where t^3 is the eigenvalue of T_L^3 and Q is the charge.

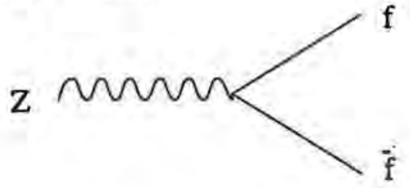
The eigenvalues t^3, Q and parameters c_V and c_A in the neutral current J_μ^{NC} are listed below ($\sin^2\theta_W \simeq 0.234$)

	t^3	Q	c_A	c_V
u, c, t	$\frac{1}{2}$	$\frac{2}{3}$	$\frac{1}{2}$	$\frac{1}{2} - \frac{4}{3}\sin^2\theta_w = 0.19$
d, s, b	$-\frac{1}{2}$	$-\frac{1}{3}$	$-\frac{1}{2}$	$-\frac{1}{2} + \frac{2}{3}\sin^2\theta_w = -0.34$
ν_e, ν_μ, ν_τ	$\frac{1}{2}$	0	$\frac{1}{2}$	$\frac{1}{2}$
e, μ, τ	$-\frac{1}{2}$	-1	$-\frac{1}{2}$	$-\frac{1}{2} + 2\sin^2\theta_w = -0.03$

The general formulae are

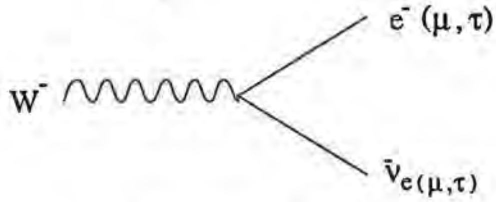
$$\begin{aligned} c_A &= t^3 \\ c_V &= t^3 - 2\sin^2\theta_W Q \end{aligned} \qquad (107)$$

This determines the $Zf\bar{f}$ vertex:-

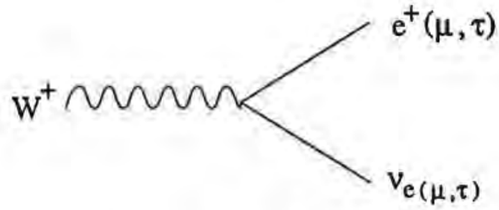


$$\frac{-ie}{\sin\theta_W \cos\theta_W}\gamma_\mu \frac{1}{2}\left(c_V^f - c_A^f \gamma_5\right)$$

along with the $W^\pm$ vertices for leptons

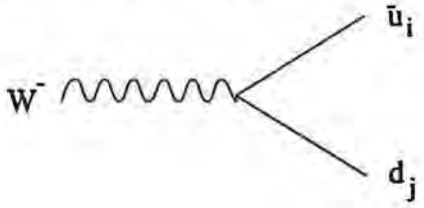


$$-\frac{i}{\sqrt{2}}\frac{e}{\sin\theta_W}\gamma^\mu \frac{1}{2}\left(1-\gamma_5\right)$$

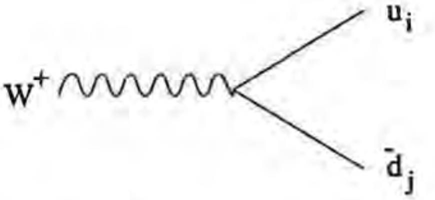


$$-\frac{i}{\sqrt{2}}\frac{e}{\sin\theta_W}\gamma^\mu \frac{1}{2}\left(1-\gamma_5\right)$$

and for quarks, including the CKM matrix,



$$-\frac{i}{\sqrt{2}}\frac{e}{\sin\theta_W}\gamma_\mu \frac{1}{2}\left(1-\gamma_5\right) V_{ij}$$



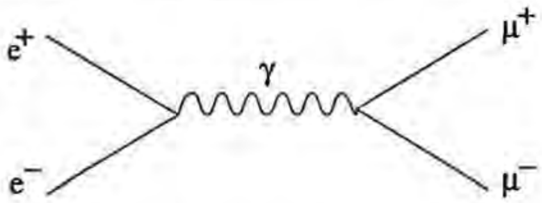
$$-\frac{i}{\sqrt{2}}\frac{e}{\sin\theta_W}\gamma_\mu\frac{1}{2}(1-\gamma_5)V^\dagger_{ji}$$

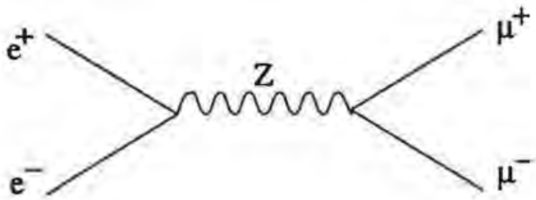
6 Electroweak Processes

In this section, we consider some simple examples of electroweak processes using the structure of the currents described above.

6.1 $\gamma - Z$ interference in $e^+ e^- \rightarrow \mu^+\mu^-$

Consider the following diagrams, which mediate electron-positron annihilation into leptons:





The amplitudes are

$$M_\gamma = -\frac{e^2}{k^2}\left(\bar{u}_\mu\gamma^\lambda u_\mu\right)\left(\bar{u}_e\gamma_\lambda u_e\right) \tag{108}$$

and

$$\begin{aligned} M_Z &= -\frac{g^2}{4\cos^2\theta_W}\left(\bar{u}_\mu\gamma^\lambda\left(c_V^\mu - c_A^\mu\gamma_5\right)u_\mu\right)\left(\frac{g_{\lambda\rho}-\frac{k_\lambda k_\rho}{m_Z^2}}{k^2-m_Z^2}\right) \\ &\qquad \times\left(\bar{u}_e\gamma^\rho\left(c_V^e - c_A^e\gamma_5\right)u_e\right) \\ &= -\frac{\sqrt{2}\,G\,m_Z^2}{s-m_Z^2}\left(c_R^\mu\bar{u}_{\mu R}\gamma^\lambda u_{\mu R} + c_L^e\bar{u}_{eL}\gamma^\lambda u_{eL}\right) \\ &\qquad \times\left(c_R^e\bar{u}_{eR}\gamma_\lambda u_{eR} + c_L^e\bar{u}_{eL}\gamma_\lambda u_{eL}\right) \end{aligned} \tag{109}$$

where $s = k^2$ and we neglect lepton masses. Recall $c_{L(R)} = c_V \pm c_A$

To calculate the cross section, we first add these amplitudes then square, i.e. $|M_\gamma + M_Z|^2$. This is electroweak interference.

The unpolarised $e^+e^- \to \mu^+\mu^-$ cross section is found by averaging over the four allowed L, R helicity combinations for e and μ. We find

$$\frac{d\sigma}{d\Omega} = \frac{\alpha^2}{4s}\left[A_0\left(1+\cos^2\theta\right) + A_1\cos\theta\right] \tag{110}$$

where

$$\begin{aligned} A_0 &= 1 + 2\Re(r)c_V^2 + |r|^2\left(c_V^2 + c_A^2\right)^2 \\ A_1 &= 4\Re(r)c_A^2 + 8|r|^2 c_V^2 c_A^2 \end{aligned} \tag{111}$$

with

$$r = \frac{s}{e^2}\frac{\sqrt{2}\,G\,m_Z^2}{s - m_Z^2 + i m_Z\Gamma_Z} \tag{112}$$

r comes from the Z propagator, modified to include the finite resonance width Γ_Z which must be included when $s \simeq m_Z^2$. In pure QED, $A_0 = 1$ and $A_1 = 0$.

This cross section is usually expressed as a forward-backward asymmetry. Define,

$$A_{FB} = \frac{F-B}{F+B}, \qquad F = \int_0^1 \frac{d\sigma}{d\Omega}\,d\Omega\,,\; B = \int_{-1}^0 \frac{d\sigma}{d\Omega}\,d\Omega \tag{113}$$

Then we have

$$A_{FB} = \frac{3}{8}\frac{A_1}{A_0}, \qquad\qquad (s \ll m_Z^2) \tag{114}$$

and

$$A_{FB} = 3\,\frac{c_V^2\,c_A^2}{(c_V^2 + c_A^2)^2} \qquad (s \simeq m_Z^2) \tag{115}$$

6.2 Z partial widths

From the $Zf\bar{f}$ vertex,

$$-i\frac{g}{\cos\theta_W}\gamma_\mu\frac{1}{2}\left(c_V^f - c_A^f\gamma_5\right) \tag{116}$$

we can calculate the decay rate,

$$\Gamma\left(Z \to f\bar{f}\right) = \frac{g^2}{48\pi\,\cos^2\theta_W}\left({c_V^f}^2 + {c_A^f}^2\right) m_Z \tag{117}$$

This enables us to compute the partial widths for the set of decays:-

$$\begin{aligned} \Gamma(Z \to \bar{\nu}_e\,\nu_e) &\equiv \Gamma_Z^0 = \frac{g^2\,m_Z}{96\pi\,\cos^2\theta_W} && = 0.17\,GeV \\ \Gamma\left(Z \to e^+e^-\right) &= \Gamma_Z^0\left(1 - 4\sin^2\theta_W + 8\sin^4\theta_W\right) && = 0.09\,GeV \\ \Gamma(Z \to \bar{u}\,u) &= 3\,\Gamma_Z^0\left(1 - \frac{8}{3}\sin^2\theta_W + \frac{32}{9}\sin^4\theta_W\right) && = 0.30\,GeV \end{aligned}$$

$$\Gamma\left(Z \to \bar{d}d\right) = 3\Gamma_Z^0\left(1 - \frac{4}{3}\sin^2\theta_W + \frac{8}{9}\sin^4\theta_W\right) = 0.39\,GeV \quad (118)$$

(The 3 in the last two expressions is the number of colours, $N_C = 3$)

The more light generations, i.e. with mass less than $m_Z/2$, the bigger the Z width. LEP measurements can therefore determine the number of light generations. The experimental value

$$\Gamma_Z\,(total) \sim 2.6GeV \quad (119)$$

confirms $N_\nu = 3$.

Cosmological Interlude – $N_\nu = 3$ from Big Bang Nucleosynthesis

As well as the LEP measurement of Γ_Z, there is good evidence for $N_\nu = 3$ from measurements of the ^{4}He abundance in the universe. This is based on primordial nucleosynthesis in the big bang model. Very roughly, the argument is as follows:-

1. Most ($\sim 90\%$) of the present day ^{4}He abundance is primordial. ^{4}He production in stars contributes $< 10\%$.

2. At high temperatures ($kT \gg 1MeV$) just after the big bang, neutrons and protons were in equilibrium through the reversible processes

$$n \to p\,e^-\bar{\nu}_e \quad (120)$$
$$n + e^+ \to p + \bar{\nu}_e \quad (121)$$
$$n + \nu_e \to p + e^- \quad (122)$$

with a neutron to proton ratio (n/p) of

$$\frac{n}{p} \sim e^{-\Delta m/kT} \quad (123)$$

where $\Delta m = m_n - m_p = 1.3MeV$.

3. When kT drops below $1MeV$ (after $t = 10sec$), the rate for $p \to n$ processes becomes much smaller than $n \to p$. When this rate falls below the expansion rate of the universe, the $p \to n$ transitions "freeze-out", fixing the n/p ratio (apart from neutron decay.)
Now, the expansion rate depends on the square root of the energy density of relativistic particles, so is greater for a larger number of light particles, viz. neutrinos with $m_\nu < 100MeV$.

So, a bigger N_ν $\Rightarrow$ faster expansion rate

$\Rightarrow$ earlier freeze-out of n/p at higher T
$\Rightarrow$ bigger freeze-out n/p ratio.

In fact, $N_\nu = 3 \Leftrightarrow n/p \sim 1/6$ at freeze-out.

4. Nucleosynthesis begins later, at around $t = 2$ mins, when the temperature is low enough for deuterons to be stable against photodisintegration. By this time, free neutron decay has reduced n/p to $1/7$.

5. Virtually all the neutrons in existence at the start of primordial nucleosynthesis end up as ^{4}He.
So, the bigger the n/p ratio the greater the abundance of ^{4}He.

The present value for ^{4}He abundance ($\approx 24\%$) rules out $N_\nu = 4$ and is consistent with $N_\nu = 3$. Further evidence comes from a detailed investigation of abundances of ^{4}He, D and ^{7}Li.

6.3 3 and 4 gauge boson vertices

Recalling that

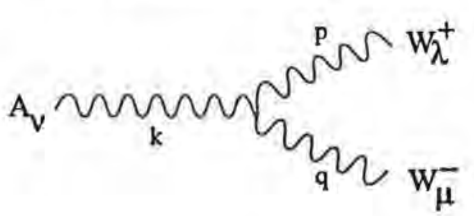
$$F^A_{\mu\nu} = \partial_\mu W^A_\nu - \partial_\nu W^A_\mu - g\,\epsilon^{ABC} W^B_\mu W^C_\nu \tag{124}$$

we see from the field strength terms in the Lagrangian

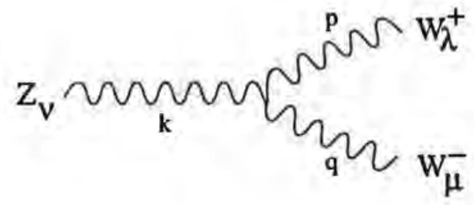
$$\mathcal{L}_{gauge} = \int dx \left[-\frac{1}{4} F^A_{\mu\nu} F^{A\mu\nu} - \frac{1}{4} F_{\mu\nu} F^{\mu\nu} \right], \tag{125}$$

where the first term corresponds to $SU(2)_L$ and the second to $U(1)_Y$, that there will be vertices with 3 and 4 gauge field propagators.

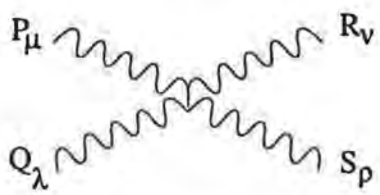
In terms of the A, $W^\pm$ and Z fields, these are:



$$e\left[(k-q)_\lambda g_{\mu\nu} + (q-p)_\nu g_{\lambda\mu} + (p-k)_\mu g_{\lambda\nu}\right]$$



$$g\,\cos\theta_W\,[(k-q)_\lambda\, g_{\mu\nu} \;+\; (q-p)_\nu\, g_{\lambda\mu} \;+\; (p-k)_\mu\, g_{\lambda\nu}]$$



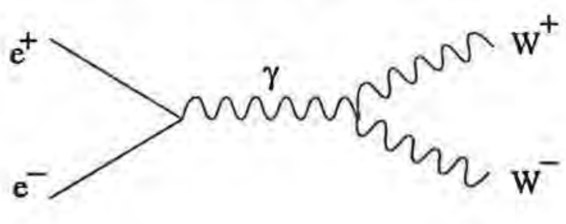
$$K_{PQRS}\,[2\,g_{\mu\nu}g\lambda\rho \;-\; g_{\mu\lambda}g\nu\rho \;-\; g_{\mu\rho}g\nu\lambda]$$

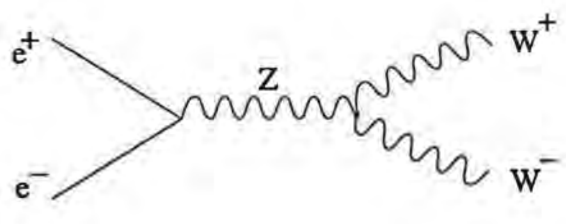
where for the different possible vertices:-

P	Q	R	S	K_{PQRS}
W^+	W^-	W^+	W^-	ig^2
A	W^+	A	W^-	$-ie^2$
Z	W^+	Z	W^-	$-ig^2\cos^2\theta_w$
A	W^+	Z	W^-	$-ie\,g\cos\theta_w$

(recall $e = g\sin\theta_W$)

At LEP 200, with e^+e^- collisions at $100 + 100GeV$, it will soon be possible to pair produce W^+W^- through the diagrams:





This will provide the first direct measurement of the 3 gauge boson coupling.

7 Spontaneous Symmetry Breaking I – Global Symmetries and Goldstone's Theorem

7.1 A global $U(1)$ model

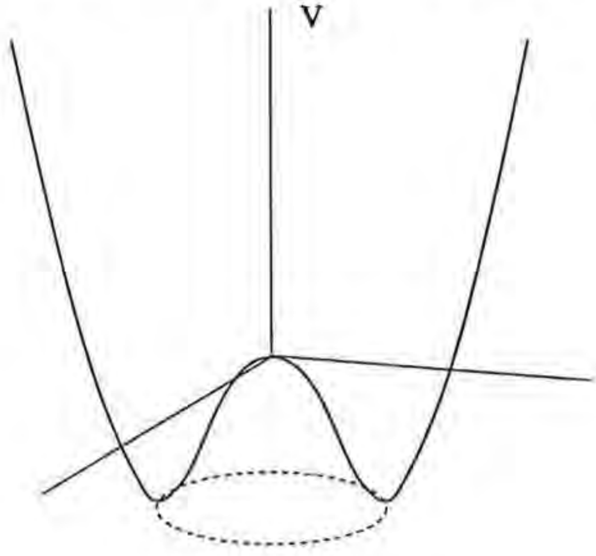
As a toy model, consider a complex scalar field with Lagrangian

$$\mathcal{L} = \partial_\mu \phi \, \partial^\mu \phi^* - V(\phi, \phi^*) \tag{126}$$

with

$$V(\phi, \phi^*) = -\mu^2 \phi^* \phi + \lambda (\phi^* \phi)^2 \tag{127}$$

We have chosen the opposite sign from usual for the quadratic term.



Plot of V over the complex ϕ plane.

Rewriting ϕ in modulus-phase form, $\phi = \frac{1}{\sqrt{2}} \rho \, e^{i\chi}$, the Lagrangian is

$$\mathcal{L} = \frac{1}{2}(\partial_\mu \rho)^2 + \frac{1}{2}\rho^2 (\partial_\mu \chi)^2 - V(\rho) \tag{128}$$

where

$$V(\rho) = -\frac{1}{2}\mu^2 \rho^2 + \frac{\lambda}{4}\rho^4 \tag{129}$$

This theory has a global $U(1)$ symmetry, i.e. invariance under $\phi \to \phi \, e^{i\alpha}$, $\alpha =$ constant, i.e. $\chi \to \chi + \alpha$. This is reflected in the potential, which depends only on ρ.

With the sign of μ^2 chosen, the minimum of $V(\rho)$ is not $\rho = 0$, but at the bottom of the rim. The minimum is not unique – there is a family of degenerate minima connected by $U(1)$ transformations.

Select one of these equivalent minima, say $\chi = 0$, $\rho = v$, then write $\rho = v + H$. In the quantum theory v is the vacuum expectation value (VEV) of ϕ, i.e.

$$\langle 0|\phi|0\rangle = v \neq 0. \tag{130}$$

Then,

$$\mathcal{L} = \frac{1}{2}(\partial_\mu H)^2 + \frac{1}{2}v^2(\partial_\mu \chi)^2 + \left(v\,H + \tfrac{1}{2}H^2\right)(\partial_\mu \chi)^2 - V(H) \tag{131}$$

where

$$V(H) = -\frac{1}{4}\lambda v^4 + \lambda v^2 H^2 + \lambda v H^3 + \frac{\lambda}{4}H^4 \tag{132}$$

At the minimum, $v^2 = \mu^2/\lambda$.

In perturbation theory about this minimum, the H field describes a massive scalar particle with $m_H^2 = 2\lambda v^2$.

Rewriting in terms of λ and m_H, and rescaling $\tilde{\chi} = v\chi$ so that, as usual for a scalar field, $\tilde{\chi}$ has dimension 1, the Lagrangian is

$$\begin{aligned}\mathcal{L} = {} & \tfrac{1}{2}(\partial_\mu \tilde{\chi})^2 + \tfrac{1}{2}(\partial_\mu H)^2 - \tfrac{1}{2}m_H^2 H^2 + \tfrac{\sqrt{2\lambda}}{m_H}H(\partial_\mu \tilde{\chi})^2 \\ & + \tfrac{\lambda}{m_H^2}H^2(\partial_\mu \tilde{\chi})^2 - \sqrt{\tfrac{\lambda}{2}}\,m_H H^3 - \tfrac{\lambda}{4}H^4 + \tfrac{1}{16}\tfrac{1}{\lambda}m_H^4 \end{aligned} \tag{133}$$

We can read off the spectrum of the quantum theory from $\mathcal{L}$. The theory has one massive scalar H – this corresponds to fluctuations up the side of the walls in the potential. Crucially, it also has one *massless* scalar $\tilde{\chi}$, corresponding to fluctuations around the circle of degenerate minima. This is known as a Goldstone boson.

There are also interaction terms, and a constant non-zero vacuum energy density. (This could be a problem if we think of including gravity in a theory with SSB.)

7.2 Goldstone's theorem

This model illustrates a general theorem. We say that a symmetry is spontaneously broken if the vacuum is not invariant under the symmetry, i.e. if a field which varies under the symmetry acquires a VEV. This field is said to be an "order parameter" in the language of statistical mechanics.

In the model above, $\mathcal{L}$ is invariant under $U(1)$, but the vacuum state has no residual invariance. $U(1)$ is broken to the identity.

In general, $\mathcal{L}$ will have a symmetry G and the vacuum will have a residual invariance under a subgroup G_0. We say the symmetry is broken from G to G_0. In that case, the space of degenerate minima is the coset manifold G/G_0.

Goldstone's theorem:

This states that corresponding to each broken generator of G (i.e. a generator in G which is not in G_0) there is a massless scalar boson in the spectrum.

The corresponding scalar field $\chi(x)$ takes values in the coset manifold G/G_0.

Proof:

We give a general, non-perturbative proof in quantum theory. Corresponding to each symmetry generator in G there is a conserved current. The Ward Identity is

$$\langle 0|T^* \partial^\mu J_\mu^a \ \Phi|0\rangle = \langle 0|\delta^a \Phi|0\rangle \tag{134}$$

where $\delta^a\Phi$ is the variation of Φ under the generator T^a of the group G.

The VEV is equal to zero for the unbroken generators, i.e. T^a in G_0. But for the broken generators, i.e. T^a in G but not in G_0, we have

$$k^\mu \langle 0|J_\mu^a(k) \ \Phi(-k)|0\rangle \neq 0 \tag{135}$$

writing the Green function in momentum space. This is true for all momenta, in particular $k_\mu = 0$.

The only way this can be true is if there exists a massless state $|\chi\rangle$ in the spectrum coupling to the broken current. Then

$$\begin{aligned} k^\mu \langle 0|J_\mu^a|\chi\rangle \ \Delta_{\chi\chi} \ \langle \chi|\Phi|0\rangle &= k^\mu i k_\mu F_\chi \ \frac{1}{k^2} \langle \chi|\Phi|0\rangle \\ &\neq 0. \end{aligned} \tag{136}$$

where $\Delta_{\chi\chi}$ is the χ propagator $\sim \frac{1}{k^2}$ and F_χ is the decay constant. Clearly there is one massless χ state for each broken current.

7.3 Chiral symmetry breaking in QCD

An important example of global spontaneous symmetry breaking occurs in QCD. Consider QCD with just two flavours u and d and neglect their masses. Since QCD is independent of flavour, there is a rotation symmetry between u and d. Also, since parity is conserved for massless quarks, we can rotate the left and right handed fields separately. So, massless QCD has a global symmetry $SU(2)_L \times SU(2)_R$.

This is spontaneously broken to the $SU(2)_V$ subgroup (the axial generators are all broken) by the appearance of a VEV $\langle 0|\bar{u}u + \bar{d}d|0\rangle$, also called a "condensate".

Since we have SSB with $G = SU(2)_L \times SU(2)_R$ and $G_0 = SU(2)_V$, Goldstone's theorem says there are 3 massless pseudoscalar bosons (since $3 = \dim G/G_0$).

These are the pions, π^+, π^-, π^0, which would be exactly massless in QCD with $m_u = m_d = 0$.

8 Spontaneous Symmetry Breaking II – Gauged Symmetries and the Higgs Mechanism

8.1 A local $U(1)$ model

Now go back to the toy model of section 7.1 and make the $U(1)$ into a local (gauge) symmetry. The Lagrangian is

$$\mathcal{L} = (D_\mu\phi)^*(D^\mu\phi) - V(\phi,\phi^*) - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} \tag{137}$$

where $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ and $D_\mu\phi = (\partial_\mu - ieA_\mu)\phi$. The potential is the same, with a non-zero vacuum expectation value for ϕ. Making the substitution $\phi = \frac{1}{\sqrt{2}}(v+H)e^{i\chi}$ we have

$$\mathcal{L} = \frac{1}{2}(\partial_\mu H)^2 + \frac{1}{2}(v+H)^2(\partial_\mu\chi - eA_\mu)^2 - V(H) - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} \tag{138}$$

with $V(H)$ as before. Now write

$$W_\mu = A_\mu - \frac{1}{e}\partial_\mu\chi \tag{139}$$

Since this is a gauge transformation, $F_{\mu\nu} = \partial_\mu W_\nu - \partial_\nu W_\mu$, independent of χ. This leaves

$$\mathcal{L} = \frac{1}{2}(\partial_\mu H)^2 + \frac{1}{2}e^2v^2W_\mu W^\mu + e^2(vH + \frac{1}{2}H^2)W_\mu W^\mu - V(H) - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} \tag{140}$$

In this form we can read off the particle content we expect the quantum theory to have:

The χ field has disappeared! So there are no massless scalar bosons.

The W_μ field is massive, with $m_W = e^2v^2$. It therefore has 3 degrees of freedom (two inherited from A and one from χ).

So, starting from a theory with a $U(1)$ gauge symmetry, we find that the spectrum in the SSB phase has a massive gauge boson. This is the Higgs Mechanism.

Going back to $\mathcal{L}$, in its original form it had 3 parameters, e, λ, μ. The final theory has a massive W and a massive H, so we can write $\mathcal{L}$ in terms of e, m_H, m_W.
(Check: $v^2 = \frac{\mu^2}{\lambda}$, then $m_H^2 = 2\lambda v^2$ and $m_W^2 = e^2v^2$).
This gives

$$\begin{aligned}
\mathcal{L} &= -\frac{1}{4}(\partial_\mu W_\nu - \partial_\nu W_\mu)^2 + \frac{1}{2}m_W^2 W_\mu W^\mu + \frac{1}{2}(\partial_\mu H)^2 \\
&- \frac{1}{2}m_H^2 H^2 + em_W H W_\mu W^\mu + \frac{1}{2}e^2H^2W_\mu W^\mu \\
&- \frac{1}{2}e\frac{m_H^2}{m_W}H^3 - \frac{1}{8}e^2\frac{m_H^2}{m_W^2}H^4 + \frac{1}{8e^2}m_H^2 m_W^2
\end{aligned} \tag{141}$$

8.2 Quantisation and renormalisation

Notice that the above description of the Higgs Mechanism was entirely at the classical level. Strictly speaking, it is no more than a plausibility argument as to what we expect in the full quantum theory.

Remember that to quantise a gauge theory, we have to start with the functional integral, introduce a gauge-fixing term, and construct the Faddeev-Popov ghosts. To obtain the physical spectrum, we have to prove that these ghosts decouple along with the unphysical components of the gauge field.

All this has to be re-done in a theory with SSB. It works and the spectrum is as described above.

Gauge invariance is essential to the renormalisation of the theory. We have to prove ('t Hooft, 1971) that SSB does not spoil renormalisation, despite the appearance of gauge boson masses.

The beauty of the Higgs Mechanism is that this is true – gauge theories with spontaneous symmetry breaking *are* renormalisable.

9 The Higgs Mechanism and Mass Generation in the $SU(2)_L \times U(1)_Y$ Model

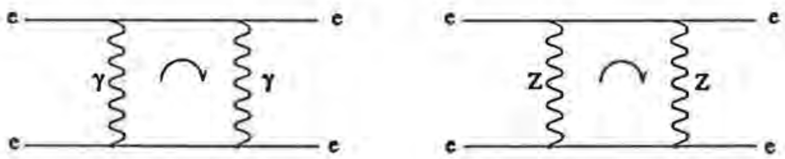
9.1 Mass generation

In the form we have presented so far, the $SU(2)_L \times U(1)_Y$ electroweak model has no mass for either the gauge bosons or fermions.

Gauge bosons:

Mass terms for $W^\pm$ or Z simply added to $\mathcal{L}$ violate the gauge symmetry. (This is also true for any gauge theory.) But the gauge symmetry is necessary for the theory to be renormalisable, and therefore predictive.

For example, consider $e^- e^-$ scattering at 1-loop. The Feynman diagrams include



The loop gives

$$\int d^4q \Delta_e \Delta_e \Delta_\gamma \Delta_\gamma \overset{q\to\infty}{\sim} \int d^4q \frac{1}{q}\frac{1}{q}\frac{1}{q^2}\frac{1}{q^2} \sim \int d^4q \frac{1}{q^6} = convergent$$

for the photon diagram, since the photon propagator (in Feynman gauge) is $\Delta_\gamma = -\frac{i}{q^2} g_{\mu\nu}$.
On the other hand, the Z propagator is $\Delta_Z = \frac{i}{q^2 - m_Z^2}(-g_{\mu\nu} + \frac{q_\mu q_\nu}{m_Z^2})$ so the loop gives

$$\int d^4 q \Delta_e \Delta_e \Delta_Z \Delta_Z \stackrel{q\to\infty}{\sim} \int d^4 q \frac{1}{q^2} = divergent$$

The divergence has to be cancelled by a counterterm

$$\mathcal{L}_{counterterm} = (div) \frac{1}{m_Z^2} \int dx\, e\, e\, e\, e$$

which is a four-Fermi interaction (dim = 6). But this introduces a new parameter. The process continues and an infinite set of higher dimension operators are induced. The theory is non-renormalisable.

We therefore need a dynamical mechanism to generate vector boson masses while keeping gauge invariance. This is achieved by the Higgs mechanism.

Fermions:

In general, we can add fermion mass terms to the Lagrangian in a gauge theory. For example, in QCD we can add quark masses, $\mathcal{L}_{mass} = \int dx\, m(\bar{q}_R q_L + \bar{q}_L q_R)$. $\mathcal{L}_{mass}$ is gauge invariant.

However, in $SU(2)_L \times U(1)_Y$, because $SU(2)_L$ is a chiral gauge theory (the group acts only on the left handed fields) fermion masses violate the gauge symmetry. For example $\mathcal{L}_{mass} = \int dx\, m(\bar{e}_R e_L + \bar{e}_L e_R)$ is *not* invariant under an $SU(2)_L$ transformation.

We therefore need a mechanism to generate fermion masses dynamically in the standard model. Remarkably, the Higgs fields can also achieve this, through Yukawa couplings.

9.2 Higgs mechanism in $SU(2)_L \times U(1)_Y$.

We need to repeat the analysis of the $U(1)$ Higgs mechanism described earlier, generalised to a non-abelian theory. The aim is to find a Higgs sector which will break $SU(2)_L \times U(1)_Y$ to $U(1)_{em}$.

The simplest choice (Weinberg and Salam, 1967) is to take

$$D_\mu \phi = (\partial_\mu + igT^A W_\mu^A + \frac{ig'}{2} Y B_\mu)\phi \tag{142}$$

where ϕ is a complex $SU(2)_L$ doublet with $Y = 1$. Then

$$\mathcal{L}_{Higgs} = (D_\mu \phi)^\dagger (D_\mu \phi) - V(\phi, \phi^\dagger) \tag{143}$$

with the potential $V(\phi,\phi^\dagger) = -\mu^2\phi^\dagger\phi + \lambda(\phi^\dagger\phi)^2$. So, remembering the relation $Q = T_L^3 + \frac{1}{2}Y$, the charge assignment is

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} \tag{144}$$

If charge conservation is to remain unbroken, only the ϕ^0 should get a vacuum expectation value. This motivates rewriting

$$\phi = \frac{1}{\sqrt{2}} e^{i\frac{1}{v}T^A\chi^A} \begin{pmatrix} 0 \\ v+H \end{pmatrix} \tag{145}$$

where χ^A, H are real fields. The potential is $V(H) = -\frac{1}{2}\mu^2(v+H)^2 + \frac{\lambda}{4}(v+H)^4$ and we have chosen $v^2 = \frac{\mu^2}{\lambda}$ to give the minimum at $H=0$.

Now substitute ϕ into $\mathcal{L}_{Higgs}$. We have

$$\phi = \frac{1}{\sqrt{2}} U \begin{pmatrix} 0 \\ v+H \end{pmatrix} \tag{146}$$

where U is an $SU(2)_L$ gauge transformation. But since $\mathcal{L}$ is gauge invariant, it will not depend on U, which can be absorbed into a trivial redefinition of the gauge fields, just as in section 8.1 for the $U(1)$ transformation $U = e^{i\chi}$.

In this so-called 'unitary gauge' the Lagrangian has the form

$$\mathcal{L}_{Higgs} = (D_\mu\phi)^\dagger(D_\mu\phi) - V(\phi,\phi^\dagger) \tag{147}$$

with

$$\phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v+H \end{pmatrix} \tag{148}$$

The Goldstone boson fields χ^A, which parametrise the space of flat directions in the potential, disappear from the spectrum. The vacuum expectation value for the scalar fields is $\langle\phi\rangle = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ v \end{pmatrix}$.

This is not invariant under $SU(2)_L$ transformations or $U(1)_Y$, since we have assigned $Y=1$ to ϕ. However it *is* invariant under $U(1)_{em}$ transformations, since

$$\begin{aligned} Q\langle\phi\rangle &= (T^3 + \frac{1}{2}Y)\langle\phi\rangle = \frac{1}{\sqrt{2}}\left(\frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} + \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right)\begin{pmatrix} 0 \\ v \end{pmatrix} \\ &= \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 0 \\ v \end{pmatrix} = 0 \end{aligned} \tag{149}$$

So $G = SU(2)_L \times U(1)_Y$ is spontaneously broken to $G_0 = U(1)_{em}$. There are dim $G/G_0 = 3$ broken generators, which implies 3 Goldstone boson χ^A. These are absorbed by the vector bosons $W^{\pm}, Z$, which acquire masses. The remaining vector boson, the photon, is still massless because $U(1)_{em}$ is unbroken.

There is one massive neutral scalar left in the physical spectrum - the Higgs boson H.

