Durham University

Durham E-Theses

Phenomenology of the Standard Model, and beyond, at high-energy colliders

Summers, David John

How to cite:

Summers, David John (1993) *Phenomenology of the Standard Model, and beyond, at high-energy colliders*, Durham theses, Durham University. Available at Durham E-Theses Online: http://etheses.dur.ac.uk/5619/

Use policy

The full-text may be used and/or reproduced, and given to third parties in any format or medium, without prior permission or charge, for personal research or study, educational, or not-for-profit purposes provided that:

- a full bibliographic reference is made to the original source
- a link is made to the metadata record in Durham E-Theses
- the full-text is not changed in any way

The full-text must not be sold in any format or medium without the formal permission of the copyright holders.

Please consult the full Durham E-Theses policy for further details.

Academic Support Office, The Palatine Centre, Durham University, Stockton Road, Durham, DH1 3LE
e-mail: e-theses.admin@durham.ac.uk Tel: +44 0191 334 6107
http://etheses.dur.ac.uk

Phenomenology of the Standard Model, and beyond, at high energy colliders

David John Summers

Department of Physics
University of Durham

The copyright of this thesis rests with the author. No quotation from it should be published without his prior written consent and information derived from it should be acknowledged.

A thesis submitted to the University of Durham for the Degree of Doctor of Philosopy

July 1993

21 OCT 1993

Abstract

I review planned searches for the so far unobserved Higgs boson of the Standard Model of High Energy Physics. In particular a light 'intermediate' mass Higgs with mass in the range $80\,\mathrm{GeV} \lesssim M_H \lesssim 130\,\mathrm{GeV}$ will be hard to detect. I suggest several methods at planned future high energy particle colliders for observing this Higgs boson.

At LEP II we have reasonable numbers of Higgs produced in association with a Z boson up to the limit imposed by phase space $M_H < \sqrt{s} - 100\,\mathrm{GeV}$. Unfortunately if the Higgs is degenerate in mass with the Z boson we have large numbers of background events from double Z production. I investigate possible methods round this background. Firstly in polarizing the initial e^+e^- beams, and secondly in studying the differing topologies of the ZH signal, and ZZ background events.

Moving on to the hadron super colliders the LHC and the SSC. These colliders typically produce very clean signals for 'heavy' Higgs. However for a light 'intermediate' mass Higgs all Higgs decays are either dominated by huge QCD backgrounds; or put very strong constraints upon our experimental apparatus. I investigate the signals and backgrounds for an alternative approach where rather than looking for the Higgs in isolation, we look for it produced in association with other heavy particles. Despite these production mechanisms having a far lower rate than isolated Higgs production they have far better signal to background ratios, which makes them look promising. Two modes in particular appear to give encouraging signals; WH production, and $t\bar{t}H$ production. Both these production modes can be detected in the isolated lepton and two photon channel.

Acknowledgements

Above all others two people in particular should be thanked, James Stirling and Nigel Glover. James has been my supervisor for the past 3 years, his tremendous physical understanding; in conjunction with his apparent total grasp of the mathematical subtlies, has been unnerving to watch. I hope one day to possess some small fraction of his talent for making even the most complicated of arguments effortlessly simple. Nigel arrived on the scene in my second year, his fresh young mind with original ideas and new techniques was a very welcome addition to the Centre for Particle Theory. For the past two years barely a day has passed when I have not chatted to him – and this has furthered my understanding of what we are trying to do, why we are trying to do it, and how we should go about it. Certainly he more than anyone else has influenced the way I approach high energy physics. His friendship as well has been very rewarding.

Without either of these two people this thesis would not have been half the thesis that it is; without them both ...

Secondly I would like to thank those people who have time and again been willing to talk shop with me, and furthered my understanding, a few people in particular spring to mind Roger Phillips, Peter Sutton and three young L3 experimenters, Jayant Shukla, Ofer Rind, and Gerhard Raven. The 3 experimenters especially have shown me a side of the subject that I knew nothing about before. I only hope that I have taught them as much about the theory of high energy physics as they have taught me about the experimental side,

"Whaddya mean? You can't tell a π^0 from a photon ... ones a hadron and the others electromagnetic radiation."

"Yes – but most π^0 decay into 2 photons; and if the π^0 is energetic enough the separation of these two photons is too small to be experimentally measured ..."

Then there have been all the other members of the group – they have put up with my non stop inane chatter with remarkably few signs of complaint ! Mark Oakden, Andrew Morgan, Francisco Astorga-Sáenz, Mathew Reader, Neil Shaban, Dunk "the bridge playing hairy hippy" Curtis, Ghadir Abu Leil, David Barclay, Foster Langbein, Ben Bullock, Adnan Bashir, Ramon Muñoz Tapia, Adrian Askew, and Ayse Kizilersu. Also the other members of staff, Mike Pennington, Mike Whalley, Alan Martin and Chris Maxwell.

Finally all the people outside of the physics department that have made my time here in Durham so much fun.

I would also like to thank SERC without whose funding I would not have been able to do this PhD.

Declaration

I declare that no material in this thesis has previously been submitted for a degree at this or any other University.

Parts of Chapter 2 have been published in the paper

"Possible enhancement of a Higgs signal at an e^+e^- collider with $\sqrt{s} = 200$ GeV using polarized beams"
Phys. Lett. B274 (1992) 209.

Chapter 3 is largly based on a talk I gave at the SSC Symposium at Madison, Wisconsin, March 1993. It also contains much work from the published papers,

"Calculation of two photon bremsstrahlung off a $Q\bar{Q}$ pair: A Possible background to detection of an intermediate mass Higgs at high-energy pp colliders"
Phys. Lett. B277 (1992) 366.

"Production of an intermediate mass Higgs boson in association with a single top quark at LHC and SSC" with W.J. Stirling
Phys. Lett. B283 (1992) 411.

"Calculation of the contributions to the signals and backgrounds for intermediate mass Higgs detection at the LHC and SSC from the $q\bar{q}$ initial state"
Phys. Lett. B306 (1993) 129.

The copyright of this thesis rests with the author. No quotation from it should be published without prior written consent and any information derived from it should be acknowledged.

Contents

Introduction

By the time it came to the edge of the Forest the stream had grown up, so that it was almost a river, and being grown-up, it did not run and jump and sparkle along as it used to do when it was younger, but moved more slowly. For it knew now where it was going, and it said to itself, 'there is no hurry. We shall get there some day.' But all the little streams higher up the in the Forest went this way and that, quickly, eagerly, having so much to find out before it was too late.

The House At Pooh Corner, A.A.Milne

I have partitioned this thesis into two distinct halves, the Chapters and the Appendices.

In the Chapters I have tried to describe the physics of a few of the ideas that have occupied my mind for the previous few years, although these Chapters are largely based on my own work these Chapters are by no means my work alone. High energy physics phenomenology has many hundreds of people round the world working on it; with each individual contributing just small pieces to the picture as a whole, and just as with a jigsaw each separate piece is meaningless by itself – it is only when slotted in to the greater picture as a whole that it gains any meaning. In my published papers I have described my contributions in detail, and if the fine details are what the reader requires he should read those papers. Here, in this thesis, I have tried to give a more complete, if less rigorous, picture of the areas I have worked on *as I have seen it*, this necessitates spending time on the ideas and suggestions of others. However in the process I hope that I have given a clear view of the important concepts in the areas where I have worked, in a reasonably readable form.

In the Appendices, which by no means should be looked as as secondary to the Chapters; I have given the mathematical basis upon which the Chapters are based. I firmly believe that in the realm of theoretical quantum theories that physical intuition is secondary to mathematical understanding. Quantum theories are not based upon physical intuition; we can not make

analogies with billiard ball physics, as we can with classical mechanics. To my mind, and having had a mathematical training I may well be biased; the only basis that quantum theories have is a mathematical one. We put forward a particular quantum model, our axioms, and working within these axioms make mathematical statements; these statements we then take as some reflection of physical reality. As such we can make comparisons between these mathematical predictions, and physical reality; and so form a 'working model' of that physical reality. I feel, we are ultimately dependent upon the underlying maths upon which our models of reality are based. In this sense I believe the mathematical framework to be more basic than any physical understanding; and indeed without the mathematical framework all physical predictions are void. Am I saying that my Chapters, in which I try to give a purely physical overview, are empty of any meaning, just some shallow reflection of the underlying mathematical structure? Well, yes and no! Although quantum predictions are mathematically based, and it is this that gives us physical predictions, these physical predictions give us a physical understanding of some aspects of our mathematical model. And this physical understanding we can apply to other processes for which the mathematical details have not yet been calculated, of course without these mathematical details our understanding is not sound. But it is not for sound predictions that we need our physical insight, no one claims, or at least no one should claim, to make a firm prediction on purely physical understanding; however our physical understanding does give us both motivation, and indeed a direction to work in. It gives us the questions to ask, which thoughts we should pursue – and which to drop.

In this sense a good solid physical understanding of the mathematical underflow is invaluable, and indispensable. We should however take care not to trust this physical intuition too far, without the backing of solid mathematical statements.

This thesis does not contain all the material that I have worked on over the period of my PhD, and indeed does not even contain all the published material. In an amazing brainstorming session with Roger Phillips, Frank Close and Jurgen Körner I played a small part in looking at methods of experimentally measuring a b quarks polarization, this resulted in a preprint and contribution to the proceedings of the B physics workshop in Edinburgh [1]. I have also worked on hard photonic radiation at LEP, although this has resulted in a paper [2], this work is far from completion. To include it in this thesis would not only be *presenting a single piece of the jigsaw*, but also as yet an *uncoloured jigsaw piece.* Also as the thesis stands there is a common thread of Higgs Physics running through it – to include other topics would have made the thesis far more a *random walk in phenomenology.*

Finally I should say a few works about the quotes used in this thesis. All quotes have been taken from A.A. Milne's brilliant books Winnie–the–Pooh and The House At Pooh

Corner. The animals that live within the pages of these books live in a remarkably simple world, much as I believe we do. However to these animals their word is full of complications and misunderstandings, indeed much as my PhD has been! And yet despite these difficulties these animals are (on the whole) thrilled by their world, and the wonders it contains – and this is, perhaps, what I find hardest to explain; quite how much I've enjoyed doing this PhD as a whole. Many is the time that I've felt totally ignorant, like Pooh; or omniscient (but wrong), like Owl; or indeed just down right fed up, like dear Eeyore. As such I hope that these quotes reflect much that I have felt while doing this PhD; and of course these quotes provide occasional light relief – the text is heavy enough going as it is ...

Chapter 1

The Standard Model

NOISE, BY POOH

Oh, the butterflies are flying,
Now the winter days are dying,
And the primroses are trying
To be seen.
And the turtle-doves are cooing,
And the woods are up and doing,
For the violets are blue-ing
In the green

Oh, the honey-bees are gumming
On their little wings, and humming
That the summer, which is coming,
Will be fun
And the cows are almost cooing,
And the turtle-doves are mooing
Which is why a Pooh is poohing
In the Sun.

For the spring is really springing;
You can see a skylark singing,
And the blue-bells, which are ringing,
Can be heard.
And the cuckoo isn't cooing,
But he's cucking and he's cooing,
And a Pooh is simply poohing
Like a bird.

The House At Pooh Corner, A.A.Milne

1.1 Introduction

In this chapter I review the basic ideas that make up the Standard Model of particle physics. This is strongly based upon the books given in Ref. 3. The more technical mathematical details are given in Appendices A and B.

1.2 The 4 Forces

We believe that there are 4 fundamental forces is the universe. Two of these, the gravitational forces, and the electromagnetic force, are manifestly obvious in the world about us. The gravitational forces is the force that pulls us to the surface of the Earth, that causes the Earth to orbit the sun, and the sun to hold its position in the Galaxy as a whole. This is

described classically in Einstein's Theory of General Relativity, and for physics on the large scale (planet sizes and larger) this force rules supreme.

The second force that we encounter in our day to day lives is the electromagnetic force, this is the force responsible for making objects solid to the touch, light that we see has its origins in this force, as well as radio waves. This force acts between electrically charged particles, and is enormously stronger than the gravitational force – indeed if it were not for the fact that most matter is, at least on macroscopic scales, electrically neutral, this force would dominate in the universe. Classically this force is described by Maxwell's equations.

The remaining two forces are not so immediately obvious in the world, largely because the nature of these forces is very short distance indeed (although this is for differing reasons as we shall see).

The first is the weak force, this is the force that is responsible for radioactive decay, it has a very short range and so can only affect physics inside the nucleus of atoms, radioactive decay being the decay of the nucleus of an atom. Although this force is called the weak force, at high energies, or short distances, its strength is similar to the electromagnetic force; only at low energies and large distances does this force appear weak.

Lastly we have the strong force, again this is a force that is only evident inside the nucleus of atoms. Indeed it is the force responsible for holding the nucleus together in the first place. As its name suggests it is the strongest of all forces, as long as we are at short enough distances.

For three of these forces, the electromagnetic, weak, and strong, a quantum theory exists; indeed it is only in the framework of quantum theories that the weak and strong forces have any meaning. Recall that they are short distance forces, and quantum theories typically hold sway over classical theories at small distance scales. On the other hand at the moment we have no satisfactory model for a quantum version of gravity, all attempts so far to quantize gravity have been plagued with deep theoretical problems.

In this thesis I will deal with the quantum theories of the electromagnetic, weak, and strong forces; and the strange world that they live in. I shall look at how we can test, and measure the predictions of these models. This necessitates working at very short distances and very high energies, in practice these energies are only achievable at high energy particle colliders. This means that I will be considering the physics potential of these high energy particle colliders.

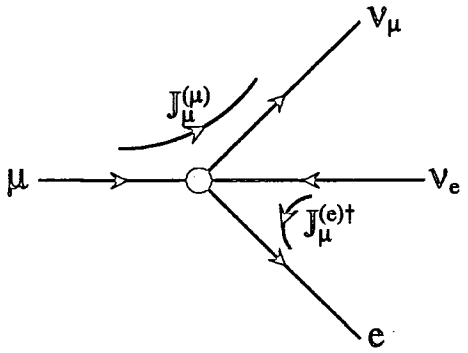


*Fig.*1.3.1 The Feynman diagram for the decay $\mu \rightarrow e\nu_\mu\bar{\nu}_e$, via the Fermi 4 point interaction. We have also show the two currents involved.

1.3 The Weak Force

The earliest attempts to describe the weak interaction, for example β decay of the neutron, were put forward by Fermi in 1934, he described the decay in terms of a 4 point interaction. This is equivalent to the interaction Lagrangian was given by,

$$\mathcal{L}_I = \frac{G_F}{\sqrt{2}} J_\mu J^{\mu\dagger} \quad , \tag{1.3.1}$$

where J_μ is the sum of the two currents involved in the decay $J_\mu^l + J_\mu^h$. If we consider a muon decaying.

$$\mu \rightarrow e\nu_\mu\bar{\nu}_e \quad . \tag{1.3.2}$$

This interaction gives rise to the matrix element,

$$\mathcal{M} = G_F(\bar{u}(\nu_\mu)\gamma^\mu u(\mu))(\bar{u}(e)\gamma_\mu u(\nu_e)) \quad , \tag{1.3.3}$$

where $\bar{u}(\nu_\mu)\gamma_\mu u(\mu)$ is the muon current, $J_\mu^{(\mu)}$, and $\bar{u}(e)\gamma_\mu u(\nu_e)$ is the electron current, $J_\mu^{(e)\dagger}$. This is shown in *Fig.*1.3.1.