To find the masses and couplings, we expand out $\mathcal{L}_{Higgs}$. In the unitary gauge, where

$$\phi = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v+H \end{pmatrix}$$

we have

$$\begin{aligned} D_\mu \phi &= \begin{pmatrix} \partial_\mu + \frac{ig}{2} W^3_\mu + \frac{ig'}{2} B_\mu & \frac{ig}{2}(W^1_\mu - iW^2_\mu) \\ \frac{ig}{2}(W^1_\mu + iW^2_\mu) & \partial_\mu - \frac{ig}{2} W^3_\mu + \frac{ig'}{2} B_\mu \end{pmatrix} \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v+H \end{pmatrix} \\ &= \frac{1}{\sqrt{2}} \begin{pmatrix} \frac{ig}{2}(W^1_\mu - iW^2_\mu)(v+H) \\ \partial_\mu H + (-\frac{ig}{2} W^3_\mu + \frac{ig'}{2} B_\mu)(v+H) \end{pmatrix} \\ &= \frac{1}{\sqrt{2}} \begin{pmatrix} \frac{ig}{2} W^+_\mu (v+H) \\ \partial_\mu H - \frac{i}{2}(g\cos\theta_W + g'\sin\theta_W) Z_\mu (v+H) \end{pmatrix} \end{aligned} \tag{150}$$

in terms of W^+, W^-, Z. Notice that, as expected, the photon field A_μ does not appear.

Substituting into $\mathcal{L}_{Higgs}$, we find the vector boson masses

$$m_W^2 = \frac{1}{4} g^2 v^2 \tag{151}$$

$$m_Z^2 = \frac{1}{4}(g^2 + g'^2)v^2 = \frac{1}{4}\frac{g^2}{\cos^2\theta_W} m_W^2 \tag{152}$$

$$m_H^2 = 2\lambda v^2 \tag{153}$$

We therefore predict the ρ parameter, originally introduced as the relative strength of the neutral and charged current interactions:-

$$\rho = \frac{m_W^2}{m_Z^2 \cos^2\theta_W} = 1 \tag{154}$$

in the Weinberg-Salam-Higgs model. This is a special property of the particular representation of Higgs field we have chosen to induce the breaking of $SU(2)_L \times U(1)_Y$. Other choices are possible – not all give $\rho = 1$ however.

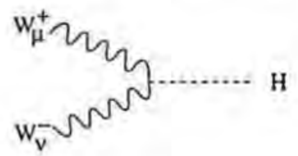
The deeper reason is that $\rho = 1$ is a prediction of a global $SU(2)$ ('custodial') symmetry implicit in $\mathcal{L}_{Higgs}$. Writing ϕ in terms of real components, $\phi^+ = \phi^1 + i\phi^2$ and $\phi^0 = \phi^3 + i\phi^4$ gives $V(\phi^\dagger\phi) = V(\phi_1^2 + \phi_2^2 + \phi_3^2 + \phi_4^2)$. The potential has an O(4) symmetry, broken to O(3) by the vacuum expectation value. In fact, the Higgs sector is a linear sigma model with coset manifold $O(4)/O(3) \sim SU(2) \times$

$SU(2)/SU(2)$. The unbroken custodial global SU(2) ensures the mass relation between m_W, m_Z that gives $\rho = 1$.

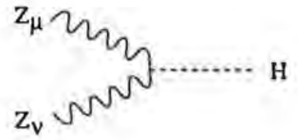
Finally, rewriting $\mathcal{L}_{Higgs}$ in terms of the parameters m_W^2, m_Z^2, m_H^2, g rather than the original set μ, λ, g', g we have

$$\begin{aligned}\mathcal{L}_{Higgs} &= \frac{1}{2}(\partial_\mu H)^2 - \frac{1}{2}m_H^2 H^2 + m_W^2 W_\mu^+ W^{-\mu} + \frac{1}{2}m_Z^2 Z_\mu Z^\mu \\ &+ g m_W H W_\mu^+ W^{-\mu} + \frac{1}{4}g^2 H^2 W_\mu^+ W^{-\mu} + \frac{1}{4}g\frac{m_Z^2}{m_W} H Z_\mu Z^\mu \\ &+ \frac{1}{8}g^2\frac{m_Z^2}{m_W^2}H^2 Z_\mu Z^\mu - \frac{1}{4}\frac{g}{\sqrt{2}}\frac{m_H^2}{m_W}H^3 - \frac{1}{32}g^2\frac{m_H^2}{m_W^2}H^4 + \frac{1}{2g^2}m_H^2 m_W^2 \end{aligned} \tag{1.55}$$

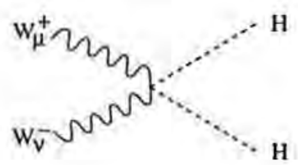
The corresponding Feynman rules for the vertices are shown in the figures.



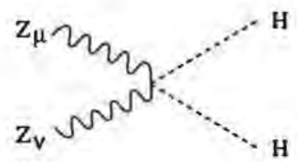
$i g m_W g_{\mu\nu}$



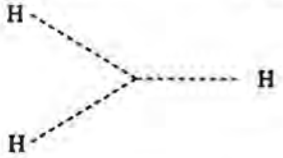
$i g \frac{m_Z^2}{m_W} g_{\mu\nu}$



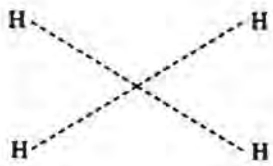
$\frac{i}{2} g^2 g_{\mu\nu}$



$\frac{i}{2} g^2 \frac{m_Z^2}{m_H^2} g_{\mu\nu}$



$-i\frac{3}{4} g^2 \frac{m_H^2}{m_W^2}$



$-i\frac{3}{2} g \frac{m_H^2}{m_W}$

9.3 Fermion masses

Yukawa interactions involving 2 fermion fields and the Higgs field can be constructed in such a way as to be $SU(2)_L \times U(1)_Y$ invariant, and so preserve renormalisability. When the Higgs field gets a vacuum expectation value, the interaction terms give rise to mass terms for the fermions.

Leptons:

Choose the $SU(2)_L \times U(1)_Y$ invariant Yukawa terms,

$$\mathcal{L}_{Yukawa} = -G_e \left[(\bar{\nu}_{eL} \;\; \bar{e}_L) \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} e_R + \bar{e}_R (\phi^- \;\; \phi^{0*}) \begin{pmatrix} \nu_{eL} \\ e_L \end{pmatrix} \right] + e \to \mu + e \to \tau \tag{156}$$

Setting

$$\begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + H \end{pmatrix} \tag{157}$$

as above, we have

$$\begin{aligned} \mathcal{L}_{Yukawa} &= -\frac{G_e}{\sqrt{2}} v (\bar{e}_L e_R + \bar{e}_R e_L) - \frac{G_e}{\sqrt{2}} (\bar{e}_L e_R + \bar{e}_R e_L) H + \mu, \tau \text{ terms} \\ &= -m_e \bar{e} e - \frac{g}{2} \frac{m_e}{m_W} \bar{e} e H + \mu, \tau \text{ terms} \end{aligned} \tag{158}$$

The Yukawa coupling (another free parameter) is traded for the lepton mass. There is also a lepton-Higgs boson vertex, proportional to m_{lepton}/m_W. This is a general feature of the model – the Higgs boson couples to particles with a strength proportional to their mass.

Quarks:

This is slightly trickier because we have to arrange masses for the upper components of the $SU(2)_L$ doublets as well.
Define

$$\phi_C = -i\tau_2 \phi^* = \begin{pmatrix} -\phi^{0*} \\ \phi^- \end{pmatrix} \to \frac{1}{\sqrt{2}} \begin{pmatrix} v + H \\ 0 \end{pmatrix} \tag{159}$$

in the unitary gauge.

Allowing for quark mixing, remembering that the $SU(2)_L$ eigenstates are $\begin{pmatrix} u_{iL} \\ d'_{iL} \end{pmatrix}$ with $d'_i = V^{CKM}_{ij} d_j$, we can write

$$\mathcal{L}_{Yukawa} = -(\bar{u}_{iL} \;\; \bar{d}'_{iL}) \phi_C \, G^u_{ij} \, u_{jR} - (\bar{u}_{iL} \;\; \bar{d}'_{iL}) \phi \, G^d_{ij} \, d_{jR} + h.c.$$

$$= -\frac{1}{\sqrt{2}}\bar{u}_{iL}G^u_{ij}u_{jR}(v+H) - \frac{1}{\sqrt{2}}\bar{d}'_{iL}V^\dagger_{ik}G^d_{kj}d_{jR}(v+H) + h.c.$$

$$= -m_u^{(i)}\,\bar{u}_i\,u_i - \frac{g}{2}\frac{m_u^{(i)}}{m_W}\bar{u}_i\,u_iH - m_d^{(i)}\,\bar{d}_i\,d_i - \frac{g}{2}\frac{m_d^{(i)}}{m_W}\bar{d}_i\,d_iH \qquad (160)$$

where we have chosen G^u_{ij} to be diagonal and G^d_{kj} such that $V^\dagger G$ is diagonal.

10 The Standard Model Lagrangian

This completes the construction of the standard model Lagrangian. The standard model is the $SU(3)_C \times SU(2)_L \times U(1)_Y$ gauge theory with quarks, leptons and the Higgs field, with Lagrangian:-

$$\mathcal{L}_{SM} = \underset{SU(2)_L}{-\frac{1}{4}F^A_{\mu\nu}F^{A\mu\nu}} \underset{U(1)_Y}{- \frac{1}{4}F_{\mu\nu}F^{\mu\nu}} \underset{SU(3)_C}{- \frac{1}{4}G^a_{\mu\nu}G^{a\mu\nu}}$$

$$+ \; i\bar{\chi}_L(\partial_\mu + ig\frac{\tau^A}{2}W^A_\mu + i\frac{g'}{2}YB_\mu)\chi_L$$

$SU(2)_L$ and $U(1)_Y$ fermion-gauge interaction on L fields

$$+ \; i\bar{\psi}_R(\partial_\mu + i\frac{g'}{2}YB_\mu)\psi_R$$

$U(1)_Y$ interaction on R fields

$$+ \; i\bar{q}(\partial_\mu + ig_{QCD}\lambda^a G^a_\mu)q$$

$SU(3)_C$ interaction on quarks.

$\lambda^a = SU(3)_C$ triplet representation generators.

G^a_μ = gluon field. q = colour triplet quarks.

$$+ \; |(\partial_\mu + ig\frac{\tau^A}{2}W^A_\mu + i\frac{g'}{2}YB_\mu)\phi|^2 \; - V(\phi^\dagger\phi)$$

Higgs sector $\Rightarrow$ W, Z masses and H interactions.

$\phi = SU(2)_L$ doublet Higgs field ($Y = 1$) V is the Higgs potential

$$- \quad (G^d \bar{\chi}_L \phi d_R + G^u \bar{\chi}_L \phi_C u_R \ + h.c.)$$

Yukawa interactions $\Rightarrow$ fermion mass

With the construction of this Lagrangian, our task in these lectures comes to a close. This is, however, more of a beginning than an end.

Many questions immediately arise. Going beyond the tree level dynamics and symmetries we used to guide us to the Lagrangian, what does the standard model actually predict and is it true? Here, the evidence for the model is strong and compelling. Perturbative radiative corrections to the tree-level predictions are impressively verified in precision electroweak experiments at LEP, and perturbative QCD, exploiting the power of the renormalisation group, is well established. Non-perturbative phenomena are much harder, but lattice gauge theories and other approaches are beginning to make serious inroads into the physics of QCD bound states. Beyond that, there are predictions, in general yet to be tested, concerning the role played by extended objects such as instantons, monopoles and strings which are implicit in the model.

The least tested and most controversial aspect of the standard model is of course the symmetry breaking, Higgs sector. Here, even the confrontation of the model with precision electroweak data provides little more than circumstantial evidence in favour of the precise mechanism presented here. Experimental confirmation of the Higgs mechanism, or indeed an alternative dynamical symmetry breaking scheme, will probably have to await the LHC.

Finally, we are led to the big questions. Assuming the standard model to be true, why is it the way it is? What determines the symmetries and the representations in which the elementary quark and lepton fields lie? What determines the parameters, nineteen in all? Aesthetic criteria, often so successful in fundamental physics, tempt us to the view that the standard model is just the low energy effective theory of a deeper, more unified theory of the fundamental interactions. But that would be another lecture course.

Acknowledgements

And finally, it is a great pleasure to thank everyone responsible for making lecturing at the School such an enjoyable and rewarding experience. In particular, I would like to thank Ken Peach for keeping the show rolling with such good humour, Ann Roberts for all the behind the scenes organisation, my fellow lecturers and tutors. Most of all, I would like to thank the students for their goodwill and enthusiasm. If you enjoyed it half as much as I did ...

Problems

1. Check that $\gamma_5^2 = 1$ and $\{\gamma_5, \gamma_\mu\} = 0$.
Show that $P_L = \frac{1}{2}(1-\gamma_5)$ and $P_R = \frac{1}{2}(1+\gamma_5)$ are projection operators, i.e.

$$P_L^2 = P_L \quad P_R^2 = P_R \quad P_L P_R = P_R P_L = 0 \quad P_L + P_R = 1$$

Consider a massless fermion with $p_\mu = (E, 0, 0, E)$. Show that $P_L u(p)$ and $P_R u(p)$ are eigenstates of helicity h with eigenvalues $-1/2$ and $1/2$ respectively.

$$h = \frac{1}{2}\frac{\underline{\sigma}.\underline{p}}{|\underline{p}|} = -\frac{1}{2}\frac{\gamma_0\gamma_5\underline{\gamma}.\underline{p}}{E}$$

2. Consider the current $J_\mu = \frac{1}{2}\bar{u}\gamma_\mu(1-\gamma_5)u$. Show that under a combined CP transformation,

$$J_\mu \rightarrow -\frac{1}{2}\bar{u}\gamma_\mu^\dagger(1-\gamma_5)u$$

Hence verify that the product $J_\mu J^{\dagger\mu}$ is CP invariant.

What happens if we have different types of fermions u_i and a current $J_\mu = \frac{1}{2}\bar{u}_i\gamma_\mu(1-\gamma_5)V_{ij}u_j$, for some matrix V?

3. Suppose that the weak charged current had the Lorentz structure

$$J_\mu^{CC} = \bar{\nu}_e\gamma^\mu(a + b\gamma_5)e \quad + \quad (\mu \rightarrow e)$$

Calculate the cross section for $\nu_\mu e^- \rightarrow \mu^-\nu_e$ and show that

$$\frac{d\sigma}{d\Omega} = \frac{G^2 s}{32\pi^2}(A^+ + A^-\cos^4\frac{\theta}{2})$$

where $A^\pm = (a^2+b^2)^2 \pm 4a^2b^2$. Neglect m_e and m_μ and assume

$$\begin{aligned} Tr\gamma.k\gamma^\mu\gamma.k'\gamma^\nu &= 4(k^\mu k'^\nu + k'^\mu k^\nu - k.k'g^{\mu\nu}) \\ Tr\gamma_5\gamma.k\gamma^\mu\gamma.k'\gamma^\nu &= 8i\epsilon^{\mu\nu\lambda\rho}k_\lambda k'_\rho \\ -\frac{1}{2}\epsilon^{\mu\nu\lambda\rho}\epsilon_{\mu\nu\alpha\beta} &= (\delta_\alpha^\lambda\delta_\beta^\rho - \delta_\beta^\lambda\delta_\alpha^\rho) \end{aligned}$$

4. The decay rate for the 2-body decay $Z \rightarrow f\bar{f}$ is

$$\Gamma = \frac{1}{2m_Z}\int D\,|M|^2 = \frac{1}{64\pi^2 m_Z}\int d\Omega\,|M|^2$$

where D denotes the phase space measure.

The $Zf\bar{f}$ vertex is $-i\frac{g}{\cos\theta_W}\gamma^\mu\frac{1}{2}(c_V^f - c_A^f\gamma_5)$.

First show that, summing over the fermion and averaging over the boson spins,

$$|M|^2 = \frac{1}{12}\frac{g^2}{\cos^2\theta_W}(c_V^{f2} + c_A^{f2})(-g_{\mu\nu})Tr\gamma^\mu\gamma.k_1\gamma^\nu\gamma.k_2$$

where k_1, k_2 are the fermion momenta and the gauge boson polarisation sum is

$$\sum_\lambda \epsilon_\mu^{(\lambda)*}\epsilon_\nu^{(\lambda)} = -g_{\mu\nu} + \frac{q_\mu q_\nu}{m_Z^2}$$

Then show that the decay rate is

$$\Gamma = \frac{1}{48\pi}\frac{g^2}{\cos^2\theta_W}(c_V^{f2} + c_A^{f2})m_Z$$

5. Using the explicit forms for c_V and c_A in the electroweak model, derive expressions for the decay rates $Z \to \nu_e\bar{\nu}_e$, $Z \to e^+e^-$, $Z \to \bar{u}u$ and $Z \to \bar{d}d$ in terms of $\sin^2\theta_W$

What is the total width of the Z in the standard model?
[$G_F = 1.2 \times 10^{-5}GeV^2$, $\sin^2\theta_W = 0.23$, $m_Z = 91GeV$]

6. Consider a Higgs theory for a general gauge group G and Higgs field ϕ. Show that the vector boson mass matrix is

$$(m^2)^{AB} = g^2\langle\phi^\dagger\rangle\{T^A, T^B\}\langle\phi\rangle$$

where $\langle\phi\rangle$ is the vacuum expectation value of ϕ and T^A is the generator of G in the representation to which ϕ belongs.

Specialise the above result to $G = SU(2)_L \times U(1)_Y$, with ϕ in an $SU(2)_L$ doublet representation with $Y = 1$ and assume the breaking conserves $U(1)_{em}$. Show that in the charged sector, $m_{W^\pm}^2 = g^2v^2$, where v is the magnitude of the VEV for ϕ, while in the neutral sector, the mass matrix for W_μ^3 and B_μ is

$$\frac{1}{4}v^2\begin{pmatrix} g^2 & -gg' \\ -gg' & g'^2 \end{pmatrix}$$

Diagonalise this to find the mass eigenstates. Show that these are the photon A_μ and Z_μ defined as

$$Z_\mu = W_\mu^3\cos\theta_W - B_\mu\sin\theta_W$$
$$A_\mu = W_\mu^3\sin\theta_W + B_\mu\cos\theta_W$$

TOPICS IN STANDARD MODEL PHENOMENOLOGY

By R G Roberts
Rutherford Appleton Laboratory

Lectures delivered at the School for Young High Energy Physicists
Rutherford Appleton Laboratory, September 1994

Topics in Standard Model Phenomenology

Dick Roberts

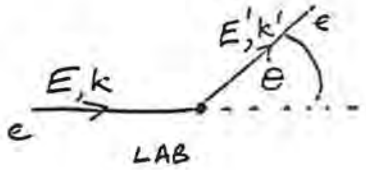
Rutherford Appleton Laboratory

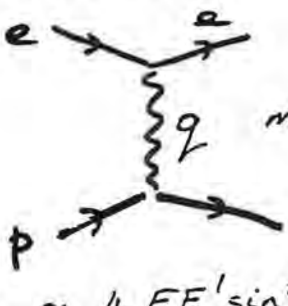
2

STRUCTURE of the PROTON

– deep inelastic scattering

First, recall ordinary **ELASTIC** electron-proton scattering.





mom. transfer $q^2 \leq 0$

$$q^2 = -(k-k')^2 \simeq 4EE'\sin^2\frac{\theta}{2}$$

$$\mathcal{M} \sim \frac{\alpha}{q^2}$$

Assume a) Point like proton
b) spinless electron
c) static limit $M \to \infty$

$$\Rightarrow \frac{d\sigma}{d\Omega} = \frac{\alpha^2}{4E^2\sin^4\frac{\theta}{2}}$$ Rutherford Cross section

spin $\frac{1}{2}$ electron $\Rightarrow \cos^2\frac{\theta}{2}$ on top

$M_p \neq \infty \Rightarrow \frac{E'}{E}\left[1 - \frac{q^2}{2M^2}\tan^2\frac{\theta}{2}\right]$ factor

$\Rightarrow$ **Mott cross-section**

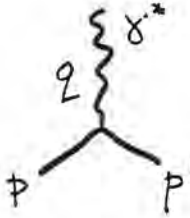
But proton isn't pointlike

Spatial charge distbn. $\rho(r)$

Amplitude modified by form-factor $F(q^2)$:

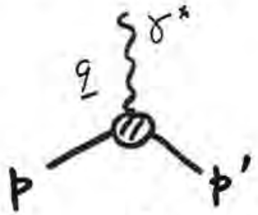
$$F(q^2) = \int d^3\underline{r}\; e^{i\underline{q}\cdot\underline{r}}\rho(r)$$

$$\Rightarrow \frac{d\sigma}{d\Omega} = \left.\frac{d\sigma}{d\Omega}\right|_{Mott} F(q^2)$$



$$j_\mu^{em} = \bar{u}(p')\gamma_\mu u(p)$$

$$\Downarrow$$



$$j_\mu^{em} = \bar{u}(p')\,O_\mu^i\, u(p)$$

conserved current $q^\mu O_\mu^i = 0$

$O_\mu^i = \gamma_\mu,\ \sigma_{\mu\nu}q^\nu$ ✓ ✓

$\sigma_{\mu\nu}\gamma_5 q^\nu,\ \gamma_5(q^2\gamma_\mu - \not{q}q_\mu)$ ×:

$$\therefore j_\mu^{em} = \bar{u}(p')\left[\gamma_\mu F_1(q^2) + \frac{i}{2M}\sigma_{\mu\nu}q^\nu F_2(q^2)\right]u(p)$$

$F_1(0)$ = charge on nucleon $\quad F_1^p(0) = 1 \quad F_1^n(0) = 0$

As a result $\left[1 - \frac{q^2}{2M^2}\tan^2\frac{\theta}{2}\right]$ 3

$$\Rightarrow \left[\left(F_1^2 - \frac{q^2}{4M^2}F_2^2\right) - \frac{q^2}{2M^2}\left(F_1 + F_2\right)^2 \tan^2\frac{\theta}{2}\right]$$

Rosenbluth formula

Experiment determines form-factors

Probe with large wavelength γ^*

- small q^2

$$F(q^2) = \int d^3r\, \rho(r)\, e^{i\underline{q}\cdot\underline{r}}$$

$$= \int d^3r\, \rho(r)\left[1 + i\,\underline{q}\cdot\underline{r} - \tfrac{1}{2}(\underline{q}\cdot\underline{r})^2 + \cdots\right]$$

$$= 1 - \frac{q^2}{6}\langle r^2\rangle + \cdots$$

Expt $\Rightarrow \langle r^2\rangle^{1/2} = 0.862 \pm 0.012$ fm

Usually write: $G_E = F_1 + \frac{q^2}{4M^2}F_2$

$$G_M = F_1 + F_2$$

$F_2(0) = \mu$ 'anomalous' mag. moment of nucleon

$\mu^p = 1.7928456(11)$

$\mu^n = -1.91304184(88)$

Both $G_E(q^2)$, $G_M(q^2)$ are 'dipole' like

$$\sim \left(1 \mp \frac{q^2}{q_0^2}\right)^{-2}$$

$q_0^2 \approx 0.7\,\mathrm{GeV}^2$

$\to Q^2$

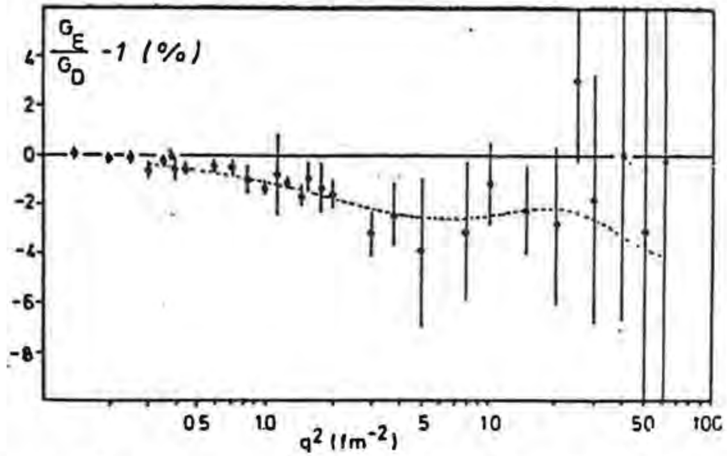


Fig. 2. The ratio G_E^p/G_D in percent deviation from 1, versus q^2. Figure taken from Simon et al. (1980).

Analogous definitions may be introduced for the electric and magnetic form factors (58) and (61).

It is customary to show the measured electric form factor G_E^p of the proton for low momentum transfers, divided by the so-called *dipole* fit:

$$G_D(q^2) = \frac{1}{\left(1 + q^2/q_0^2\right)^2} \quad \text{with} \quad q_0^2 = 18.23\ \mathrm{fm}^{-2}. \tag{67}$$

This dipole dependence of the form factors $G_E^p(q^2)$ and $G_M^p(q^2)/\mu^p$ on q^2 is only of historical interest in so far as early data seemed to be in good agreement with this

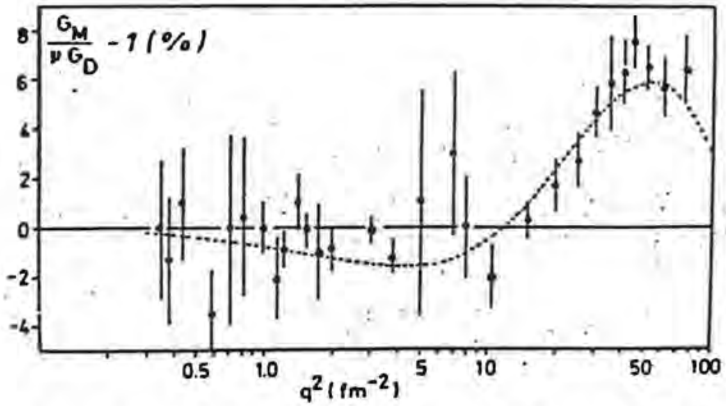


Fig. 3. The ratio $G_M^p/\mu^p G_D$ versus q^2. Figure taken from Simon et al. (1980).

they calculate values for G_M^p which have approximately the correct normalization, within an overall uncertain-

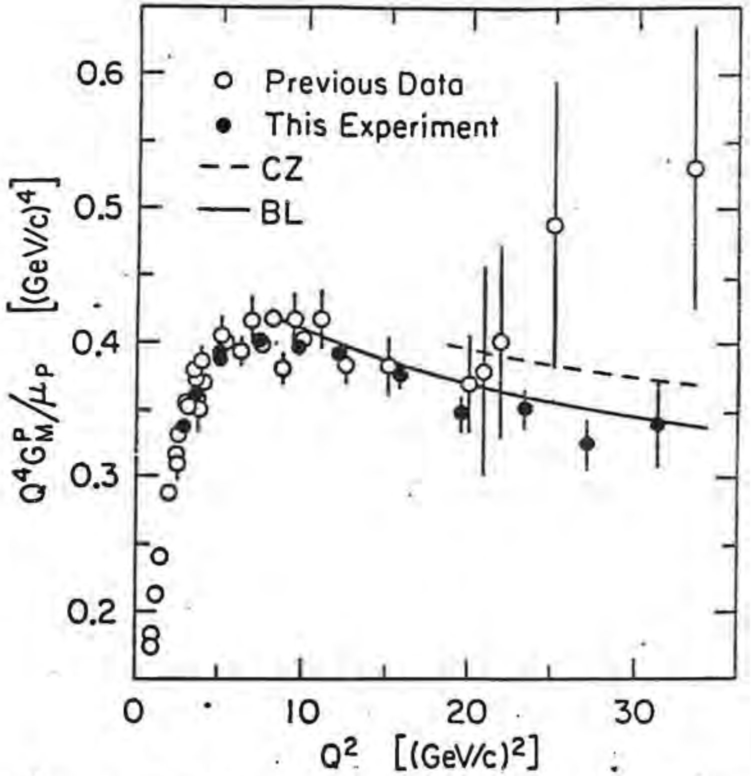


FIG. 2. Extracted values of $Q^4G_M^p/\mu_p$ vs Q^2. Open circles show previous data as given in Ref. 1. Solid circles show the results of this experiment. The curves show the perturbative QCD predictions of Refs. 6 (BL) and 7 (CZ) for $\Lambda_{QCD} = 100$ MeV.

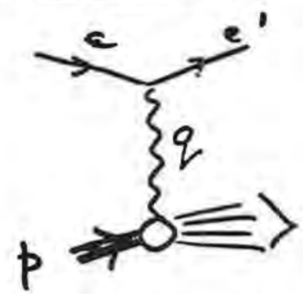
Message : Probing the nucleon with a 'long' wavelength reveals ~~structure~~. spatial distbn.

Structure $\Longleftrightarrow$ strong (ie power law) dependence on q^2

Now carry on, with smaller wavelengths

let $Q^2 = -q^2 \longrightarrow$ large

Deep inelastic scattering



$ep \to e' X$

or $\nu p \to \mu X$

$\nu p \to \nu X$

ie γ, Z^0, W exchange.

$$Q^2 = -q^2 \longrightarrow \infty$$

and $\nu = \frac{p \cdot q}{M} \longrightarrow \infty$

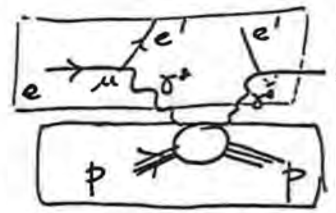
so that $x = \frac{Q^2}{2M\nu}$ finite $0<x<1$

Inclusive Cross-section $k_0' \frac{d\sigma}{d^3k'}$ only final electro detected

5

$$k_0' \frac{d\sigma}{d^3k'} = \frac{2M}{s} \frac{\alpha^2}{Q^4} \, l^{\mu\nu} \, W_{\mu\nu}$$

$l^{\mu\nu}$



$W_{\mu\nu}$ hadronic tensor

Imag. part of forward $\gamma^* p$ amplitude

For e.m. current 2 structure functions

$$W_{\mu\nu}(p, q) = -W_1 \left(g_{\mu\nu} + \frac{1}{Q^2} q_\mu q_\nu\right) + W_2 \left(p_\mu + \frac{p \cdot q}{Q^2} q_\mu\right)\left(p_\nu + \frac{p \cdot q}{Q^2} q_\nu\right)$$

cons. current $\quad q^\mu W_{\mu\nu} = 0$

(weak current : extra str. fn. W_3)

Thus the structure functions measured from the inclusive cross-section

Expect what ? $\quad M W_1(x, Q^2) \equiv F_1$
$\qquad\qquad\qquad\quad \nu W_2(x, Q^2) \equiv F_2$

Do they fall rapidly with Q^2 ?

No — scaling observed

$F_1 \, F_2$ — depend only on x

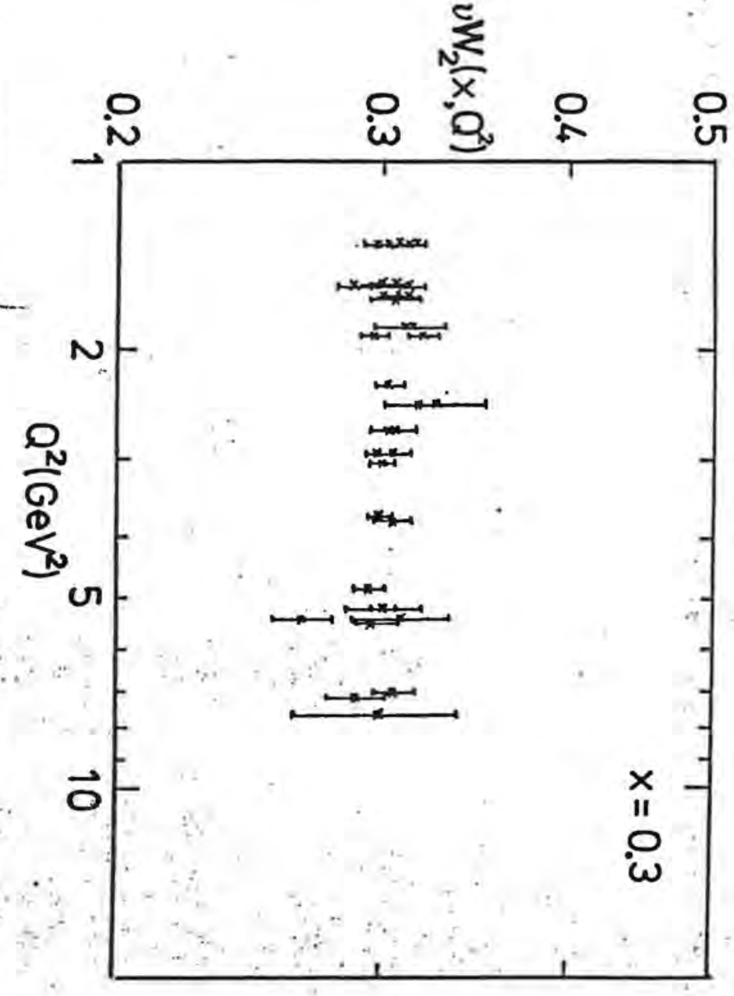
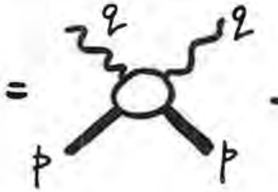


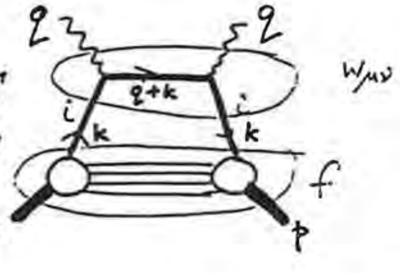
Fig. 1. 3

MB 1.3

$\Rightarrow$ Virtual photon interacts with 'point-like' objects $\Rightarrow$

Quark Parton Model

$W_{\mu\nu} =$ 

$\longrightarrow \sum_i$ 

$$= \sum e_i^2 \int d^4k \; w_{\mu\nu}(q,k)\; f(p,k)\; \delta\left[(k+q)^2 - m^2\right]$$

on-shell

$$= \frac{1}{4} e_i^2 \, \mathrm{Tr}\left[\not{k}\gamma_\mu(\not{k}+\not{q})\gamma_\nu\right]$$

$$d^4k = \frac{\pi}{2}\frac{dx}{x}\, dk^2 dk_T^2$$

$$\Downarrow$$

$$\frac{1}{2M\nu}\,\delta(\xi_i - x)$$

$$\xi_i = k^+/p^+$$

$$q_i(x) = \frac{\pi}{4}\int dk_T^2\, f(p.k)$$

probability of finding a quark of type i with a fraction x of proton momentum

at $\mu=\nu=2$ $\quad W_{22} \Rightarrow W_1$

$$\Rightarrow \quad MW_1(p,q)\;(=F_1) = \sum_i e_i^2\, q_i(x)$$

scaling

Taking $\mu=\nu=0$ $\qquad W_{00} = \underbrace{\frac{\nu^2}{MQ^2}}_{\to\infty}\left[-MW_1 + \frac{1}{2x}\nu W_2\right]$

From QPM, W_{00} finite

$$\Rightarrow \quad MW_1 = \frac{1}{2x}\nu W_2$$

$$2x\,F_1(x) = F_2(x)$$

CG relatio

σ_T, σ_L assoc. with transverse / longitudinal photon

$$F_1 \sim M\nu\,\sigma_T$$

$$F_2 \sim Q^2(\sigma_L + \sigma_T)$$

$$\therefore \quad \frac{F_1}{F_2} \sim \frac{M\nu}{Q^2}\,\frac{\sigma_T}{\sigma_L+\sigma_T}$$

$$\Rightarrow \quad \frac{F_2}{2xF_1} = \frac{\sigma_L + \sigma_T}{\sigma_T}$$

$$\Rightarrow \quad R = \frac{\sigma_L}{\sigma_T} = 0 \quad \text{in QPM}$$

QPM $\Rightarrow$ $\quad F_2^p(x) = \sum_i e_i^2 \; x\, q_i(x)$

e.g. $$F_2^p(x) = x\left[\frac{4}{9}\left(u(x)+\bar{u}(x)+c(x)+\bar{c}(x)\right) + \frac{1}{9}\left(d(x)+\bar{d}(x)+s(x)+\bar{s}(x)\right)\right]$$

Convention u, d, s etc. refer to proton

In a neutron simply $u \leftrightarrow d$

Valence & Sea Quarks

$$u(x) - \bar{u}(x) = u_v(x) \quad \text{valence up}$$

$$d(x) - \bar{d}(x) = d_v(x) \quad \text{valence down}$$

$$\int_0^1 dx\, u_v(x) = 2 \qquad \int_0^1 dx\, d_v(x) = 1$$

write

$$u(x) = u_v(x) + u_{sea}(x)$$
$$\bar{u}(x) = \; u_{sea}(x)$$

infinite !