However the Lagrangian (1.3.1) conserves parity, and in 1956, it was conclusively shown in the weak decay,

$$^{60}\mathrm{Co} \rightarrow {}^{60}\mathrm{Ni}^* + e^- + \bar{\nu}_e \quad , \tag{1.3.4}$$

that the weak interaction does not conserve parity. Subsequent work showed that the fermion current should read,

$$J_\mu^{(\mu)} = \bar{u}(\nu_\mu)\gamma_\mu(1-\gamma_5)u(\mu) \quad , \tag{1.3.5}$$

and so the matrix element reads,

$$\mathcal{M} = \frac{G_F}{\sqrt{2}}(\bar{u}(\nu_\mu)\gamma^\mu(1-\gamma_5)u(\mu))(\bar{u}(e))\gamma_\mu(1-\gamma_5)u(\nu_e) \quad . \tag{1.3.6}$$

However if we keep the 4 point interaction we run into serious difficulty, as the energy of collision is increased in say the process,

$$p e^- \to n \bar{\nu} \quad , \tag{1.3.7}$$

the cross-section grows rapidly without limit, however a very basic assumption about the world in which we live is that the physics is unitary, *i.e.*, that probability is conserved; and this implies that the cross-section for (1.3.7) must grow slower than,

$$\sigma < \text{constant}(\ln s)^2 \quad . \tag{1.3.8}$$

This is known as the Froissart bound [4], and unfortunately process (1.3.7) grows quicker than this bound. Also if we calculate in this theory beyond the leading order we find that the theory is plagued with infinities, and for each extra order in perturbation theory we acquire new infinities (this follows because the coupling G_F is dimensionful).

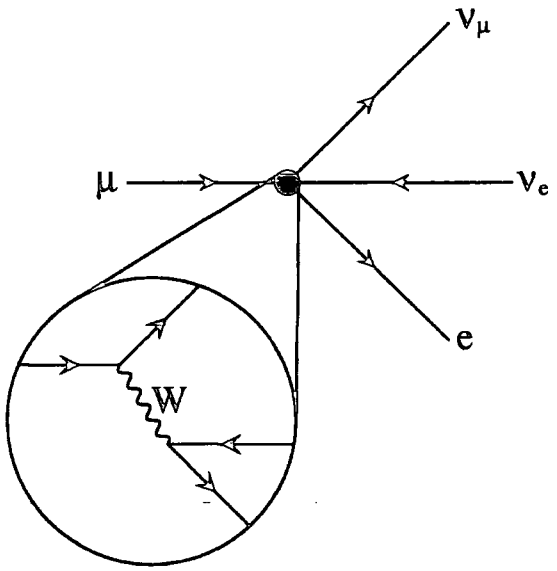


*Fig.*1.3.2 The Feynman diagram for the Intermediate Vector Boson Model, the Fermi 4 point interaction when looked at closely enough is due to the exchange of a massive vector boson, the W.

The first of these problems, the bad high energy behaviour, was solved by the Intermediate Vector Boson Model. In this model we remove the Fermi 4 point interaction, and instead say that both currents involved couple to some super heavy vector boson, the W, this is shown in *Fig.*1.3.2. This gives the matrix element,

$$\mathcal{M} = -(\bar{u}(\nu_\mu) g_W \gamma^\mu (1-\gamma_5) u(\mu)) \frac{-g_{\mu\nu} + k_\mu k_\nu / M_W^2}{k^2 - M_W^2} (\bar{u}(e)) g_W \gamma^\nu (1-\gamma_5) u(\nu_e) \ , \tag{1.3.9}$$

where k is the 4 momentum flowing down the W. Now in the limit that k is small relative to M_W, as for example in muon decay, this matrix element reduces to the Fermi 4 point matrix element (1.3.6) if,

$$\frac{G_F}{\sqrt{2}} = \left(\frac{g_W}{m_W}\right)^2 \quad . \tag{1.3.10}$$

However when k becomes large (1.3.9) has a very different behaviour to (1.3.6) and drops off as $1/k^2$. This gives the Intermediate Vector Boson model good high energy behaviour, at least for processes like $p + e^+ \rightarrow n\bar{\nu}$.

Although we have cured some problems in the Intermediate Vector Boson model, we have by no means cured all difficulties. Although many processes that previously had bad high energy behaviour now have good high energy behaviour we have created a whole new set of process like,

$$\nu\bar{\nu} \rightarrow W^+W^- \quad , \tag{1.3.11}$$

that now acquire bad high energy behaviour. This means that again our model is not renormalizable. When we calculate beyond leading order, all process that we calculate have infinite cross-section, and even if we renormalize away this infinity, we get extra new infinities at each extra order in perturbation theory. As well as being unaesthetic this means that any answer that we obtain is essentially arbitrary, as renormalizing in a different way gives a different answer.

1.4 Electro-Weak Theory

These problems were finally solved by Steven Weinberg and Abdus Salam in their unified theory of weak and electromagnetic interactions [5].

In this model weak and electromagnetic interactions are based on an $SU(2) \otimes U(1)$ gauge theory. We have the covariant derivative define as,

$$D_\mu = \partial_\mu - igT^i A^i_\mu - ig'Y/2B_\mu \quad , \tag{1.4.1}$$

where T^i are the generators of $SU(2)$, and Y the generator of $U(1)$. Y is known as the *weak hypercharge*. The electric charge of particles, Q, is given by,

$$Q = T_3 + Y/2 \quad . \tag{1.4.2}$$

To start off with all gauge bosons, which we wish to associate with photons and W bosons, are forced to be massless. However we know that if we are to model the weak interaction we must have massive bosons. This is achieved by spontaneous symmetry breaking, the Higgs

mechanism. First we add an extra complex doublet scalar field with weak hypercharge $Y = 1$, and $T = \frac{1}{2}$, to the model, this is called the Higgs field.

$$\Phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} \qquad Y(\Phi) = 1 \qquad T(\Phi) = \frac{1}{2} \quad . \tag{1.4.3}$$

We then arrange that the vacuum state of this field spontaneous breaks the symmetry of the Lagrangian by choosing that the Higgs Lagrangian is given by,

$$\mathcal{L}_H = (D_\mu \Phi)^\dagger (D^\mu \Phi) - V(\phi) \tag{1.4.4}$$

$$V(\Phi) = -\mu^2 \Phi^\dagger \Phi + \lambda (\Phi^\dagger \Phi)^2 \quad , \tag{1.4.5}$$

where $\mu^2 > 0$ implies that the minimum of the potential is at $|\Phi| = v/\sqrt{2}$, where,

$$v = \left(\frac{\mu^2}{\lambda} \right)^{1/2} \quad . \tag{1.4.6}$$

Now we have a gauge freedom in how we choose this vacuum, and so we can perform rotations in $SU(2)$ space. This means that we can choose Φ's representation to be,

$$\Phi(x) = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v + \eta(x) \end{pmatrix} \quad . \tag{1.4.7}$$

This particular choice of gauge is called the unitary gauge. In this gauge we just have one physical neutral CP even scalar associated with the η field above, this particle is known as the Higgs boson. Expanding the Higgs Lagrangian in terms of this η field we find that its mass is,

$$M_H = (2\mu^2)^{1/2} \quad , \tag{1.4.8}$$

unfortunately this parameter μ^2 is totally unknown, and so we can not make a solid prediction for the mass of this fundamental scalar. Recent negative searches at LEP I have shown that [6],

$$M_H > 62.5\,\text{GeV} \qquad \text{at the 95\% confidence level.} \tag{1.4.9}$$

Also if the Higgs particle is too heavy then in the high energy limit process like,

$$WW \to WW \quad , \tag{1.4.10}$$

break unitarity before the physical effect of the Higgs boson can cure this problem (for lighter Higgs the process $WW \to H \to WW$ gives the process (1.4.10) good high energy behaviour). This means that we expect,

$$M_H \lesssim 1\,\text{TeV} \quad . \tag{1.4.11}$$

If we now look at the terms arising from the covariant derivative term $(D_\mu\Phi)^\dagger(D^\mu\Phi)$ in $\mathcal{L}_H$. The terms that are quadratic in the gauge fields A^i and B, and have no dependence on η, read,

$$\mathcal{L}_{VMM} = \frac{v^2}{8}\left(g^2\left[(A^1_\mu)^2 + (A^2_\mu)^2\right] + (gA^3_\mu - g'B_\mu)^2\right) \tag{1.4.12}$$

$$\equiv M_W^2 W^+_\mu W^{-\mu} + \frac{1}{2}M_Z^2 Z_\mu Z^\mu \quad , \tag{1.4.13}$$

where in the last line we have defined for the charged vector mesons,

$$W^\pm_\mu = (A^1_\mu \mp iA^2_\mu)/\sqrt{2} \tag{1.4.14}$$

$$M_W^2 = g^2v^2/4 \quad . \tag{1.4.15}$$

The means that we seem to have two massive charged vector bosons, a W^+ and a W^-. These play the part of the massive vector boson in the Intermediate Vector Boson model. Looking at the second term in (1.4.13) we find,

$$\begin{aligned}\frac{1}{2}M_Z^2 Z_\mu Z^\mu &= \frac{v^2}{8}\begin{pmatrix} A^3_\mu & B_\mu\end{pmatrix}\begin{pmatrix} g^2 & -gg' \\ -gg' & g'^2\end{pmatrix}\begin{pmatrix} A^3_\mu \\ B_\mu\end{pmatrix} \\ &= \frac{1}{2}\begin{pmatrix} Z_\mu & A_\mu\end{pmatrix}\begin{pmatrix} M_Z^2 & 0 \\ 0 & 0\end{pmatrix}\begin{pmatrix} Z_\mu \\ A_\mu\end{pmatrix}\end{aligned} \quad , \tag{1.4.16}$$

where we have defined,

$$\begin{aligned} Z_\mu &= \cos\theta_W A^3_\mu - \sin\theta_W B_\mu \quad , \\ A_\mu &= \sin\theta_W A^3_\mu + \cos\theta_W B_\mu \quad , \end{aligned} \tag{1.4.17}$$

$$\tan\theta_W = g'/g \quad , \tag{1.4.18}$$

$$M_Z^2 = v^2(g^2 + g'^2)/4 \quad . \tag{1.4.19}$$

This means that we have two neutral vector bosons, one massive, and one massless. The massless neutral vector boson we associate with the photon, and the massive neutral vector boson is a new particle, the Z^0 boson.

Now we know that when we calculate using this Weinberg–Salam model that at low energies we should reproduce the Fermi 4 point interaction. This tells us that,

$$\frac{G_F}{\sqrt{2}} = \frac{g^2}{8M_W^2} \quad , \tag{1.4.20}$$

and this with (1.4.15) tell us that,

$$v = 2^{-1/4}G_F^{-1/2} \simeq 246\,\text{GeV} \quad . \tag{1.4.21}$$

Now neutral current exchanges tell us a similar story, and these give a value for,

$$\sin^2\theta_W \simeq 0.23 \quad , \tag{1.4.22}$$

also the massless neutral gauge boson must couple with the same strength as the photon, and this gives,

$$g\sin\theta_W = g'\cos\theta_W = e \quad , \tag{1.4.23}$$

Using this information M_W and M_Z can be predicted to be,

$$M_W = 78\,\mathrm{GeV} \quad , \tag{1.4.24}$$

$$M_Z = 89\,\mathrm{GeV} \quad . \tag{1.4.25}$$

Now when the W and Z were discovered at UA1 and UA2 at CERN and turned out to have masses of,

$$M_W^{\mathrm{UA1}} = 80.9 \pm 1.5 \pm 2.4\,\mathrm{GeV} \quad M_W^{\mathrm{UA2}} = 81.0 \pm 2.5 \pm 1.3\,\mathrm{GeV} \quad , \tag{1.4.26}$$

$$M_Z^{\mathrm{UA1}} = 95.6 \pm 1.4 \pm 2.9\,\mathrm{GeV} \quad M_Z^{\mathrm{UA2}} = 91.9 \pm 1.3 \pm 1.4\,\mathrm{GeV} \quad , \tag{1.4.27}$$

this was stunning verification of the Weinberg–Salam model.

In everyday life the 'weak' part of the Weinberg–Salam model, in otherwords the physics of W's and Z's, is much weaker than the electromagnetic part, this is because W's and Z's are very massive indeed on the scale of day to day physics. Now the uncertainty principle tells us that we can 'borrow' energy, but the more we borrow the less time we can borrow it for. To produce a W or a Z at low energies we need to borrow a great deal of energy – and this must be paid back within very short time scales. However at higher energies, when we no longer have to borrow energy, the weak force becomes as strong as the electromagnetic force.

1.5 The Fermionic Sector

So far I have not mentioned the role of fermions in the Weinberg–Salam model, and indeed this sector is distinctly different to the sector that we would normally get from using a covariant derivative in the Dirac equation. The underlying reason for this difference is that we don't see any right handed neutrinos, but only left handed neutrinos. This is because for the weak current parity is maximally broken, and adding a covariant derivative to the Dirac equation would conserve parity. With this in mind in the Weinberg–Salam model we start

by writing the Dirac equation in such a way that the different chirality states of the fermions are manifest. If we consider the Lagrangian for the massless Dirac equation,

$$\begin{aligned} \mathcal{L}_F &= \bar{\psi} i\gamma^\mu \partial_\mu \psi \\ &= \bar{\psi}_l i\gamma^\mu \partial_\mu \psi_l + \bar{\psi}_r i\gamma^\mu \partial_\mu \psi_r \quad , \end{aligned} \tag{1.5.1}$$

where we have written the left handed chirality separately from the right handed chirality. Now we can assign the left handed chirality and the right handed chirality *different* weak hypercharge, Y, and $SU(2)$ representations, T. We choose the values shown in *Table* 1.5.1.

$\begin{matrix}\nu_l \\ e_l\end{matrix}$	$Y = \quad 1$	$T = 1/2$	$T_3 = \begin{matrix}1/2 \\ -1/2\end{matrix}$
e_r	$Y = -2$	$T = \quad 0$	$T_3 = \quad 0$
$\begin{matrix}u_l \\ d_l\end{matrix}$	$Y = 1/3$	$T = 1/2$	$T_3 = \begin{matrix}1/2 \\ -1/2\end{matrix}$
u_r	$Y = 4/3$	$T = \quad 0$	$T_3 = \quad 0$
d_r	$Y = -2/3$	$T = \quad 0$	$T_3 = \quad 0$

Table 1.5.1 The Y hypercharge, and T $SU(2)$ representation of the fermions and quarks. Particles grouped in boxes fall in the same $SU(2)$ representation. The three generations of families have the same hypercharge and $SU(2)$ representation, ie $e = \mu = \tau$, $\nu_e = \nu_\mu = \nu_\tau$, $u = c = t$, and $d = s = b$.