Currents

Neutral Current γ Z

Charged Current W^+, W^-.

e.g.

Using different DIS processes

e.g. $F_2^{\nu p} = 2x\left[d(x) + s(x) + \bar{u}(x) + \bar{c}(x)\right]$

Associated with weak current $(Z^0, W^{\pm})$

– extra str. fn. F_3

arises from interference of V, A parts of current.

In QPM e.g.

$$xF_3^{\nu p}(x) = x\left[d(x) + s(x) - \bar{u}(x) - \bar{c}(x)\right]$$

$$xF_3^{\bar{\nu} p}(x) = x\left[u(x) + c(x) - \bar{d}(x) - \bar{s}(x)\right]$$

$$\therefore\; xF_3^{\nu p}(x) + xF_3^{\bar{\nu} p}(x) = x\left[u(x) - \bar{u}(x)\right] + x\left[d(x) - \bar{d}(x)\right]$$

$$= x\,u_v(x) + x\,d_v(x)$$

$\Rightarrow$ extract valence, sea etc distbns. from data

and get sum rules

e.g. $\int_0^1 dx\left(F_3^{\nu p}(x) + F_3^{\bar{\nu} p}(x)\right) = 3 \times \left(1 - \frac{\alpha_s}{\pi}\right)$

$S(x)$ = sea

$u(x) = u_v(x) + S(x)$
$d(x) = d_v(x) + S(x)$
$\bar{u}(x) = \bar{d}(x) = S(x)$ assumed !
$s(x) = \bar{s}(x) \simeq \frac{1}{2} S(x)$

- distbns at $Q^2 = 10$ figure.

$\int dx\, u_v(x) = 2 \qquad \int dx\, d_v(x) = 1$

$\int dx\, S(x) = \infty \qquad S(x) \sim \frac{1}{x}(1-x)^{10}$

$\sum_q \int_0^1 dx\, x\,(q(x) + \bar{q}(x)) \simeq 0.5$

Momentum of quarks
$\approx \frac{1}{2}$ mom. of proton.

- rest is gluons
- not directly measurable in DIS.

If $\bar{u} = \bar{d}$
$\int_0^1 dx \left[F_2^{\mu p}(x) - F_2^{\mu n}(x)\right] = \frac{1}{3}\int_0^1 dx\,(u_v - d_v) = \frac{1}{3}$

but NMC get $\int^1 dx \left[F_2^p - F_2^n\right] \simeq 0.23$

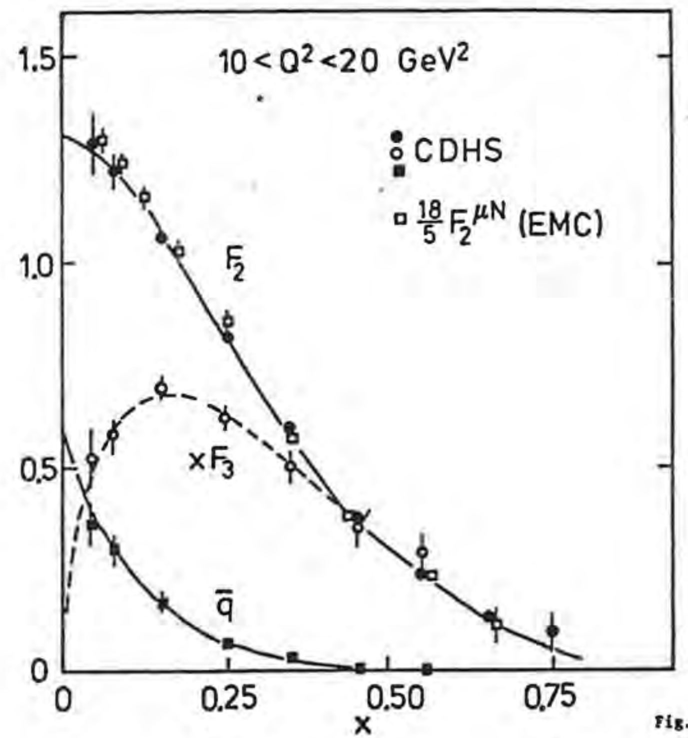


Fig. 3. 8

$x V(x) \to 0 \quad x \to 0$
$x S(x) \not\to 0$

QCD + DIS

— "improved" parton model

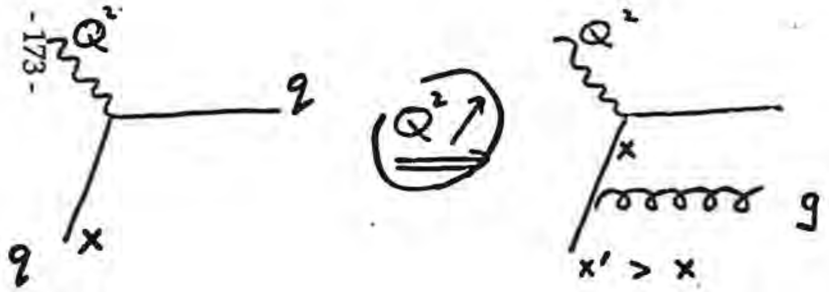
Scaling $\equiv$ indepent of Q^2

$\rightarrow$ $\log Q^2$ corrections

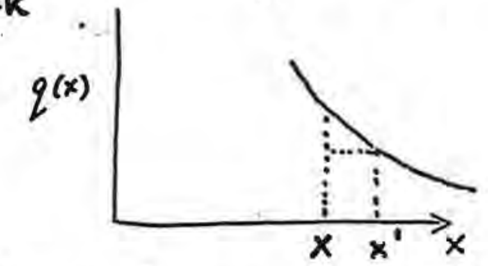
Where do they come from?

Increasing Q^2 = decreasing the 'wavelength' of the 'probing' photon

$\rightarrow$ reveals more substructure in the proton.



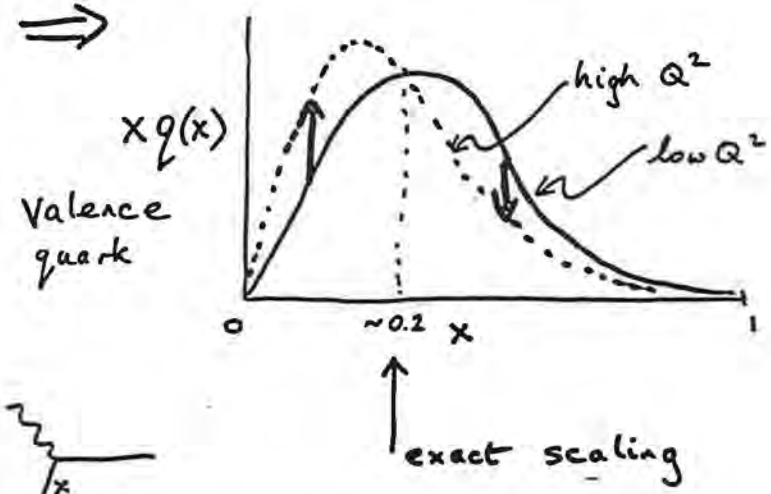
momentum of struck quark smaller than original quark



Because $q(x) \searrow$ at large x, the $q(x)$ distribution will fall as Q^2 increases.

Also a negative contribution from quarks " which had mom. frac. x <u>before</u> radiating a gluon with mom. frac. $x' < x$.

$\Longrightarrow$



$$q(x,Q^2) = \int_0^1 dz \int_0^1 dx' \; q(x',Q^2)\, P_{qq}(z,Q^2)\, \delta(x - zx$$

$\Downarrow$ $z = $

P_{qq} ≡ Probability that the quark, on radiating a gluon, degrades its mo by a fraction

$$\text{QPM}: \quad P_{qq}(z,Q^2) = \delta(1-z)$$

12

QCD

$$P_{qq}(z, Q^2 + dQ^2) = \delta(z-1) + \frac{\alpha_s}{2\pi} P_{qq}(z) \frac{dQ^2}{Q^2}$$

$$t = \ln Q^2$$

$$\Rightarrow \frac{dq(x,t)}{dt} = \frac{\alpha_s(t)}{2\pi} \int_0^1 dz \int_0^1 dx' \, q(x';t) \, P_{qq}(z) \, \delta(x - zx')$$

$$\longrightarrow \frac{\alpha_s(t)}{2\pi} \int_x^1 \frac{dx'}{x'} \, q(x';t) \, P_{qq}\left(\frac{x}{x'}\right)$$

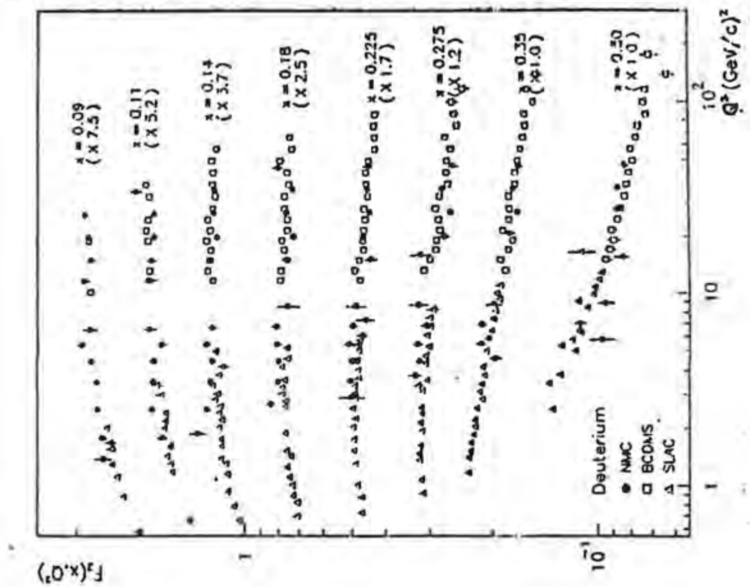
$P_{qq}(z)$ quark-quark "splitting function"

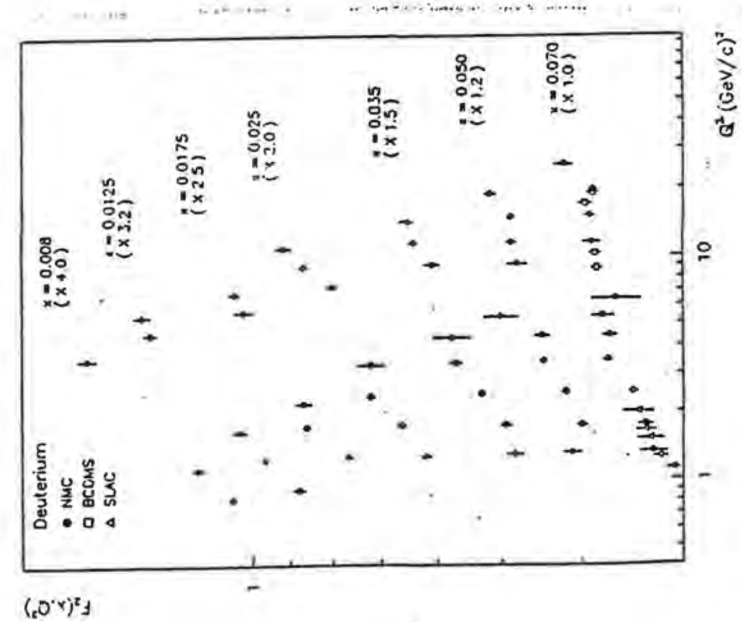
$$= \frac{4}{3}\left[\frac{1+z^2}{(1-z)_+} + \frac{3}{2}\delta(1-z)\right]$$

$$\int_0^1 dz \, P_{qq}(z) = 0 \qquad \therefore \int_0^1 dx \, q_v(x) = \text{const with } Q^2.$$

Remember: Asymptotic Freedom $\alpha_s(t) = \dfrac{4\pi}{\beta_0 \ln \frac{Q^2}{\Lambda^2}}$

$Q^2 \longrightarrow \infty \Rightarrow$ naive quark model





Pert. QCD

$\Longrightarrow$ How things change with Q^2

('short' distance physics)

Non pert. QCD

$\Longrightarrow$ Starting values at $Q^2 = Q_0^2$

('long' distance physics)

... lattice ....

Q^2 dep. Full Altarelli-Parisi Eqns: ...

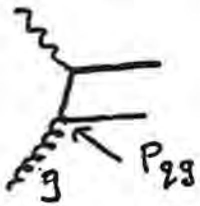
$$\frac{d}{dt}\begin{pmatrix} q(x,t) \\ g(x,t) \end{pmatrix} = \frac{\alpha_s(t)}{2\pi}\int_x^1 \frac{dx'}{x'}\begin{pmatrix} P_{qq}(\frac{x}{x'}) & P_{qg}(\frac{x}{x'}) \\ P_{gq}(\frac{x}{x'}) & P_{gg}(\frac{x}{x'}) \end{pmatrix}\begin{pmatrix} q(x',t) \\ g(x',t) \end{pmatrix}$$

" Leading Order "

Note

1) 3 other splitting functions

$\frac{d}{dt}q(x,t)$ also depends on



2) The size of α_s determines how 'fast' α_s runs

$$\frac{d\,\alpha_s(t)}{dt} = -\frac{\beta_0}{4\pi}\,\alpha_s^2(t) + \dots$$

$\therefore$ Scaling Violations are directly controlled by Λ

$\Rightarrow$ Only data can tell us value of Λ

– Data on DIS structure functi

Present Situation EMC, BCDMS, SLAC, NMC .....

much better than 2-3 years ago

$\Longrightarrow \Lambda_{\overline{MS}} \simeq 200$ MeV

3) Life is tough

a) Leading Order not good enough

Suppose $\sigma(Q^2) = \underbrace{\sigma_0 + \sigma_1 \frac{\alpha_s}{4\pi}}_{L.O.} + \sigma_2\left(\frac{\alpha_s}{4\pi}\right)^2 +$

$$\alpha_s = \frac{4\pi}{\beta_0 \ln Q^2/\Lambda^2}$$

Now let $\Lambda \to \Lambda' = \frac{1}{K}\Lambda$

$$\alpha_s \to \alpha_s' = \frac{4\pi}{\beta_0 \ln Q^2/\Lambda^2}\left[1 - \frac{\ln K^2}{\ln Q^2/\Lambda^2}\right]$$

$$= \alpha_s - \frac{\beta_0}{4\pi}\ln K^2\, \alpha_s^2$$

$\therefore\ \sigma(Q^2)$, to L.O. unchanged!

$\therefore\ \Lambda_{L.O.}$ meaningless.

b) include next-to-leading order corrections

$$\frac{d\,\alpha_s(t)}{dt} = -\beta_0 \frac{\alpha_s^2}{4\pi} - \beta_1 \frac{\alpha_s^3}{(4\pi)^2}$$

$$P_{qq}(z) = P_{qq}^{(0)}(z) + \frac{\alpha_s}{2\pi} P_{qq}^{(1)}(z) \quad \text{etc.}$$

$$\alpha_s(Q^2) = \frac{4\pi}{\beta_0 \ln Q^2/\Lambda^2} - 4\pi\beta_1 \frac{\ln\ln Q^2/\Lambda^2}{(\beta_0 \ln Q^2/\Lambda^2)^2}$$

Procedure depends on how you do renormalise! Different schemes

minimal subtraction, MS, $\overline{MS}$

momentum ", MOM,

minimal sensitivity, PMS

. . . . .

So decide on <u>one</u> scheme and stick to

Consensus: $\overline{MS}$ $\qquad \Lambda_{\overline{MS}} \simeq 200$ MeV

c) Scale Dependence (Renorm. scale)

$$\sigma(Q^2) = \sigma_0 + \sigma_1 \frac{\alpha_s(Q^2)}{4\pi} + \sigma_2\left(\frac{\alpha_s(Q^2)}{4\pi}\right)^2 + \ldots$$

But $$\alpha_s^{-1}(Q^2) = \alpha_s^{-1}(\mu^2) + \frac{\beta_0}{4\pi}\ln\frac{\mu^2}{Q^2} + O(\alpha_s)$$

$$\simeq \alpha_s^{-1}(\mu^2)\left[1 + \frac{\beta_0}{4\pi}\ln\frac{\mu^2}{Q^2}\cdot\alpha_s(\mu^2)\right]$$

$$\sigma(Q^2) = \sigma_0 + \frac{\alpha_s(\mu^2)}{4\pi}\sigma_1 + \left[\frac{\alpha_s(\mu^2)}{4\pi}\right]^2\left(\sigma_2 - \beta_0\sigma_1 \ln\frac{\mu^2}{Q^2}\right) + O(\alpha_s^3)$$

Problem: What do we choose for μ?

$\mu = Q$, $\frac{1}{2}Q$, $2Q$ ..... ?

Reflection of fact that we have a <u>TRUNCATED</u> series.

Choosing $\mu \iff$ guessing size of σ_3

d) Factorisation Scale Dependence

Leading Order

$$F_2(x,Q^2) = \sum_i e_i^2 \, x q_i(x,Q^2) \equiv \mathcal{F}(x,Q^2)$$

But at LO + NLO

$$F_2(x,Q^2) = \sum_i \int_x^1 \frac{dx'}{x'} \mathcal{F}(x',M^2)\, C_q\left(\frac{x}{x'}\right) + \int_x^1 \frac{dx'}{x'} x' G(x',M^2)\, C_g\left(\frac{x}{x'}\right)$$

$$C_q(z) = \delta(1-z) + \frac{\alpha_s(Q^2)}{4\pi} C_q^{(1)}(z)$$

$$C_g(z) = \frac{\alpha_s(Q^2)}{4\pi} C_g^{(1)}(z)$$

Not important in DIS - but in other processes it is!

Summary:

BAD NEWS - In Perturbative QCD phenomenology, have to worry about renormalisation schemes, renormalisation scales, factorisation scales, ...

GOOD NEWS - After fitting DIS + other processes (see shortly) we now have very precise knowledge of the structure of the proton - i.e. we know the $xq_i(x,Q^2)$ over a wide range of x, Q^2

$i = u_v, d_v, \bar{u}, \bar{d}, s, c$

$i = c, b$ generated by photon - gluon fusion

Progress very much due to third generation expts on DIS - very precise measurements e.g. NMC n/p.

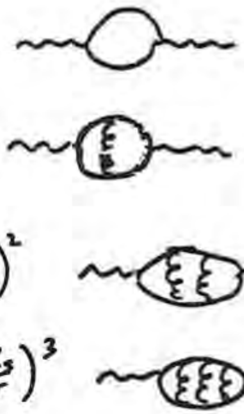
GOOD/BAD NEWS

Problems with choice of scales due to ignorance of higher order corrections. Only one QCD correction been calculated to $O(\alpha_s^3)$

$\sigma_{TOT}(e^+e^- \to \text{hadrons})$

$$R = \frac{\sigma_{TOT}(e^+e^- \to \text{had.})}{\sigma(e^+e^- \to \mu^+\mu^-)} = 3\sum_i e_i^2\, \tilde{\sigma}(Q^2)$$

$$\tilde{\sigma}(Q^2) = 1 + \frac{\alpha_s}{\pi} + \dots + 1.4092\left(\frac{\alpha_s}{\pi}\right)^2 + \dots - 12.8046\left(\frac{\alpha_s}{\pi}\right)^3 + \dots$$



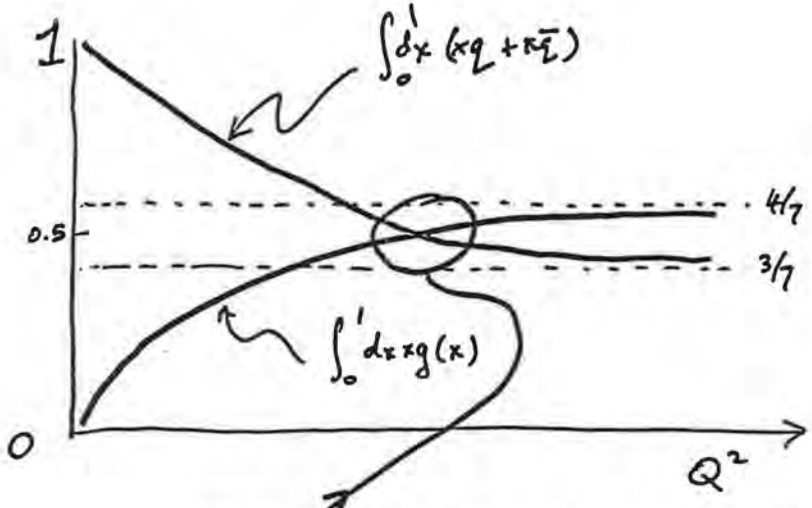
was + 60 ~~until recently ago~~

calculation has taken several years + lots of computer time

b) The gluon

DIS tells little about $x\,g(x)$

$$QCD \Rightarrow \int_0^1 dx \left[\left\{\sum_i x q_i(x) + x\bar{q}_i(x)\right\} + x g(x)\right] = \text{const} = 1$$



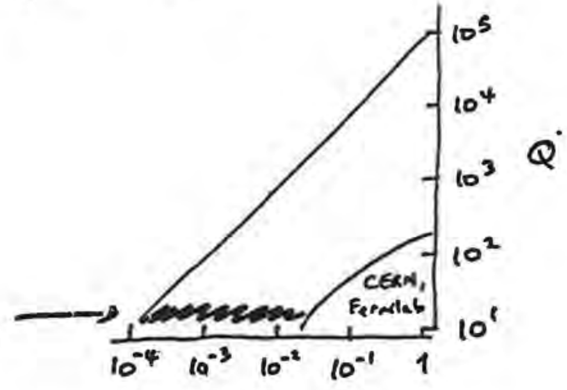
we seem to be here $Q^2 \sim 5 \to 200$ G

ie 50% of proton momentum carried by gluons

$\boxed{x g(x)}$ Small x - gluons important.

$\therefore$ HERA will reveal gluon information

interesting new region!



Small x STRUCTURE FUNCTIONS and HERA.

Can QCD predict x-dependence as well Q^2-dependence

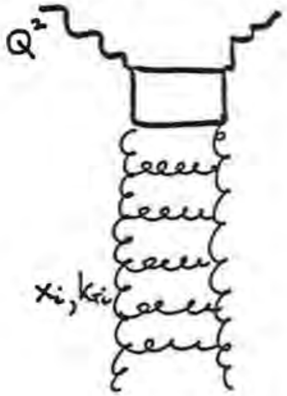
YES — at small x

$$W^2 = Q^2\left(\frac{1}{x} - 1\right)$$

So TWO large scales Q^2 and W^2

As $x \to 0$ $\quad \overline{P_{gg} \sim \frac{1}{z}}$

Gluons dominate.



A-P sums leading $\alpha_s \ln Q^2$ terms and $\alpha_s \ln Q^2 \ln\frac{1}{x}$ terms

k_{Ti} are ordered

x_i strongly ordered

$$\Rightarrow \quad xg(x) \sim \exp\left[K\sqrt{\ln\frac{1}{x}\ \ln\ln Q^2}\,\right]$$

A-P sums only some of leading $\ln\frac{1}{x}$ terms !

Leading Log $LL(Q^2)$ contributions arise from region $Q^2 \gg k_{nT}^2 \gg \ldots \gg k_{1T}^2$

Going also to small x, there will also be large $\log\left(\frac{1}{x}\right)$ term

The behaviour $xg(x, Q^2) \sim \exp\left[2\left\{\frac{3\alpha_s}{\pi}\log Q^2 \ln\frac{1}{x}\right\}^{\frac{1}{2}}\right]$

involves the double log summation DLL

Requires both $\quad Q^2 \gg k_{nT}^2 \gg \ldots \gg k_{1T}^2$

$\quad x \ll x_{n-1} \ll \ldots\ldots \ll x_1$

But if we want small x, moderate Q^2

must sum leading $\alpha_s \log\frac{1}{x}$ ie keep full Q^2 dependence

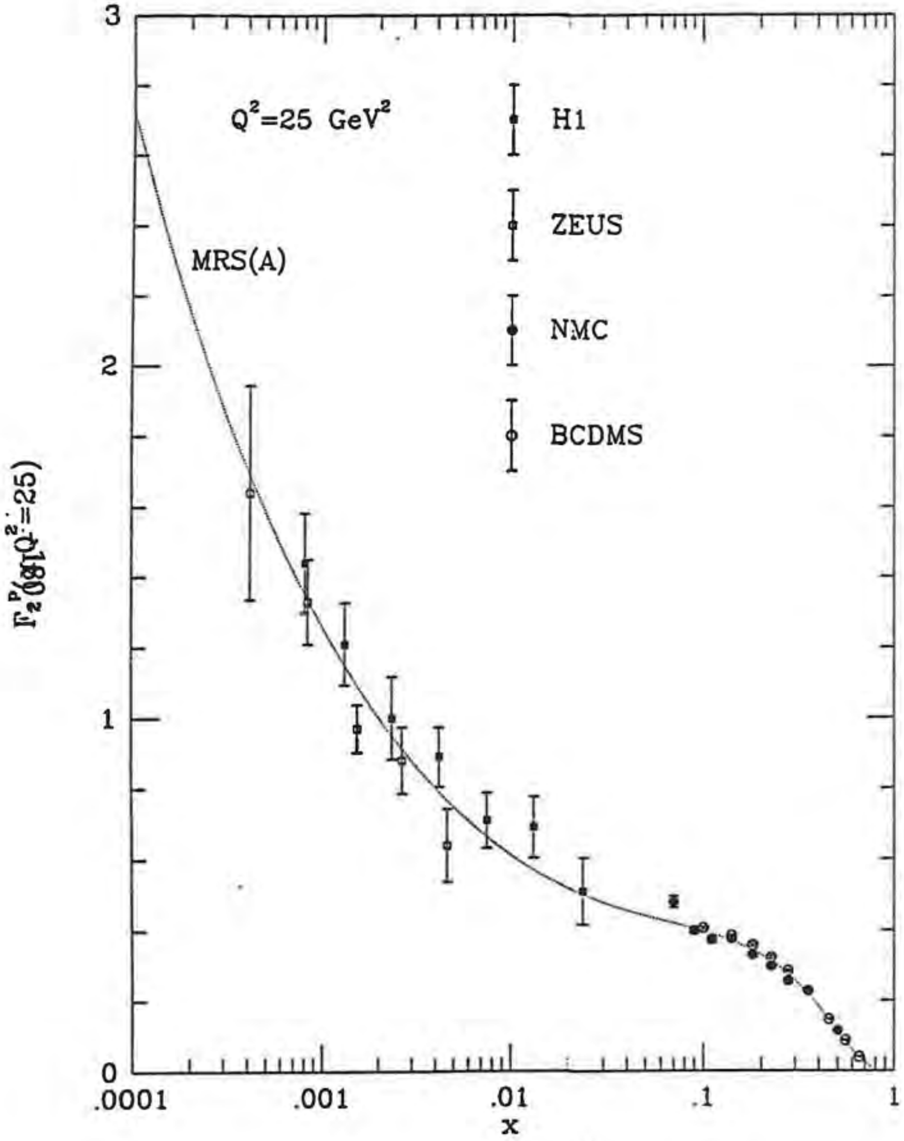
BFKL (Lipatov) equation does this

— now too restrictive. Integrate over all k_{Ti}

Solution $\quad xg(x) \sim x^{-\lambda}$

$$\lambda = \frac{12}{\pi}\ln 2\,\alpha_s \approx \frac{1}{2}$$

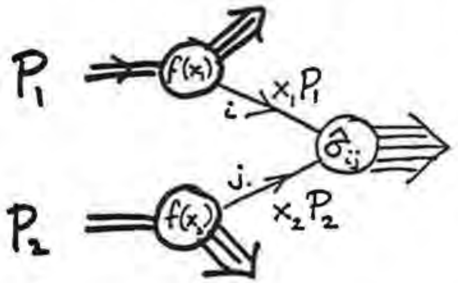
Famous prediction of singular gluon – hence F_2 expected to rise steeply at $x \to 0$



Q^2=25 GeV^2
MRS(A)
H1
ZEUS
NMC
BCDMS
$F_2^p(x, Q^2=25)$
x
3
2
1
0
.0001
.001
.01
.1
1

QCD Parton Model in Hadron - Hadron Collisions

Large number of processes in hadron colliders can be understood via :-



Two hadrons collide, final state produced by HARD scattering of the partons, i and j, in the hadrons.

HARD $\equiv$ at least one momentum is large

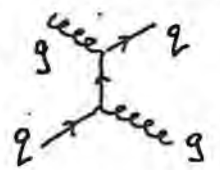
Subprocess cross section $\hat{\sigma}_{ij}$ is $\propto \alpha_s$

Universal formula:

$$\sigma(P_1, P_2) = \sum_{ij} \int dx_1 dx_2 \; f_i(x_1, \mu_1) \, f_j(x_2, \mu_2)$$
$$* \; \hat{\sigma}_{ij}\left(x_1 P_1, x_2 P_2, \alpha_s(\mu), Q\right)$$

$f_i(x, \mu) \equiv$ the quark probability distrib. fn
x = fraction of parent hadron mom
μ = factorisation scale

$\hat{\sigma}(p_1, p_2, Q) \equiv$ partonic cross section
e.g.



$Q \equiv$ characteristic scale of the hard scattering
e.g. m_W, m_t, p_T^{jet} ....
(depends on process)

Perturbation theory

$$\Rightarrow \quad \hat{\sigma} = c^{(0)} \alpha_s^k \left[1 + \sum_{j=1}^{n} c^{(i)} \alpha_s^j \right]$$

$c^{(k)}$ depend on $\hat{s}, \hat{t}, \hat{u}$

Note: main feature of universal formula = factorisation

Leading Order ($n=0$) $\quad \hat{\sigma}$ = normal parton scattering cross section calculated just as in QED.

But in higher orders

$\sigma(P_1 P_2)$ has 2 parts

(i) depending on short distances (large mom)

$\boxed{\text{(ii) depending on long distances}}$ (small mom)

$\longrightarrow$ all shoved into $f_i(x,\mu)$

$\downarrow$

Same $f_i(x,\mu) = q_i(x,\mu)$ we had in DIS <u>even</u> with higher orders!

182 Factorisation proved to ALL orders.

$\Rightarrow$ QCD is reliable calculational tool

μ - fac. scale is arbitrary - can be chosen $= Q$.

$\hat{\sigma}(\hat{s},\hat{t})$ will vary depending on process

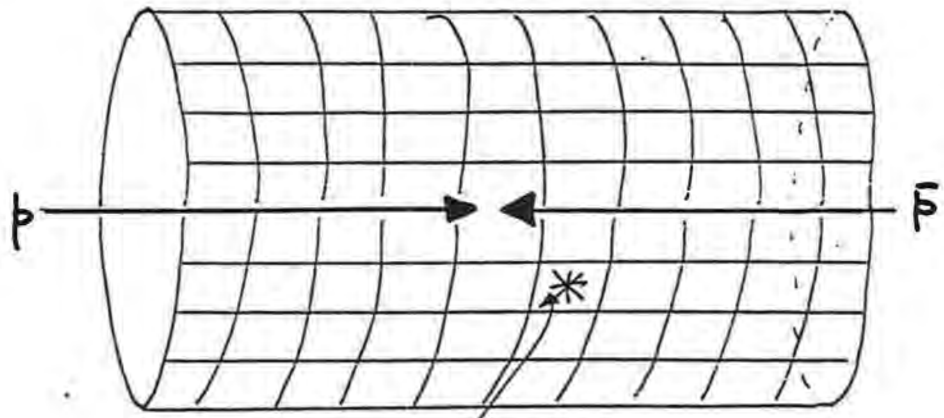
Remaining factors same -

Example Large p_T Jet Production

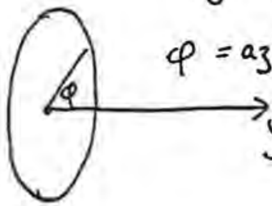
If partonic sub-process is

$+ \ldots$

The two final state partons observed as 2 jets of hadrons



Signature of a jet = LARGE energy deposition in a localised group of calorimeter cells



φ = azimuth angle

y, rapidity

No "best" definition of a jet – important for theory & experiment to use the SAME definition.

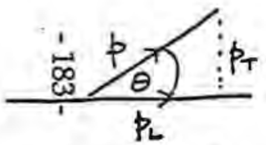
Commonly used definition:

Cluster of transverse energy E_T in a cone of size ΔR

with $(\Delta R)^2 = (\Delta y)^2 + (\Delta \varphi)^2$

---

a aside: Rapidity – a useful variable!