So for example in the leptonic sector we have,

$$\begin{aligned} \mathcal{L}_l = &\left(\bar{\psi}_{\nu_l} \quad \bar{\psi}_{e_l}\right) i\gamma^\mu \left(\partial_\mu - i\frac{g}{2}\tau^a A^a_\mu + i\frac{g'}{2}B_\mu\right)\begin{pmatrix}\psi_{\nu_l} \\ \psi_{e_l}\end{pmatrix} \\ &+ \bar{\psi}_{e_r} i\gamma^\mu \left(\partial_\mu + ig' B_\mu\right)\psi_{e_r} \end{aligned} \quad , \tag{1.5.2}$$

Using this set up we correctly get the $V - A$ form of the W coupling to fermion fields, although the Z coupling is more complicated due to mixing between the A^a_μ and B_μ fields. We find that the Z coupling to fermions is,

$$g_{Zf\bar{f}} = \frac{g}{cos\theta_W}(T_3 - Q\sin^2\theta_W) \quad , \tag{1.5.3}$$

and as left and right handed fermions have differing T_3 charge the $Zf\bar{f}$ coupling is also parity violating.

Now usually in the Dirac equation we introduce the mass of the fermion through terms like,

$$\begin{aligned} \mathcal{L} &= -m\bar{\psi}\psi \\ &= -m(\bar{\psi}_l\psi_r + \bar{\psi}_r\psi_l) \quad . \end{aligned} \tag{1.5.4}$$

However if we put ψ_l and ψ_r in different $SU(2)$ multiplets then we are not allowed to construct terms like $\bar{\psi}_l\psi_r$, as this is no longer a $SU(2)$ invariant. To get round this difficulty we use the Yukawa mechanism where we introduce couplings between the Higgs field and fermions. We modify the Lagrangian with terms like,

$$\mathcal{L}_{\text{Yukawa}} = -f^{(e)}\bar{\psi}_{l_l}\Phi\psi_{e_r} - f^{(u)}\bar{\psi}_{q_l}\tilde{\Phi}\psi_{u_r} - f^{(d)}\bar{\psi}_{q_l}\Phi\psi_{d_r} + \text{h.c.} \quad , \tag{1.5.5}$$

where,

$$\tilde{\Phi} = -2iT^2\Phi^* \qquad \text{with } Y(\tilde{\Phi}) = -1 \quad . \tag{1.5.6}$$

Substituting (1.4.7) we find,

$$\mathcal{L}_{\text{Yukawa}} = -f^{(e)}\psi_e(v+\eta)\psi_e - f^{(u)}\psi_u(v+\eta)\psi_u - f^{(d)}\psi_d(v+\eta)\psi_d \quad , \tag{1.5.7}$$

and so fermions have acquired a mass, fv; however we also learn that the fermions coupling to the Higgs particle associated with the η field is f. So the Yukawa mechanism tells us that a Higgs bosons coupling to fermions is given by,

$$g_{Hf\bar{f}} = \frac{m_f}{v} \quad . \tag{1.5.8}$$

1.6 Mass eigenstates vs. electroweak eigenstates – the CKM matrix

In the above I assumed that the states that the Higgs field couples to are the same states that the gauge particles couple too, in fact this does not necessarily have to happen. Between particle families the form of the theory is fixed in the Weinberg–Salam model; however within particle families we can have a unitary rotation between the different members of the family. Usually the mass eigenstates are taken to define the fermions, then the weak eigenstates are taken to be,

$$\begin{aligned} \begin{pmatrix} e' \\ \mu' \\ \tau' \end{pmatrix} &= S_l \begin{pmatrix} e \\ \mu \\ \tau \end{pmatrix} & \begin{pmatrix} \nu'_e \\ \nu'_\mu \\ \nu'_\tau \end{pmatrix} &= S_\nu \begin{pmatrix} \nu_e \\ \nu_\mu \\ \nu_\tau \end{pmatrix} \\ \begin{pmatrix} u' \\ c' \\ t' \end{pmatrix} &= S_u \begin{pmatrix} u \\ c \\ t \end{pmatrix} & \begin{pmatrix} d' \\ s' \\ b' \end{pmatrix} &= S_d \begin{pmatrix} u \\ c \\ t \end{pmatrix} \quad . \end{aligned} \tag{1.6.1}$$

Any so the Lagrangian for the Dirac equation becomes,

$$\mathcal{L}_F = \bar{\psi}S^\dagger i\gamma^\mu\partial_\mu S\psi \quad . \tag{1.6.2}$$

Now if the $\bar{\psi}$ field is the same field as the ψ field then clearly the $S^\dagger$ cancels on the S, as S is unitary, and this means that for neutral gauge bosons the unitary transformation is

unobservable. Also as all three neutrinos are massless, S_ν can be an arbitrary unitary matrix; choosing $S_\nu = S_l$ ensures that we can never observe either S_ν nor S_l; however in the quark sector we have the term,

$$\mathcal{L}_F = \bar{\psi}_u S_u^\dagger i\gamma^\mu D_\mu S_d \psi_d \quad , \tag{1.6.3}$$

when we have W exchange, in this case the product $S_u^\dagger S_d$ is observable, and is known as the Cabibbo–Kobayashi–Maskawa matrix, U. Experimentally this matrix is far from the identity matrix, it is parametrised as,

$$U = \begin{pmatrix} U_{ud} & U_{us} & U_{ub} \\ U_{cd} & U_{cs} & U_{cb} \\ U_{td} & U_{ts} & U_{tb} \end{pmatrix} \quad , \tag{1.6.4}$$

and we experimentally we know that approximately,

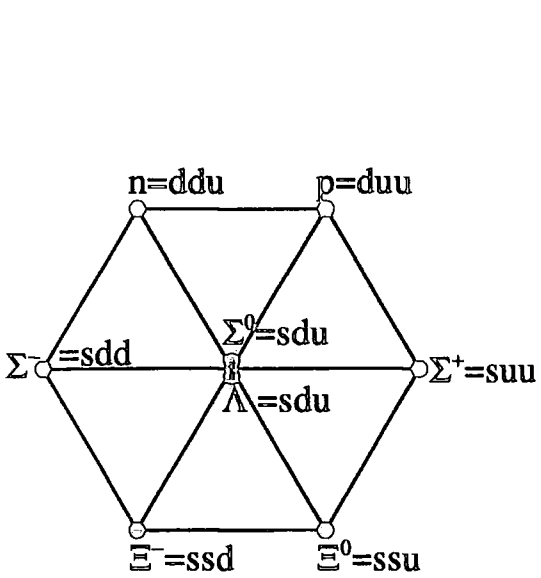
$$U = \begin{pmatrix} |U_{ud}| = 0.973 & |U_{us}| = 0.23 & |U_{ub}| \simeq 0 \\ |U_{cd}| \simeq 0.24 & |U_{cs}| \simeq 0.97 & |U_{cb}| \simeq 0.06 \\ |U_{td}| \simeq 0 & |U_{ts}| \simeq 0 & |U_{tb}| \simeq 1 \end{pmatrix} \quad . \tag{1.6.5}$$

We could also assign different rotations to the left and right handed fermions, however within the Standard Model we choose the same rotation for both.

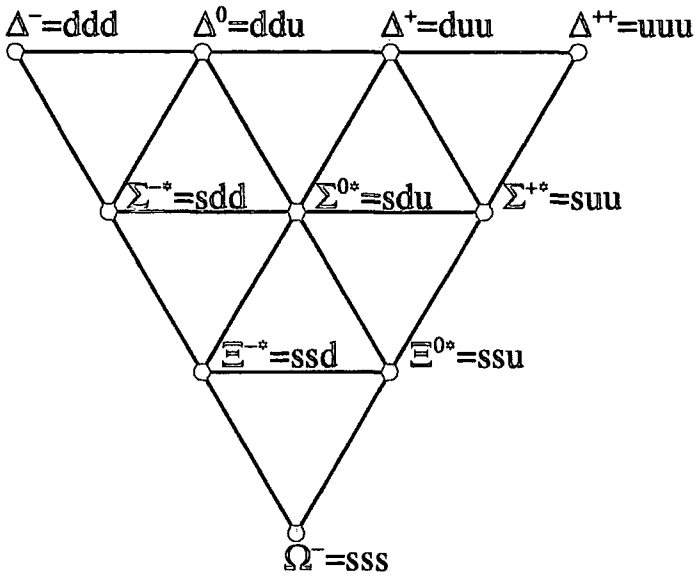
1.7 The Strong Force – QCD

In the 1950's particle accelerators were producing new hadrons every other day. This was great for the experimentalists, always something new and unexpected round the corner – however for the theoretician this was a nightmare; how can you describe something that is changing daily ? Gradually patterns began to emerge, and the quark model was put forward. In this model the hadrons are not taken as the fundamental building blocks, but are taken to be made up of quarks and antiquarks. We show two of these patterns, and the underlying quark content that gives them their structure in *Fig.*1.7.1.

Although the quark model explained the structure of these patterns it raised many difficulties. These states were very tightly bound, in a way that could not be accounted for by electromaganetism or the weak force. Also there were problems with spin statistics. These quarks had spin 1/2, and so obeyed Fermi–Dirac statistics; this means that you can not place two identical quarks in the same state; however in the quark model in the particles Δ^-, Δ^{++} and Ω^- we have respectively three identical d, u and s quarks; also all these quarks have their spins aligned in the same direction to give the overall spin of 3/2. The wavefunction of these particles is also symmetric and so we have 3 fermions in an identical state! This meant that for many people the quark model was not taken as a physical model, but merely a convenient way of describing the group structure of the hadronic patterns. Soon deep inelastic events at



a) The octet of spin 1/2 baryons



b) The decuplet of spin 3/2 baryons

*Fig.*1.7.1 Hadron patterns, here we have plotted I_3, the third component of isospin, against Y the hypercharge. Also shown is the underlying quark structure.

SLAC showed that hadrons were not pointlike, but were made up of smaller particles, called partons.

It was natural to associate these partons with the quarks of hadronic patterns; however this leaves the problem of how the hadrons held themselves together, and how they evaded Fermi–Dirac statistics. The solution finally hit upon was to say that the three quarks in hadrons are not identical; that they carry some extra label, or charge, that distinguishes them from each other. As baryons are made up of 3 quarks we clearly require at least 3 labels; this extra charge is called colour, and the 3 extra labels are r, g and b, standing for red, green and blue. Now this colour is never actually seen on the hadrons themselves, and so nature arranges itself so that all asymptotic particles are colourless; that is they contain equal amounts of r, g and b.

Now turning to the force that holds these hadrons together, well an "obvious" choice is a field that couples to the colour content of quarks; then the fact that all asymptotic particles are colourless would explain why this force is not observed on long distance scales. So adding this force that couples to colour as a gauge field we must use a $SU(3)$ gauge group as we have 3 colours. This $SU(3)$ gauge field theory of the strong force that holds hadrons together is known as Quantum Chromo Dynamics, or QCD.

Now we can ask, *Why should quarks only come in 3 colours? Why not 4 or 5 ?*; well if there were more than the 3 colours, then we could not arrange for observed baryons, of 3

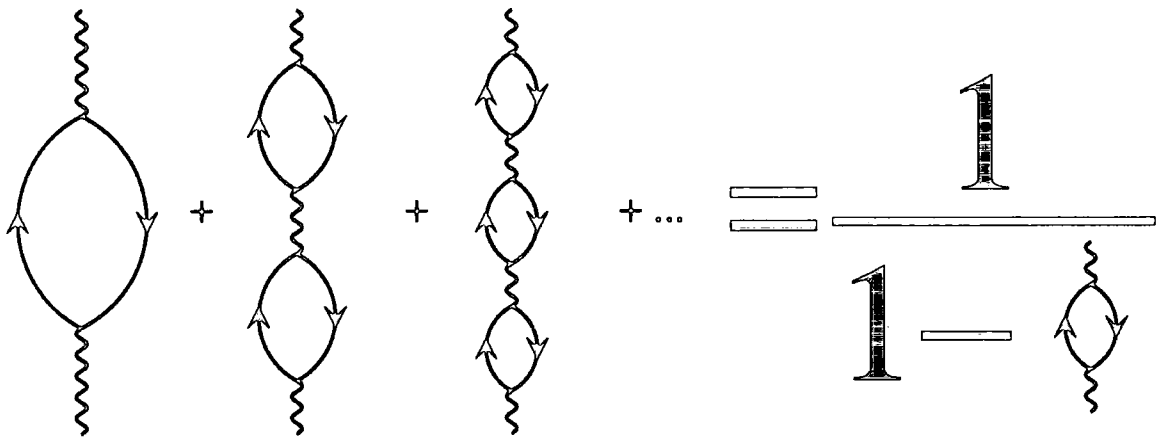


*Fig.*1.8.1 The bubble diagrams that give corrections to the photon propagator, these can be resummed to give symbolically the form on the right.

quarks, to be colour singlets, *i.e.*, colourless. Also if we look at the process,

$$e^+e^- \to \gamma^* \to q\bar{q} \quad , \tag{1.7.1}$$

and compare this too

$$e^+e^- \to \gamma^* \to \mu^+\mu^- \quad , \tag{1.7.2}$$

then we find (at lowest order at least),

$$\frac{\sigma(e^+e^- \to \gamma^* \to q\bar{q})}{\sigma(e^+e^- \to \gamma^* \to \mu^+\mu^-)} = N_c \frac{e_q^2}{e_\mu^2} \quad . \tag{1.7.3}$$

Now we can not detect these final quarks, as they *hadronize* into a collection of hadrons; however we can form the ratio, R,

$$R \equiv \frac{\sigma(e^+e^- \to \gamma^* \to \text{hadrons})}{\sigma(e^+e^- \to \gamma^* \to \mu^+\mu^-)} = N_c \sum_{\text{quarks}} \frac{e_q^2}{e_\mu^2} \quad , \tag{1.7.4}$$

where we sum over the quarks that the virtual photon is kinematically allowed to decay to. If this is done then the experimental evidence points very strongly to there being just 3 colours, *i.e.*, $N_c = 3$.

1.8 Running couplings

At lowest order the strength of the gauge field coupling is a fixed constant, however as we move to higher orders in perturbation theory this simple set up is modified. In particular the bubble insertion diagrams shown in *Fig.*1.8.1 can significantly alter the photon propagator. Now as each bubble is identical to every other bubble (it has the same momenta flowing in

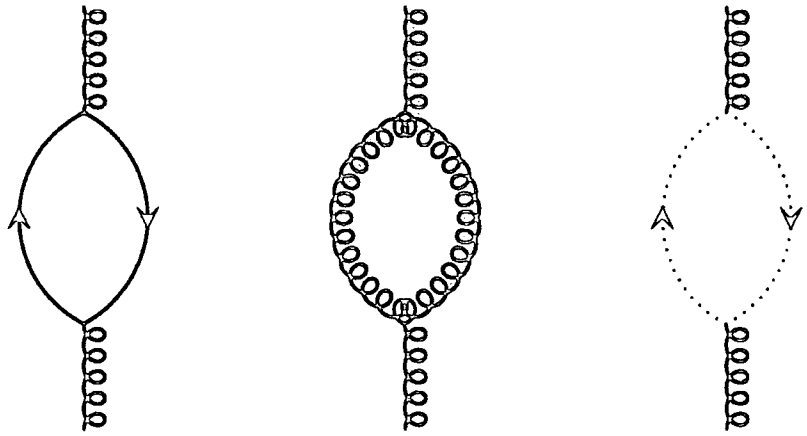


*Fig.*1.8.2 The bubble Feynman diagrams that contribute to the running of α_S. As well as fermion loops we also have gluon loops and Faddeev–Popov ghosts.

and out of it), these string of bubbles form a geometric series, and so can be resummed. If we do this we get corrections to the photon propagator, these corrections can be rewritten as a modified coupling constant. We find,

$$\alpha_{em}(Q^2) = \frac{\alpha_{em}(\mu^2)}{1 - \alpha_{em}(\mu^2)/3\pi \log(Q^2/\mu^2)} \quad . \tag{1.8.1}$$

Where Q^2 is the momentum flowing down the photon leg, and μ^2 is some reference scale at which we have measured α_{em}. This means that as we measure α at higher energies that its strength appears to get larger; and so if we measure α at the electron mass, *i.e.*, at $Q^2 = m_e^2$ we find $\alpha(m_e^2) = 1/137$, however by the Z mass we find $\alpha(M_Z^2) = 1/128$.

If we now play the same game with QCD we find the bubble diagrams shown in *Fig.*1.8.2, in this case we also have contributions from gluon and ghost loops as well as fermion loops – now the gluon loops enter with a different sign than the fermion loops; and in fact dominate the fermion loops. We find,

$$\alpha_S(Q^2) = \frac{\alpha_S(\mu^2)}{1 + \alpha_S(\mu^2)/12\pi(33 - 2n_f)\log(Q^2/\mu^2)} \quad , \tag{1.8.2}$$

where n_f is the number of active fermion flavours in the loop. Notice that the overall sign in front of the $\log(Q^2/\mu^2)$ has flipped; this means that α_S runs in the opposite direction to α_{em} and so as we go to lower Q^2, α_S increases. In particular at some low energy α_S diverges; calling this energy Λ we can rewrite (1.8.2) as,

$$\alpha_S(Q^2) = \frac{12\pi}{(33 - 2n_f)\log(Q^2/\Lambda^2)} \quad . \tag{1.8.2}$$

Does this mean that at some Q^2 that the strong coupling constant diverges? Well no, we have used perturbation theory to obtain (1.8.2), and if α_S becomes large our perturbation theory

breaks down. All we can really say is that at low Q^2 perturbation theory is not valid; and we must use other methods.

However this does go some way to giving us understanding of why we do not see asymptotic coloured states. If we create a colour—anti-colour pair as they separate from each other spatially α_S increases in strength, until it becomes strong enough to pull a $q\bar{q}$ pair out of the vacuum. This continues to happen until all the various quarks and gluons have arranged themselves into *spatially compact* colourless states; which we see as hadrons.

Chapter 2

Higgs Boson Phenomenology at e^+e^- colliders

Eeyore, the old grey Donkey, stood by the side of the stream, and looked at himself in the water.

'Pathetic,' he said. 'That's what it is. Pathetic.'

He turned and walked slowly down the stream for twenty yards, splashed across it, and walked slowly back on the other side. Then he looked at himself in the water again.

'As I thought,' he said. 'No better from this side. But nobody minds. Nobody cares. Pathetic, that's what it is.'

Winnie–the–Pooh, A.A.Milne

2.1 Introduction

In this chapter I describe searches for the Standard Model Higgs boson at LEP I and LEP II. The sections on polarizing the e^+e^- beams, and topological searches at LEP II are my own work and has been partally published in Ref. 7, with the rest of the Chapter being review of other peoples work.

The Higgs Boson by its very nature only couples strongly to massive particles, which means W's, Z's, and t quarks. Hence in any process where we hope to produce Higgs particles at a reasonable rate we require other heavy particles to be present, either as real external particles or as virtual internal particles.

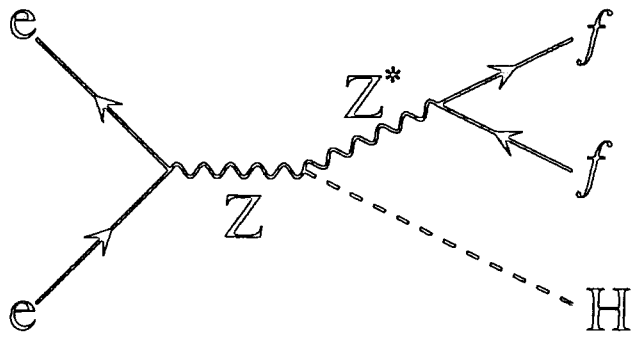


*Fig.*2.2.1 Higgs production at LEP I via $Z \rightarrow HZ^*$

2.2 Higgs Boson Phenomenology at LEP I

Now LEP I is a Z factory, and so we could hope that it would give a large source of Higgs particles from processes like *Fig.*2.2.1,

$$e^+e^- \rightarrow Z \rightarrow Z^*H \quad . \tag{2.2.1}$$

This process, *Fig.*2.2.1, can be easily calculated and we find[8, 9],

$$\frac{1}{\Gamma(Z^0 \rightarrow f\bar{f})}\frac{d\Gamma(Z^0 \rightarrow Hf\bar{f})}{dx} = \frac{g^2(12 - 12x + x^2 + 8y^2)\sqrt{x^2 - 4y^2}}{192\pi^2 \cos^2\theta_W((x-y^2)^2 + \gamma^2)} \quad , \tag{2.2.2}$$

where,

$$x = \frac{2E_H}{m_{Z^0}}, \qquad y = \frac{m_H}{m_{Z^0}}, \qquad \gamma = \frac{\Gamma_{Z^0}}{m_{Z^0}} \quad . \tag{2.2.3}$$

When the virtual Z is well off mass shell we can ignore the Z width Γ_{Z^0}, this happens when the Higgs has $M_H > \Gamma_{Z^0}$. In this limit the partial cross-section (2.2.2) can be integrated exactly and we get[9],

$$\begin{aligned}\frac{\mathrm{Br}(Z \rightarrow Hf\bar{f})}{\mathrm{Br}(Z \rightarrow f\bar{f})} = \frac{g^2}{192\pi^2\cos^2\theta_W}\Bigg[&\frac{3y(y^4 - 8y^2 + 20)}{\sqrt{4-y^4}}\cos^{-1}\left(\frac{y(3-y^2)}{2}\right) \\ &- 3(y^4 - 6y^2 + 4)\ln y - \frac{1}{2}(1-y^2)(2y^4 - 13y^2 + 47)\Bigg] \quad ,\end{aligned} \tag{2.2.4}$$

This is shown in *Fig.*2.2.2. Also of importance is the process $Z \rightarrow H\gamma$ which proceeds via the Feynman diagrams shown in *Fig.*2.2.3. This process proceeds via loops and so would seem to be far less important than $Z \rightarrow Z^*H \rightarrow f\bar{f}H$, however it wins out by only having two particles in the final state and so as M_H increases its available phase space does not drop so rapidly and so this process dominates for large M_H. The decay width for this process is given by

$$\frac{\Gamma(Z^0 \rightarrow H\gamma)}{\Gamma(Z^0 \rightarrow \mu^+\mu^-)} = \frac{\alpha^2}{8\pi^2}\left(1 - \frac{M_H^2}{M_{Z^0}^2}\right)^3 \frac{|A_F + A_W|^2}{1 + (1 - 4\sin^2\theta_w)^2} \quad , \tag{2.2.5}$$

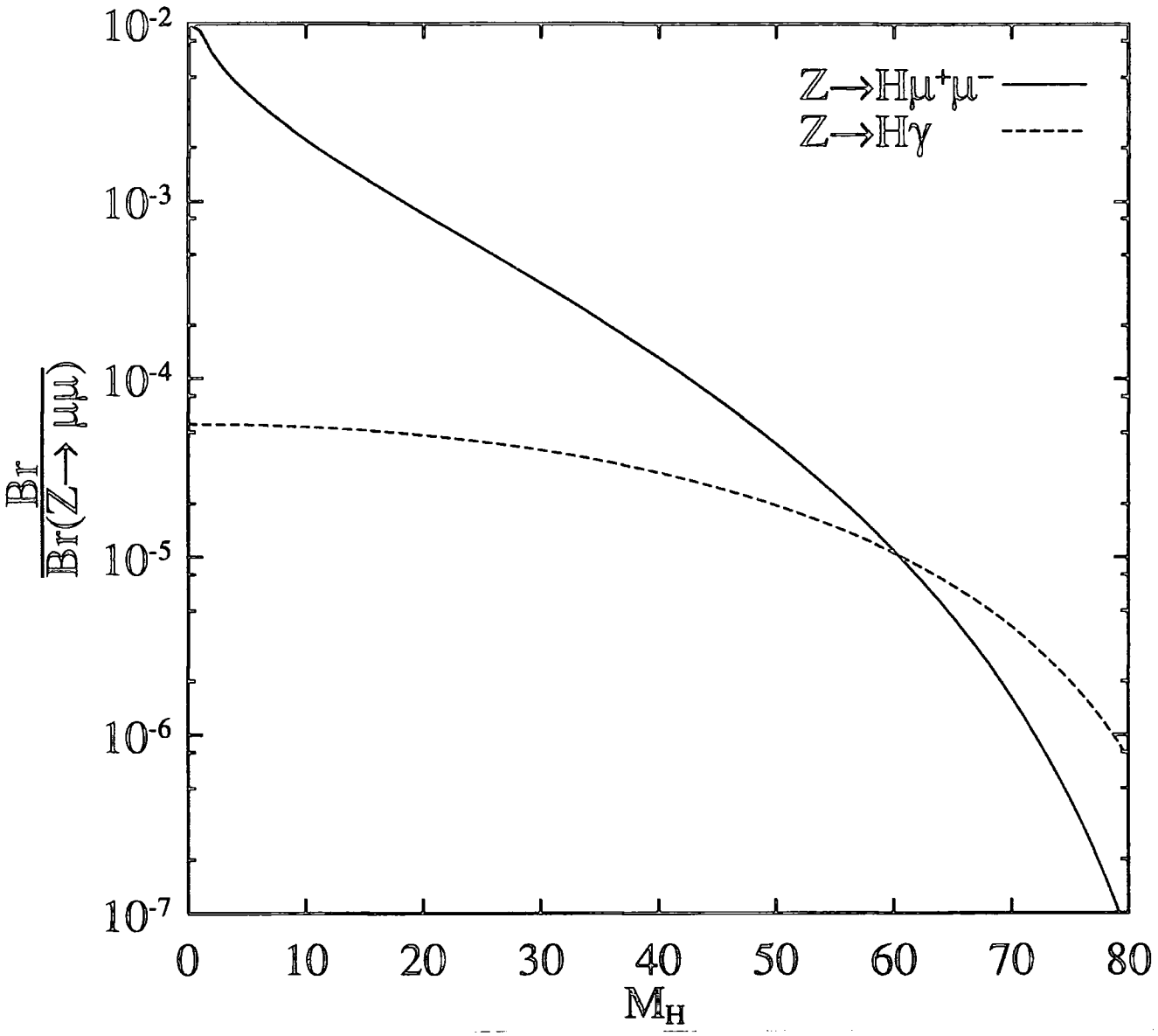


*Fig.*2.2.2 The ratios $\frac{\mathrm{Br}(Z\to H\mu^+\mu^-)}{\mathrm{Br}(Z\to\mu^+\mu^-)}$ and $\frac{\mathrm{Br}(Z\to H\gamma)}{\mathrm{Br}(Z\to\mu^+\mu^-)}$ at LEP I

where the A_F terms come from fermion loops, *Fig.*2.2.3.c, and the A_W terms come from W loops, *Fig.*2.2.3.a,b. Expressions for these A_F and A_W can be found in Appendix G [9], however all we need to know is that the W loops dominate the fermion loops and that,

$$A_W \simeq -\left[9.50 + 0.65\frac{M_H^2}{m_W^2}\right] \quad . \tag{2.2.6}$$

This width is also shown in *Fig.*2.2.2, where we have used the exact expressions for A_W and A_F rather than the approximation shown above.

Because the Higgs boson couples proportionally to mass it will typically decay to the heaviest particle available, and so a Higgs with $M_H > 11$ GeV will typically decay to a $b\bar{b}$ pair.

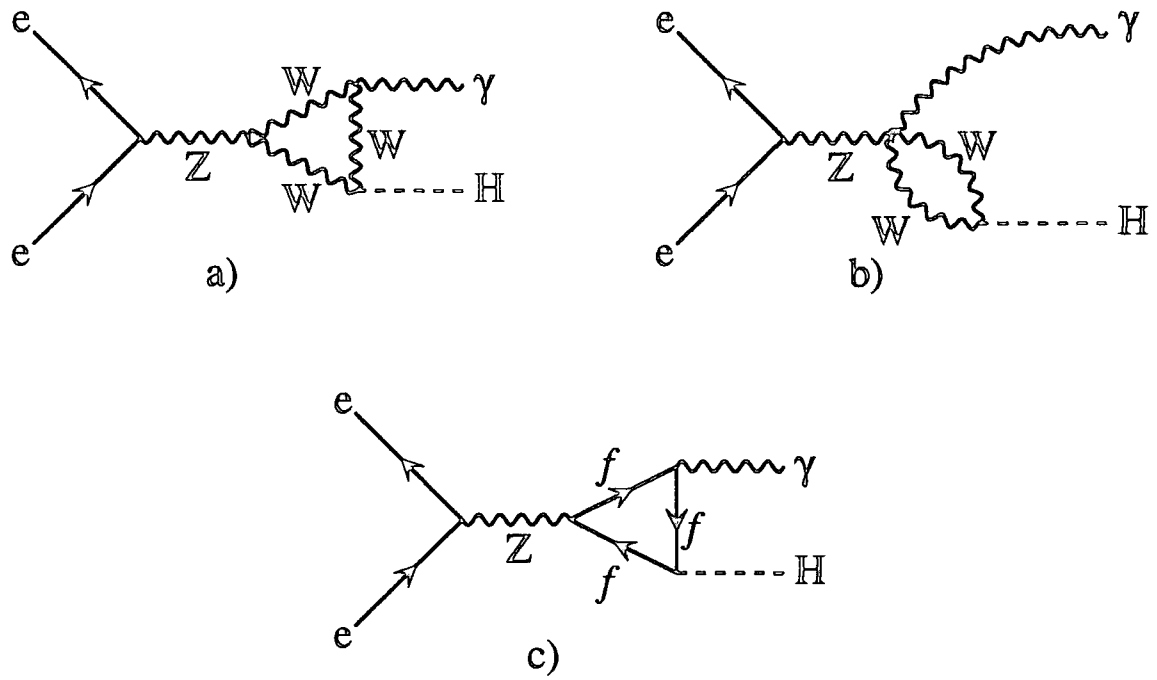


*Fig.*2.2.3 The Feynman diagrams for $Z \rightarrow H\gamma$.
The W loops (a,b) dominate the fermion loops (c).