Rapidity of a particle $y = \frac{1}{2} \ln \left[\frac{E+p_L}{E-p_L} \right] = \sinh^{-1}\left(\frac{p_L}{\chi}\right)$

$\chi = \sqrt{m^2 + p_T^2}$

$\therefore p = (\chi \cosh y, p_T, \chi \sinh y) \xrightarrow[m \to 0]{} p_T(\cosh y, \underline{1}, \sinh y)$

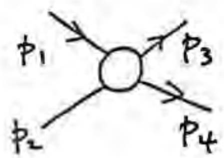
Under longitudinal Lorentz transfms (boosts)

$y \to y' = y + \text{const.}$ ie range of y same in com system / lab system

Also if $m \ll p$, $y \simeq \frac{1}{2} \ln \left[\tan \frac{\theta}{2} \right] \equiv$ pseudo-rapidity

useful in cosmic-ray physics.

Two jet Production



$\equiv$ subprocess

$$E_3 \frac{d\sigma}{dp_3} E_4 \frac{d\sigma}{dp_4} = \frac{1}{2\hat{s}} \frac{1}{16\pi^2} \sum |M|^2 \delta^4(p_1+p_2-p_3-p_4)$$

8 graphs of $O(\alpha_s^2)$

$$\Rightarrow \frac{d^3\sigma}{dy_3\, dy_4\, dp_T^2} = \frac{1}{16\pi \hat{s}^2} \sum_{\substack{i,j\\k,l}} \frac{1}{x_1} f_i(x_1,\mu) \frac{1}{x_2} f_j(x_2,\mu) \frac{1}{1+\delta_{kl}} * |M(ij \to kl)|^2 .$$

– see table of $|M|^2$

Note:

1) biggest for elastic $ab \to ab$

2) $p_T^2 = \frac{\hat{s}}{4} \sin^2 \hat{\theta} = \frac{\hat{u}\,\hat{t}}{\hat{s}}$

So poles in $\hat{t}, \hat{u}$ correspond to $\cos \hat{\theta} = \pm 1$ ie $p_T = 0$

Table 9.1. *Squared matrix elements for 2 → 2 subprocesses in QCD (averaged over spin and color): q and q' denote distinct flavors of quark, $g_s^2 = 4\pi\alpha_s$ is the coupling squared. For identical final partons, integrate only half the solid angle.*

Subprocess	$\lvert\mathcal{M}\rvert^2/g_s^4$	$\lvert\mathcal{M}(90^\circ)\rvert^2/g_s^4$
$qq' \to qq'$, $q\bar{q}' \to q\bar{q}'$	$\frac{4}{9}\frac{\hat{s}^2+\hat{u}^2}{\hat{t}^2}$	2.2
$qq \to qq$	$\frac{4}{9}\left(\frac{\hat{s}^2+\hat{u}^2}{\hat{t}^2}+\frac{\hat{s}^2+\hat{t}^2}{\hat{u}^2}\right)-\frac{8}{27}\frac{\hat{s}^2}{\hat{u}\hat{t}}$	3.3
$q\bar{q} \to q'\bar{q}'$	$\frac{4}{9}\frac{\hat{t}^2+\hat{u}^2}{\hat{s}^2}$	0.2
$q\bar{q} \to q\bar{q}$	$\frac{4}{9}\left(\frac{\hat{s}^2+\hat{u}^2}{\hat{t}^2}+\frac{\hat{t}^2+\hat{u}^2}{\hat{s}^2}\right)-\frac{8}{27}\frac{\hat{u}^2}{\hat{s}\hat{t}}$	2.6
$q\bar{q} \to gg$	$\frac{32}{27}\frac{\hat{u}^2+\hat{t}^2}{\hat{u}\hat{t}}-\frac{8}{3}\frac{\hat{u}^2+\hat{t}^2}{\hat{s}^2}$	1.0
$gg \to q\bar{q}$	$\frac{1}{6}\frac{\hat{u}^2+\hat{t}^2}{\hat{u}\hat{t}}-\frac{3}{8}\frac{\hat{u}^2+\hat{t}^2}{\hat{s}^2}$	0.1
$qg \to qg$	$\frac{\hat{s}^2+\hat{u}^2}{\hat{t}^2}-\frac{4}{9}\frac{\hat{s}^2+\hat{u}^2}{\hat{u}\hat{s}}$	6.1
$gg \to gg$	$\frac{9}{4}\left(\frac{\hat{s}^2+\hat{u}^2}{\hat{t}^2}+\frac{\hat{s}^2+\hat{t}^2}{\hat{u}^2}+\frac{\hat{u}^2+\hat{t}^2}{\hat{s}^2}+3\right)$	30.4

Integrate over p_4 to get single jet cross section.

$$E_J \frac{d\sigma}{d^3p_J} = \frac{1}{16\pi^2 s} \sum_{\substack{ij\\kl}} \int_0^1 \frac{dx_1}{x_1} \int_0^1 \frac{dx_2}{x_2} f_i(x_1,\mu) f_j(x_2,\mu)$$

$$* \sum |M(ij \to kl)|^2 \frac{1}{1+\delta_{kl}} \delta(\hat{s}+\hat{t}+\hat{u})$$

Comparison with experiment

Normalisation of cross section is fixed by $\alpha_s^2(Q^2)$

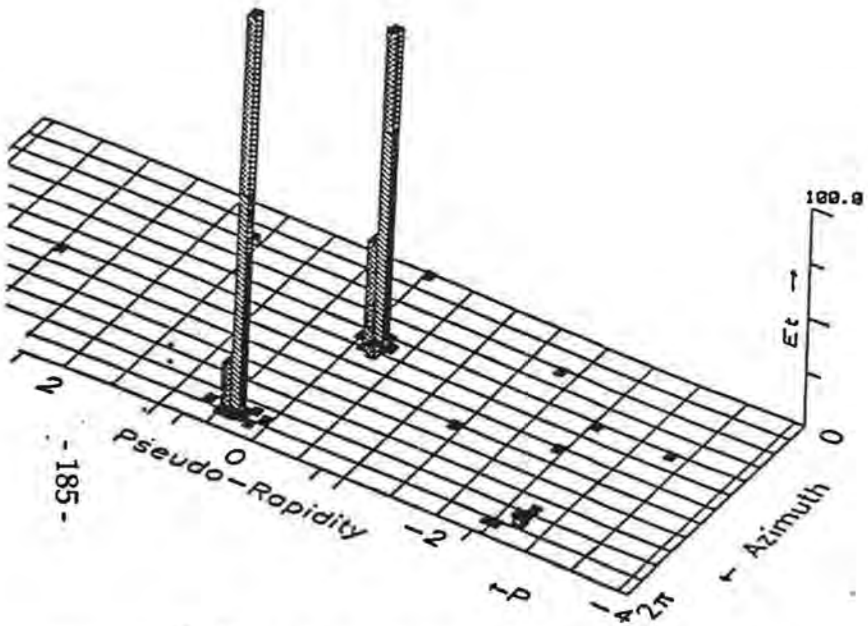
So result is sensitive to value of $\Lambda_{\overline{MS}}$ and Q

This uncertainty virtually disappeared thanks to $O(\alpha_s^3)$ calculation

Thus no free parameters left in comparison with data.

CDF, (Fermilab) $\sqrt{s} = 1.8$ TeV

- clear unambiguous identification of jet



3: One of the highest E_t jet events seen at the CDF detector, with
ariant mass of approximately 630 GeV.

l; only above 15 GeV one begins to have some confidence in clustering eans of jet identification. Figure 4 shows a typical two jet event at the t transverse energies seen at CDF. The existence of clusters of energy $\eta - \phi$ plot is unmistakeable. Figure 4 shows a typical four jet event. figure the jet energies range from 40 to 60 GeV, and, although still identifiable as isolated clusters, it is clear that they are substantially r in extent than the jets in figure 4.

e identification of jets is done using a clustering algorithm based on neter data in which one seeks to identify separated clusters of energy. lgorithms have been evaluated for their ability to find clusters in dense . The two most commonly used algorithms will be termed the *nearest or* and the *fixed cone* algorithms. The nearest neighbor algorithm uses tower above a threshold (2 GeV), and finds all contiguous towers nding the seed tower with energies above a shoulder threshold (0.2

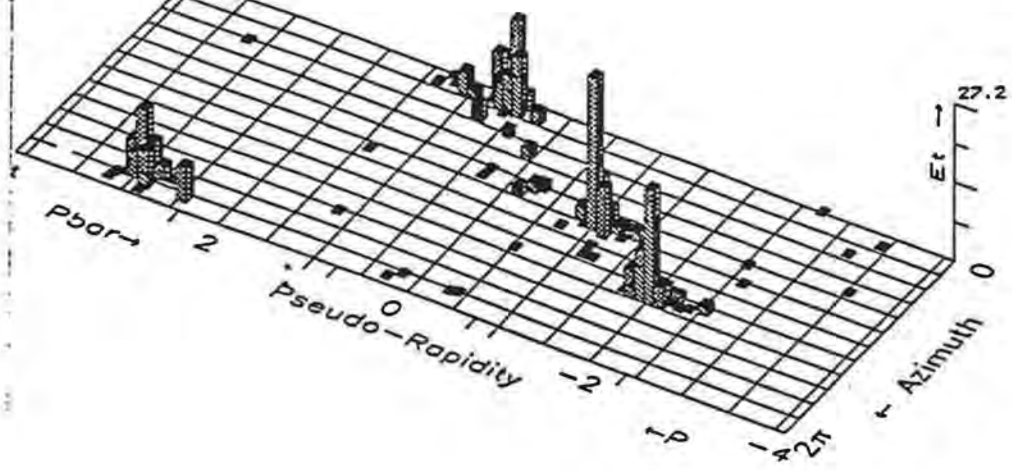


Figure 4: A four jet event as seen in the CDF detector. All 4 jets have E_t's in excess of 40 GeV.

fixed fraction, set to 2.0 by default. The merged towers are themselves used as seeds for a search of contiguous towers. The parent-daughter test allows jets to be defined around local maxima in the transverse energy. If a tower is not merged, it becomes the seed of another cluster.

The cone algorithm starts from the clusters defined in the contiguous tower algorithm before the merging step, and uses the E_t weighted centroid as the center of a circle (or cone) in $\eta - \phi$ space, with a radius $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$. All the towers above the shoulder threshold inside this circle are included in the cluster. The centroid is recomputed from these towers, a new circle is drawn and a new list of towers is generated. This process is iterated until the list of towers inside the cone is stable upon successive iterations. The fixed cone algorithm has the advantage that it is most closely related to the algorithms used by theorists to regulate collinear ingularities in the calculation of gluon *bremstrahlung* [5],[6].

The two other algorithms which we have investigated are the E_t depen-

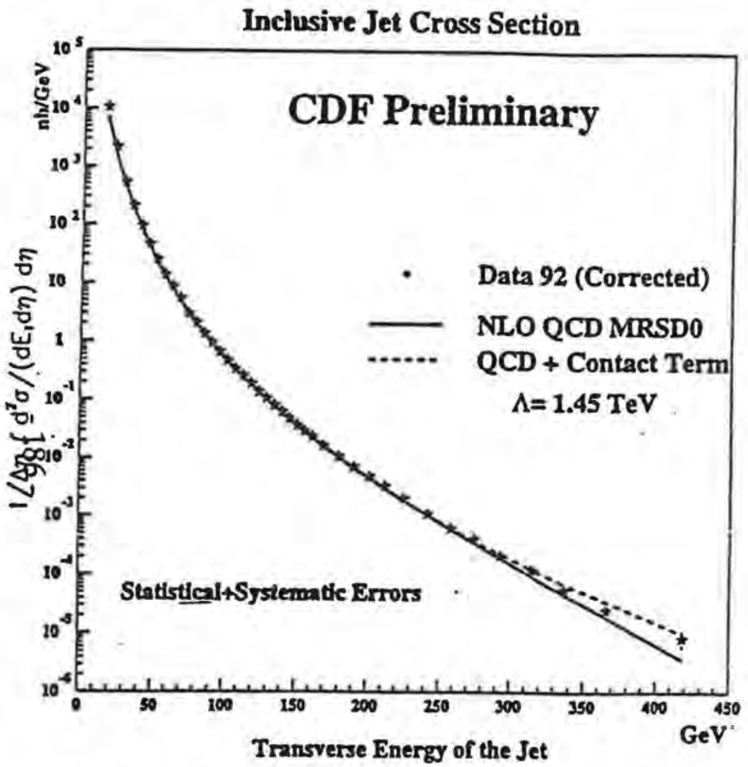


Figure 2: Comparison of Jet Cross Section with NLO QCD and with QCD+contact term predictions.

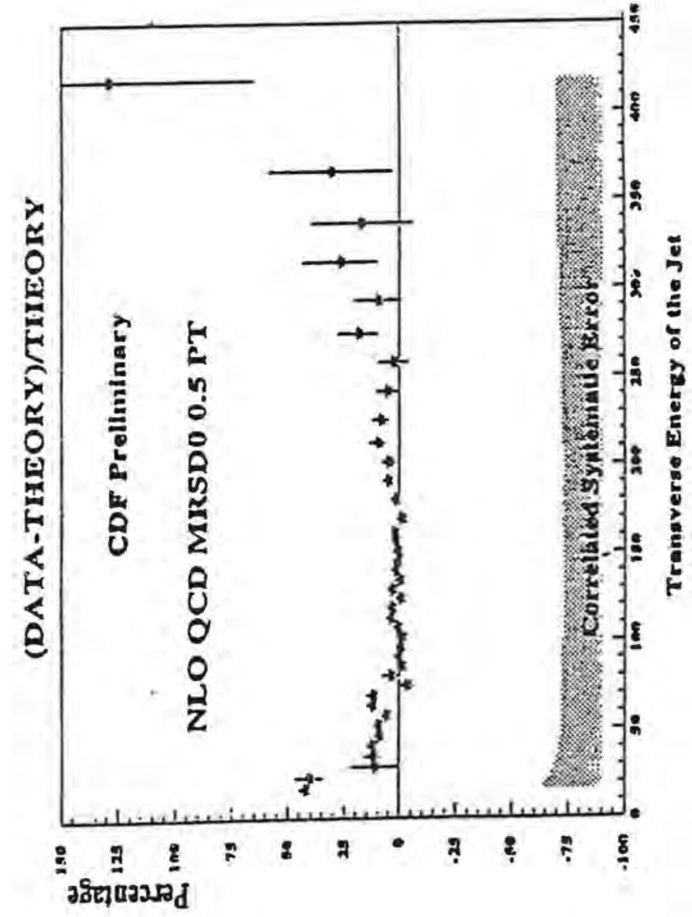


Figure 3: Comparison of Jet Cross Section with NLO QCD.

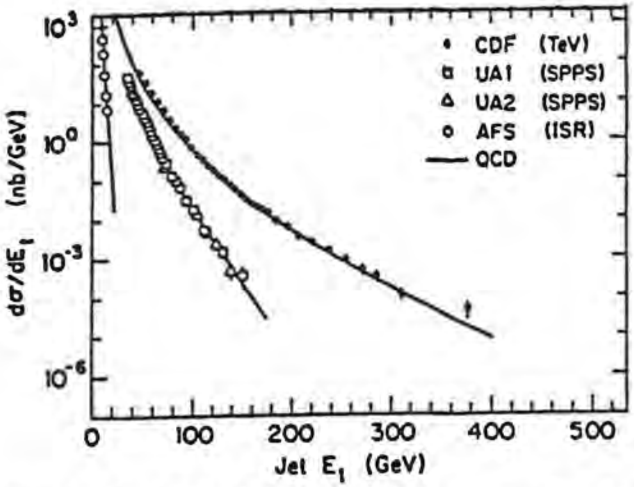


Figure 4: Inclusive jet E_t spectrum from the ISR, SPPS, and the Tevatron. Also shown is the prediction of a leading order QCD calculation made with the choice $Q^2 = E_t^2/2$ for all values of $\sqrt{s}$. The normalizations are absolute.

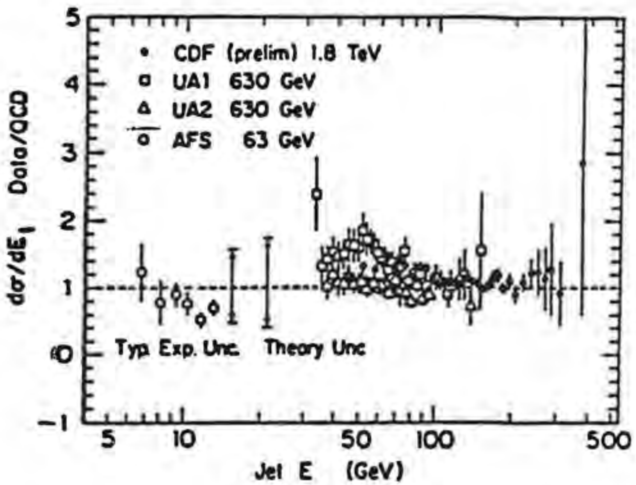


Figure 5: The ratio data/QCD for the data shown in figure 4. The error bars shown illustrate the typical uncertainty associated with both the theoretical calculations at leading order and the exper-

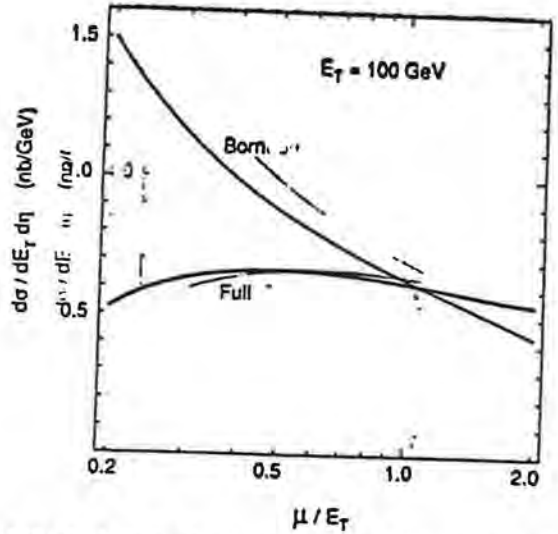


FIG. 2. Inclusive jet cross section $d\sigma/dE_T d\eta$ vs the ratio μ/E_T for $\sqrt{s}=1800$ GeV, $E_T=100$ GeV, $\eta=0$, and $R=0.6$.

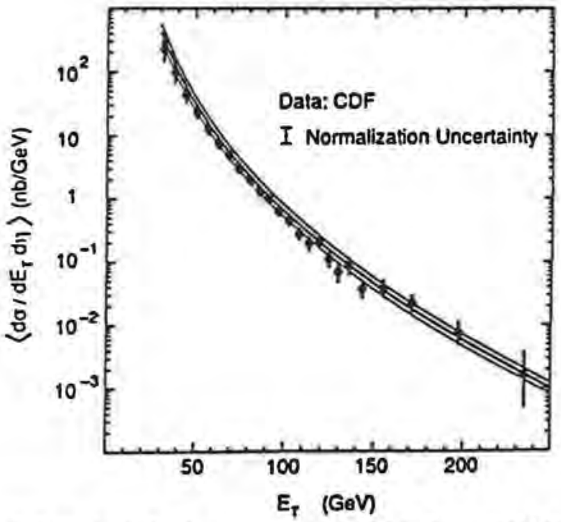


FIG. 3. Inclusive jet cross section $d\sigma/dE_T d\eta$ averaged over $0.1<|\eta|<0.7$ vs E_T at $\sqrt{s}=1800$ GeV for $R=0.6$ and $\mu=0.5E_T$. The data are from the CDF Collaboration (Ref. 1). The experimental error bars include both the energy-dependent systematic errors and the statistical errors as given in Ref. 1. The magnitude of the energy-independent systematic uncertainty is indicated by the single isolated error bar.

The $\hat{s}, \hat{t}, \hat{u}$ dependence of $|M|^2$ can be explored, measuring the angular distribution of the JET.

$$\frac{d\hat{\sigma}_{q\bar{q}}}{d\cos\hat{\theta}} = \frac{\pi \alpha_s^2}{2\hat{s}} \frac{4}{9} \left[\frac{\hat{s}^2 + \hat{u}^2}{\hat{t}^2} + \frac{\hat{s}^2 + \hat{t}^2}{\hat{u}^2} \right]$$

and the significant processes are $q\bar{q}$, qq, qg

elastic scattering

$$\hat{s} + \hat{t} + \hat{u} = 0 \qquad \hat{t} = \tfrac{1}{2}\hat{s}\,(1-\cos\hat{\theta})$$
$$\hat{u} = \tfrac{1}{2}\hat{s}\,(1+\cos\hat{\theta})$$

Putting $\chi = (1+\cos\hat{\theta})/(1-\cos\hat{\theta})$

-188-

$$\frac{d\hat{\sigma}_{q\bar{q}}}{d\chi} \sim \frac{1}{(1+\chi)^2} \left[\frac{1}{\chi^2} + \frac{1}{\chi} + 1 + \chi + \chi^2 \right]$$

(removes the $1/\sin^4\frac{\theta}{2}$ ...)

very similar for $gg \to gg$, $qg \to qg$

compare with CDF $\dfrac{d\sigma}{d\chi}$

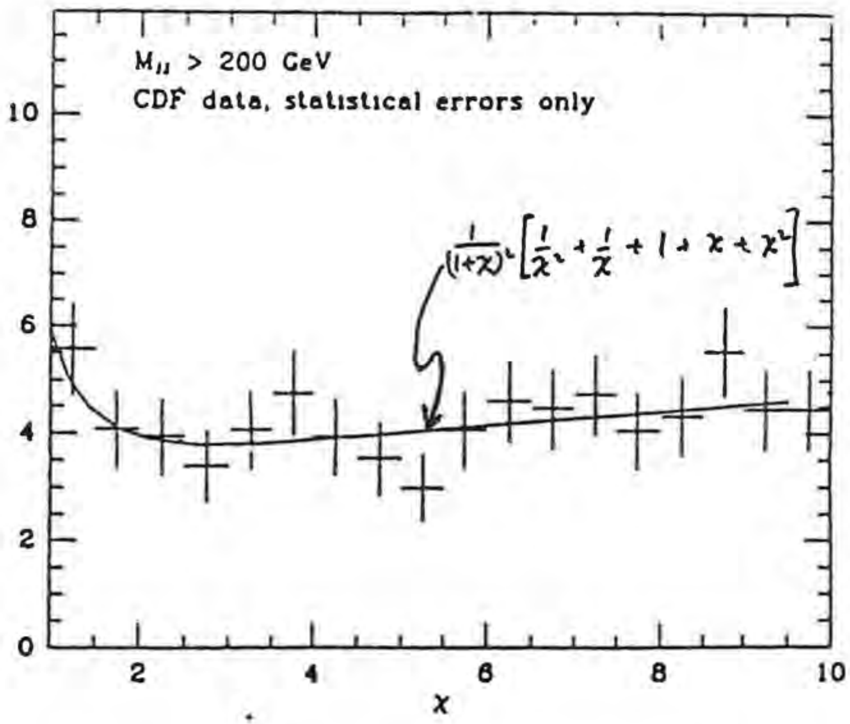
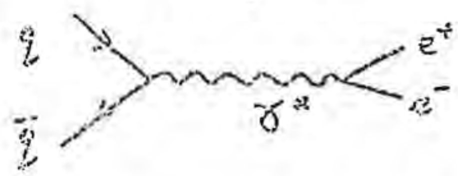


Figure 30: χ distribution from the CDF collaboration compared with the leadir order QCD prediction

[illegible] Production of Dileptons [illegible]

The Drell-Yan mechanism

Partonic sub-process



$q^2 = M^2_{e^+e^-} > 0$

$$\hat{\sigma}(q\bar{q} \to e^+e^-) = e_q^2 \frac{4\pi\alpha^2}{3\hat{s}} * \underbrace{3 \cdot \frac{1}{9}}_{\text{colour factor}}$$

3 = summing over colours

$\frac{1}{9}$ averaging over initial colour states.

-189-

So

$$\frac{d\hat{\sigma}}{dM^2_{ee}} = \frac{1}{3} e_q^2 \frac{4\pi\alpha^2}{3M^2} \delta(\hat{s} - M^2)$$

and

$$\frac{d\sigma}{dM^2}(AB \to e^+e^-) = \sum_i \int dx_1 dx_2 \left[q(x_1,\mu)\, \bar{q}(x_2,\mu) + (1 \leftrightarrow 2) \right] * \frac{d\hat{\sigma}}{dM^2}$$

Experimentally measure M (dilepton mass) and rapidity of dilepton pair y

$\Longrightarrow x_1, x_2$ can be found. (Problem!)

For $y = 0$ $\quad x_1 = x_2 = \tau = \frac{M}{\sqrt{s}}$

If we have pp collisions, dominant contribution is $\left[\frac{8}{9} u(x) + \frac{2}{9} d(x) + \frac{2}{9} S(x)\right] \cdot \underbrace{S(x)}_{\text{sea-quark}}$

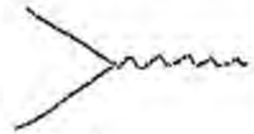
$\therefore$ SHAPE of the dilepton cross section very sensitive to shape of $S(x)$

— better way of fixing $S(x)$ than DIS

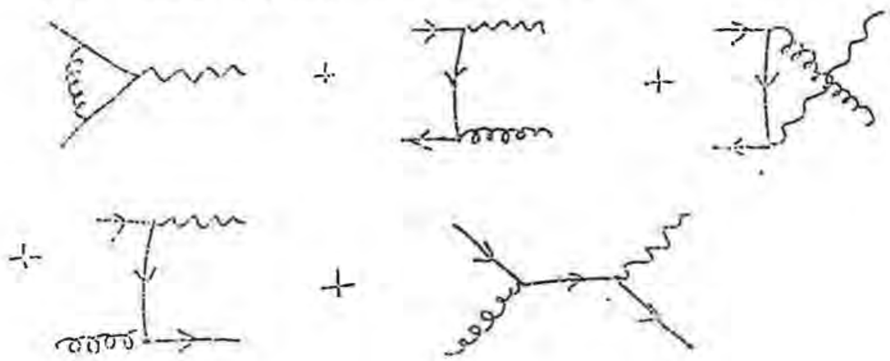
[illegible]

Next-to-leading order contributions

As well as



we have $O(\alpha_s)$ corrections :-



π^2 ... is not small

Including $O(\alpha_s)$ we have

$$\frac{d\sigma}{dM^2} = \frac{4\pi\alpha^2}{3M^2}\int_0^1 dx_1\, dx_2\, \delta(z x_1 x_2 - \tau) \times$$

$$\times \left\{ \left[\sum_i e_i^2 \left(q_i(x_1,\mu)\, \bar{q}_i(x_2,\mu) + (1 \leftrightarrow 2) \right)\right]\left[\delta(1-z) + \frac{\alpha_s(\mu)}{2\pi} f_q(z)\right]\right.$$

$$\left. + \left[\sum_i e_i^2 \left(g(x_1,\mu)\left(q_i(x_2,\mu) + \bar{q}_i(x_2,\mu)\right) + (1 \leftrightarrow 2)\right)\right] \frac{\alpha_s(\mu)}{2\pi} f_g(z) \right\}$$

$$f_q(z) = \frac{4}{3}\left[\delta(1-z)\left(1 + \frac{4\pi^2}{3}\right) - \ldots \right]$$

$$f_g(z) = \frac{1}{2}\left[\ldots \right]$$

Roughly: $\frac{d\sigma}{dM^2}$ is multiplied by a 'K-factor'

$$K \approx 1 + \frac{\alpha_s}{2\pi} \cdot \frac{4}{3}\left(1 + \frac{4}{3}\pi^2\right)$$

$$\approx 1.8 \quad !!$$

to the shape of the sea distribution q_s. A satisfactory description of the new, precise E605 data[9] is obtained from "BCDMS" parton distributions if $\eta_s = 7.2$ (see curve B' of Fig. 2) whereas the "EMC distributions" require a larger value of η_s to compensate for the flatter shape of $F_2^N(x)$. We have seen that such a large value of η_s would worsen the agreement with the deep inelastic data (χ^2 would increase by more than 70) and as a compromise we choose $\eta_s = 6.9$ for the "EMC" distributions. Then the description of the "EMC" deep inelastic set is not optimal

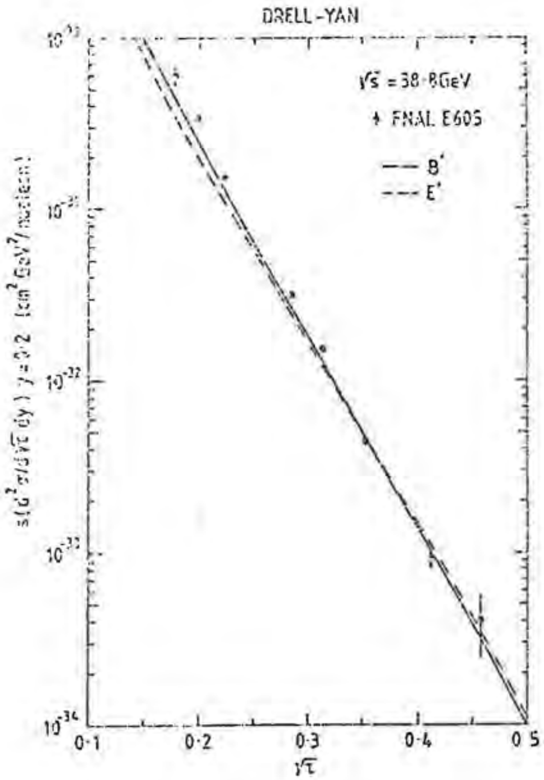


Fig. 2. The cross-section $s\, d^2\sigma/d\sqrt{\tau}dy$ at $y = 0.2$ for $pN \to \mu^+\mu^- X$ at $\sqrt{s} = 38.8$ GeV (with $N = 0.547n + 0.453p$). The data are from experiment E605.[9] The curves are the next-to-leading order values calculated using parton distributions E' (dashed curve) and B' (solid curve), respectively. The $O(\alpha_s^2)$ correction (which is not included) is estimated to increase the prediction by about $+O(20\%)$ approximately independent of $\sqrt{\tau}$.

More precisely, at fixed target energies
$K \approx 1.5 - 1.8$

But for W, Z prod. at colliders, [illegible]
$K \approx 1.3$

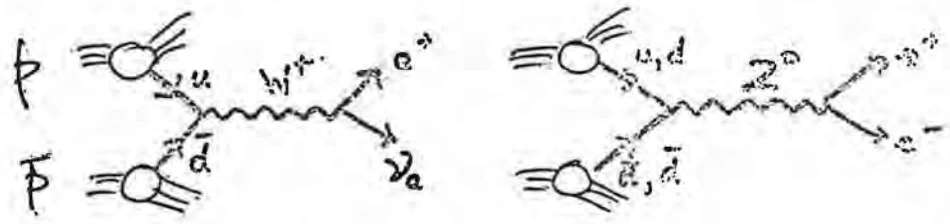
What about $O(\alpha_s^2)$ corrections ?

Experiments suggest $K_{expt} > K_{theory}^{O(\alpha_s)}$

probably another 30% from $O(\alpha_s^2)$

W, Z hadroproduction

Basically, Drell-Yan mechanism, [illegible]



The partonic cross-sections are:

$$\hat{\sigma}(q\bar{q}' \to W) = \frac{\pi}{3}\sqrt{2}\, G_F M_W^2 |V_{qq'}|^2\, \delta(\hat{s} - M_W^2)$$

← CKM mat. elt.

$$\hat{\sigma}(q\bar{q} \to Z^0) = \frac{\pi}{3}\sqrt{2}\, G_F M_Z^2 (V_q^2 + A_q^2)\, \delta(\hat{s} - M_Z^2)$$

$$V_q = t_{3L}(q) - 2Q_q \sin^2\theta_W$$
$$A_q = t_{3L}(q)$$

$O(\alpha_s)$ QCD correction same as before

Relevant x here is $\frac{M_{W,Z}}{\sqrt{s}}$

$= 0.12$ CERN
$= 0.044$ Fermilab
$= 0.005$ LHC
$= 0.002$ SSC

provides crucial test of small-x parton distributions – see fig.

What is measured is actually

$\sigma(p\bar{p} \to W) \times B.R.(W \to e\nu)$ ①

or $\sigma(p\bar{p} \to Z) \times B.R.(Z \to e^+e^-)$ ②

Before CDF ($m_t > 90$ GeV)

ratio $R = \frac{\text{no. of decays } W \to e\nu}{\text{〃 〃 〃 } Z \to e^+e^-}$

was used to give a limit on top quark mass. Also to limit N_ν before LEP

Fig ⇒ UA2, CDF data consistent with $N_\nu = 3$, $m_t > 90$ GeV

Ang. distr. of leptons in W, Z decay

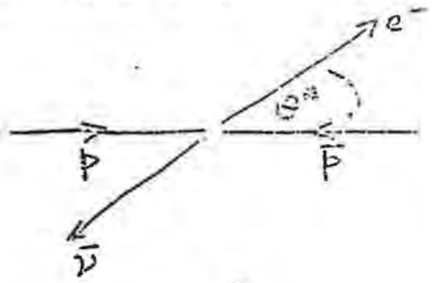
For $d + \bar{u} \longrightarrow e^- + \bar{\nu}$

we have

$$\overline{\sum} |M(d\bar{u} \to e^-\bar{\nu})|^2 = 64 \left(\frac{G_F M_W^2}{\sqrt{2}}\right)^2 |V_{ud}|^2 * \frac{(p_u \cdot p_e)^2}{\left[\left((p_u + p_d)^2 - M_W^2\right)^2 + M_W^2 \Gamma^2\right]}$$

similarly for $u + \bar{d} \to e^+ + \nu$

Define θ^*



$\therefore (p_u \cdot p_e)^2 = (1 + \cos\theta^*)^2$

ie outgoing electron prefers to go along p dirn.

(Basically ang. mom. conservation)

(W couples to −ve hel. fermions, +ve hel. anti-fermions)

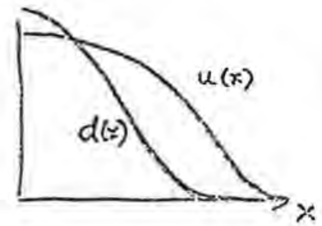
This lepton. asymmetry is seen.