In looking at *Fig.*2.2.2 we should recall that for the mode $Z \rightarrow Z^*H \rightarrow f\bar{f}H$ we have only included $f = \mu$, whereas in practice $f = e$ and $f = \nu$ also give usable signals; looking like *jets l l* and *jets + missing energy* respectively; this increases the branching ratio by a factor of 7. The above signals, *jets l l* and *jets + missing energy*, have small backgrounds and so the detection limit is only given by the requirement that we have enough events to give a reasonable signal. The other decays of the virtual Z^* are typically to jets, and we are swamped by a massive *4 jet* QCD background. On the other hand the $Z \rightarrow H\gamma \rightarrow$ *jets* γ signal suffers from the background,

$$Z \rightarrow q\bar{q}\gamma \rightarrow \textit{jets}\ \gamma \quad , \tag{2.2.7}$$

because the Higgs has a very narrow width the photon in the signal process is very monochromatic, which in principle separates the signal from the background (2.2.7); however this is clearly dependent upon our E_γ resolution. In practice the limited number of events in the $Z \rightarrow H\gamma$ and the lack of experimental resolution mean that we can not use this as a signal to detect the Higgs – however having discovered the Higgs it would give a useful test of the size of the HWW coupling.

At the moment each LEP I experiment has about 1,000,000 hadronic Z decays each, this leads to about 1,500,000 $Z \rightarrow e^+e^-, \mu^+\mu^-, \nu\bar{\nu}$ events in total, and so we are sensitive to Higgs with $\frac{Br(Z\rightarrow Hf\bar{f})}{Br(Z\rightarrow f\bar{f})} > 10^{-5}$ if we are to have about 10 signal events; from *Fig.*2.2.2

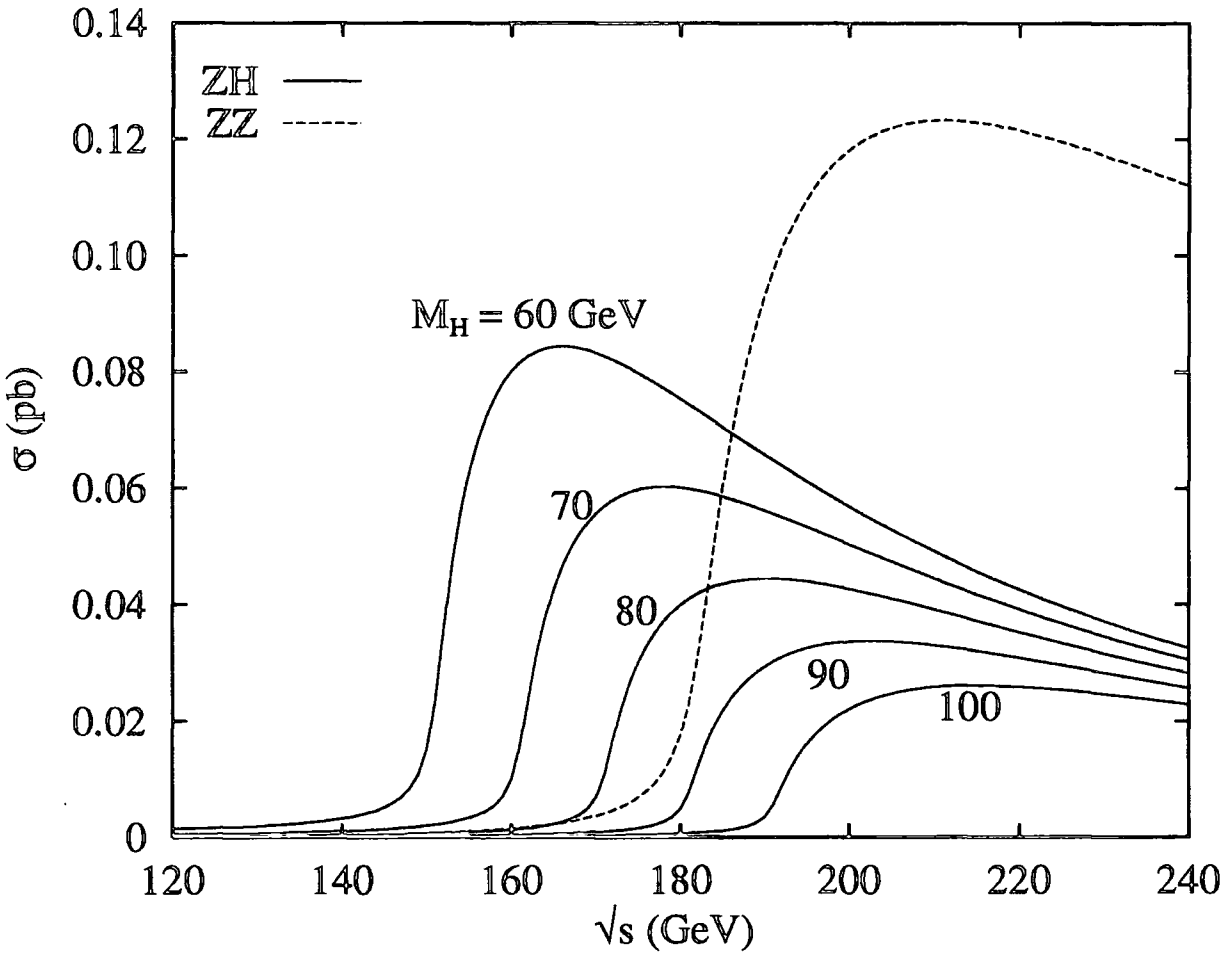


*Fig.*2.2.4 The cross-section for the process $e^+e^- \rightarrow ZH \rightarrow l^+l^-$ *jets* and $e^+e^- \rightarrow ZZ \rightarrow l^+l^-$ *jets* for $l = e, \mu$

this means that we are sensitive to Higgs with $M_H < 60\,\text{GeV}$. A more detailed analysis based upon the data of all 4 LEP I experiments gives $M_H > 62.5\,\text{GeV}$ at the 95% confidence level[6]. From *Fig.*2.2.2 it is clear the the $Z \rightarrow Z^*H$ branching ratio is falling very rapidly as a function of M_H falling by an order of magnitude between $M_H = 60\,\text{GeV}$ and $M_H = 70\,\text{GeV}$; this means that to add an extra $10\,\text{GeV}$ to the Higgs mass bound we need 10 times the integrated luminosity, and as the mass bound $M_H > 62.5\,\text{GeV}$ is based upon several years running at LEP I it is clear that the Higgs search is close to the end of the road at LEP I

2.3 Higgs Searches at LEP II

If LEP I is a Z factory then LEP II is a W^+W^- factory, and so we could think of looking for Higgs Bosons in W decays, however as the W is lighter than the Z this would give us no advantage over Z decays, and so no advantage over what we have learnt at LEP I. Another

approach is to still use the Bjorken mechanism, *Fig*.2.2.1, however rather than looking for a Z decaying to a Higgs and an off shell Z looking for an off mass shell Z decaying to an on mass shell Z and H, *i.e.*, the process $e^+e^- \to Z^* \to ZH$ rather than $e^+e^- \to Z \to Z^*H$. By just changing the Z which is off mass shell we keep the same number of on and off mass shell propagators and so the rate for this second process at LEP II is similar to the rate at LEP I, although we need far greater $\sqrt{s}$ energy, at LEP I we just require the energy to produce a Z, *i.e.*, $\sqrt{s} = M_Z$, whereas at LEP II we require the energy to produce both a Z and a H, *i.e.*, $\sqrt{s} = M_Z + M_H$, this means that to discover a 60 GeV Higgs at LEP II we will require $\sqrt{s} > 150\,\text{GeV}$. This is shown in *Fig*.2.2.4 where we plot the cross-section for the process,

$$e^+e^- \to ZH \to \begin{array}{l} jets \\ \hookrightarrow l^+l^- \end{array} \quad , \tag{2.3.1}$$

with $l = e, \mu$, for various Higgs masses as a function of $\sqrt{s}$. I have included the finite width effects of the Z and Higgs which mean that the cross-section does not drop abruptly to zero as we decrease the $\sqrt{s}$ beneath the the threshold for producing the Z and Higgs, but extends beyond it. As an approximate rule we can see that to produce a reasonable number of events we require,

$$\sqrt{s} \gtrsim M_H + 100\,\text{GeV} \quad , \tag{2.3.2}$$

that is the rest mass of the Higgs plus the rest mass of the Z plus a small amount of extra energy to give us some phase space (at threshold the volume of phase space vanishes and we have zero rate). Now this process has a relatively clean signal in the,

$$e^+e^- \to ZH \to \begin{array}{l} jets \\ \hookrightarrow l^+l^- \end{array} \quad , \tag{2.3.3}$$

$$e^+e^- \to ZH \to \begin{array}{l} jets \\ \hookrightarrow \nu\bar{\nu} \end{array} \quad , \tag{2.3.4}$$

decay modes of the Z, the former decay (2.3.3) giving a better signal than the latter decay (2.3.4) due to less severe backgrounds and a better M_H resolution. These processes have been put through full event simulations in Ref. 10 and as long as we have enough $\sqrt{s}$ energy the conclusion is that with a reasonable integrated luminosity, $\mathcal{L} = 500\,\text{pb}^{-1}$, that a Higgs boson with mass up to $M_H = 80\,\text{GeV}$ can be detected, however for masses beyond $M_H = 80\,\text{GeV}$ we start getting a large background from the processes, *Fig*.2.3.1,

$$e^+e^- \to ZZ \to \begin{array}{l} jets \\ \hookrightarrow l^+l^- \end{array} \quad , \tag{2.3.5}$$

$$e^+e^- \to ZZ \to \begin{array}{l} jets \\ \hookrightarrow \nu\bar{\nu} \end{array} \quad , \tag{2.3.6}$$

This can be seen in *Fig*.2.2.4, where as well as the ZH cross-section we have also shown the ZZ cross-section for the process (2.3.5), clearly as we pass the threshold for ZZ production

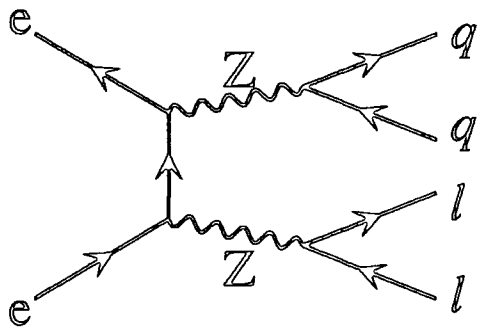


*Fig.*2.3.1 The Feynman diagram for the background process $e^+e^- \rightarrow ZZ \rightarrow l^+l^-$ *jet jet.*

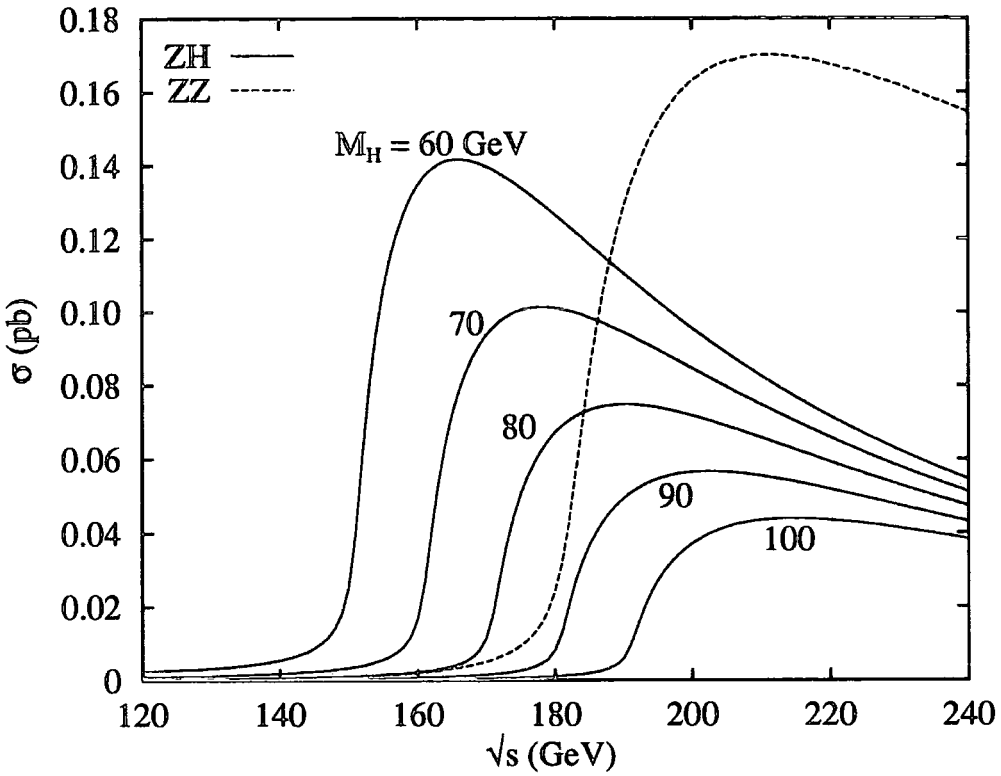
its cross-section rapidly becomes far larger than the ZH cross-section. Now this is not a problem if $M_H \neq M_Z$ because we can reconstruct the invariant mass of the Higgs peak and separate it from the Z peak, for example by reconstructing the *jet jet* invariant mass, or looking at the mass that recoils from the l^+l^- pair. If $M_H \approx M_Z$ then in principle this is still not a problem, because the Higgs has a much narrower width, $\mathcal{O}(\,\mathrm{MeV})$, than the Z width, $\mathcal{O}(\,\mathrm{GeV})$; and so the Higgs peak should show as a very sharp peak on top of a broader Z peak. However life is not so simple, in practice we can not reconstruct the Higgs/Z invariant mass to an arbitrary precision; but have limits imposed by our experimental resolution. In practice the experimental resolution on invariant masses is only likely to be about 5 GeV[10], and with this resolution the Higgs peak becomes obscured beneath the Z peak for Higgs masses larger than $M_H > 80\,\mathrm{GeV}$.

2.4 Enhancing the Signal at LEP II

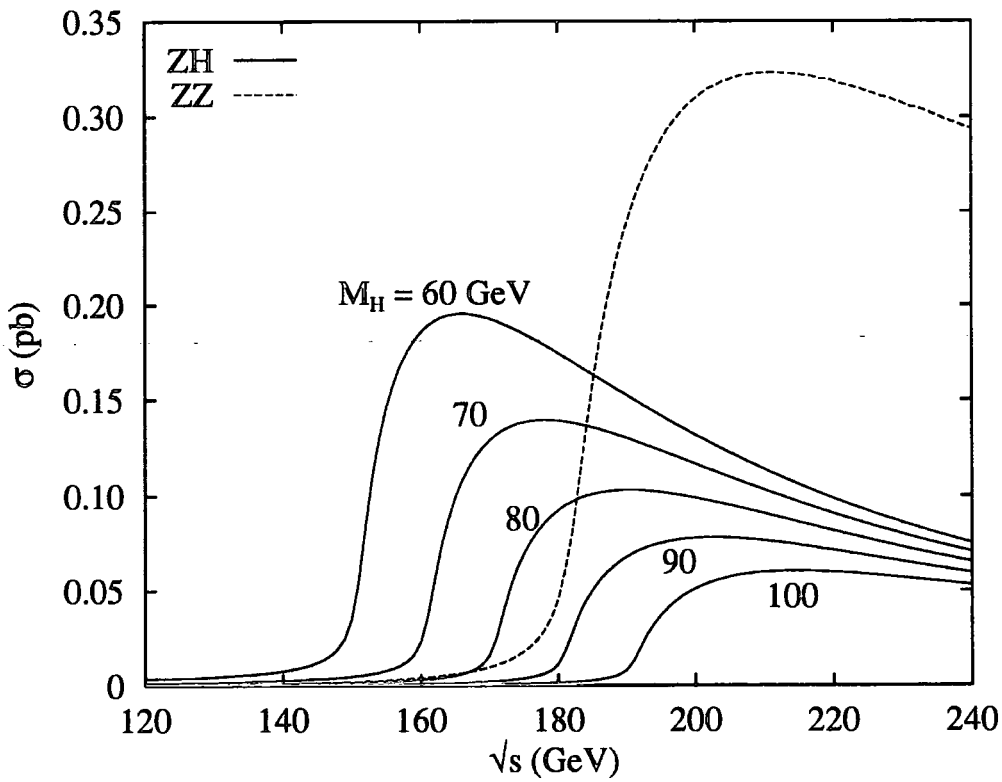
If the ZH signal is obscured by the ZZ background is there any way in which we can enhance the signal or decrease the background?

2.4.1 Polarized e^+e^- beams

The ZH signal, *Fig.*2.2.1, and the ZZ background, *Fig.*2.3.1, have different numbers of Z's coupling to the e line, now as a Z coupling isn't parity conserving we can change the relative importance of these two processes by polarizing the e^+e^- beams. In particular a Z coupling to a $+ve$ chirality line is proportional to $g_V - g_A$, and coupling to a $-ve$ chirality line is proportional to $g_V + g_A$; and so for electrons the Z couples more strongly to $-ve$ chirality electrons. This can be seen in *Fig.*2.4.1 where we have plotted the cross-section for



a) For $+ve$ chirality e^+e^- beams



b) For $-ve$ chirality e^+e^- beams

*Fig.*2.4.1 The cross-section for the process $e^+e^- \rightarrow ZH \rightarrow l^+l^-$ *jet jet* and $e^+e^- \rightarrow ZZ \rightarrow l^+l^-$ *jet jet* for $l = e, \mu$ with polarized e^+e^- beams.

both $-ve$ chirality e^+e^- beams and $+ve$ chirality e^+e^- beams ‡ †. The first thing to note is that the cross-sections are about a factor of 2 larger than the unpolarized cross-section, this is nothing other than the fact that for unpolarized beams half the electrons and positrons have zero cross-section because they have opposite chirality; or in the language of Feynman diagrams for polarized beams we no longer need to average over the initial spins of the electrons. When we have taken this factor of 2 into account we can see that for $-ve$ chirality beams we have enhanced the ZH signal but the ZZ background is enhanced by a larger amount because of the two Z's that couple to the electron line; and in the $+ve$ chirality case we reduce both cross-sections, suppressing the ZZ background more than the ZH signal. Unfortunately this isn't exactly what we want, with $-ve$ chirality beams we have a larger signal, but a worse signal to background ratio, and for $+ve$ chirality signals the opposite; unfortunately these two effects almost totally cancel each other out. If we look at the case with $\sqrt{s} = 200\,\text{GeV}$ and $M_H = M_Z$ then we find the results shown in *Table* 2.4.1,

	ZZ (pb)	ZH (pb)	S/B
Unpolarized	0.1181	0.03062	0.26
+ve Chirality	0.1630	0.05151	0.32
−ve Chirality	0.3095	0.07098	0.23

Table 2.4.1 The cross-sections for the processes $e^+e^- \to ZH, ZZ \to l^+l^-$ *jet jet* with $l = e, \mu$ and $M_H = M_Z$ at $\sqrt{s} = 200\,\text{GeV}$, and also the signal to background ratio (S/B).

If we observe N events this has a statistical 1sd error of $\sqrt{N}$ events, and so to separate the signal from the background by m standard deviations we require,

$$\mathcal{L}\sigma_S = \mathcal{L}(\sigma_{S+B} - \sigma_B) = m\sqrt{\mathcal{L}\sigma}$$

$$\Rightarrow \qquad \mathcal{L} = \frac{m^2\sigma}{\sigma_S^2} \tag{2.4.1}$$

‡ Throughout this thesis I use the word *chirality* to mean the spin of a fermion in the direction of the flow of that fermion. This is distinct from the *helicity* of the fermion, that is the spin of the fermion in the direction of the momentum of the particle. For fermions chirality and helicity have the same meaning; however for antifermions whose fermion line travels backwards in time chirality and helicity have opposite meanings, that is a positive chirality positron has negative helicity. This definition of chirality is useful because it is defined for internal off shell fermions. In *Fig*.2.3.1 the t channel exchanged electron either travels forward or backward in time depending upon the momenta of the other particles; as such its helicity flips when its direction in time flips. Whereas its chirality is independent of whether it is propagating forward or backward in time, and in particular is identical (in this case) to the chirality of the initial e^+ or e^-.

† Note that the cross-section for opposite chirality e^+e^- beams vanishes because an electrons coupling to a Z conserves chirality because the Z is a vector boson.

where $\sigma = \sigma_B$ if there is no Higgs and we want to know the luminosity we require to say there is no Higgs; and $\sigma = \sigma_{S+B}$ if there is a Higgs and we want to know the luminosity required to see it. For $m = 1$ and $\sigma = \sigma_{S+B}$ this corresponds to the integrated luminosities shown in *Table* 2.4.2. If we wish to separate the Higgs signal from the background by m standard deviations we require m^2 as many events.

	Luminosity (pb^{-1})
Unpolarized	160
+ve Chirality	81
−ve Chirality	76

Table 2.4.2 Integrated luminosity required to separate the $e^+e^- \to ZH \to l^+l^- + jets$ signal from the $e^+e^- \to ZZ \to l^+l^- + jets$ background by 1 standard deviation for $M_H = M_Z$ with $\sqrt{s} = 200\ GeV$.

Quite clearly the only gain from using polarized beams is the factor of 2 that we get from not averaging over the initial spins. If we can only polarize one of the two beams then because only like chirality electron positron pairs scatter we have the same effect as polarizing both the beams, however of course we then need to average over the unpolarized beams polarization and this introduces a factor of 2 and so is comparable to the unpolarized beams.

2.4.2 b tagging.

Because the Higgs boson couples to other particles proportional to their mass it decays predominantly to heavy particles, and so for a Higgs with $M_H = M_Z$ we find the dominant branching ratios shown in *Table* 2.4.3[11],

Particle	Br
b	85%
c	4%
τ	8%

Table 2.4.3 The dominant branching ratios of a Standard Model Higgs with $M_H = M_Z$

So the decay $H \to b\bar{b}$ dominates Higgs decays. Now the Z branching ratio to a $b\bar{b}$ pair is only $\mathrm{Br}(Z \to b\bar{b}) = 15\%$; or more importantly only 21% of Z hadronic decays are $Z \to b\bar{b}$ whereas 92% of Higgs hadronic decays are $H \to b\bar{b}$. So if we can b tag the jets from the Z or Higgs decay we can vastly increase the Higgs signal in relation to the Z background.

	ZZ (pb)	ZH (pb)	S/B	$\mathcal{L}$ (pb^{-1})
Without b tagging	0.118	0.031	0.26	160
With b tagging	0.026	0.028	1.08	69

Table 2.4.4 A comparison of Higgs detection both with and without b tagging in the detection mode $e^+e^- \rightarrow ZH, ZZ \rightarrow l^+l^-$ *jet jet* with $l = e, \mu$, for $M_H = M_Z$ and $\sqrt{s} = 200\,\text{GeV}$. The integrated luminosity,$\mathcal{L}$, is the luminosity required to separate the signal from the background by 1 standard deviation.

If we could achieve 100% b tagging efficiency, with no misidentification of other quark jets as being b quark jets then we find the results show in *Table* 2.4.4

Clearly there is a great gain to be had from efficiently tagging the jets as b jets, however we must keep the tagging efficiency high as we are likely to have very few events available, and wish to waste as few as possible. All 4 experiments at LEP have recently installed vertex detectors, which are capable of tagging b jets with high efficiency; and will help tremendously in Higgs searches.

Also tagging on b's means we can be more adventurous in the Higgs signals we look for, we can start to consider process like,

$$e^+e^- \rightarrow ZH \rightarrow b\bar{b} \quad \hookrightarrow q\bar{q} \tag{2.4.2}$$

If we don't tag on b jets then this process suffers from a tremendous background from,

$$e^+e^- \rightarrow WW \rightarrow q\bar{q}' \quad \hookrightarrow q\bar{q}' \tag{2.4.3}$$

However if we tag on b jets this background goes away as W's very rarely decay to b quarks, because the decay $W \rightarrow tb$ is kinematically forbidden and the CKM matrix entries for V_{bc} and V_{bu} are very small. We need to take great care in a process like this that we don't misidentify jets as b jets, because only a slight contamination from the WW background swamps the signal because the original WW is so much larger than the ZH signal. Also we need to take care that we know accurately how efficiently we are tagging the b jets, as we are detecting the Higgs only on the total event rate; and any lack of knowledge in what we expect the event rate to be (both with and without the Higgs signal) reflects strongly on our ability to detect the Higgs. These problems can be largely overcome by tagging 2 (or more) jets as b quark jets. For more details see the work of Nick Brown in Ref. 12.

2.4.3 Tau decays of Higgs

Another option for detecting the ZH signal is to look for Higgs decays other than the $H \rightarrow$ *jet jet*. This initially sounds like a daft idea, as the only sizable non jet decay of the

Higgs is the decay $H \to \tau^+\tau^-$, and this only has a branching ratio of about 8%, this means that to have any events at all we must go for picking up most of the Z decays, this can only mean looking for the hadronic decays of Z's which have a branching ratio of 70%. We might also worry about how we detect the taus, as the majority of taus decay hadronically into a jet, however these jets are very modest jets in comparison to QCD jets, have very few tracks and a low invariant mass (equals m_τ); and so can easily be distinguished from QCD jets. Now if we look for this process we get signals from two sources,

$$e^+e^- \to ZH \to \tau^+\tau^- \quad , \qquad (2.4.4)$$
$$\hookrightarrow q\bar{q}$$

$$e^+e^- \to ZH \to q\bar{q} \qquad (2.4.5)$$
$$\hookrightarrow \tau^+\tau^-$$

Whereas we only get a sizable number of background events from the process,

$$e^+e^- \to ZZ \to \tau^+\tau^- \quad . \qquad (2.4.6)$$
$$\hookrightarrow q\bar{q}$$

Note that we don't get any background from WW decays as we would if we looked for both Z and Higgs decaying hadronically.

For this process we find the cross-sections shown in *Table* 2.4.5

Process	σ (pb^{-1})
$e^+e^- \to ZH \to \tau^+\tau^-$ $\hookrightarrow$ *jet jet*	0.0288
$e^+e^- \to ZH \to$ *jet jet* $\hookrightarrow \tau^+\tau^-$	0.0153
$e^+e^- \to ZZ \to \tau^+\tau^-$ $\hookrightarrow$ *jet jet*	0.0591

Table 2.4.5 The cross-sections for the process $e^+e^- \to \tau^+\tau^-$ *jet jet* for $M_H = M_Z$ and $\sqrt{s} = 200\,\text{GeV}$.

Now if $M_H = M_Z$ then we do not distinguish process (2.4.4) from process (2.4.5), this gives as a signal to background ratio of 0.74, and so to separate the ZH signal from the ZZ background by 1sd requires integrated luminosity $\mathcal{L} = 53\,\text{pb}^{-1}$.

For more details see the work of P.Janot in Ref. 13.

2.5 Topological searches at LEP II

All the methods suggested so far to detect a Higgs with $M_H = M_Z$ at LEP II rely upon accurate theoretical predictions for how many of a certain type of event that we expect to see; there is no bump in an invariant mass distribution, other than the expected bump at M_Z, and we are very dependent upon our theoretical model of how the universe works. In this section a different approach will be used; which, at least in principle, is far more model independent.

If we consider the process $e^+e^- \rightarrow q\bar{q}l^+l^-$ then counting the number of degrees of freedom that the final state has we find,

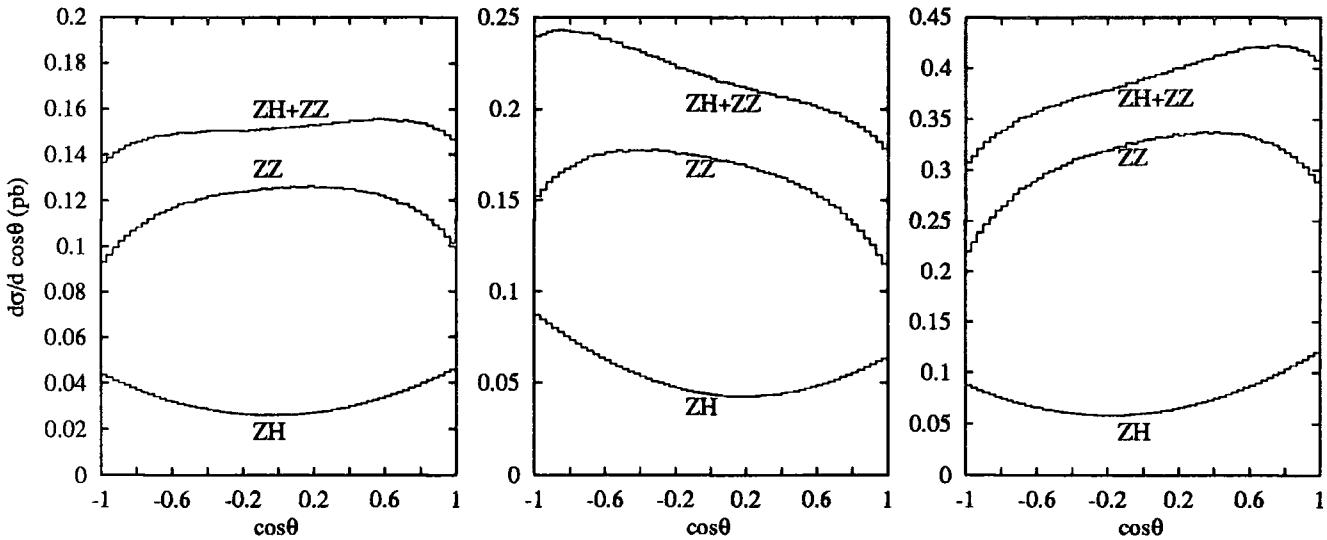
$$\begin{aligned} 4 \times 4 \text{ mom} = 16 \text{ d.o.f} &\xrightarrow{\substack{4 \text{ on mass shell} \\ \text{constraints}}} 12 \text{ d.o.f} \\ &\xrightarrow{\substack{4 \text{ mom} \\ \text{conservation}}} 8 \text{ d.o.f} \\ &\xrightarrow{\substack{\text{An arbitrary} \\ \text{azimuthal angle}}} 7 \text{ d.o.f} \end{aligned} \qquad (2.5.1)$$

Now each of these 7 degrees of freedom in principle contains information that describes the process, two of these degrees of freedom we can take to be the invariant masses of the $q\bar{q}$ pair and the l^+l^- pair, and so if we are trying to distinguish a ZH signal with $M_H = M_Z$ from a ZZ background these can not give us any information. The 5 remaining degrees of freedom we can take to be the 4 forward-backward $\cos\theta$ angles with respect to the beam direction of each of the 4 particles, and the angle between the plane defined by $q\bar{q}$ pair and the plane defined by the l^+l^- pair.