Finally, asymmetry of the $W^\pm$ in the longitudinal direction:

There is an asymmetry between the W^+ and W^- simply because $u(x) \neq d(x)$ in the proton

$$p \Longrightarrow \begin{matrix} u \longrightarrow W^+ \longleftarrow \bar{d} \\ d \longrightarrow W^- \longleftarrow \bar{u} \end{matrix} \Longleftarrow \bar{p}$$

In proton:



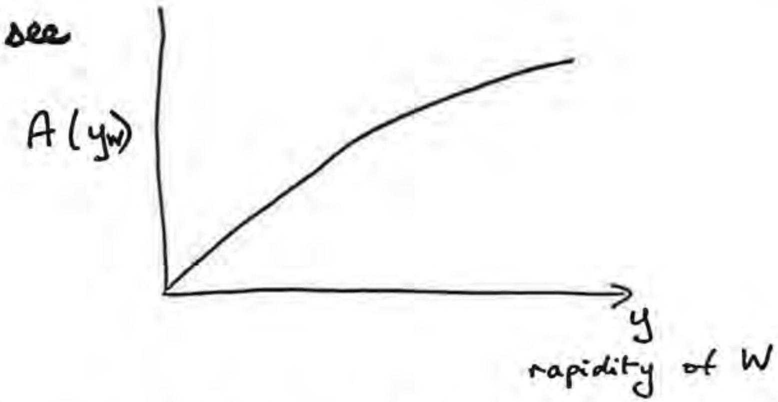
$\therefore$ W^+ prefers to follow the proton

W^- " " " " antiproton

Ideally, should measure

$$A(y) = \frac{\frac{d\sigma}{dy}(W^+) - \frac{d\sigma}{dy}(W^-)}{\frac{d\sigma}{dy}(W^+) + \frac{d\sigma}{dy}(W^-)}$$

and see



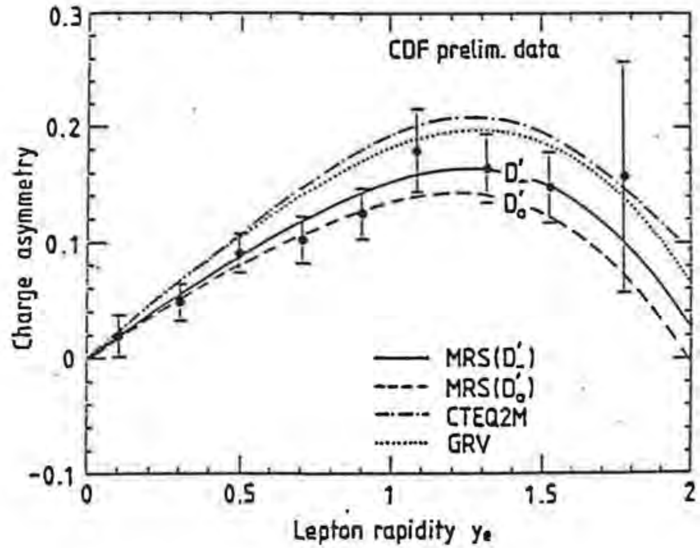
But can only see the electron (positron) — asymmetry $A(y_e)$ is measured.

$(1+\cos\theta^*)^2$ lepton ang. distbn dilutes the $W^\pm$ asymmetry. — CDF data.

probes smaller x as y increases.

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

At SSC, can ∴ probe $x \sim 10^{-5}$!!



Production of Heavy Quarks

Decays of heavy quarks.

Take free top quark in S.M.

Consider

For $m_t >> m_W$

$$\Gamma(t \to b W) = \frac{G_F m_t^3}{8\pi\sqrt{2}} |V_{tb}|^2 \simeq 170\,\mathrm{MeV}\left[\frac{m_t}{M_W}\right]^3$$

So if top really very heavy, width > hadronic scale

$\Rightarrow$ top decays <u>before</u> it hadronises.

All direct searches for top, assume something about B.R. → leptons.

$$p\bar{p} \longrightarrow t\bar{t} + X$$

→ 3 jets ($\bar{t} \to W^- \bar{b}$, $\to \ell\bar{\nu}$)

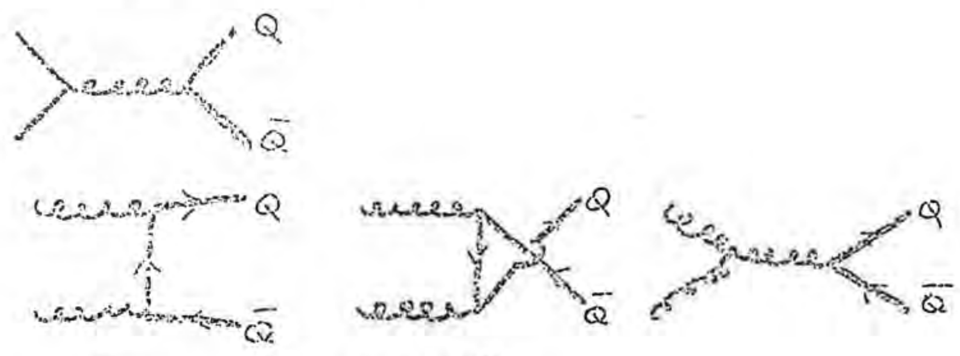
→ $b e\nu$ = jet + $e\nu$

B.R. = 1/9 in SM.

signal for top : look for $e\nu$ + n jets [illegible]

[illegible]

Theory of $Q\bar{Q}$ production



Graphs to $O(\alpha_s^2)$.

$1 + 2 \to 3 + 4$

Process	$\overline{\sum} \lvert M^2 \rvert^2 / g^4$
$q\bar{q} \to Q\bar{Q}$	$\frac{4}{9}\left(\tau_1^2 + \tau_2^2 + \frac{\rho}{2}\right)$
$gg \to Q\bar{Q}$	$\left(\frac{1}{6\tau_1\tau_2} - \frac{3}{8}\right)\left(\tau_1^2 + \tau_2^2 + \rho - \frac{\rho^2}{4\tau_1\tau_2}\right)$

with

$$d\hat{\sigma}_{ij} = \frac{1}{2\hat{s}} \frac{d^3p_3}{(2\pi)^3 2E_3} \frac{d^3p_4}{(2\pi)^3 2E_4} (2\pi)^4 \delta(p_1+p_2-p_3-p_4) \overline{\sum} |M_{ij}|^2$$

$$\Rightarrow \frac{d\sigma}{dy_3\,dy_4\,d^2p_T} = \frac{1}{16\pi^2\hat{s}^2}\sum_{ij} x_1 f_1(x_1,\mu)\, x_2 f_2(x_2,\mu)\,\overline{\sum}|M_{ij}|^2$$

for the full hadronic cross-section

$y_{3,4}$ rapidities of the $Q, \overline{Q}$.

$$p_1 = \frac{\sqrt{s}}{2}(x_1, 0, 0, x_1)$$
$$p_2 = \frac{\sqrt{s}}{2}(x_2, 0, 0, -x_2)$$
$$p_3 = (m_T\cosh y_3, p_T, 0, m_T\sinh y_3)$$
$$p_4 = (m_T\cosh y_4, -p_T, 0, m_T\sinh y_4)$$

with $m_T = \sqrt{m^2+p_T^2}$

Inversely
$$x_1 = \frac{m_T}{\sqrt{s}}(e^{y_3} + e^{y_4})$$
$$x_2 = \frac{m_T}{\sqrt{s}}(e^{-y_3} + e^{-y_4})$$
$$\hat{s} = 2m_T^2(1+\cosh\Delta y)$$

with $\Delta y = y_3 - y_4$

- then go back evaluate $\overline{\sum}|M_{ij}|^2$ in terms of $m^2, m_T^2, \Delta y$.

However,

4:

Note

1) Propagators in the graphs

e.g. $(p_1+p_2)^2 = 2p_1\cdot p_2 = 2m_T^2(1+\cosh\Delta y)$

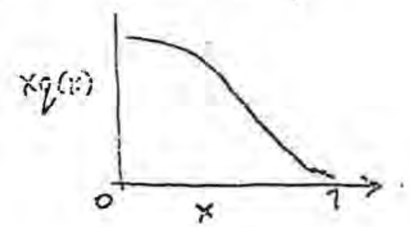
all off-shell by at least m^2

Contrast: light quark production, lower cut-off in props. $< \Lambda^2$

$\therefore$ Good chance that pert. QCD works for heavy quarks, σ controlled by α_s

2) Cross section suppressed for $p_T^2 \gg m^2$.

3) As m^2 increases, relevant $x \simeq \frac{m}{\sqrt{s}}$ also increases,



$\therefore$ cross section falls.

Dominant contribution when $p_T \simeq m$.

m sets the scale e.g. in $xg(x,\mu)$

Higher Order corrections and the predictions made for heavy quark production.

Graphs considered so far $O(\alpha_s^2)$

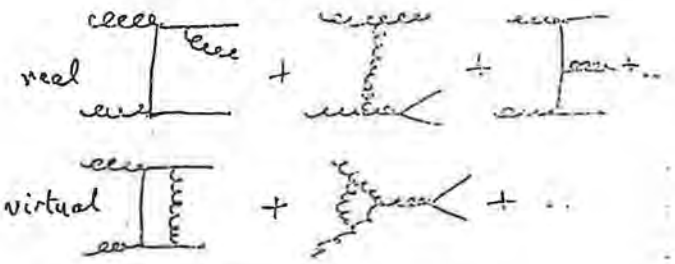
ie $\quad \hat{\sigma}_{ij}(\hat{s}, m^2) = \dfrac{\alpha_s^2(\mu)}{m^2} \tilde{\mathcal{F}}_{ij}\left(\rho, \dfrac{\mu^2}{m^2}\right)$

We have

$$\tilde{\mathcal{F}}_{ij}\left(\rho, \frac{\mu^2}{m^2}\right) = \tilde{\mathcal{F}}_{ij}^{(0)} + 4\pi\alpha_s(\mu)\left[\tilde{\mathcal{F}}_{ij}^{(1)}(\rho) + \bar{\tilde{\mathcal{F}}}_{ij}^{(1)}(\rho)\ln\frac{\mu^2}{m^2}\right]$$

from 4 graphs already seen.

contributions from



K. Ellis et al (1990)

Scale parameter μ. [illegible]

[illegible]

- 196 -

[illegible] pronounced dependence on μ [illegible] the pert. Q.D. [illegible]!

Example: hadroproduction of $t\bar{p}$ with $m_t = 120$ GeV.

Comparison of $O(\alpha_s^2)$, $O(\alpha_s^3)$ - Fig.

1) Calculation is 'stabalised' by including NLO

2) $O(\alpha_s^2)$ sufficient if $\mu \approx \frac{1}{2} m_t$

3) Similar exercise for bottom production less encouraging! - Fig

Limit on m_t from

$p\bar{p} \rightarrow t + \bar{t} + X \quad$ CDF.

Using $O(\alpha_s^3)$ can predict cross-section as a function of m_t. - fig.

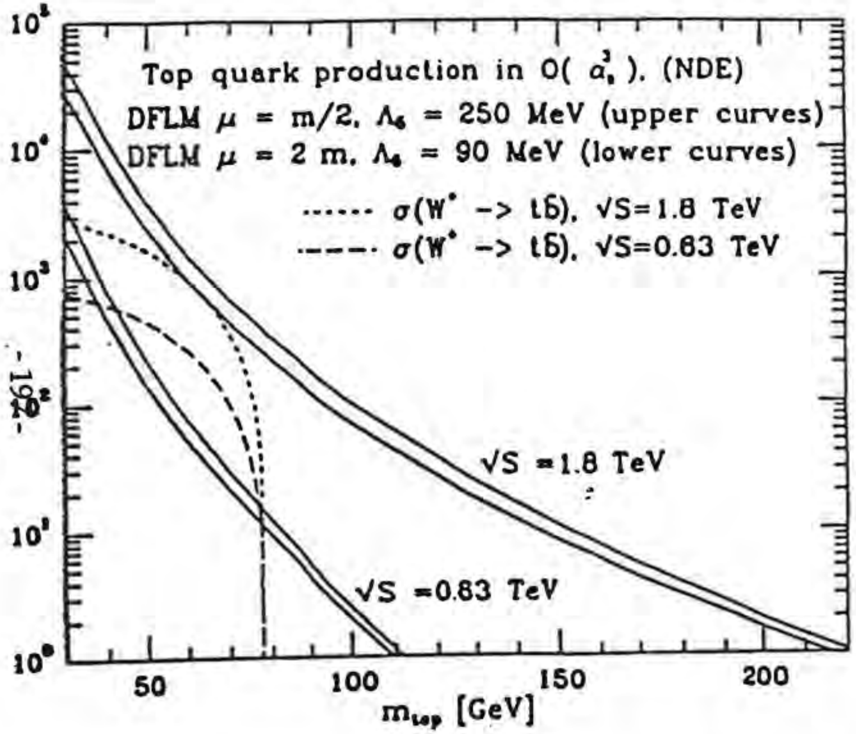


Figure 56: The cross section for top quark production at CERN and FNAL

ion, but the background is larger due to the process $p\bar{p} \to W + \text{jets}$. This backgroun

April 1994

FERMILAB-PUB-94/097-E
CDF/PUB/TOP/PUBLIC/2561

Evidence for Top Quark Production in $\bar{p}p$ Collisions at $\sqrt{s} = 1.8$ TeV

F. Abe,[13] M. G. Albrow,[7] D. Amidei,[16] J. Antos,[28] C. Anway-Wiese,[4] G. Apollinari,[26] H. Areti,[7] P. Auchincloss,[25] M. Austern,[14] F. Azfar,[21] P. Azzi,[20] N. Bacchetta,[18] W. Badgett,[16] M. W. Bailey,[24] J. Bao,[34] P. de Barbaro,[25] A. Barbaro-Galtieri,[14] V. E. Barnes,[24] B. A. Barnett,[12] P. Bartalini,[23] G. Bauer,[15] T. Baumann,[9] F. Bedeschi,[23] S. Behrends,[2] S. Belforte,[23] G. Bellettini,[23] J. Bellinger,[33] D. Benjamin,[32] J. Benlloch,[15] J. Bensinger,[2] D. Benton,[21] A. Beretvas,[7] J. P. Berge,[7] S. Bertolucci,[8] A. Bhatti,[26] K. Biery,[11] M. Binkley,[7] F. Bird,[29] D. Bisello,[20] R. E. Blair,[1] C. Blocker,[29] A. Bodek,[25] V. Bolognesi,[23] D. Bortoletto,[24] C. Boswell,[12] T. Boulos,[14] G. Brandenburg,[9] E. Buckley-Geer,[7] H. S. Budd,[25] K. Burkett,[16] G. Busetto,[20] A. Byon-Wagner,[7] K. L. Byrum,[1] C. Campagnari,[7] M. Campbell,[16] A. Caner,[7] W. Carithers,[14] D. Carlsmith,[33] A. Castro,[20] Y. Cen,[21] F. Cervelli,[23] J. Chapman,[16] M.-T. Cheng,[28] G. Chiarelli,[8] T. Chikamatsu,[31] S. Cihangir,[7] A. G. Clark,[23] M. Cobal,[23] M. Contreras,[5] J. Conway,[27] J. Cooper,[7] M. Cordelli,[8] D. P. Coupal,[29] D. Crane,[7] J. D. Cunningham,[2] T. Daniels,[15] F. DeJongh,[7] S. Delchamps,[7] S. Dell'Agnello,[23] M. Dell'Orso,[23] L. Demortier,[26] B. Denby,[23] M. Deninno,[3] P. F. Derwent,[16] T. Devlin,[27] M. Dickson,[25] S. Donati,[23] R. B. Drucker,[14] A. Dunn,[16] K. Einsweiler,[14] J. E. Elias,[7] R. Ely,[14] E. Engels, Jr.,[22] S. Eno,[5] D. Errede,[10] S. Errede,[10] Q. Fan,[25] B. Farhat,[15] I. Fiori,[3] B. Flaugher,[7] G. W. Foster,[7] M. Franklin,[9] M. Frautschi,[18] J. Freeman,[7] J. Friedman,[15] H. Frisch,[5] A. Fry,[29] T. A. Fuess,[1] Y. Fukui,[13] S. Funaki,[31] G. Gagliardi,[23] S. Galeotti,[23] M. Gallinaro,[20] A. F. Garfinkel,[24] S. Geer,[7] D. W. Gerdes,[16] P. Giannetti,[23] N. Giokaris,[26] P. Giromini,[8] L. Gladney,[21] D. Glenzinski,[12] M. Gold,[18] J. Gonzalez,[21] A. Gordon,[9] A. T. Goshaw,[6] K. Goulianos,[26] H. Grassmann,[6] A. Grewal,[21] G. Grieco,[23] L. Groer,[27] C. Grosso-Pilcher,[5] C. Haber,[14] S. R. Hahn,[7] R. Hamilton,[9] R. Handler,[33] R. M. Hans,[34] K. Hara,[31] B. Harral,[21] R. M. Harris,[7] S. A. Hauger,[6] J. Hauser,[4] C. Hawk,[27] J. Heinrich,[21] D. Hennessy,[6] R. Hollebeek,[21] L. Holloway,[10] A. Hölscher,[11] S. Hong,[16] G. Houk,[21] P. Hu,[22] B. T. Huffman,[22] R. Hughes,[25] P. Hurst,[9] J. Huston,[17] J. Huth,[9] J. Hylen,[7] M. Incagli,[23] J. Incandela,[7] H. Iso,[31] H. Jensen,[7] C. P. Jessop,[9] U. Joshi,[7] R. W. Kadel,[14] E. Kajfasz,[7a] T. Kamon,[30] T. Kaneko,[31] D. A. Kardelis,[10] H. Kasha,[34] Y. Kato,[19] L. Keeble,[30] R. D. Kennedy,[27] R. Kephart,[7] P. Kesten,[14] D. Kestenbaum,[9] R. M. Keup,[10] H. Keutelian,[7] F. Keyvan,[4] D. H. Kim,[7] H. S. Kim,[11] S. B. Kim,[16] S. H. Kim,[31] Y. K. Kim,[14] L. Kirsch,[2] P. Koehn,[25] K. Kondo,[31] J. Konigsberg,[9] S. Kopp,[5] K. Kordas,[11] W. Koska,[7] E. Kovacs,[7a] W. Kowald,[6] M. Krasberg,[16] J. Kroll,[7] M. Kruse,[24] S. E. Kuhlmann,[1] E. Kuns,[27] A. T. Laasanen,[24] S. Lammel,[4] J. I. Lamoureux,[33] T. LeCompte,[10] S. Leone,[23] J. D. Lewis,[7] P. Limon,[7] M. Lindgren,[4] T. M. Liss,[10] N. Lockyer,[21] O. Long,[21] M. Loreti,[20] E. H. Low,[21]

Glasgow '94

For each event, there are 2 ways of assigning the tagged b-jet, 3 ways of choosing another jet to be the other b, and 2 solutions for the neutrino longitudinal momentum:

12 possible configurations per event.

Select solution with smallest χ^2. If Mtop > 260 GeV/c^2, select next higher χ^2 solution.

Check procedure on Monte Carlo $t\bar{t}$ events.

As a function of top mass, fit the data to a sum of top and background mass distributions. The largest likelihood is for

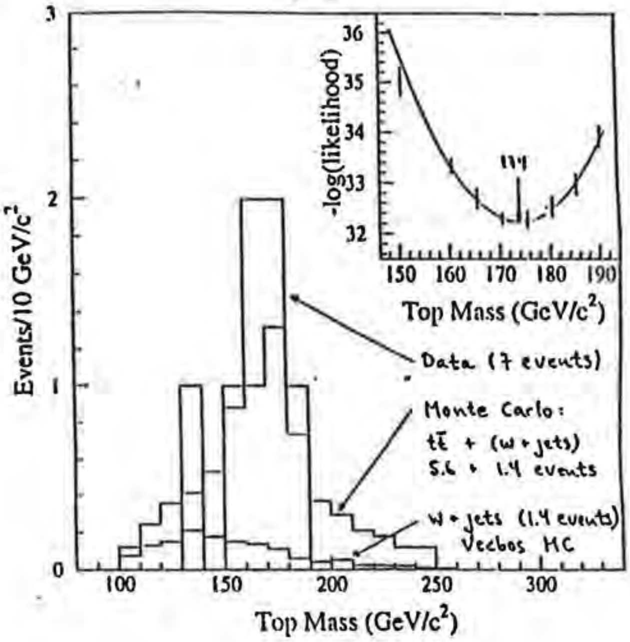
$M_{top} = 174 \pm 10 ^{+13}_{-12}$ GeV/c^2 (7 events)

where the systematic error includes the mass shift from removing two events at random.

For comparison, the average mass of the 7 events is 166 GeV/c^2, while the average of the 6 highest mass events is 172 GeV/c^2.

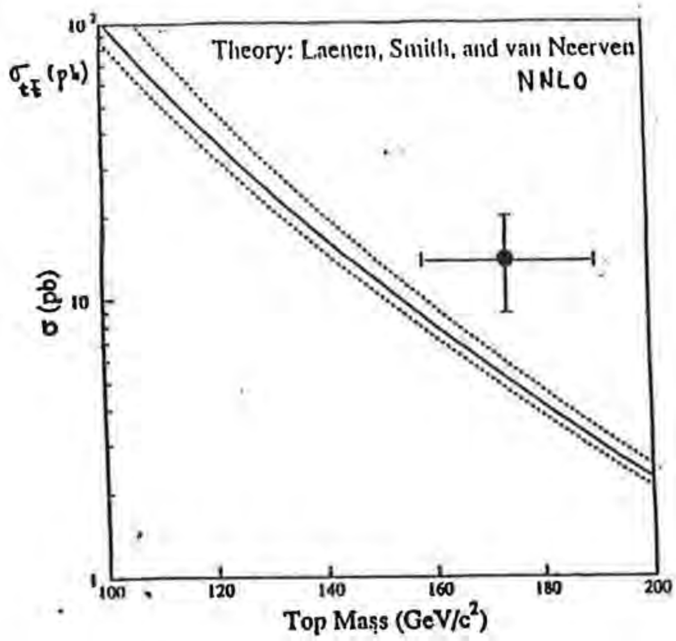
A fit with the W + jets spectrum alone is 50 times less likely.

Top Mass Fit to Seven W + jets Events



Glasgow '94

Top Mass and Cross Section



MASS DETERMINATION

In the b-tagged W + ≥ 3 jets sample, require fourth jet with E_t > 8 GeV and η < 2.4. This leaves 7 events (out of 10), with expected background of $1.4^{+2.0}_{-1.1}$ events.

Correct observed jet energies back to parton level (use Monte Carlo to find correction).

Fit each of the 7 events in turn with SQUAW:

$$\bar{p}p \to t_1 + t_2 + X$$

$$t_1 \to b_1 + W_1$$

$$t_2 \to b_2 + W_2$$

$$W_1 \to e + \nu \quad (\text{or } \mu + \nu)$$

$$W_2 \to j_1 + j_2$$

11*4 = 44 parameters, 20 equations and 26 measured (or known) quantities = 2 constraint fit (think of $M_{jj} = M_W$ and $M_{t1} = M_{t2}$).

W, Z Physics

The Lagrangian for (fermion-gauge) sector of the Standard Model

$$\mathcal{L} = \frac{1}{2\sqrt{2}}\, g\,(1-\gamma_5)\left[W_+^{\mu} J_{-\mu} + W_-^{\mu} J_{+\mu}\right]$$

$$+\left[(g\cos\theta_W + g'\sin\theta_W)(1-\gamma_5)J_3^{\mu} - g\sin\theta_W\, J_{em}^{\mu}\right] Z_\mu$$

$$+\left[\underline{g'\cos\theta_W}\, J_{em}^{\mu} + (\underline{g'\cos\theta_W - g\sin\theta_W})(1-\gamma_5)J_3^{\mu}\right] A_\mu$$

$$J_{\pm}^{\mu} = 2(J_1^{\mu} \mp i J_2^{\mu}) \qquad \text{charged currents}$$

$$J^{\mu} = J_{em}^{\mu} - J_3^{\mu} \qquad U(1) \text{ current}$$

$\rightarrow A_\mu$ = photon

∴ unification condition

$$e = g'\cos\theta_W = g\sin\theta_W$$

$$\mathcal{L}_{eff}^{weak} = \frac{G_F}{\sqrt{2}}\left[J_{\pm}^{\mu} J_{\mp\mu} + \rho\, J_{NC}^{\mu} J_{NC\,\mu}\right]$$

∴ neutral W.I. part of $\mathcal{L}$

$$= \mathcal{L}_{NC} = \frac{e\rho}{2\cos\theta_W \sin\theta_W}\, J_{NC}^{\mu} Z_\mu$$

with
$$J_{NC}^{\mu} = 2\left((1-\gamma_5)J_3^{\mu} - \sin^2\theta_W\, J_{em}^{\mu}\right)$$

for each fermion f we get

$$J_{NC}^{\mu} = 2\left[\bar{f}\gamma^{\mu} t_3 \tfrac{1}{2}(1-\gamma_5) f - \sin^2\theta_W\, \bar{f}\gamma^{\mu} Q_f f\right]$$

$$\equiv 2\bar{f}\left[\gamma^{\mu} V_f + \gamma^{\mu}\gamma_5 A_f\right] f$$

Vector : $V_f = \frac{1}{2}(t_3)_{f_L} - Q_f \sin^2\theta_W$

axial vector : $A_f = -\frac{1}{2}(t_3)_{f_L}$

Note: <u>both</u> left and right handed fermions couple to Z.

$$J_{NC}^{\mu} = 2\bar{f}\left[\underbrace{(V_f - A_f)}_{g_f^L}\tfrac{1}{2}(1-\gamma_5)\gamma^{\mu} + \underbrace{(V_f + A_f)}_{g_f^R}\tfrac{1}{2}(1+\gamma_5)\gamma^{\mu}\right]$$

In minimal (ONE HIGGS DOUBLET) $SU(2) \times U(1)$

3 fundamental parameters

$$g \qquad g' \qquad v = \langle 0 | \Phi | 0 \rangle$$

but usually choose

$$\alpha_{em} \qquad G_F \qquad \sin^2\theta_W$$

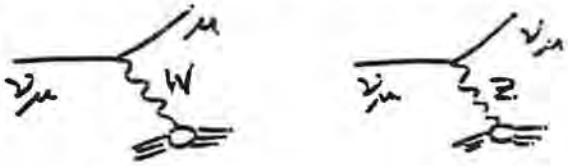
Josephson effect $\Rightarrow$ $\alpha_{em}^{-1} = 137.03604(11)$

μ lifetime $\Rightarrow$ $G_F = 1.16637(2) \times 10^{-5}\ \mathrm{GeV}^{-2}$

$\sin^2\theta_W$ — can be measured in several ways.

Most precise use neutrino scattering.

① DIS

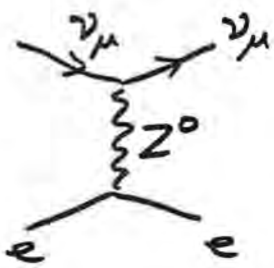


CC NC

$$\frac{\sigma_{\nu N}^{NC}}{\sigma_{\nu N}^{CC}} = \frac{1}{2} - \sin^2\theta_W + \frac{20}{27}\sin^4\theta_W$$

CHARM + CDHS $\Rightarrow$ $\sin^2\theta_W = 0.231 \pm \underset{\text{exp}}{0.004} \pm \underset{\text{theor.}}{0.005}$

② νe scattering



$$\frac{\sigma(\nu_\mu e^-)}{\sigma(\bar{\nu}_\mu e^-)} = 3\,\frac{1 - 4\sin^2\theta_W + \frac{16}{3}\sin^4\theta_W}{1 - 4\sin^2\theta_W + 16\sin^4\theta_W}$$

SHOW THIS

CHARM (II) $\Rightarrow$ $0.239 \pm \underset{\text{stat}}{0.009} \pm \underset{\text{sys.}}{0.007}$ no rad. corr.

0.237 " " with rad. corr.

note (i) radiative corrections small

(ii) relevant 'scale' for $\sin^2\theta_W$ is q^2 in Z exch.

<u>$\sin^2\theta_W$</u>, <u>m_{top}</u>, <u>M_{Higgs}</u> <u>+ Radiative Correcti</u>

$\sin^2\theta_W$ appears in several places.

1) $M_W / M_Z = \cos\theta_W$

2) $$\frac{G_F}{\sqrt{2}} = \frac{e^2}{8\sin^2\theta_W M_W^2}$$

3) Unification Condition

$$e = g \sin\theta_W = g' \cos\theta_W$$

4) Neutral Current

$$J^{\mu}_{NC} = 2\left(J^{\mu}_{3} - \sin^2\theta_W \, J^{\mu}_{em}\right)$$

Q. Does it matter which we take as definition of $\sin^2\theta_W$?

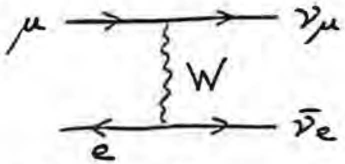
A. No — if we work to only lowest order (tree level)

Yes — if we include higher orders (loop graphs)

$\sin^2\theta_W$, Radiative Corrections and m_{top}, m_{Higgs}

Consider $\mu \to e \, \nu_\mu \, \bar{\nu}_e$

Relevant tree graph



$$\mathcal{M} = i\left(\frac{g}{2\sqrt{2}}\right)^2 \frac{J^{(\mu)}_{cc} \, J^{(e)}_{cc}}{q^2 - M_W^2}$$

$$J^{(\mu)}_{cc} = \bar{\nu}_\mu \gamma_\mu (1-\gamma_5)\mu$$

But, old Fermi model $\Rightarrow$

$$\mathcal{M} = i \frac{G_\mu}{\sqrt{2}} J^{(\mu)}_{cc} \, J^{(e)}_{cc}$$

For $q^2 \ll M_W^2 \Rightarrow \dfrac{G_\mu}{\sqrt{2}} = \dfrac{g^2}{8M_W^2} = \dfrac{e^2}{8 s_W^2 M_W^2}$

$$\Rightarrow \quad M_W^2 = \frac{\pi\alpha}{\sqrt{2}\,G_\mu} \cdot \frac{1}{s_W^2}$$

$$M_Z^2 = \frac{\pi\alpha}{\sqrt{2}\,G_\mu} \cdot \frac{1}{s_W^2 c_W^2}$$

Take. $\alpha^{-1} = 137.030604(1)$
$G_\mu = 1.16637(2) \times 10^{-5}\ \mathrm{GeV}^{-2}$
$s_W^2 = 0.231 \pm 0.006$ νN

		EXP
$\Rightarrow$	$M_W = 77.6 \pm 1.0$ GeV	80.6 ± 0.4
	$M_Z = 88.5 \pm 0.8$ GeV	91.16 ± 0.03

higher orders must be included ! $\Leftarrow$ deviation of more than 2σ

$$s_W^{o\,2} = 1 - \frac{M_W^2 + \delta M_W^2}{M_Z^2 + \delta M_Z^2}$$
$$= s_W^2 + c_W^2\left(\frac{\delta M_Z^2}{M_Z^2} - \frac{\delta M_W^2}{M_W^2}\right)$$

So
$$\frac{G_\mu}{\sqrt{2}} = \frac{e^2}{8 s_W^2 M_W^2}\left[1 + 2\frac{\delta e}{e} - \frac{c_W^2}{s_W^2}\left(\frac{\delta M_Z^2}{M_Z^2} - \frac{\delta M_W^2}{M_W^2}\right) - 2\right.$$

all renormalised quantities.

$$= \frac{e^2}{8 s_W^2 M_W^2}\left[1 + \Delta r\right]$$

Bare quantities

e^o g^o M_W^o M_Z^o
s_W^o

$$s_W^{o\,2} = 1 - \frac{M_W^{o\,2}}{\rho M_Z^{o\,2}}$$

Renormalised quantities

e g M_W M_Z
s_W

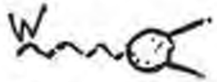
$$e^o = e + \delta e$$

$$g^o = g + \delta g$$

$$M_W^{o\,2} = M_W^2\left(1 + \frac{\delta M_W^2}{M_W^2}\right)$$

$$M_Z^{o\,2} = M_Z^2\left(1 + \frac{\delta M_Z^2}{M_Z^2}\right)$$

$\Sigma^\gamma(q^2) = q^2\Pi^\gamma(q^2)$ photon vac. poln.