Now the $\cos\theta$ distributions for the particles are identical to the $-\cos\theta$ distributions for the antiparticles, by the CP invariance of the ZH signal and ZZ background processes; and so we only have two independent sets of $\cos\theta$ distributions, one for the quarks and one for the leptons, however we do get two entries in these distributions for each event we detect, one for the particle and one for the antiparticle.

So if we plot the particles $\cos\theta$ and the antiparticles $-\cos\theta$ then we find the $\cos\theta$ distributions shown in *Fig.*2.5.1, clearly there is a difference in the angular distributions for the ZH signal and ZZ background. The angular distributions for the ZH signal are fairly easy to understand, the lepton distribution is just the standard $1 + \cos^2\theta$ type distribution for $e^+e^- \rightarrow Z \rightarrow l^+l^-$, with a bit of smearing as the decaying Z is not at rest. The quark distribution is approximately isotropic because the Higgs that the quarks originate from is spin 0; the slight drop at $\cos\theta = \pm 1$ again arises because the Higgs is not produced at rest, and so the $q\bar{q}$ system has a slight Lorentz boost. The ZZ angular distributions differ from the ZH angular distributions because the Z's can now be produced with longitudinal polarization.

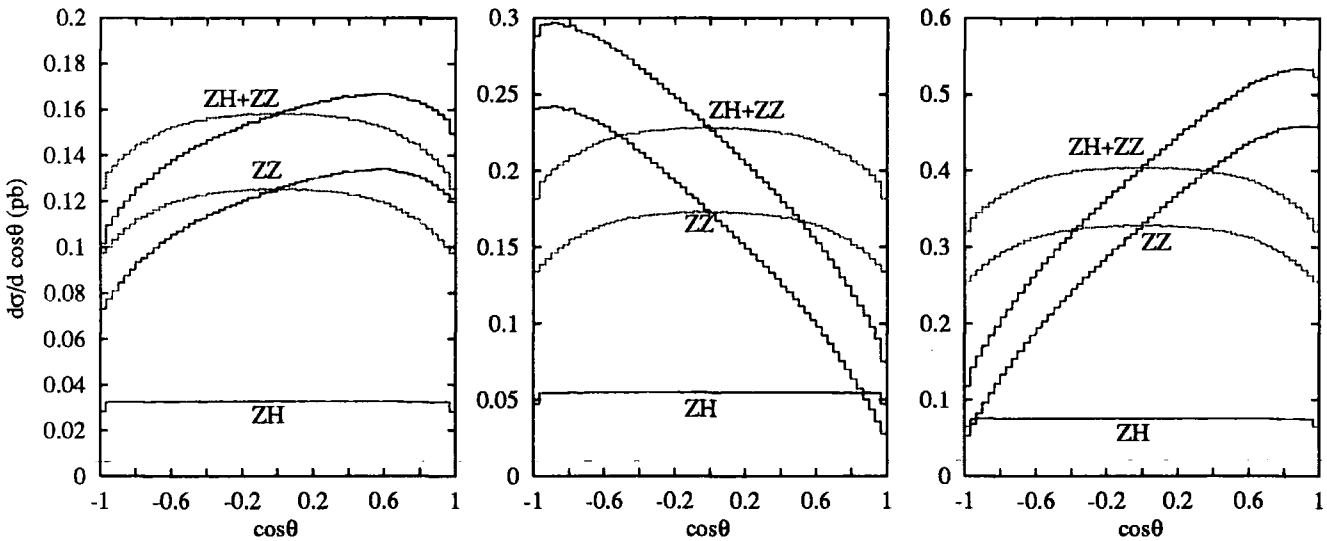
The largest difference between the ZZ and ZH angular distributions arises for the quarks, *i.e.*, the directions of the jets; as the ZZ background has a large forward backward asymmetry.



a) l^- distribution for unpolarized beams

b) l^- distribution for $+ve$ chirality beams

c) l^- distribution for $-ve$ chirality beams



d) q distribution for unpolarized beams

e) q distribution for $+ve$ chirality beams

f) q distribution for $-ve$ chirality beams

*Fig.*2.5.1 The $\cos\theta$ distributions for the process $e^+e^- \to ZH, ZZ \to l^+l^-q\bar{q}$, for the q $(= -\bar{q})$ distributions I have also plotted the symmetrised distribution, as in experiment we can not distinguish quarks from antiquarks, and so we only have access to the symmetrised curves.

Unfortunately we can't distinguish quark jets from antiquark jets, and the distribution for antiquarks being identical to the distribution for quarks under the transformation $\cos\theta \leftrightarrow -\cos\theta$ means that we can't measure any forward backward asymmetry for the jets; but can only measure the symmetrised distribution. This means that the greatest sensitivity to the differences in the ZH signal and ZZ background comes from the lepton distributions. From now on I will just concentrate on the lepton angular distributions.

The question now arises as to how to extract these angular distributions from the data; an optimal method is given in Appendix F based upon orthogonal polynomials and Monte Carlo integration. A convenient choice of polynomials to parametrise the angular distributions is,

$$\frac{1}{\sigma}\frac{d\sigma}{d\cos\theta} = \sum_n a_n P_n(\cos\theta) \quad , \tag{2.5.2}$$

where the $P_n(\cos\theta)$ are Legendre polynomials. Using this choice of polynomials, and the methods of Appendix F we extract the values for the a_n shown in *Table* 2.5.1. The ZH case only has the a_n different from zero by a statistically significant amount for $n = 0, 1, 2$ whereas the ZZ case is far more complicated having significant values for $n = 0, 1, 2, 4, 6$. It may seem that I have unnecessarily complicated the ZZ distributions by looking at the $\cos\theta$ distributions in the lab frame, where we have contributions from $\cos^6\theta$ terms. Whereas if we were to define θ_Z as the angle relative to the Z direction we only get contributions from $\cos^n\theta_Z$ for $n = 0, 1, 2$ terms (see Appendix H for more details). Although this is true if we have no experimental cuts the situation becomes far more complicated with the addition of experimental cuts. In particular in any experiment we have a cut that keeps us separated from the beam direction; if we define θ as relative to the beam direction then it is clear that this cut just removes an area about $\cos\theta = \pm 1$; however if we define θ_Z relative to the produced Z direction then this beam direction cut removes events from *all* $\cos\theta$ directions, and in particular this will drastically change the form of the $\cos\theta_Z$ distributions and complicate the distribution from its $a + b\cos\theta_Z + c\cos^2\theta_Z$ form. Also as the cut from the beam direction is altered the $\cos\theta$ distributions just have events removed or added about $\cos\theta = \pm 1$, whereas the entire $\cos\theta_Z$ distributions are altered. Also having $\cos^n\theta$ terms for $n = 0, 1, 2, 4, 6$ should not worry us; because, as I will show, we can redefine our orthogonal polynomials so only two different distributions describe both the ZZ and ZH distributions.

Now we can also use the methods of Appendix F experimentally to extract the values of the a_n, however if we wish purely to differentiate the ZH signal from the ZZ background, by performing rotations in space generated by the $P_n(\cos\theta)$ polynomials we can arrive at a single polynomial that is optimally efficient in differentiating the ZH signal from the ZZ background.

n	ZZ	ZH	ZZ+ZH
0	1.0000	1.0000	1.0000
1	(347±5)×10^{-4}	(347±6)×10^{-4}	(347±6)×10^{-4}
2	(−1512±6)×10^{-4}	(3976±6)×10^{-4}	(−383±7)×10^{-4}
3	(−5±6)×10^{-4}	(4±6)×10^{-4}	(−3±7)×10^{-4}
4	(−380±6)×10^{-4}	(3±6)×10^{-4}	(−301±7)×10^{-4}
5	(3±7)×10^{-4}	(−5±7)×10^{-4}	(1±8)×10^{-4}
6	(−33±7)×10^{-4}	(−6±8)×10^{-4}	(−27±9)×10^{-4}
7	(−13±9)×10^{-4}	(4±9)×10^{-4}	(−10±10)×10^{-4}

For unpolarized electrons

n	ZZ	ZH	ZZ+ZH
0	1.0000	1.0000	1.0000
1	(−1152±5)×10^{-4}	(−2227±6)×10^{-4}	(−1410±6)×10^{-4}
2	(−1515±6)×10^{-4}	(3976±6)×10^{-4}	(−197±6)×10^{-4}
3	(−8±6)×10^{-4}	(4±6)×10^{-4}	(−5±7)×10^{-4}
4	(−382±6)×10^{-4}	(3±6)×10^{-4}	(−290±7)×10^{-4}
5	(5±7)×10^{-4}	(−6±7)×10^{-4}	(2±7)×10^{-4}
6	(−36±7)×10^{-4}	(−5±8)×10^{-4}	(−28±8)×10^{-4}
7	(−11±9)×10^{-4}	(7±9)×10^{-4}	(−6±10)×10^{-4}

For $+ve$ chirality electrons

n	ZZ	ZH	ZZ+ZH
0	1.0000	1.0000	1.0000
1	(1136±5)×10^{-4}	(2216±6)×10^{-4}	(1337±6)×10^{-4}
2	(−1511±6)×10^{-4}	(3976±6)×10^{-4}	(−488±7)×10^{-4}
3	(−4±6)×10^{-4}	(5±6)×10^{-4}	(−2±7)×10^{-4}
4	(−379±6)×10^{-4}	(3±6)×10^{-4}	(−308±7)×10^{-4}
5	(2±7)×10^{-4}	(−5±7)×10^{-4}	(1±8)×10^{-4}
6	(−31±7)×10^{-4}	(−6±8)×10^{-4}	(−26±9)×10^{-4}
7	(−14±9)×10^{-4}	(3±9)×10^{-4}	(−11±10)×10^{-4}

For $-ve$ chirality electrons

Table 2.5.1 The values for the a_n for the processes $e^+e^- \to ZH, ZZ \to l^+l^-q\bar{q}$ for $M_H = M_Z$ and $\sqrt{s} = 200\,\text{GeV}$. The errors arise from the Monte Carlo errors in estimating equation (2.5.2).

We wish to differentiate the ZZ angular distribution,

$$P_{ZZ} = \sum a_i^{ZZ} P_i(\cos\theta) \quad , \tag{2.5.3}$$

from the $ZH + ZZ$ angular distribution,

$$P_{ZH+ZZ} = \sum (a_i^{ZH} + a_i^{ZZ}) P_i(\cos\theta) \quad . \tag{2.5.4}$$

Now if we write,

$$P_{ZZ} = Q_2 + \lambda_{ZZ} Q_1 \qquad P_{ZH+ZZ} = Q_2 + \lambda_{ZH+ZZ} Q_1 \quad , \tag{2.5.5}$$

with Q_1 and Q_2 orthogonal; then,

$$P_{ZH} = \sum a_i^{ZH} P_i(\cos\theta) = (\lambda_{ZH+ZZ} - \lambda_{ZZ}) Q_1 \quad , \tag{2.5.6}$$

and so if we define,

$$Q_1 = \sum a_i^{ZH} P_i(\cos\theta) \quad , \tag{2.5.7}$$

and then find Q_2 by Gram-Schmidt orthogonalisation with P_{ZZ}. If we now use Q_1 and Q_2 in place of our orthogonal polynomials P_1 and P_2 (and, of course, redefine the higher P_i by Gram-Schmidt orthogonalisation so all the polynomials are orthogonal) then we have a set of polynomials that is optimal in distinguishing the ZZ distribution from the $ZH + ZZ$. We expect the coefficient of Q_1 to take one of two values (separated by 1 from (2.5.6,2.5.7)), one of which corresponds to just the ZZ background, and the other the ZH signal plus the ZZ background. The coefficient of Q_2 has some finite (known) value, and the coefficients of all the remaining polynomials we expect to be zero.

At this stage we can theoretically calculate how we expect the experimental error on the measurement of the coefficient of Q_1 to decrease as we observe more events (see Appendix F for more details) and so estimate how may events we require to separate the ZZ background from the $ZH + ZZ$ signal-plus background. If we do this we find the number of events required to separate the signal from the background by 1 sd shown in *Table* 2.5.2, to separate the signal from the background by m sd we require m^2 as many events. These events correspond to the integrated luminosity shown in *Table* 2.5.2.

	No. of events	Luminosity (fb^{-1})
Unpolarized	250	2.0
+ve Chirality	200	1.1
−ve Chirality	320	1.0

Table 2.5.2 The number of events, and the equivalent integrated luminosity, required the separate the $ZH + ZZ$ signal plus background from the ZZ background by 1sd; with $M_H = M_Z$ and $\sqrt{s} = 200\,\text{GeV}$

2.5.1 Azimuthal correlations

The only remaining possible topological source of information that we can use to differentiate between the ZH signal and the ZZ background is the azimuthal angle between the plane defined by the l^+l^- pair and the plane defined by the $q\bar{q}$ pair. If we could differentiate quark and antiquark jets then this angle would take values between $0° - 180°$ as we could define a "top" and a "bottom" to each plane, however we can't experimentally differentiate quark from gluon jets, and so we can only measure the angle in the range $0° - 90°$, the range $90° - 180°$ being mapped onto $90° - 0°$. We show the distributions for the ZH signal and ZZ background in *Fig.*2.5.2. In this figure the y axis has been expanded many times, if the differential cross-section were plotted from zero then both the ZH signal and ZZ background look flat, indeed the ZH azimuthal angle correlation is entirely flat, because the Higgs is spin zero and so can communicate no information about this angle between the l^+l^- and the $q\bar{q}$ pair. The ZZ correlation, on the other hand, is not entirely flat as the magnified graph shows; there is a very slight tendency for the l^+l^- plane to be flat relative to the $q\bar{q}$ plane. In practice this slight correlation is not enough to experimentally distinguish the ZH signal from the ZZ background; this is quite apparent from *Fig.*2.5.2, to plot this figure we have used the equivalent of many millions of events, and still the ZZ background distribution shows many "wiggles". At most we will only have a few hundred events experimentally and this effect is totally unobservable.