$\Sigma^W(q^2)$

$\Sigma^Z(q^2)$

Contributions to Δr

Consider e.g. photon propagator

fermion, mass m

$$D^\gamma_{\mu\nu} = -i g_{\mu\nu}\frac{1}{q^2} + i g_{\mu\nu}\frac{1}{q^2}\Sigma^\gamma(q^2)\frac{1}{q^2}$$
$$= -i g_{\mu\nu}\frac{1}{q^2}\left\{1 - \Pi^\gamma(q^2)\right\}$$

$$q^2\Pi^\gamma(q^2) = \Sigma^\gamma(q^2) = \frac{\alpha}{3\pi}\left\{ q^2\left(\Delta - \ln\frac{m^2}{\mu^2}\right) - (q^2 + 2m^2)\int_0^1 dx \ln\left[\frac{x^2q^2 + xq^2 + m^2}{m^2} - i\epsilon\right] - q^2/3 \right]$$

artefact of dim. reg

Special Cases

(i) $q^2 = 0$ $\quad \Pi^\gamma(0) = \frac{\alpha}{3\pi}\left(\Delta - \ln\frac{m^2}{\mu^2}\right)$

(ii) $|q^2| >> m^2$ $\boxed{\text{light fermions}}$

$$\hat{\Pi}^\gamma(q^2) = \Pi^\gamma(q^2) - \Pi^\gamma(0) = \frac{\alpha}{3\pi}\left[\frac{5}{3} - \ln\frac{|q^2|}{m^2} + i\theta(q^2)\right]$$

(i) $|q^2| << m^2$ $\boxed{\text{heavy fermions}}$

$$\hat{\Pi}^\gamma(q^2) = \frac{\alpha}{3\pi}\cdot\frac{q^2}{5m^2}$$

Now

$$\Delta r = 2\frac{\delta e}{e} - \frac{c_w^2}{s_w^2}\left(\frac{\delta M_z^2}{M_z^2} - \frac{\delta M_w^2}{M_w^2}\right) - 2\frac{\delta g}{g}$$

m_t, m_H dependence here

$$\frac{\delta e}{e} = \frac{1}{2}\Pi^\gamma(0) - \frac{s_w}{c_w}\frac{\Sigma^{\gamma Z}(0)}{M_z^2}$$

$$\Pi^\gamma(0) = -\mathrm{Re}\,\hat{\Pi}^\gamma(M_z^2) - \frac{\alpha}{3\pi}\frac{c_w^2 - s_w^2}{4s_w^2}\ln c_w^2$$

$$\hat{\Pi}^\gamma(M_z^2) = \frac{\alpha}{3\pi}\sum_f Q_f^2\left(\frac{5}{3} - \ln\frac{M_z^2}{m_f^2} + i\pi\right)$$

summing over $e, \mu, \tau, u, d, s, c, b$

$$\Rightarrow \mathrm{Re}\,\hat{\Pi}^\gamma(M_z^2) = \begin{matrix}-0.0602\\ \pm 0.0009\end{matrix}$$

Same contribution to Δr as the running of α

$$\alpha(M_z^2) = \frac{\alpha}{1 + \mathrm{Re}\,\Pi^\gamma(M_z^2)}$$

ie $\alpha^{-1}(M_z^2) = 128.8$ not 137

This is the light fermion contrib. to Δr

<u>Heavy quark</u> (top) contributions

Keep only singular parts, and quadratic te

m_t^2 enters in $\frac{\delta M_z^2}{M_z^2} - \frac{\delta M_w^2}{M_w^2}$

Since ρ is defined as $\rho = \frac{M_w^2}{M_z^2 c_w^2}$

$\rho^0 = \frac{M_w^{0\,2}}{M_z^{0\,2} c_w^{0\,2}} = 1$ by defn.

$$\therefore \rho = \rho^0\left(1 + \left(\frac{\delta M_z^2}{M_z^2} - \frac{\delta M_w^2}{M_w^2}\right) + \cdots\right.$$

$$= 1 + \Delta\rho$$

$$\frac{\delta M_z^2}{M_z^2} - \frac{\delta M_w^2}{M_w^2} = \frac{3\alpha}{16\pi s_w^2 c_w^2}\frac{m_t^2}{M_z^2} = \frac{3\sqrt{2}G_\mu}{16\pi^2}m_t^2$$

We have

$$\Delta r = \Delta\alpha - \frac{c_w^2}{s_w^2}\Delta\rho + (\Delta r)_{\text{remainder}}$$

$$(\Delta r)^{top}_{remainder} = \frac{\alpha}{4\pi s_W^2}\left(\frac{c_W^2}{s_W^2} - \frac{1}{3}\right) \ln \frac{m_t}{M_Z}$$

$$(\Delta r)^{Higgs}_{remainder} \cong \frac{\alpha}{16\pi s_W^2} \cdot \frac{11}{3}\left(\ln \frac{M_H^2}{M_W^2} - \frac{5}{6}\right)$$

Errors!

Bardin, Riemann, Rieman

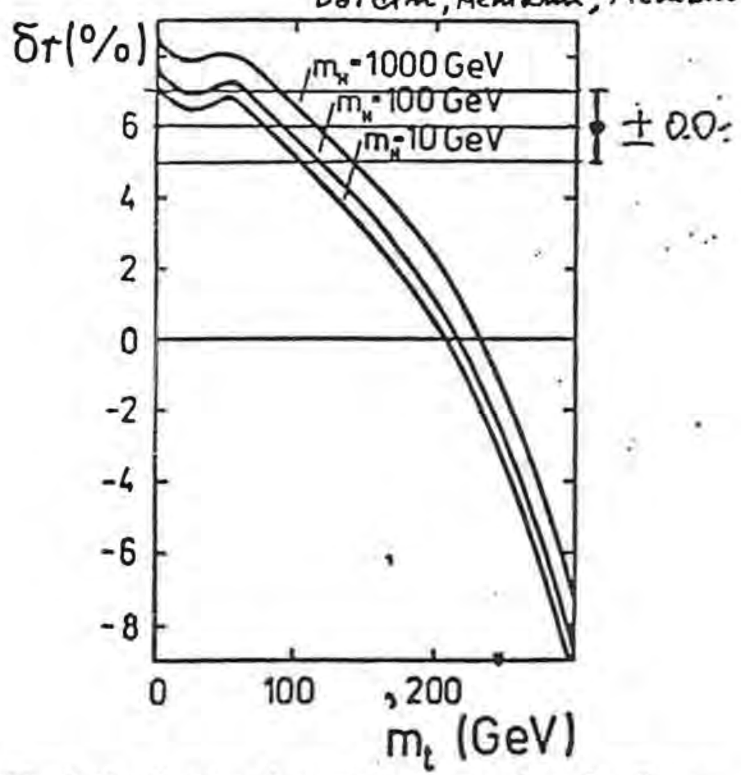


Fig. 1. The electroweak correction factor δr as function c

cross sections and to define Born approxi
The δr perturbatively connects α and (
$\alpha \equiv \alpha(0)$ is a truly low energy coupling c
... if measured from a low

Summary

Δr ⟵ light fermions
+6% contrib.

top quark

$$\sim -\frac{c_W^2}{s_W^2}\frac{3\sqrt{2}G}{16\pi^2} m_t^2$$

$$\sim -4.2\times 10^{-6}\left(\frac{m_t^2}{1\,\mathrm{GeV}^2}\right)$$

$$\sim -0.07 \quad (m_t = 130\,\mathrm{GeV})$$

Higgs
only a weak logarithmic
dependence

Altarelli SLAC Conf. Aug 89

$\therefore$ $\sin^2\theta_W$ measured from $\frac{M_W}{M_Z}$

is sensitive to value of m_t

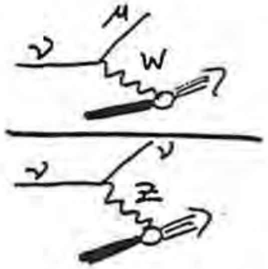
Latest values $\qquad M_W / M_Z$

CDF	0.8764 ± 0.0032
UA2	0.8813 ± 0.0055

$$\Longrightarrow \quad \sin^2\theta_W = 0.2275 \pm 0.0052$$

BUT : $\sin^2\theta_W$ measured in deep inelastic scattering only weakly dependent on m_t.

$$R_\nu = \frac{\sigma^{NC}}{\sigma^{CC}} =$$



$$= \frac{1}{2} - s_W^2 + \frac{5}{9} s_W^4 \left(1 + \frac{\sigma^{\bar\nu N}}{\sigma^{\nu N}}\right)$$

$$\Longrightarrow \quad s_W^2 = 0.236 \pm .005 \pm .005 \quad \text{CHARM}$$

$$= .2275 \pm .005 \pm .005 \quad \text{CDHS}$$

$M_H = 100$
$M_t = 100$

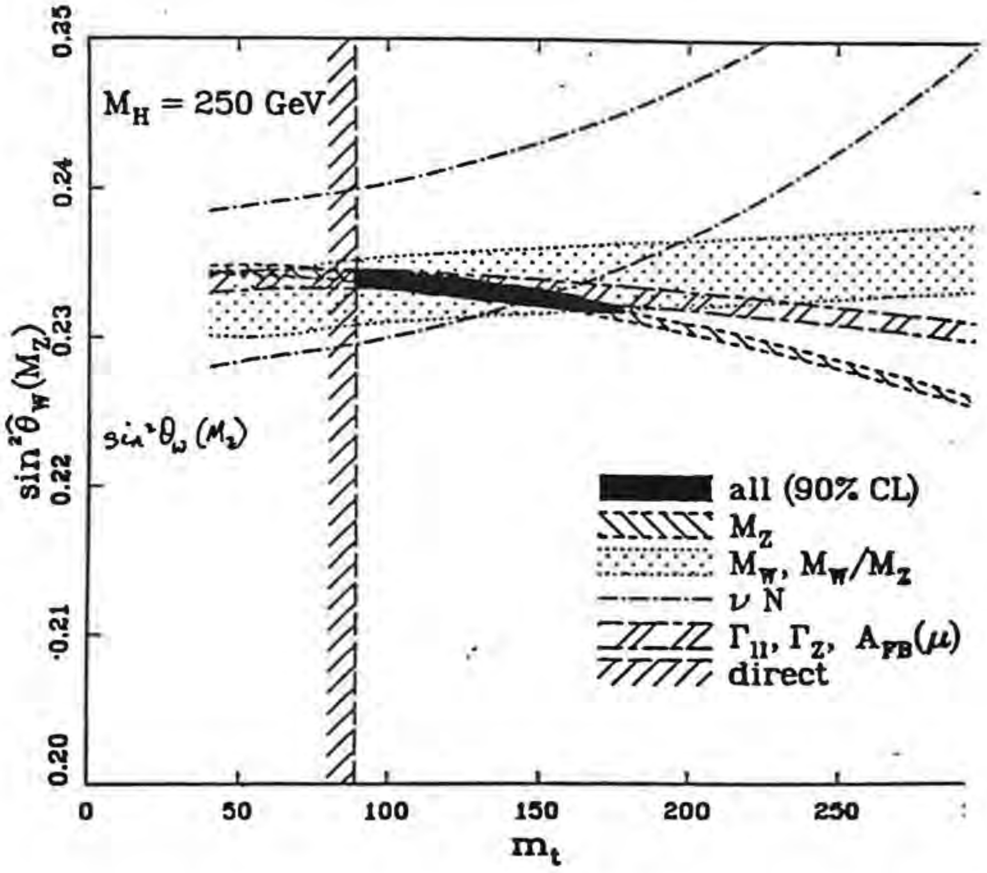
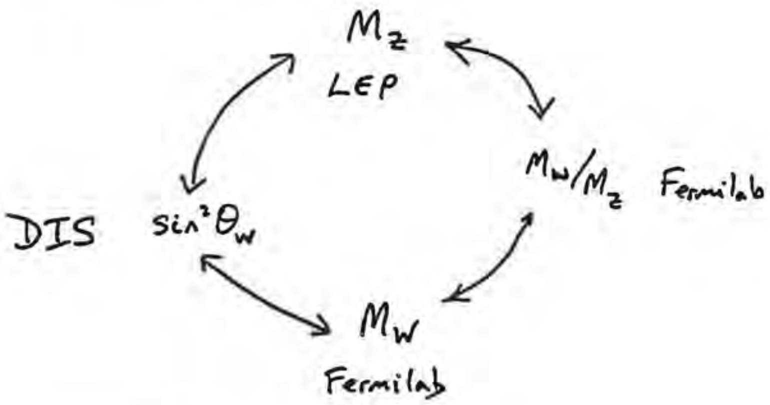


Figure 63: One σ uncertainty is $\sin^2\hat\theta_W(M_Z)$ as a function of m_t f $\sin^2\hat\theta_W(M_Z)$ determined from various inputs for $M_H = 250$ GeV. T direct lower limit $m_t > 89$ GeV and the 90% CL fit to all data are al shown. From [156].

So measuring



$\Longrightarrow$ LIMITS ON M_{TOP}

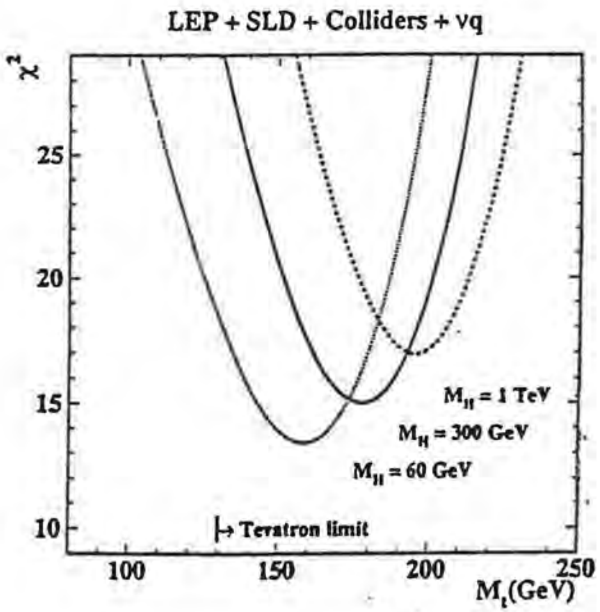
Together with information on Γ_Z, A_{FB} ...
(see later....)

$\Longrightarrow$

	LEP	LEP + Collider and ν data	LEP + Collider and ν data + A_{LR} from SLC
M_t (GeV)	$173^{+12}_{-13}\,^{+18}_{-20}$	$171^{+11}_{-12}\,^{+18}_{-19}$	$178^{+11}_{-11}\,^{+18}_{-19}$
$\alpha_s(M_Z^2)$	$0.126 \pm 0.005 \pm 0.002$	$0.126 \pm 0.005 \pm 0.002$	$0.125 \pm 0.005 \pm 0.002$
$\chi^2/(d.o.f.)$	7.6/9	7.7/11	15/12
$\sin^2\theta_{eff}^{lept}$	$0.2322 \pm 0.0004\,^{+0.0001}_{-0.0002}$	$0.2323 \pm 0.0003\,^{+0.0001}_{-0.0002}$	$0.2320 \pm 0.0003\,^{+0.}_{-0.0002}$
$1 - M_W^2/M_Z^2$	$0.2249 \pm 0.0013\,^{+0.0003}_{-0.0002}$	$0.2250 \pm 0.0013\,^{+0.0003}_{-0.0002}$	$0.2242 \pm 0.0012\,^{+0.0003}_{-0.0002}$
M_W (GeV)	$80.28 \pm 0.07\,^{+0.01}_{-0.02}$	$80.27 \pm 0.06\,^{+0.01}_{-0.01}$	$80.32 \pm 0.06\,^{+0.01}_{-0.01}$

* included : $\Delta\alpha = 0.0009 \rightsquigarrow \Delta M_t = 6$ GeV
need improvements $e^+e^- \to$ hadrons $\sqrt{s} \approx 1 - 10$ GeV
new experiment : CMD-2 (PA1 : B Khazin)

: not included yet : unc. on ew. corrs
present estimate : $\Delta M_t \approx 1-3$ GeV



total error quoted

common syst. subtracted

M_Z [GeV]

ALEPH
91.1915±.0052

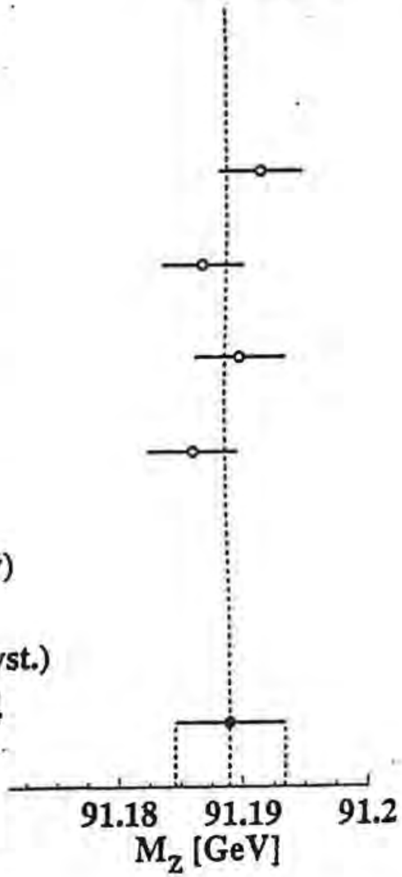
DELPHI
91.1869±.0052

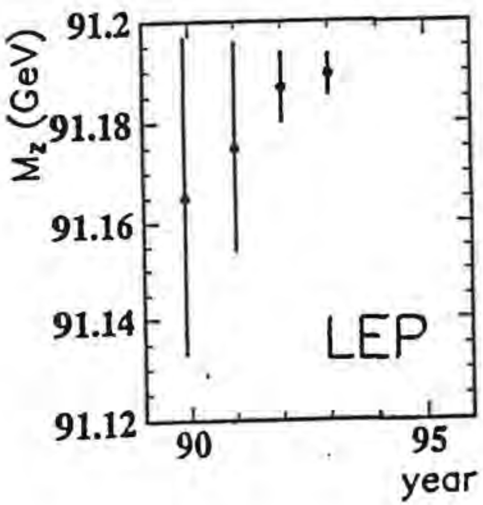
L3
91.1900±.0054

OPAL
91.1862±.0054

common syst.:
4 MeV (LEP energy)
χ^2/D.O.F.=0.5

LEP (incl. comm. syst.)
91.1888±.0044





Glasgow '94

Γ_Z [GeV]

ALEPH
2.4959±.0061

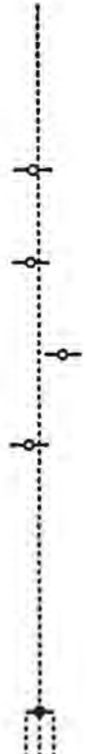
DELPHI
2.4951±.0059

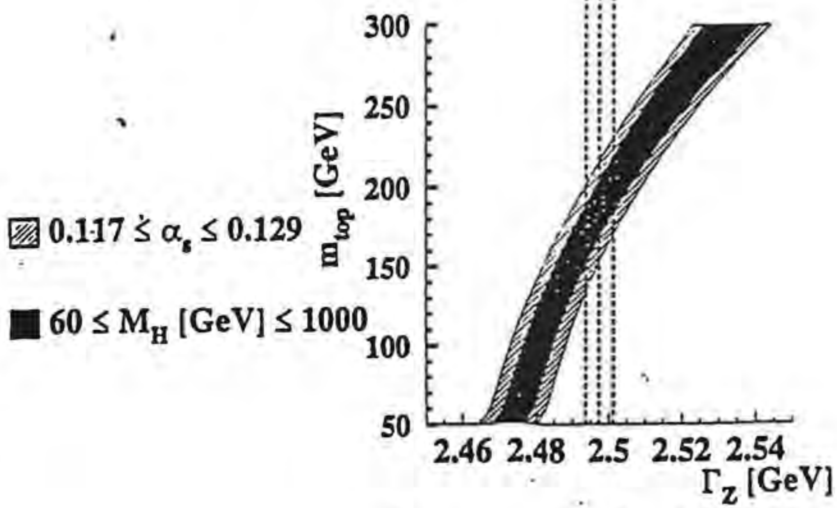
L3
2.5040±.0058

OPAL
2.4946±.0061

common syst.:
2.7 MeV (LEP energy)
χ^2/D.O.F.=0.7

LEP (incl. comm. syst.)
2.4974±.0038





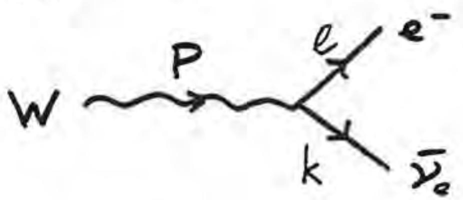
Glasgow '94

WIDTHS of W, Z

$\boxed{W}$ $\quad W^- \longrightarrow e^- \bar{\nu}_e$

$$M = \frac{g}{\sqrt{2}}\, \epsilon_\mu(P)\, \bar{e}(\ell)\, \gamma^\mu \tfrac{1}{2}(1-\gamma_5)\, \nu(k)$$

$g = e/\sin\theta_W$ poln. vector of W   electron spinor   neutrino spinor



Put lepton masses = 0

Sum over fermion spins, average over W poln.

$$\frac{1}{3}\sum_{\text{spins}} |M|^2 \longrightarrow -\frac{g^2}{6}\left(g_{\mu\nu} - \frac{P^\mu P^\nu}{M_W^2}\right) \mathrm{Tr}\left[\not{\ell}\,\gamma_\mu \not{k}\,\gamma_\nu \tfrac{1}{2}(1-\gamma_5)\right]$$

$$= -\frac{g^2}{6}\underbrace{\left(g^{\mu\nu} - \frac{P^\mu P^\nu}{M_W^2}\right)}_{\text{symm.}} \tfrac{1}{2}\times 4\left[\ell_\mu k_\nu + k_\mu \ell_\nu - g_{\mu\nu}\, k.\ell\right]$$

+ anti-symm. part

$$= -\frac{g^2}{3}\left[\ell.k + k.\ell - 4k.\ell - \frac{(P.\ell)(P.k)}{M_W^2} - \frac{(P.k)(P.\ell)}{M_W^2} + P^2\frac{k.\ell}{M_W^2}\right]$$

M_W^2

$$= -\frac{g^2}{3}\left[-k.\ell - \frac{2(P.k)(P.\ell)}{M_W^2}\right]$$

Take $\quad P = (M_W, \underline{0})$

$\qquad \ell = (k, 0, 0, k)$

$\qquad k = (k, 0, 0, -k)$

$\ell.k = \frac{1}{2}M_W^2 \quad , \quad P.k = P.\ell = \frac{1}{2}M_W^2 \qquad\qquad M_W = 2k$

$$\frac{1}{3}\sum_{\text{spins}} |M|^2 = \frac{g^2}{3} M_W^2$$

$$= \frac{8M_W^2}{3}\frac{G_F}{\sqrt{2}}\,.\,M_W^2 = \frac{8G_F}{3\sqrt{2}} M_W^4$$

$$\Gamma(W^- \to e^- \bar{\nu}_e) = \frac{k}{32\pi^2 M_W^2}\int |M|^2 d\Omega$$

$$= \frac{\frac{1}{2}M_W}{32\pi^2 M_W^2}\,.\,\frac{8G_F}{3\sqrt{2}}M_W^4\,.\,4\pi$$

$$= \frac{G_F}{\sqrt{2}}\,.\,\frac{M_W^3}{6\pi} = 224 \text{ MeV}$$

$$= \Gamma_W^0$$

Similarly $\quad W^- \to \mu^- \bar{\nu}_\mu \;,\; \tau\,\bar{\nu}_\tau \;,\;$ ie $W^- \to \ell\bar{\nu} = 3\Gamma$

$W^- \to \bar{u}\, d \;=\; 3\,\Gamma_W^0 \qquad$ (3 from color)

$W^- \to \bar{c}\, s \;=\; 3\,\Gamma_W^0 \qquad$ ($m_{\text{charm}} = 0$)

$$\Rightarrow \Gamma_W^{TOT} = 9\,\Gamma_W^0 = 2.0 \text{ GeV}$$

TABLE 8.1. W-boson partial widths and branching fractions for $M_W = 80.6$ GeV and $m_t = 45$ GeV.

Decay	Partial Width	Branching Fraction (%)
$W \to e\bar{\nu}_e$	0.229	9.14
$\to \mu\bar{\nu}_\mu$	0.229	9.14
$\to \tau\bar{\nu}_\tau$	0.229	9.13
$W \to \bar{u}d$	0.680	27.14
$\to \bar{c}s$	0.679	27.08
$\to \bar{t}b$	0.387	15.46
$W \to \bar{u}s$	0.035	1.41
$\to \bar{c}d$	0.035	1.41
$\to \bar{c}b$	0.001	0.05
$\to \bar{t}s$	0.001	0.03
$\to \bar{u}b$	0.000	0.00
$\to \bar{t}d$	0.000	0.00

Total Width: $\Gamma(W) = 2.51$ GeV

Fig. 8.2 fractions

$\boxed{2}$ Take $Z \to e^+e^-$

$$\mathcal{L} = g_Z \, J^\mu_{NC} \, Z_\mu$$

$$g_Z = \frac{e}{\cos\theta_W \sin\theta_W}$$

$$J^\mu_{NC} = J_3^\mu - \sin^2\theta_W J^\mu_{em}$$

$$= (t_3)_{f_L} - \sin^2\theta_W \, Q_{f_L}$$

$$= -\tfrac{1}{2} - \sin^2\theta_W \quad \text{for } e^-_L$$

$$= 0 - \sin^2\theta_W \quad \text{for } e^-_R$$

$$\therefore \mathcal{M} = -i g_Z \, \epsilon_\mu(P) \Big[\overline{e^-_L} \gamma^\mu e^-_L \left(-\tfrac{1}{2} + \sin^2\theta_W\right) + \overline{e^-_R} \gamma^\mu e^-_R \sin^2\theta_W \Big]$$

Just like $W \to e\bar{\nu}$, simply make the subst.

$$\left(\frac{g}{\sqrt{2}}\right)^2 \longrightarrow \left(\frac{g}{\cos\theta_W}\right)^2 \left[\left(-\tfrac{1}{2} + \sin^2\theta_W\right)^2 + \left(\sin^2\theta_W\right)^2 \right]$$

$$= \frac{g^2}{\cos^2\theta_W} \left[\tfrac{1}{4} - \sin^2\theta_W + 2\sin^4\theta_W \right]$$

and $M_W \longrightarrow M_Z$

$$\therefore\ \Gamma(Z \to e^+e^-) = \frac{G_F}{\sqrt{2}} \frac{M_Z^3}{12\pi}\left[1 - 4x_W + 8x_W^2\right]$$

$$x_W = \sin^2\theta_W$$

$$\equiv \Gamma_Z^0 \left(1 - 4x_W + 8x_W^2\right)$$

Also

$$\Gamma(Z \to \nu_e \bar{\nu}_e) = \Gamma_Z^0\, 1$$

$$\Gamma(Z \to u\bar{u}) = 3\Gamma_Z^0\left(1 - \tfrac{8}{3}x_W + \tfrac{32}{9}x_W^2\right)$$

$$\Gamma(Z \to d\bar{d}) = 3\Gamma_Z^0\left(1 - \tfrac{4}{3}x_W + \tfrac{8}{9}x_W^2\right)$$

Add up: 1st generation:

$$\Gamma_Z^{(1)} = \Gamma_Z^0\, 8\left(1 - 2x_W + \tfrac{32}{3}x_W^2\right)$$

2nd generation: same $\Gamma_Z^{(2)} = \Gamma_Z^{(1)}$

3rd generation: no $Z \to t\bar{t}$!

$$\therefore\ \Gamma_Z^{(3)} = \Gamma_Z^0\left(5 - 8x_W + \tfrac{32}{3}x_W^2\right)$$

$$\Gamma_Z^0 = \frac{G_F}{\sqrt{2}} \cdot \frac{M_Z^3}{12\pi} = 0.160 \text{ GeV}$$

$$\therefore\ \Gamma_Z^{TOT} = \Gamma_Z^0\left[16\left(1 - 2x_W + \tfrac{32}{3}x_W^2\right) + \left(5 - 8x_W + \tfrac{32}{3}x_W^2\right)\right]$$

$= 2.3$ GeV (neglecting masses)

$\longrightarrow 2.500 \pm 0.042$ including masses [illegible]

(depends on m_t)

LEP $\Rightarrow 2.492 \pm 0.017$ GeV

figs.

TABLE 8.2. *Z boson partial widths and branching fractions for $M_Z = 91.9$ GeV and $m_t = 45$ GeV.*

Decay	Partial Width	Branching Fraction (%)
$Z \to \nu_e \bar{\nu}_e$	0.170	6.59
$\to \nu_\mu \bar{\nu}_\mu$	0.170	6.59
$\to \nu_\tau \bar{\nu}_\tau$	0.170	6.59
$Z \to e\bar{e}$	0.085	3.32
$\to \mu\bar{\mu}$	0.085	3.32
$\to \tau\bar{\tau}$	0.085	3.31
$Z \to d\bar{d}$	0.392	15.23
$\to s\bar{s}$	0.392	15.23
$\to b\bar{b}$	0.389	15.08
$Z \to u\bar{u}$	0.305	11.82
$\to c\bar{c}$	0.304	11.81
$\to t\bar{t}$	0.029 ×	1.11

Total Width: $\Gamma(Z) \equiv 2.58$ GeV $\longrightarrow$ 2.52 GeV

window on possible new physics. Any new particle with non-trivial

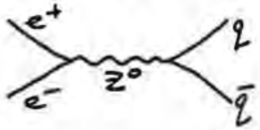
SM 2.500 ± 0.042

Z^0 line shape.

$$\Gamma_{had} \equiv \Gamma(Z \to hadrons)$$
$$\Gamma_{lep} \equiv \Gamma(Z \to \ell\bar{\ell})$$
$$\Gamma_{inv} \equiv \Gamma_Z^{TOT} - (\Gamma_{had} + \Gamma_e + \Gamma_\mu + \Gamma_\tau)$$

In S.M. $N_{fam} = N_\nu = \dfrac{\Gamma_{inv}}{\Gamma_\nu}$

3 parameter (model indep) fit to line shape



$$\sigma_{had} = \sigma^0_{had} \frac{s\Gamma_Z^2}{(s-M_Z)^2 + (s\Gamma_Z/M_Z)^2}(1+\delta_{rad}(s))$$

M_Z, Γ_Z, σ^0_{had} 3 params

$\Rightarrow$ $M_Z = 91.156 \pm .009 \pm .026$ LEP average
$\Gamma_Z = 2.492 \pm .017$ GeV

2 parameter fit in context of S.M.

Assume SM widths in terms of N_ν

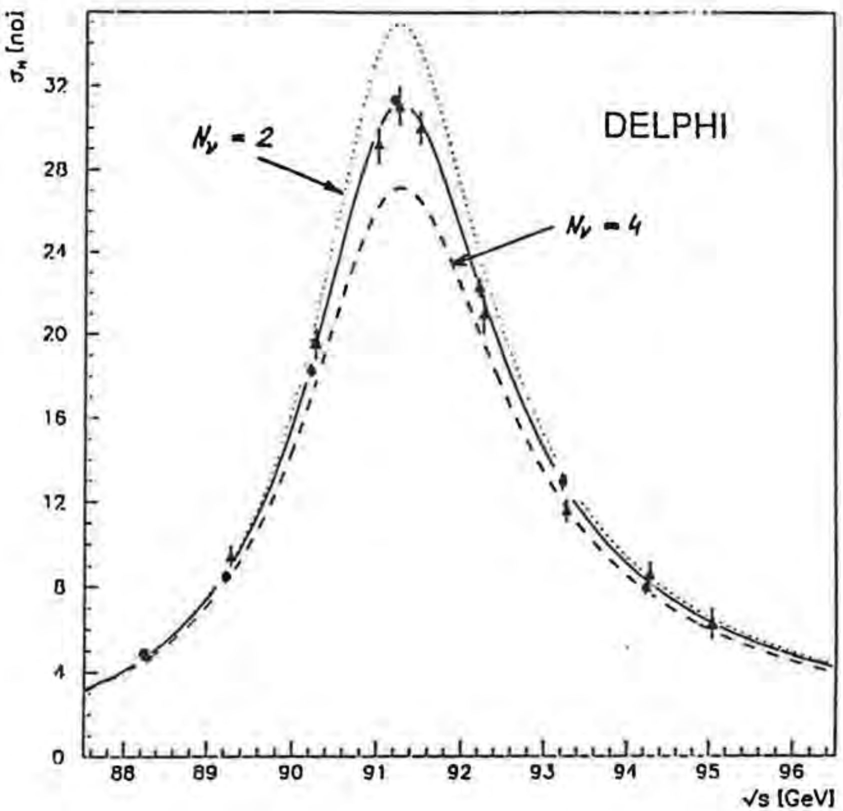
$$\sigma^0_{had} = \frac{12\pi}{M_Z^2} \cdot \frac{\Gamma_e \Gamma_{had}}{\Gamma_Z^2}$$

M_Z, N_ν 2 parameters.

Peak height very sensitive to N_ν (drops 13% as $N_\nu = 3 \to 4$)

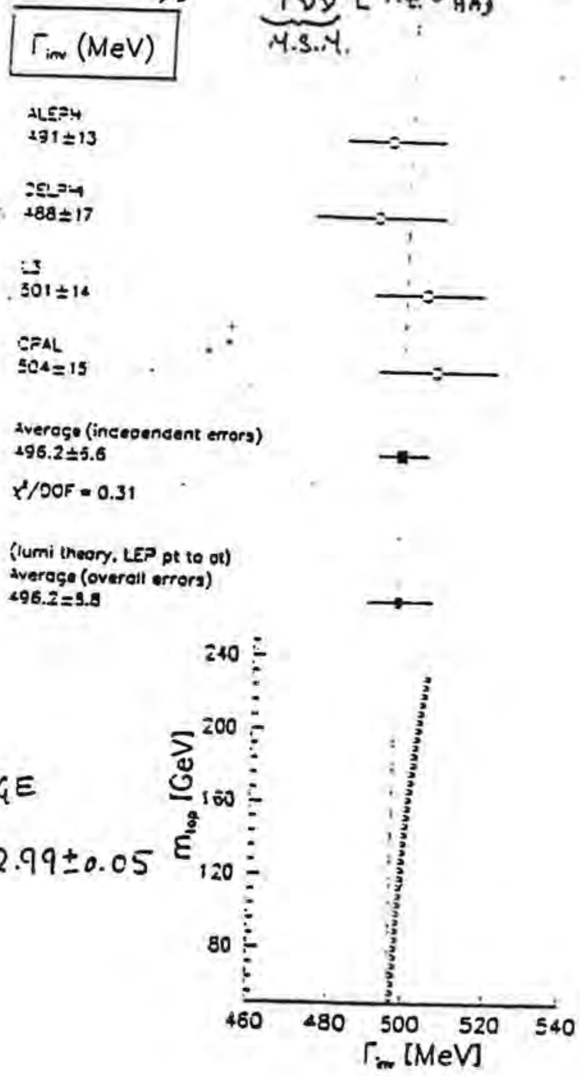
figs. tables

$Z \to q\bar{q}$



$$\Gamma_{inv} = \Gamma_Z - \Gamma_{HAD} - 3\,\Gamma_{l^+l^-}$$

$$N_\nu = \frac{\Gamma_{inv}}{\Gamma_{\nu\bar\nu}} = \underbrace{\frac{\Gamma_{l^+l^-}}{\Gamma_{\nu\bar\nu}}}_{S.M.}\left[\frac{12\pi R_Z}{M_Z^2\,\sigma^0_{HAD}} - R_Z - 3\right]$$



AVERAGE

$N_\nu = 2.99 \pm 0.05$

Electroweak effects in e^+e^-

Recall: Neutral current coupling to fermions f

$$J^\mu_{NC} \equiv 2\bar f\left(\gamma^\mu v_f + \gamma^\mu\gamma_5 a_f\right) f$$

$$\equiv 2\bar f\left[\,g_f^L\,\tfrac{1}{2}(1-\gamma_5)\gamma^\mu + g_f^R\,\tfrac{1}{2}(1+\gamma_5)\gamma^\mu\right] f$$

$$g_f^L = v_f - a_f \qquad\qquad g_f^R = v_f + a_f$$

In the S.M. $\quad v_f = \tfrac{1}{2}(t_3)_{f_L} - Q_f \sin^2\theta_W$

$\qquad\qquad\qquad a_f = -\tfrac{1}{2}(t_3)_{f_L}$

So

$v_e = v_\mu = v_\tau = -1 + 4\sin^2\theta_W \qquad$ (small!)

$a_e = a_\mu = a_\tau = -1$

$v_d = v_s = v_b = -1 + \frac{4}{3}\sin^2\theta_W$

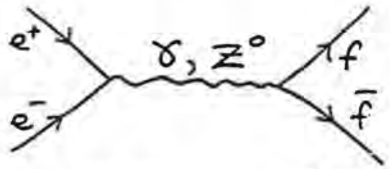
$a_d = a_s = a_b = -1$

$v_u = v_c = v_t = +1 - \frac{8}{3}\sin^2\theta_W$

$a_u = a_c = a_t = +1$

Consider $e^+e^- \to f\bar{f}$:

$$\frac{d\sigma}{d\Omega} = \frac{\alpha^2}{4s}\left[a\,(1+\cos^2\theta) + 2b\cos\theta\right]$$



3 contributions $\quad \chi_1(s) \quad$ interference between γ & Z°

$\chi_2(s) \quad$ Pure Z°

$1 \quad$ Pure γ

$$\chi_1(s) = \kappa\,\frac{s(s-M_Z^2)}{(s-M_Z)^2+\Gamma_Z^2 M_Z^2}$$

$$\chi_2(s) = \kappa^2\,\frac{s^2}{(s-M_Z)^2+\Gamma_Z^2 M_Z^2}$$

$$\kappa = \frac{G_F M_Z^2}{\sqrt{2}\,8\pi\alpha}$$

Below LEP, $\frac{s}{M_Z^2}$ small $\therefore 1 \gg \chi_1 \gg \chi_2$

$$a = Q_e^2 Q_f^2 + 2Q_eQ_f v_e v_f \chi_1 + (v_e^2+a_e^2)(v_f^2+a_f^2)\chi_2$$

$$b = 0 \quad 2Q_eQ_f a_e a_f \chi_1 + 4 v_e v_f a_e a_f \chi_2$$

integrate this only gives our familiar $\sigma = \frac{4\pi\alpha^2}{3s}$

1st 'LOW' energy $\quad$ PEP, PETRA

$e^+e^- \to f\bar{f}$ $\quad$ well below Z° pole

Weak effects seen in angular distributions.

X-section : $\sigma = \frac{4\pi\alpha^2}{3s} R_f$

$$R_f = Q_e^2 Q_f^2 + 2\kappa Q_e Q_f v_e v_f \ldots$$

Asymmetry: $A(\theta) = \dfrac{\frac{d\sigma}{d\Omega}(\theta) - \frac{d\sigma}{d\Omega}(\pi-\theta)}{\text{''} + \text{''}}$

$$= \frac{2\cos\theta}{1+\cos^2\theta} A_f$$

$$A_f = \frac{b}{a}$$

$$\simeq \frac{3}{2}\,\frac{a_e a_f}{Q_e Q_f}\,\chi_1$$

Nice decoupling

R_f measures vector charges

A_f measures axial charges

A_f nice = ratio

Note:

1) Results true for <u>all</u> flavours of light fermions (except electron!)
For quarks, $\times 3$ for colour.

2) At low energies, Pure Z^0 term $\sim 1\%$
$\sin^2\theta_W$ enters here
$\therefore$ poor way to measure $\sin^2\theta_W$

figs: Results from PETRA.

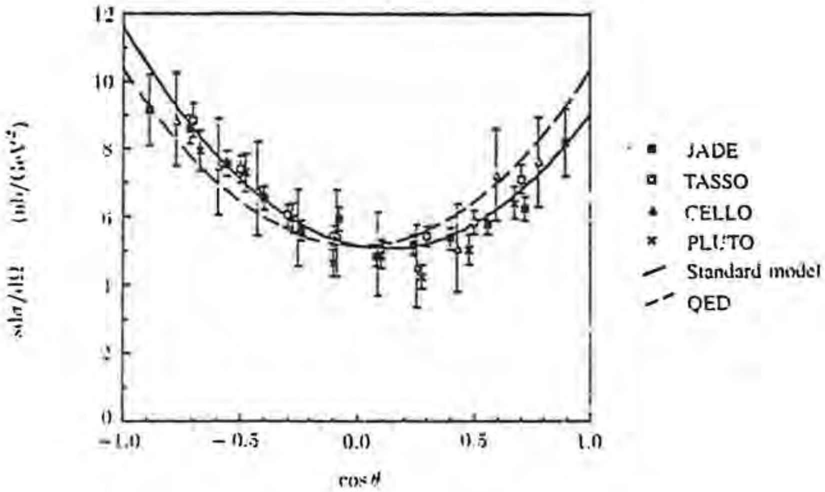


Figure 34. A sample of μ pair angular distributions from four PETRA experiments

3) For heavy quarks, mass effects should be included.

Asym. $A_{q\bar{q}} = \frac{3}{2} \frac{a_e q}{Q_e Q_q} \chi_1 \longrightarrow$ more complicated expression with M_q

Take q = bottom. $a_q = -1$, $a_e = -1$, $Q_q = -\frac{1}{3}$

$$A_{b\bar{b}} = 3 A_{\mu^+\mu^-} \rightarrow 2.7 A_{\mu^+\mu^-} \text{ with } m_b \neq 0$$

fig. - expt.

For S.M. with $(t_3)_b = -\frac{1}{2}$ $\quad A_{b\bar{b}} = -0.252$

Expt $\Rightarrow -0.228 \pm .060$

$\Rightarrow (t_3) = +\frac{1}{2}$ must exist!

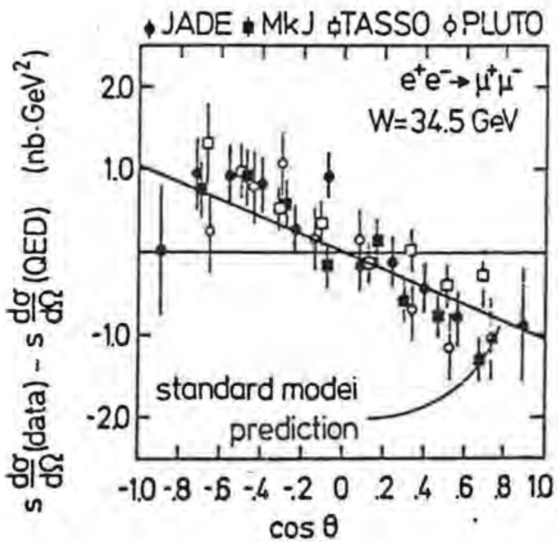


Figure 35. The extracted cosθ term in the angular distribution.

2nd 'High' energy — near the Z^0

$e^+e^- \to f\bar{f}$

$(v_e + a_e\gamma_5)\gamma^\mu$ $\qquad \frac{1}{D_Z} \qquad$ $(v_f + a_f\gamma_5)\gamma_\mu$

$\parallel$

$g_L^{e}\tfrac{1}{2}(1-\gamma_5)\gamma^\mu + g_R^e\tfrac{1}{2}(1+\gamma_5)\gamma^\mu \quad g_L^f\tfrac{1}{2}(1-\gamma_5)\gamma_\mu + g_R^f\tfrac{1}{2}(1+\gamma_5)\gamma_\mu$

$Zf\bar{f}$ coupling conserves helicity

$\Rightarrow \therefore Z \leftarrow e_L^- e_R^+ ;\ C_L^e$ or $e_R^- e_L^+ ,\ C_R^e$

$\longrightarrow f_L^- f_R^+ ,\ C_L^f$ or $f_R^- f_L^+ ,\ C_R^f$

Angular Momentum conservation $\Rightarrow$

In FORWARD dirn. C_L^e for e^+e^- goes only to C_L^f

C_R^e " " " " " C_R^f for f^+f^-

So if $C_{L,R}^i = v_i \mp a_i$ we have

$$\frac{d\sigma}{d\Omega}(e_L^- e_R^+) \propto (C_L^e)^2\left[(C_L^f)^2(1+\cos\theta)^2 + (C_R^f)^2(1-\cos\theta)^2\right]$$

$$\frac{d\sigma}{d\Omega}(e_R^- e_L^+) \propto (C_R^e)^2\left[(C_R^f)^2(1+\cos\theta)^2 + (C_L^f)^2(1-\cos\theta)^2\right]$$

Check: Add to get unpolarised cross-secti

$$\frac{d\sigma^{\text{unpol.}}}{d\Omega} \to a(1+\cos^2\theta) + b\cos\theta$$

$$\left(\frac{1}{D_Z}\right)^2 = \chi_Z$$

Two types of asymmetry.

$A_{L,R} \qquad\qquad A_{F,B}$

$$A_{L,R} = \frac{\sigma(e_R^- e_L^+) - \sigma(e_L^- e_R^+)}{\text{"} + \text{"}} = \frac{(C_R^e)^2 - (C_L^e)^2}{(C_R^e)^2 + (C_L^e)^2}$$

$$= \frac{-2v_e a_e}{(v_e^2 + a_e^2)}$$

$$\simeq -\frac{2v_e}{a_e}$$

$$\simeq 2(-1 + 4\sin^2\theta_W)$$

small !

Need 10^5 events to get $\delta\sin^2\theta_W \approx 0.001$

$$A_{F,B} = \frac{\sigma(\cos\theta > 0) - \sigma(\cos\theta < 0)}{\sigma(\cos\theta > 0) + \sigma(\cos\theta < 0)}$$

$$= -\frac{3}{4} \cdot \frac{(C_L^e)^2 - (C_R^e)^2}{(C_L^e)^2 + (C_R^e)^2} \cdot \frac{(C_L^f)^2 - (C_R^f)^2}{(C_L^f)^2 + (C_R^f)^2}$$

$$= 3 \, \frac{v_e a_e}{(v_e^2 + a_e^2)} \cdot \frac{v_f a_f}{(v_f^2 + a_f^2)}$$

Thus for f = lepton

$$A_{F,B} = 3 \, \frac{v_e^2 a_e^2}{(v_e^2 + a_e^2)^2}$$

Common to use $g_V = \frac{1}{2} v$ $\quad g_A = \frac{1}{2} a$

In. S.M. $\quad g_A = -\frac{1}{2} \quad g_V = -\frac{1}{2}(1 - 4\sin^2\theta_W)$

Thus $A_{F,B} \sim 3 \dfrac{g_V^2}{g_A^2}$ small

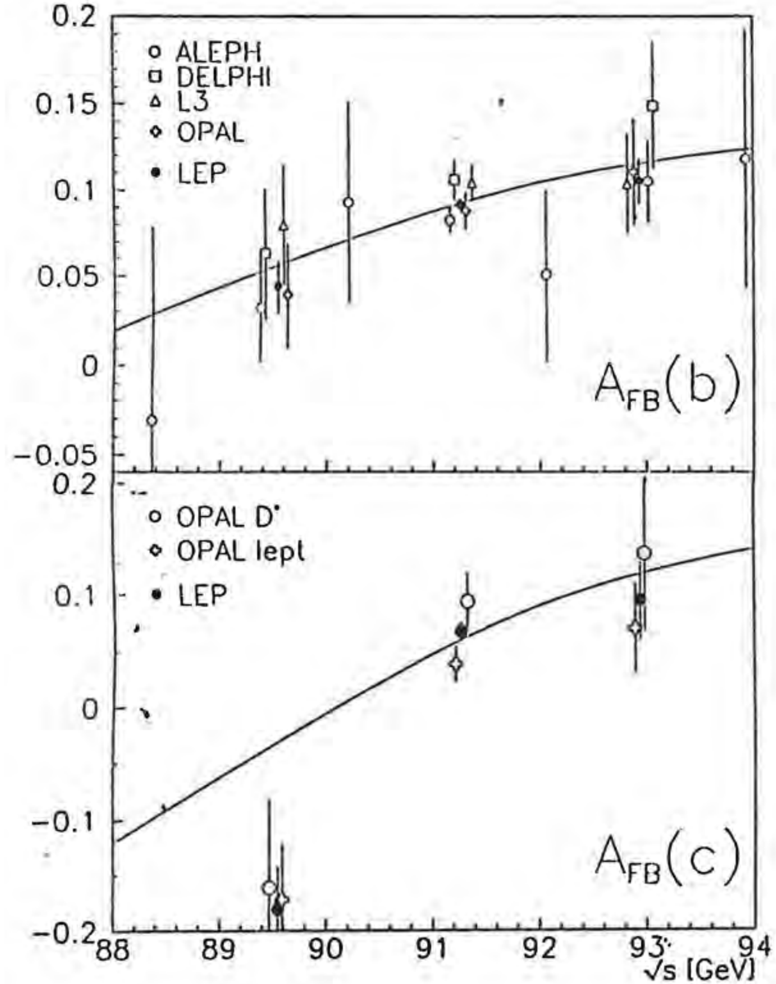
Note

$$\Gamma_{lept} = \frac{G_F M_Z^3}{6\pi\sqrt{2}} (g_V^2 + g_A^2)$$

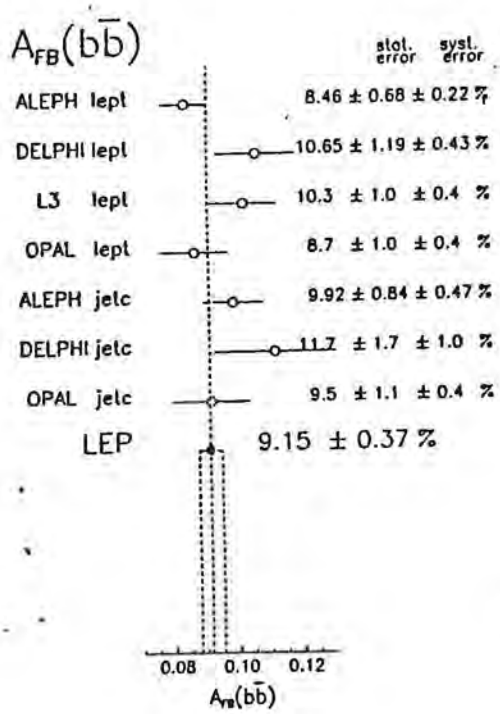
$$A_{F,B}^{lept} = 3 \, \frac{g_V^2 g_A^2}{(g_V^2 + g_A^2)^2}$$

allows extraction of $g_V, g_A, \sin^2\theta_W$

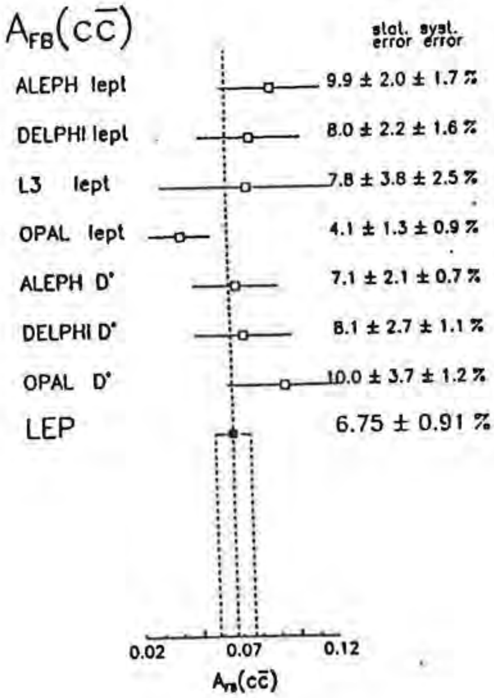
figs.



measured asymmetry at $\sqrt{s} = 91.26$ GeV

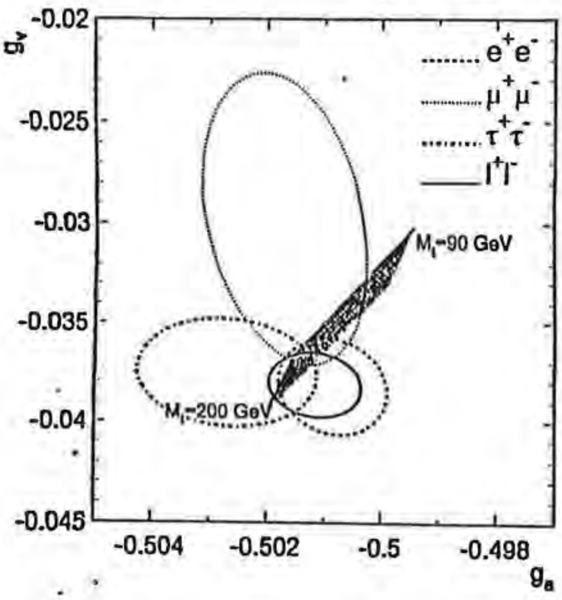


measured asymmetry at $\sqrt{s} = 91.26$ GeV



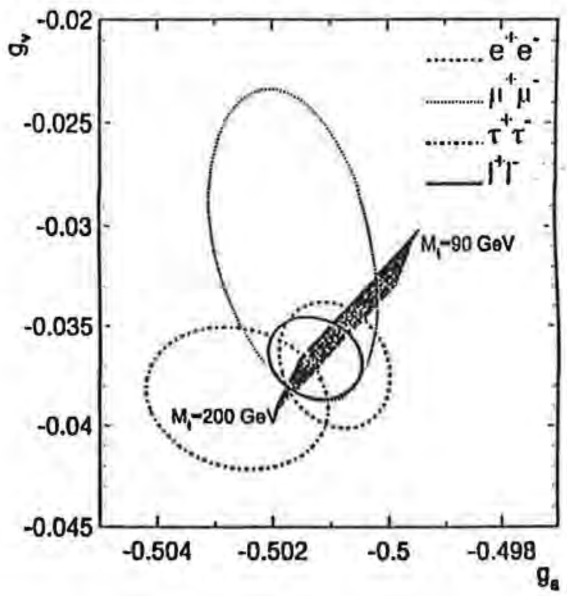
LEP + SLC

68% C.L.



LEP only

68% C.L.



Summary of e.w. precision measurem...

		errors summer !
a) LEP		
line-shape and lepton asymmetries:		
M_Z [GeV]	91.1888 ± 0.0044	0.007
Γ_Z [GeV]	2.4974 ± 0.0038	0.007
σ_h^0 [nb]	41.49 ± 0.12	0.14
R_ℓ	20.795 ± 0.040	0.049
$A_{FB}^{0,\ell}$	0.0170 ± 0.0016	0.0018
+ correlation matrix		
τ polarization:		
$\mathcal{A}_\tau$	0.143 ± 0.010	0.014
$\mathcal{A}_e$	0.135 ± 0.011	0.025
b and c quark results:		
$R_b = \Gamma_{b\bar{b}}/\Gamma_{had}$	0.2202 ± 0.0020	0.0027
$R_c = \Gamma_{c\bar{c}}/\Gamma_{had}$	0.1583 ± 0.0098	0.014
$A_{FB}^{0,b}$	0.0967 ± 0.0038	0.006
$A_{FB}^{0,c}$	0.0760 ± 0.0091	0.015
+ correlation matrix		
$q\bar{q}$ charge asymmetry:		
$\sin^2\theta_{eff}^{lept}$ from $\langle Q_{FB}\rangle$	0.2320 ± 0.0016	unchange
b) $p\bar{p}$ and νN		
M_W [GeV] (CDF, CDF prel., D0 prel., UA2)	80.23 ± 0.18	0.39 (CD
$1 - M_W^2/M_Z^2(\nu N)$	0.2256 ± 0.0047	unchange
c) SLC		
$\sin^2\theta_{eff}^{lept}$ from $\mathcal{A}_e$	0.2294 ± 0.0010	new !

5

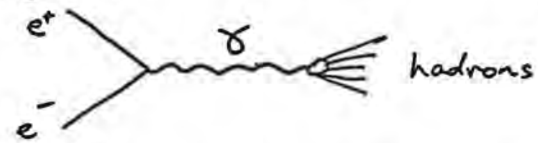
JETS IN E^+E^-
(QCD at LEP)

1) Before LEP

15 years of measuring $e^+e^- \to$ hadrons
$\qquad\qquad\qquad\qquad\quad \to \ell^+\ell^-$

at many machines. SPEAR, PETRA, PEP. TRISTAN

Energy dependence dominated by



$\Rightarrow \frac{1}{s}$ + resonances $\rho, \omega, \varphi, J/\psi, \psi',$
$\Upsilon, \Upsilon', \ldots$

fig.

Remove $\frac{1}{s}$ factor by defining

$$R = \frac{\sigma(e^+e^- \to \text{hadrons})}{\sigma(e^+e^- \to \gamma \to \mu^+\mu^-)}$$

fig. – note for $W = \sqrt{s} \geq 6$ GeV
$R \approx$ const. ~ 4

for a process mediated by the exchange of a single time-like photon varies like
this represents a considerable variation in the magnitude of the cross-secti
dependence alone. Other properties of the reaction (e.g. multiplicity, event s
vary with energy and these will be discussed in the following sections.

3.1.1 Total cross-section

A collection of total cross-section measurements made at many different a
the last 15 years is shown in Figure 6. On this log−log plot, a $1/s$ depend
a line with a slope of −1, and apart from the resonances this is at least r
above $s = 100$ GeV^2, the energy region we are concerned with here. The 1/
the basis of the exchange of a zero mass particle in the s-channel. This i
exchange at all energies and for Z^0 exchange well above the Z^0 pole.

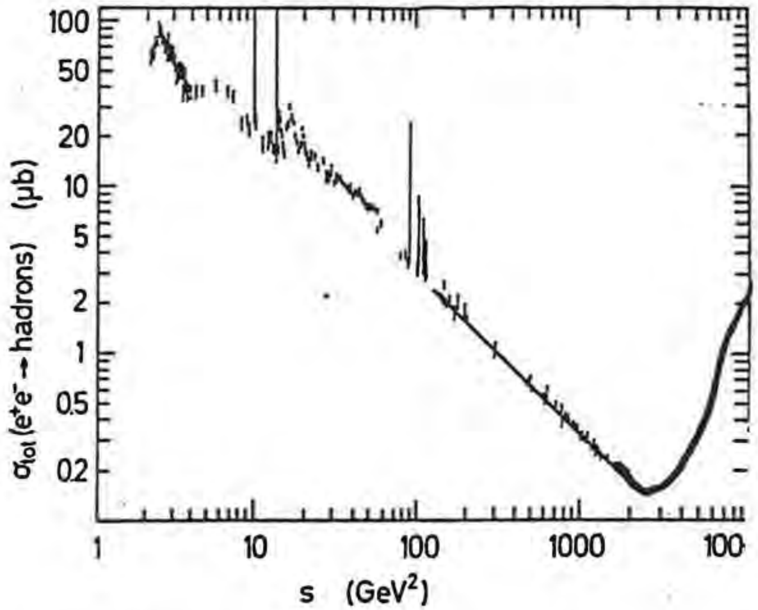
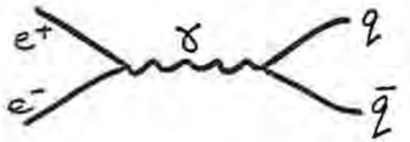


Figure 6. A summary of total cross-section measurements

In the QPM



$$R = K_{QCD} \sum_q 3 R_q$$

with $R_q = e_q^2 \, f(\beta)$

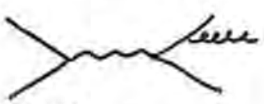
↑ threshold velocity term ≈ 1 for light quarks

If $\sqrt{s} > 4 m_b^2$, sum over 5 quarks

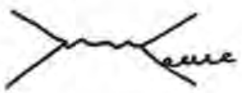
$$\sum 3 e_q^2 = 3 \times \tfrac{11}{9} \approx 4$$

K_{QCD} ?

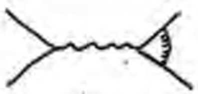
1st order $O(\alpha_s)$ correction



$2q + g$



$2q + g$



$2q$

$$\Rightarrow K_{QCD} = 1 + \frac{\alpha_s}{\pi}$$

$\Rightarrow$ 3 jet final state
2 quarks + gluon jet

At 'low energies' – PETRA, 3 jets identified

(at LEP, 3 jets spectacular ! – fig)

Calculating $O(\alpha_s^2)$ harder $2q + 2g$

" $O(\alpha_s^3)$ even harder ! $2q + 3g$

$$R = 1 + \left(\frac{\alpha_s}{\pi}\right) + (1.985 - 0.115 N_f)\left(\frac{\alpha_s}{\pi}\right)^2 - 12.8\left(\frac{\alpha_s}{\pi}\right)^3 + \ldots$$

Analysing pre-LEP data (R. Marshall)

$\Rightarrow \alpha_s (s = 1000\ \mathrm{GeV}^2) = 0.135 \pm 0.012 \pm 0.010$

or $\Lambda_{\overline{MS}} = 220 \pm^{150}_{120}$ MeV

– very difficult exercise but best method theoretically !

Standard method at $\sqrt{s} \lessapprox 30-50$ Ge

to get $\Lambda_{\overline{MS}}$, α_s uses

SHAPE VARIABLES

THRUST, SPHERICITY, SPHEROCITY, APLANARITY, ACOPLANARITY, It is conventional to express the cross-section as the ratio R, defined thus:

$$R = \frac{\text{measured yield } e^+e^- \rightarrow \text{hadrons}}{\text{theoretical yield } e^+e^- \rightarrow \gamma \rightarrow \mu^+\mu^-}$$

These measurements are summarised in Figure 7, taken from the Review of Pa Properties (Anguilar-Benitez et al 1986), where results from lower energy machine included for reference.

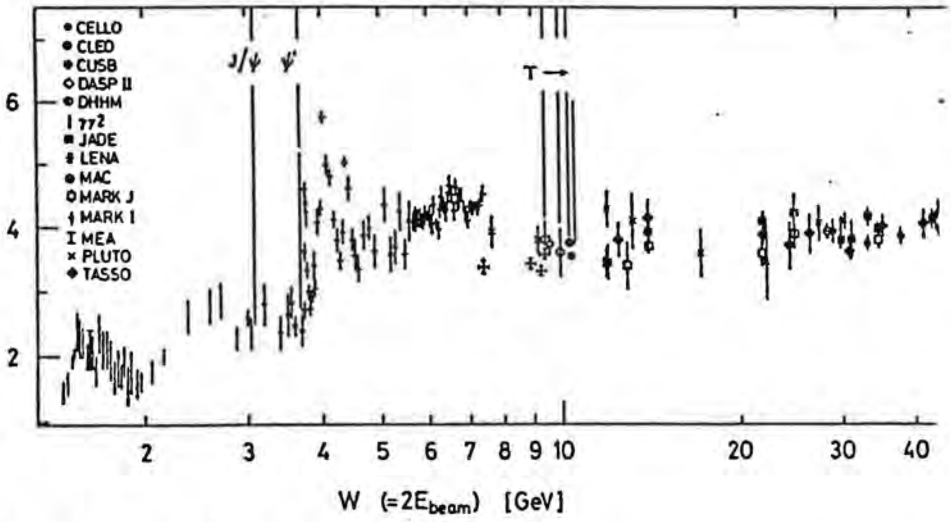


Figure 7. A summary of R measurements

(3.1)

Particle
ines are

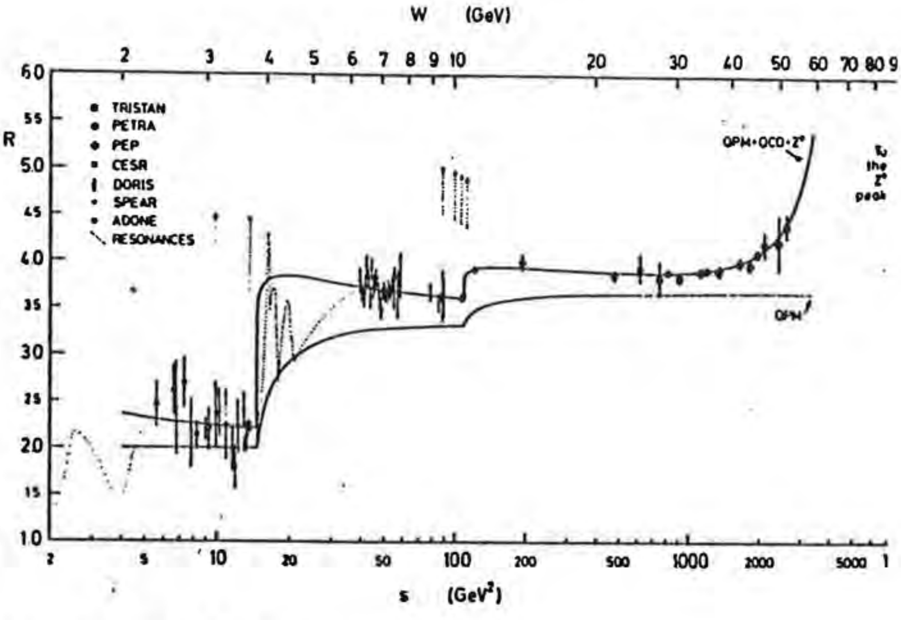
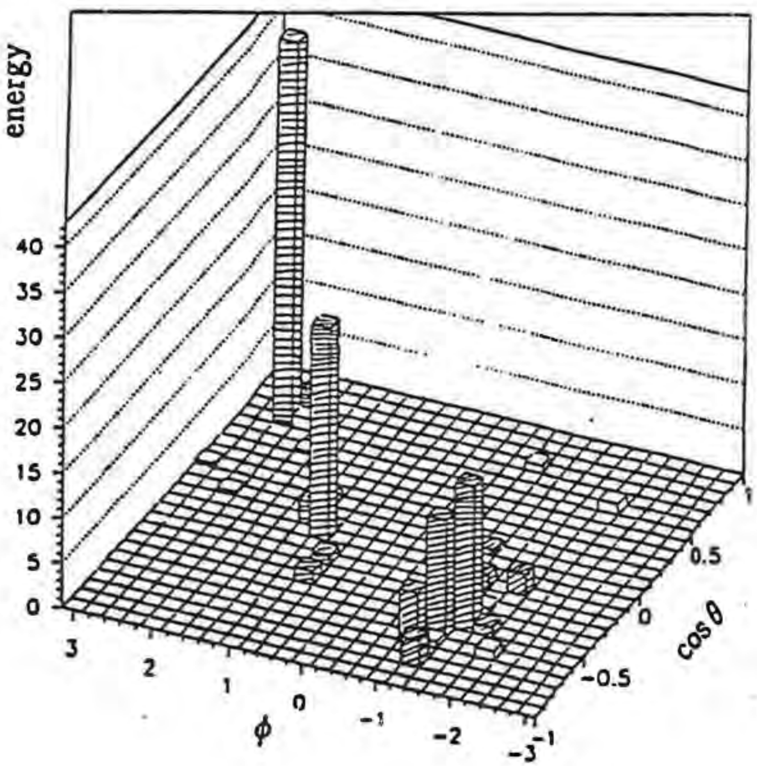


Figure 8. A comparison of R measurements with global fit

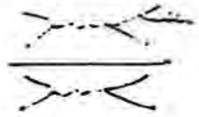
3-jet event



2) After LEP

All those variables + energy-energy correlations

$\Rightarrow$ estimates for α_s

based on 

Popular method at LEP

$\boxed{\text{JET CLUSTER}}$ ANALYSIS

1st question : How to define a JET ?

should satisfy

1) Defined at ANY order of pert. theory.

2) Gives FINITE cross sections at any order.

3) Must be INsensitive to hadronisation

4) Must be SIMPLE to implement

(exp. + theor.

Int. mass algorithms

= cluster algorithm based on

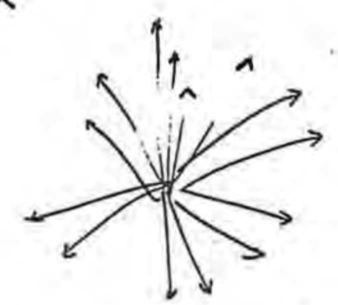
invariant mass cut-off.

E, E0, JADE, P, P0

equivalent to $O(\alpha_s)$

different to $O(\alpha_s^2)$

Start with



Choose a value of y_{cut}.

1) $y_{ab} = (p_a + p_b)^2 / s$

Look at all possible pairs of hadrons

Denote the pair with smallest y_{ab}, y_{ij}

Test : (a) if $y_{ij} < y_{cut}$

replace by a <u>single</u> pseudo-hadron

with mom. $p^{\mu}_{ij} = p^{\mu}_i + p^{\mu}_j$

(b) if $y_{ij} > y_{cut}$ leave

Go To 1)

Eventually $\Longrightarrow$

<u>all</u> y_{ab} invt. masses larger than y_{cut}

No. of clusters = jet multiplicity of event

Associated with multiplicity, n

cross-section σ_n which

depends on y_{cut}.

Theory $\Rightarrow$

$$R_2 \equiv \frac{\sigma_2}{\sigma_{tot}} = 1 + C_{2,1}(y_{cut})\,\alpha_s(\mu) + C_{2,2}\left(y_{cut}, \frac{\mu^2}{s}\right)\alpha$$

$$R_3 \equiv \frac{\sigma_3}{\sigma_{tot}} = C_{3,1}(y_{cut})\,\alpha_s(\mu) + C_{3,2}\left(y_{cut}, \frac{\mu^2}{s}\right)\alpha$$

$$R_4 \equiv \frac{\sigma_4}{\sigma_{tot}} = C_{3,3}\left(y_{cut}, \frac{\mu^2}{s}\right)\alpha_s^2$$

$\mu \equiv$ unknown scale factor ($\approx s$?,

$$C_{2,2} = C_{2,1}\, 2\pi\beta_0 \ln\frac{\mu^2}{s} + C_2'$$

$$C_{3,2} = C_{3,1}\, 2\pi\beta_0 \ln\frac{\mu^2}{s} + C_3'$$

$$C_{3,3} = C_4'$$

All LEP groups do this analysis

– 2 parameters $\Lambda_{\overline{MS}}$, scale μ^2

First – data on $\sigma_2, \sigma_3, \sigma_4$
as function of y_{cut}.

Second – dependence on scale μ^2.

Ideally, best μ^2 when least sensitive

$\Rightarrow$ values of $\mu << M$ – fig.

At low y_{cut}, data favour small μ^2
$\mu^2 = .0017\, M_Z^2$, small $\Lambda_{\overline{MS}} \sim 110\,MeV$

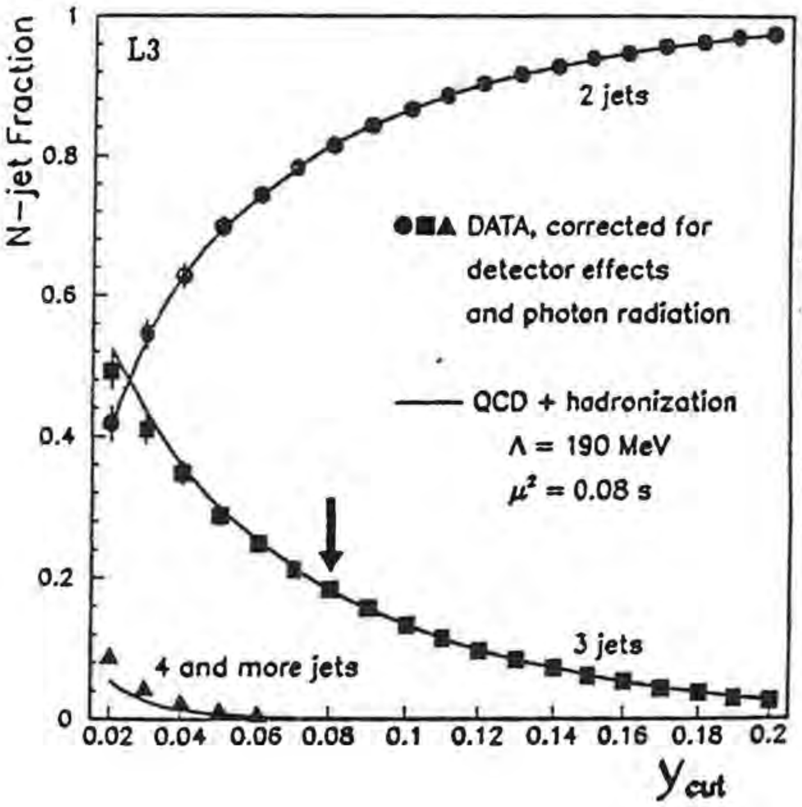
Fixing $\mu^2 = M_Z^2 \Rightarrow \Lambda_{\overline{MS}} = 230\ MeV$

$R_3 \equiv \frac{\sigma_3}{\sigma_{tot}} = 0.183 \pm 0.003$ fig.
for $y_{cut} = 0.08$

Third – s dependence
– Take $y_{cut} = 0.08$ & do analysis
PETRA, PEP, LEP
– evidence for $\ln E$ dependence
= evidence for running of α_s ! fig

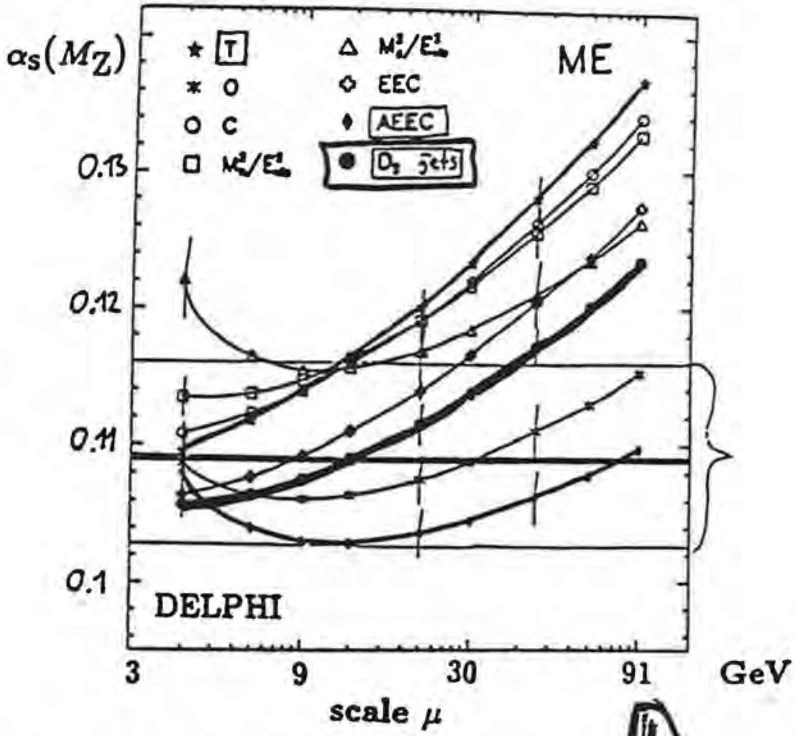
α_s from jet multiplicities

ALEPH, DELPHI, L3, OPAL, MK II



α_s from event topology

ALEPH, $\boxed{\text{DELPHI}}$, L3, OPAL, MK II

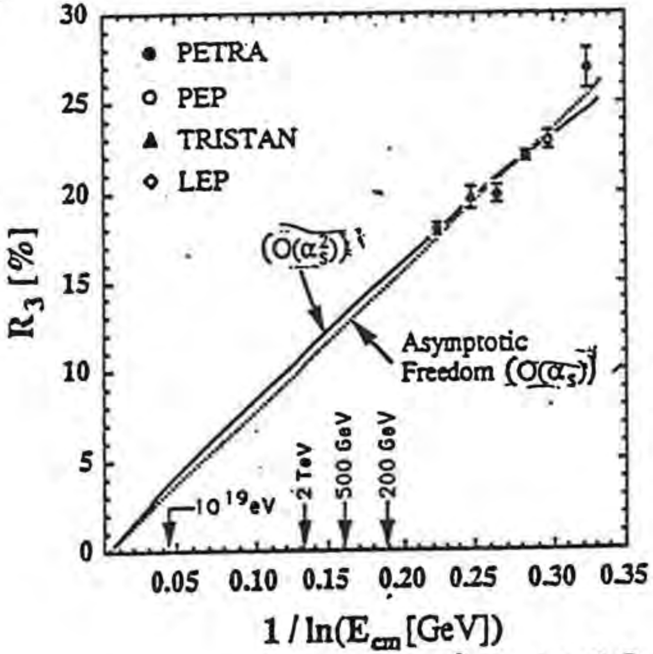


Average for two methods of hadronization correction:

$$\boxed{\alpha_s(M_Z) = 0.111^{+0.007}_{-0.006}}$$

15

FROM
S. BETHKE



MORE
NATURAL

$$\alpha_s = \frac{1}{\beta_0 \ln(E/\Lambda)} - \frac{\beta_1}{\beta_0^3} \frac{\ln[\ln(E/\Lambda)]}{\left[\ln\left(\frac{E}{\Lambda}\right)\right]^2}$$

ONLY THIS
TERM TESTED

NOT FOR
NEAR
FUTURE

78

Fourth – extracted value of $\alpha_s(M_z)$ – fig.

Fifth – compare with other values of $\alpha_s(M_z)$ from all processes !! – fig

Conclude: $\alpha_s(M_z) = 0.117 \text{ "}\pm\text{" } 0.005$

Q. Why the fuss?

Who cares about the precise value of $\alpha_s(M_z)$?

α_S from jet multiplicities

ALEPH, DELPHI, $\boxed{\text{L3}}$, OPAL, MK II

3-jet fraction at $y_{cut} = 0.08$ (26 GeV)
(corrected for detector effects and photon radiation)

$$\sigma_{3-jets}/\sigma_{tot} = 18.4\ \% \pm 0.9\ \%$$

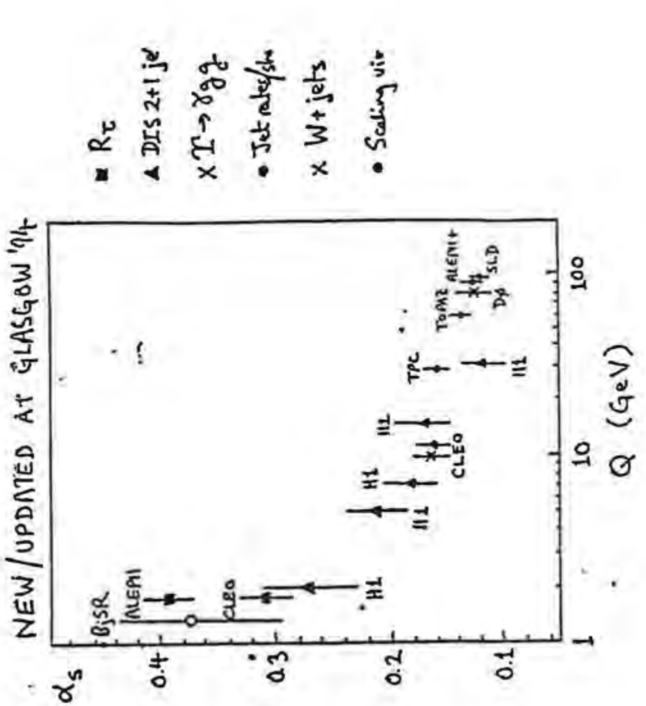
Comparison to analytical 2[nd] order QCD calculation (+ hadronization correction)

$$\boxed{\alpha_s(\sqrt{s} = M_Z) = 0.115 \pm 0.005\,(\text{exp.})\,^{+0.012}_{-0.010}\,(\text{theor.})}$$

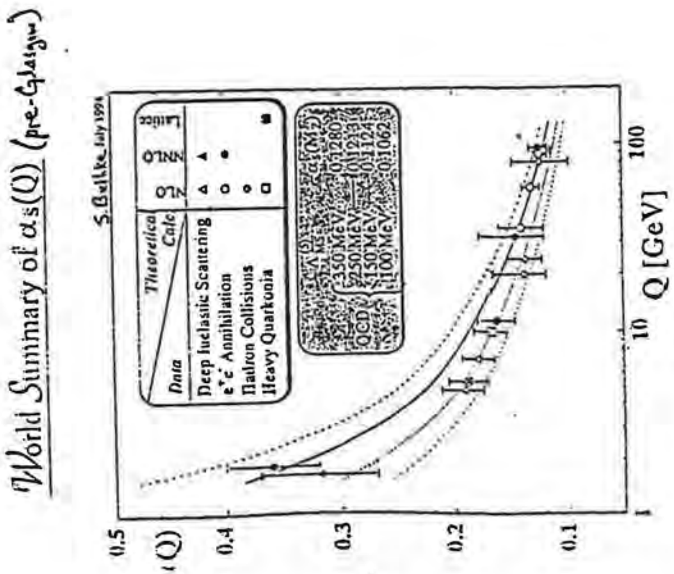
Theoretical error dominated by unknown higher order corrections, estimated by a variation of the renormalization scale μ in the range 3 – 91 GeV.

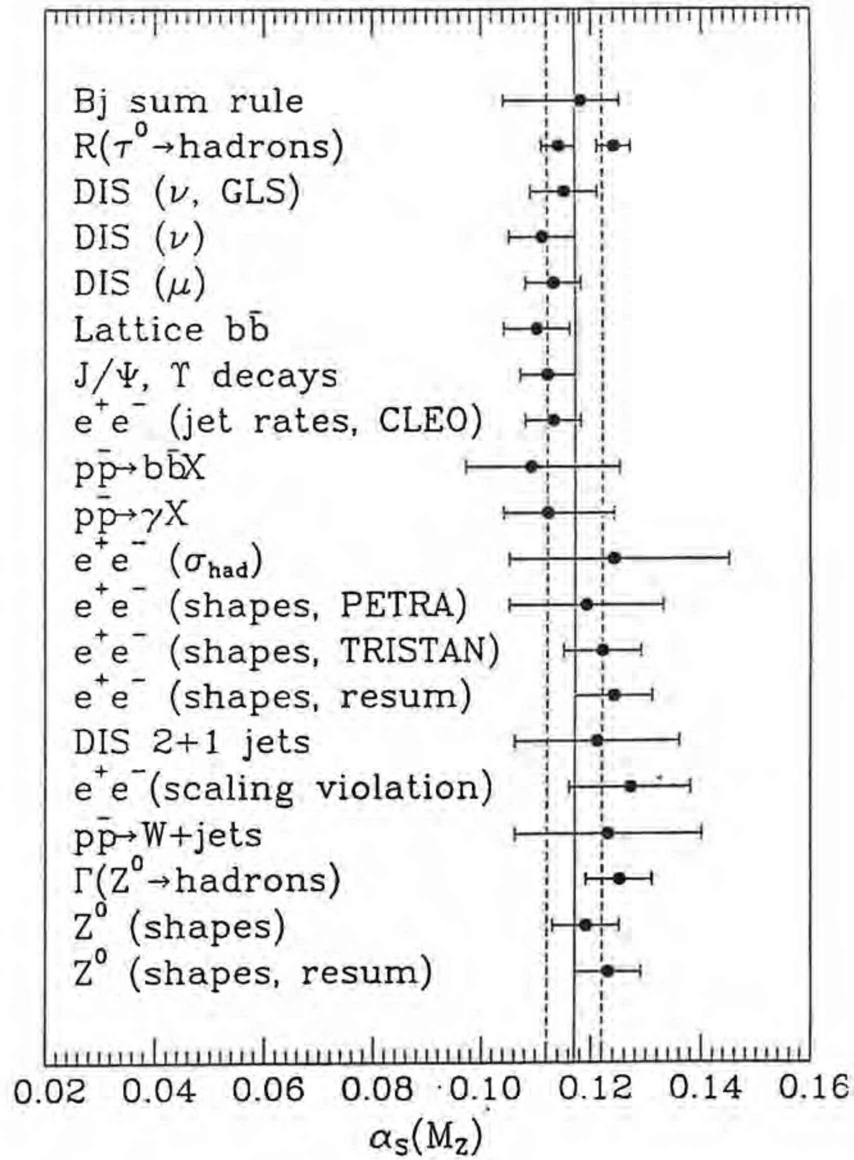
Jet rate becomes independent of the scale μ when calculated to all orders. Therefore the variation of α_s with μ is an estimate of uncalculated higher order corrections.

NEW/UPDATED AT GLASGOW '94



World Summary of $\alpha_s(Q)$ (pre-Glasgow)





All α_1 α_2 α_3 run

$\alpha_3 \equiv \alpha_s$ of SU(3) $\qquad \alpha_s(M_Z) = 0.110 \pm$?

What are $\alpha_1(M_Z)$, $\alpha_2(M_Z)$?

$$\alpha_{em}^{-1}(M_Z) = 128.8$$

and $\sin^2\theta_W = 0.2336 \pm 0.0018$

$$\alpha_2 = \frac{\alpha_{em}}{\sin^2\theta_W} \qquad \alpha_1 = \frac{5}{3}\frac{\alpha_{em}}{\cos^2\theta_W}$$

?

In GUTS, $\quad g_i^2 \, Tr(T_i^2)$ = same for each subgroup.

$$\Rightarrow e^2 Tr(Q^2) = g_2^2 Tr(T_3^2) = g_3^2 Tr(T_c^2)$$

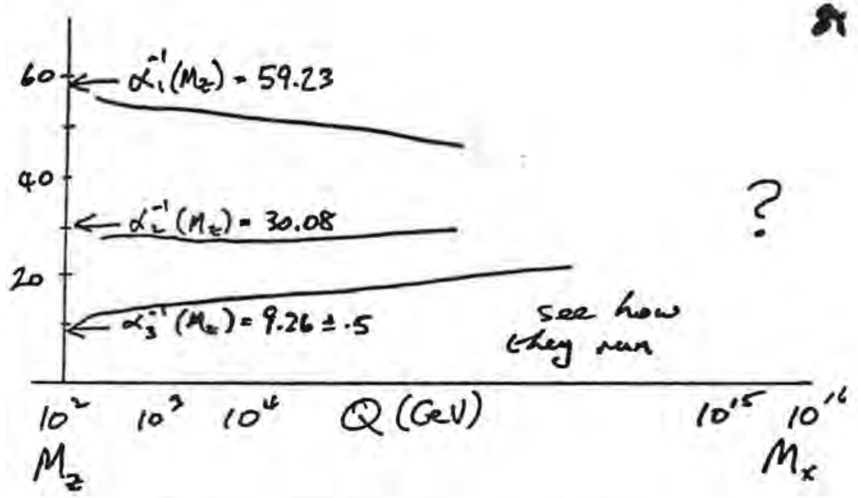
e.g. in SU(5) all 15 particles of each generation

$$\Rightarrow Tr(Q^2) = \frac{16}{3} \qquad Tr(T_3^2) = 2 \qquad Tr(T_c^2) = 2$$

$$\Rightarrow \frac{e^2}{g_2^2} = \frac{3}{8} \quad \Rightarrow \sin^2\theta_W = \frac{3}{8}, \quad g_1^2 = \underline{\frac{3}{5}} g_2^2$$

At GUT scale M_X we expect

$$\alpha_1^{-1}(M_X) = \alpha_2^{-1}(M_X) = \alpha_3^{-1}(M_X)$$



If α_i meet — GUT .....

① Values of $\alpha_1, \alpha_2, \alpha_3$ $\quad Q = M_Z$ given by EXPERIMENT

② Slopes $\frac{d\alpha^{-1}}{d\ln Q}$ at $Q = M_Z$ given by THEORY

$$\alpha_i^{-1}(Q) = \alpha^{-1}(M_X) + \frac{b_i}{2\pi} \ln\frac{M_X}{Q}$$

ie $\displaystyle \frac{d\alpha_i^{-1}}{d\ln Q} = -\frac{b_i}{2\pi}$ $\qquad$ + 2 loop

③ If they do meet at some M_X

$\Rightarrow$ the given GUT is OK

No NEW PHYSICS between $M_Z \to M_X$

In GUT,

$$\alpha_i^{-1}(\mu) = \alpha^{-1}(M_x) + \frac{b_i}{2\pi} \ln \frac{M_x}{\mu}$$

$$\boxed{\frac{d\alpha_i^{-1}}{d\ln\mu} = -\frac{b_i}{2\pi}} \qquad + 2 \text{ loop terms}$$

The b_i get contributions from types

gauge particles

quarks/leptons $\qquad N_{fam}$

Higgs $\qquad N_{Higgs}$

Take (i) minimal SU(5)

$$b_i = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = \begin{pmatrix} 0 \\ -22/3 \\ -11 \end{pmatrix} + N_{fam}\begin{pmatrix} 4/3 \\ 4/3 \\ 4/3 \end{pmatrix} + N_{Higgs}\begin{pmatrix} 1/10 \\ 1/6 \\ 0 \end{pmatrix}$$

minimal $\Rightarrow$ $N_{Higgs} = 1$

expt $\Rightarrow$ $N_{fam} = 3$ (top included)

$$\Rightarrow \quad b_1 = 4.1 \qquad b_2 = -3\tfrac{1}{6} \qquad b_3 = -7$$

Taking $\dfrac{\alpha_1^{-1} - \alpha_2^{-1}}{\alpha_2^{-1} - \alpha_3^{-1}} = \dfrac{b_1 - b_2}{b_2 - b_3}$

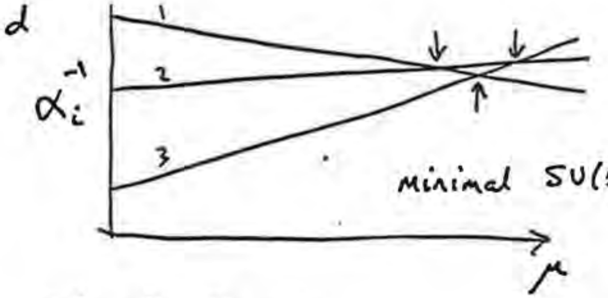
indep. of μ

So with minimal SU(5)

$$\frac{b_1 - b_2}{b_2 - b_3} = \frac{4.1 + 3\frac{1}{6}}{-3\frac{1}{6} + 7} = 1.90$$

At M_Z $\qquad \dfrac{\alpha_1^{-1} - \alpha_2^{-1}}{\alpha_2^{-1} - \alpha_3^{-1}} = \underline{\underline{1.4 \pm 0.04}}$

no good



(ii) Take minimal SUSY

more particles in the loops!

gauginos

squarks/sleptons

$$b_i = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = \begin{pmatrix} 0 \\ -6 \\ -9 \end{pmatrix} + N_{fam}\begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix} + N_{Higgs}\begin{pmatrix} 3/10 \\ 1/2 \\ 0 \end{pmatrix}$$

minimal ⇒ $N_{Higgs} = 2$

⇒ $b_1 = 6.6$, $b_2 = 1$, $b_3 = -3$

So with minimal SUSY extension of SM

$$\frac{b_1 - b_2}{b_2 - b_3} = \frac{6.6 - 1}{1 + 3} = 1.4 \quad !$$

Thus we can have unification with minimal SUSY where $M_{SUSY} \sim M_Z$!!

But ....

1./ Should include effects from 2-loops.

⇒ Unification if $M_{SUSY} \approx 1$ TeV

2./ M_{SUSY} is not a <u>single</u> scale

$M_{\tilde{g}} > M_{\tilde{W}}$

$M_{\tilde{q}} > M_{\tilde{\ell}}$

⇒ Unification only for larger SUSY scales ~ 7 TeV

3./ Doing it properly.

2 basic parameters M_0 = common scalar mass at M_X

$M_{1/2}$ = common gaugino mass at M_X

⇒ fine tuning problem for 'large' M_0, $M_{1/2}$ when trying to break electroweak symmetry.

⇒ forces us to smaller $M_0, M_{1/2}$

" " " larger $\alpha_s(M_Z)$ ≈ 0.118

⇒ LSP mass down to around M_Z !!!

MSSM

To each particle, SUSY partner

quarks (spin $\frac{1}{2}$)	—	squarks (scalars)
leptons (")	—	sleptons (")
gauge bosons (spin 1)	—	gauginos (spin $\frac{1}{2}$)
Higgs (scalars)	—	higgsinos (spin $\frac{1}{2}$)

In GUTS – major problem
– mass hierarchy problem.

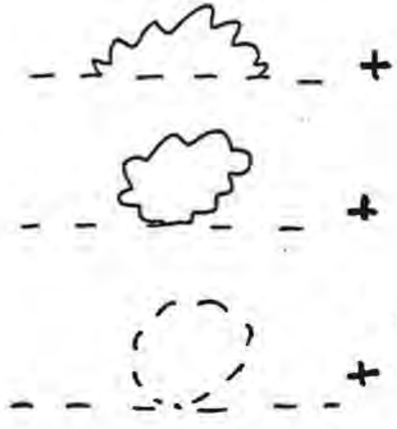
Given a single scale $M_X \sim 10^{16}$ GeV, how can you generate very light particles W, Z (but non-zero mass)
& $v = \langle 0 | H | 0 \rangle \sim 200$ GeV ?

SUSY goes a long way to solve this problem.

If we do have a small scale, e.g. a mass of a scalar, at M_X, radiative corrections would, in general, quickly give corrections of $O(M_X)$ to that scale.

— But not in SUSY.

Contributions to scalar masses (in SUSY)



in EXACT SUSY
these contribs cancel
& scalar masses are degenerate.

All left-handed fermion fields of the SM are promoted to chiral superfields

$$Q = \begin{pmatrix} U \\ D \end{pmatrix}, \quad U^c, D^c$$

$$L = \begin{pmatrix} N \\ E \end{pmatrix}, \quad E^c$$

In MSSM need TWO Higgs doublets

$$H_1 = \begin{pmatrix} H_1^0 \\ H_1^- \end{pmatrix} \qquad H_2 = \begin{pmatrix} H_2^+ \\ H_2^0 \end{pmatrix}$$

H_1 gives masses to $t_3 = -\frac{1}{2}$ partners, e.g. b, τ.

H_2 gives masses to $t_3 = +\frac{1}{2}$ partners, e.g. t

Thus in SM 1 Higgs doublet
= 4 deg. of freedom
$3 \to W^\pm, Z$
leaving 1 neutral Higgs

cf. in MSSM 2 Higgs doublets
= 8 deg. of freedom
$3 \to W^\pm, Z$
leaving 5 Higgs to be found

2 charged $h^\pm$
3 neutral h^0 (light), h^0 (heavy), A (CP odd)

One important parameter $\tan\beta = \frac{v_2}{v_1}$

$v_i = \langle 0 | H_i | 0 \rangle$

v_2 ass. with Higgs which gives mass to top
v_1 " " " " " " bottom

ie usually $\tan\beta >> 1$

Attractive Scenario :

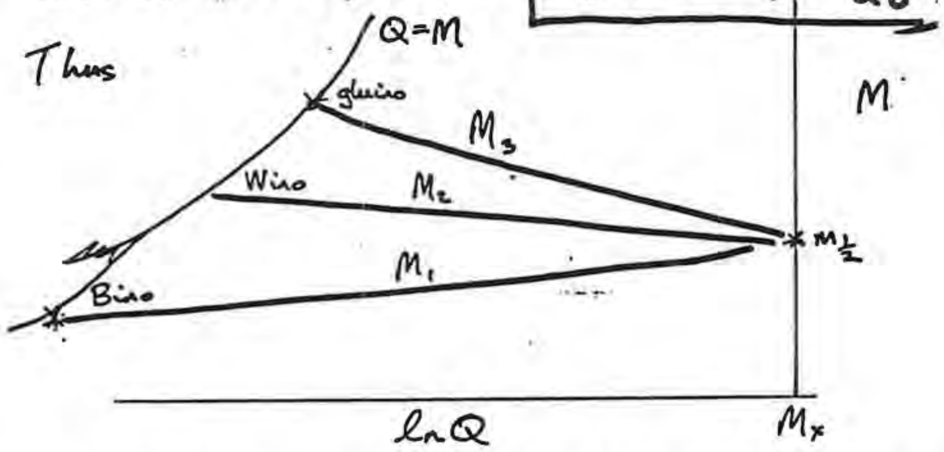
Begin at $Q = M_x$ with a "softly"-broken form of SUSY.

In particular, common mass of gauginos $m_{\frac{1}{2}}$.

Masses of gauginos at low energy – obtained by evolution — same graphs by which the $\alpha_i(Q)$ evolve !

$$\boxed{M_i(Q) = m_{\frac{1}{2}} \frac{\alpha_i(Q)}{\alpha_U}}$$

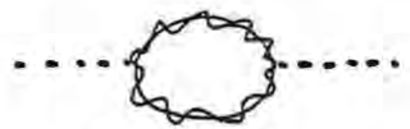
Thus



Photino (combination of W^0, B^0) is lightest gaugi[no]
probably also the "LSP".

Masses of scalars [squarks, sleptons, Higgs]

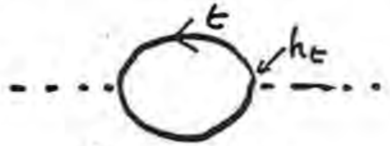
bit more complicated



gaugino loop — gives +ve contrib. to scalar mass

gluino doesn't contrib. to sleptons, Higgs
" does " " squarks

ie $M_{\tilde{q}} > M_{\tilde{\ell}}, M_H$



fermion loop — gives −ve contrib. to scalar masses

Since $h_t >>$ all other Yukawa couplings.

then stop squark is **LIGHTEST** squark.

It could be that M_{stop} is only just heavier than M_{top} !

Masses of Higgs.

H_2 couples to top-type squarks/quarks
H_1 " " bottom-type squarks/quarks

∴ Only H_2 gets the very −ve contribution

(provides a mechanism for electroweak symm. Breaking

Thus one Higgs is "light"
all others heavy

ie $M_{h^0} \sim O(M_Z)$

after quantic radiative corrections we expect $M_{h_0} \approx 110 \to 140$ GeV

c.f. SM no prediction for Higgs Mass.

Thus. **LHC, SSC** predicted to find h_0 & **SUSY** spectrum !!! ...

Fermion - Higgs Sector of S.M.

Fermion masses $m_u, m_d, m_s \ldots$
$m_e, m_\mu, \ldots$

Quark Mixing — Yukawa Couplings

Mass Matrices

CKM matrix — Cabibbo angle ......

Too Many Parameters in S.M. (≥ 17)

..... beyond S.M.

Generally, fermion masses put in by hand.

e.g. in QED, add to Lagrangian term

$$-m\,\bar{\psi}_L\,\psi_R$$

to give mass to electron.

Electroweak S.M. is a <u>chiral</u> theory.

$\bar{\psi}_L \psi_R$ is not invt. under $SU(2)\otimes U(1)$

$SU(2)\otimes U(1)$ invariant way to give masses to up quark, for example, is via the YUKAWA coupling

$$h_u\,(\bar{u}_L\ \bar{d}_L)\begin{pmatrix}\varphi^0\\ \varphi^-\end{pmatrix} u_R$$

We have up quarks — 3 generations
down

and leptons — 3 generations

So in $\mathcal{L}$ we have

$$-\mathcal{L} = U'_{\alpha\beta}\,(\bar{u}'_{L\alpha}\ \bar{d}'_{L\alpha})\begin{pmatrix}\varphi^0\\ \varphi^-\end{pmatrix} u'_{R\beta}$$

$$+\ D'_{\alpha\beta}\,(\bar{u}'_{L\alpha}\ \bar{d}'_{L\alpha})\begin{pmatrix}\varphi^+\\ \varphi^0\end{pmatrix} d'_{R\beta}$$

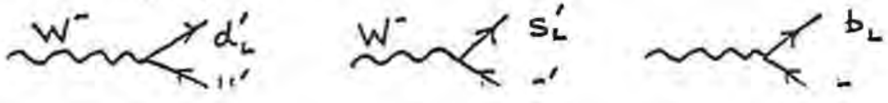
$$+\ L'_{\alpha\beta}\,(\bar{\nu}'_{L\alpha}\ \ell'_{L\alpha})\begin{pmatrix}\varphi^+\\ \varphi^0\end{pmatrix} \ell'_{R\beta}$$

α, β generation index $1 \to 3$

$'$ denotes quantities in the GAUGE basis

ie u'_L d'_L are the quantities appearing in the charged weak current

$$J^\mu_{weak}(\text{quark}) = \bar{u}'_{L\alpha}\ \tfrac{1}{2}\gamma^\mu(1-\gamma_5)\ d'_{L\alpha}$$



3.

Matrices $U'_{\alpha\beta}$ $D'_{\alpha\beta}$ $E'_{\alpha\beta}$
NOT diagonal

We wish to transform to MASS basis
where U D E ARE diagonal.
each diagonal elt $= h_q = \dfrac{\text{mass of quark}}{v/\sqrt{2}}$

v = vev. of Higgs

After SSB & dropping $\frac{v}{\sqrt{2}}$ factors,

$$-\mathcal{L} = \bar{u}'_L U' u'_R + \bar{d}'_L D' d_R + \bar{l}'_L E' l_R$$

$$= \bar{u}'_L V_L^{u\dagger} V_L^u U' V_R^{u\dagger} V_R^u u'_R + \ldots\ldots + \ldots\ldots$$

$$= \bar{u}_L U u_R + \bar{d}_L D d_R + \bar{l}_L E l_R$$

$$V_L^u, V_R^u \ ; \ V_L^D, V_R^D \ ; \ V_L^L \ V_R^L$$

are UNITARY MATRICES

Note Mass matrix U' or U
have <u>same</u> eigenvalues m_u m_c m_t

Similarly D', D
E', E

Thus $u_L = V_L^u u'_L$ etc

$$\underset{\text{diag.}}{U} = V_L^u \underset{\text{non-diag.}}{U'} V_R^{u\dagger} \quad \text{etc}$$

Meanwhile, back in Charged weak current

$$J^\mu_{weak}\ (\text{quark}) = \bar{u}'_L d'_L W^\mu$$ $W^\mu = \frac{1}{2}\gamma^\mu(1-$

$$= \bar{u}_L V_L^u V_L^{D\dagger} d_L W^\mu$$

$$= \bar{u}_{L\alpha} V^u_{L\,\alpha\gamma} V^{D\dagger}_{L\,\gamma\beta} d_{L\beta} W^\mu .$$ α, β, γ generation indices

$$= \bar{u}_{L\alpha} V^{CKM}_{\alpha\beta} d_{L\beta} W^\mu$$

Thus in the MASS basis, there is MIXING between the quark eigenstates. CKM matrix is the product $V_L^u V_L^{D\dagger} = V^{CKM}$
only V_L involved.

$\therefore$ Intimate connection between quark masses and CKM matrix elements

Usual convention, put mixing entirely into down states, defining $d'' = V^{CKM} d$

V^{CKM} 3×3 unitary matrix
- characterised by 3 real quantities + 1 phase.

If only 2 generations $V^{CKM} \to V^c$
$$= \begin{pmatrix} \cos\theta_c & \sin\theta_c \\ -\sin\theta_c & \cos\theta_c \end{pmatrix}$$

1 real parameter – no phase
Cabibbo angle θ_c (small)

$d \longrightarrow W u \qquad \propto \cos\theta_c$

$c \longrightarrow W s \qquad \propto \cos\theta_c$

$c \longrightarrow W d \qquad \propto \sin\theta_c$ "Cabibbo suppressed"

3 generations $\longrightarrow$ exists a PHASE

$\longrightarrow$ CP violation allowed in S.M. !

S.M. has a lot of parameters.

$g \quad g' \quad v = \langle 0|H|0\rangle$ / or α_{em}, G_F, $\sin\theta_W$

$m_u \quad m_d \quad m_e$

$m_c \quad m_s \quad m_\mu$

$m_t \quad m_b \quad m_\tau$

V^{CKM} 4 parameters

Λ_{QCD}

People suggest forms for mass matrices e.g.

$$U' = \begin{bmatrix} 0 & C & 0 \\ C & 0 & B \\ 0 & B & A \end{bmatrix} \qquad A \gg B \gg C$$

$$D' = \begin{bmatrix} 0 & Fe^{i\varphi} & 0 \\ Fe^{-i\varphi} & E & 0 \\ 0 & 0 & D \end{bmatrix} \qquad D \gg E \gg F$$

$$V_L^U = \begin{pmatrix} c_2 & s_2 & 0 \\ -s_2 & c_2 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ 0 & c_3 & s_3 \\ 0 & -s_3 & c_3 \end{pmatrix}$$

$$V_L^D = \begin{pmatrix} c_1 & -s_1 & 0 \\ s_1 & c_1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ 0 & c_4 & s_4 \\ 0 & -s_4 & c_4 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ 0 & e^{i\varphi} & 0 \\ 0 & 0 & e^{i\varphi} \end{pmatrix}$$

$$V_{CKM} = \begin{pmatrix} c_1c_2 - s_1s_2e^{-i\varphi} & s_1 + c_1s_2e^{-i\varphi} & s_2(s_3 - s_4) \\ -c_1s_2 - s_1e^{-i\varphi} & \cdots & \cdots \\ \cdots & \cdots & \cdots \end{pmatrix}$$

angles s_1, s_2, s_3, s_4 expressed in terms of A, B, C, D, E, F

but A, B, C, D, E, F are expressed in terms of eigenvalues of U', D' ie quark masses

$\therefore$ elements of CKM – simple functions of quark masses

Most famous = GSTO relation

$$V_{12}^{CKM} = \sin\theta_c = \sqrt{\frac{m_d}{m_s}}$$

ie no. of indep. parameters in S.M. reduced.

STRING THEORY may suggest patterns of the U', D', E'