2.6 Conclusions

At LEP I the only usable production mode of the Higgs is,

$$e^+e^- \rightarrow Z \rightarrow Z^*H \rightarrow f\bar{f}H \quad . \tag{2.6.1}$$

Using this mode the current mass limit on the Higgs is[6],

$$M_H > 62.5\,\text{GeV} \qquad \text{at the 95\% confidence level.} \tag{2.6.2}$$

However the cross-section for producing the Higgs via this mechanism (2.6.1) is dropping very rapidly as a function of M_H by $M_H = 60\,\text{GeV}$ and so as a result the Higgs mass limit will not be improved a great deal at LEP I.

Moving onto LEP II we still producing reasonable numbers of Higgs bosons by the similar mechanism,

$$e^+e^- \rightarrow Z^* \rightarrow ZH \quad . \tag{2.6.3}$$

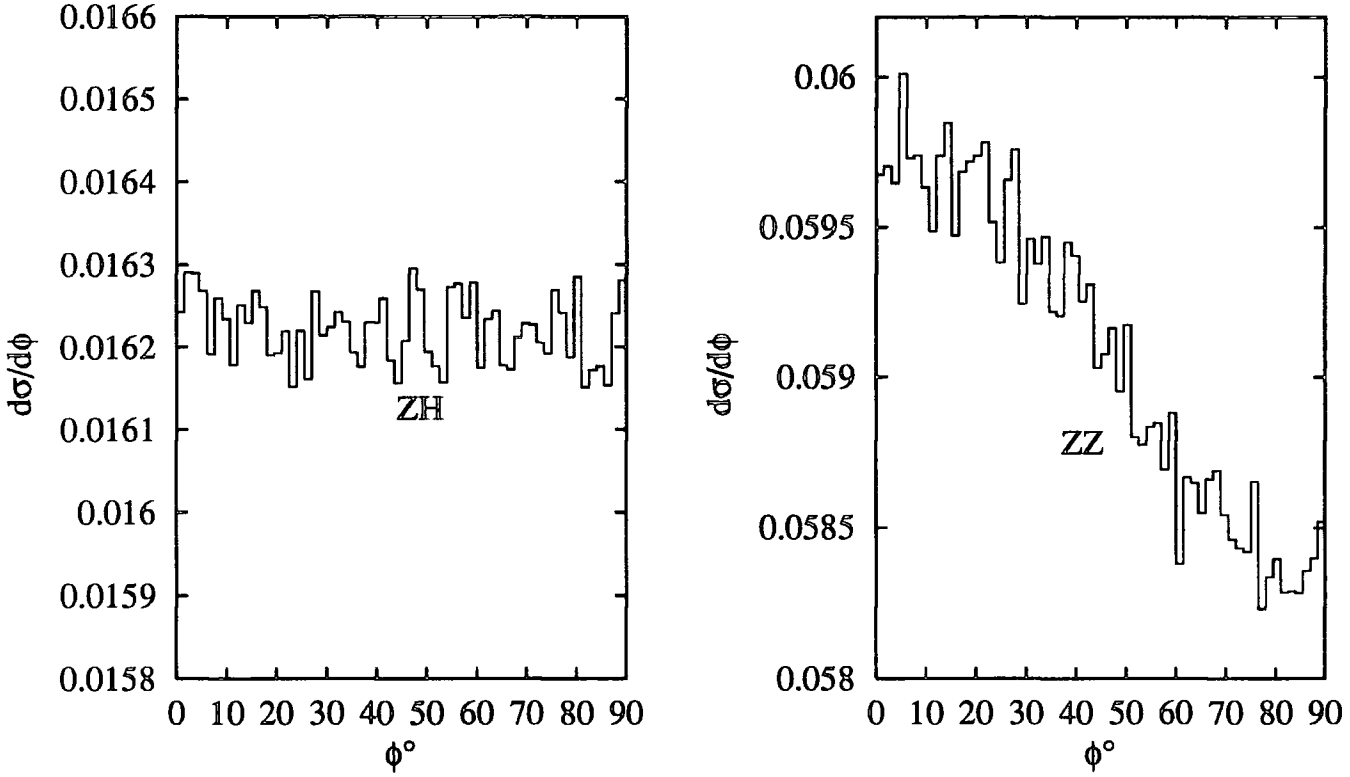


*Fig.*2.5.2 The azimuthal angle,ϕ , between the plane defined by the l^+l^- pair and the plane defined by the $q\bar{q}$ pair for the processes $e^+e^- \to ZH, ZZ \to l^+l^-q\bar{q}$ with unpolarized electrons, $M_H = M_Z$ and $\sqrt{s} = 200\,\text{GeV}$. Note that the y scale has been expanded to show the effect, the *wiggles* are the errors in the Monte Carlo integral.

This produces reasonable numbers of events for Higgs masses up to the limits in phase space, that is for,

$$M_H \lesssim \sqrt{s} - 100\,\text{GeV} \quad , \tag{2.6.4}$$

however for Higgs masses larger than $M_H > 80\,\text{GeV}$ we encounter a very large background from,

$$e^+e^- \to ZZ \quad . \tag{2.6.5}$$

There are many methods of overcoming the background, most of which involve increasing the ratio of the ZH signal events to the ZZ background events, for example b tagging, and the tau decay of the Higgs. These methods which enhance the signal relative to the background can be viable on modest integrated luminosities, say 500pb^{-1} for a 3 sd effect; however they require very exact theoretical calculations about what we *expect* to observe.

On the other hand topological searches, which don't try to enhance the signal relative to the background, but instead concentrate on the actual topology of the events observed tend to require far more events to be able to separate the ZH signal from the ZZ background,

this means that they will probably not be a useful method of discovering the Higgs boson. However these methods are relatively robust, and make far fewer demands upon theoretical calculations about what we *expect* to see at future e^+e^- colliders. They also are very valuable in telling us about the spins of particles produced.

In recent months it has become clear that the *maximum* energy of LEP II is unlikely to be larger than $\sqrt{s} > 180\,\mathrm{GeV}$, this restricts Higgs detection to the mass range,

$$M_H \lesssim 80\,\mathrm{GeV} \quad . \tag{2.6.6}$$

Chapter 3

Higgs Physics at Hadron Colliders

LINES WRITTEN BY A BEAR OF VERY LITTLE BRAIN

On Monday, when the sun is hot
I wonder to myself a lot:
'Now is it true, or is it not,
'That what is which and which is what ?'

On Tuesday, when it hails and snows,
The feeling on me grows and grows
That hardly anybody knows
If those are these or these are those.

On Wednesday, when the sky is blue.
And I have nothing else to do,
I sometimes wonder if it's true
That who is what and what is who.

On Thursday, when it starts to freeze
And hoar-frost twinkles on the trees,
How very readily one sees
That these are whose – but whose are these?

On Friday –

Winnie–the–Pooh, A.A.Milne

3.1 Introduction

In this chapter I extend Standard Model Higgs Boson searches to the pp hadron colliders the LHC and the SSC. The material on heavy Higgs detection ($M_H > 2M_Z$); and for an intermediate mass Higgs in its decays $H \to \gamma\gamma$ and $H \to ZZ^*$ is review material. The sections on associate production of a Higgs boson with either a W, $t\bar{t}$, or single t quarks, are my own work and have been published in Ref. 14,15,16.

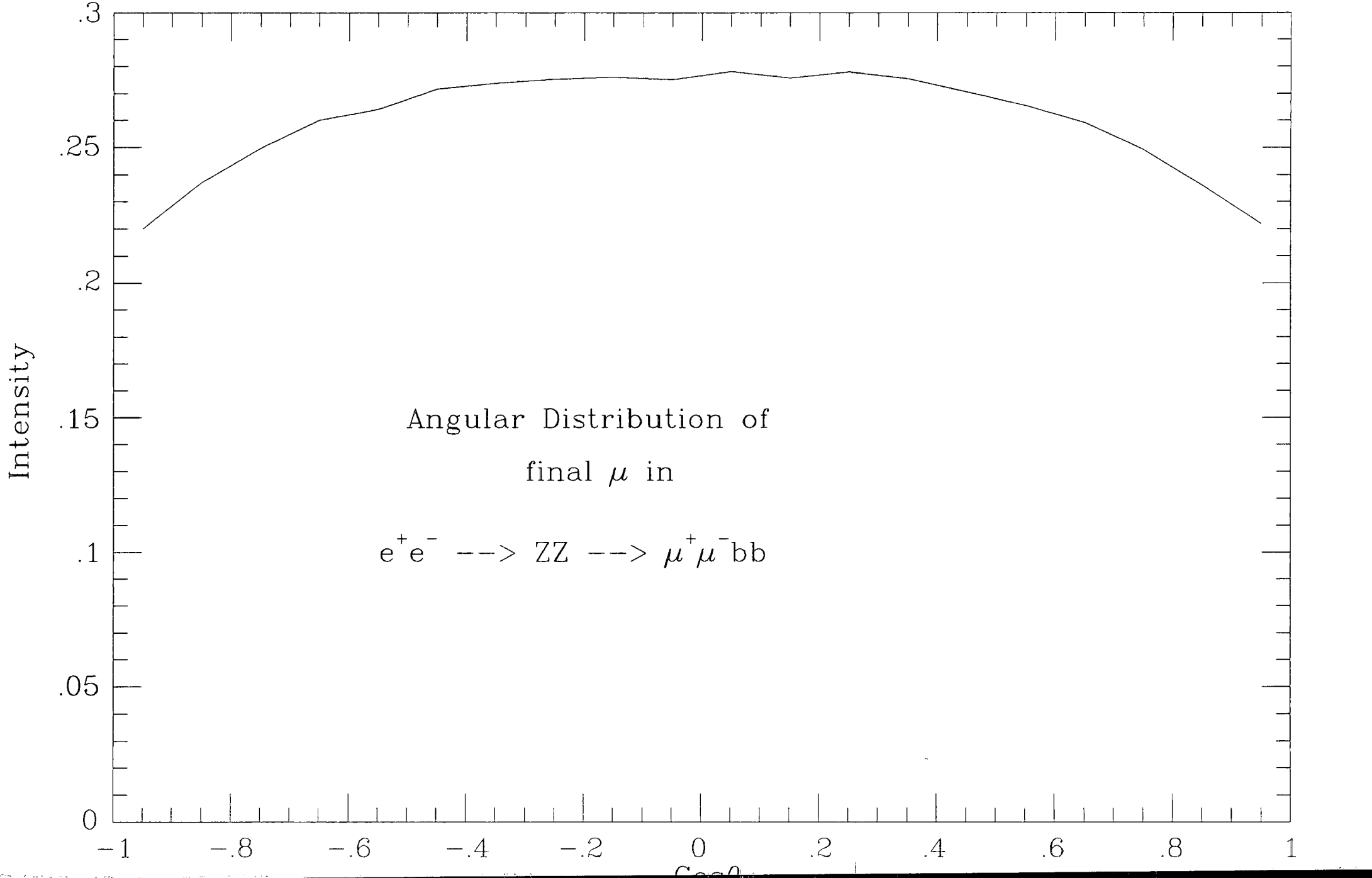
Physics at high energy hadron colliders is very different from physics at high energy e^+e^- colliders. At e^+e^- colliders we are restricted to fairly low integrated luminosity (LEP I has

$\mathcal{O}(100\,\mathrm{pb}^{-1})$, LEP II plans to have $\mathcal{O}(500\,\mathrm{pb}^{-1})$), but have a fairly *clean* environment. We know the total $\sqrt{s}$ of the machine, and know what the rest frame is. At hadron colliders the story is far different; we either collide protons on antiprotons, *e.g.*, at the Tevatron, or collide protons on protons, which is the plan for the future experiments at the LHC and the SSC. However these are not the fundamental objects that undergo the hard scattering, during the scattering process these hadrons break up into the more fundamental quarks and gluons that make up the hadrons, and then these quarks and protons undergo the hard scattering. How we describe these quarks and gluons, or partons, in the subject of Appendix E Parton distributions. That the partons undergo the hard scattering rather than the protons affects very much the environment of hadron colliders. The partons only carry some (*unknown*) fraction of the parent proton momenta, and so we do not know either the effective $\sqrt{s}$ nor the rest frame of the process. The partons do travel in approximately the same directions as the parent protons, and so we do know that the partons collide with almost no transverse momenta, $p_\perp$, and so we have some information about the frame in which the partons collide. Also when the partons that undergo the hard scattering emerge from the proton they leave the remains of the proton which, although largely collinear with the initial beam direction, still mean that our measured events have some underlying event from the proton remnants. Theory has remarkably little to say on these underlying events, and we must view this as a background. At e^+e^- colliders as the electrons are colourless the basic event is a QED scattering, which keeps our final states relatively *clean*; whereas at hadron colliders the colliding partons are coloured and so the dominant processes are QCD initiated, this also adds to the complexity of the measured event.

All these facts make precision measurements, and precision predictions, far harder at hadron colliders than at e^+e^- colliders. This means that many signals for *interesting* physics are dwarfed by *uninteresting* physics, and this very much changes the signals we look for. As an example if we look for a light Higgs boson in its $H \to b\bar{b}$ decay (which is effectively the *only* Higgs decay that we look for at e^+e^- colliders) at hadron colliders then we find that its signal is dwarfed by a huge QCD continuum background from processes like,

$$pp \to ggX \to b\bar{b}X \qquad . \tag{3.1.1}$$

Instead we tend to look for the rarer Higgs decay channels; like $H \to \gamma\gamma$ which for a light Higgs has a branching ratio of only about $\mathrm{Br}(Z \to \gamma\gamma) = \mathcal{O}(10^{-3})$. For these rarer decays the backgrounds are far smaller and far more under control. This means that typically our cross-section will be smaller that at e^+e^- machines. Hadron machines make up for this by having huge integrated luminosity in comparison to e^+e^- machines; the SSC plans to have an integrated luminosity of $\mathcal{O}(10^4\,\mathrm{pb}^{-1})$ per annum and the LHC plans to have at least this large an integrated luminosity, and quite possibly $\mathcal{O}(10^5\,\mathrm{pb}^{-1})$ per annum.


.3
.25
.2
.15
.1
.05
0
Intensity
Angular Distribution of
final μ in
e^+e^- --> ZZ --> $\mu^+\mu^-$bb
−1
−.8
−.6
−.4
−.2
0
.2
.4
.6
.8
1
cosθ

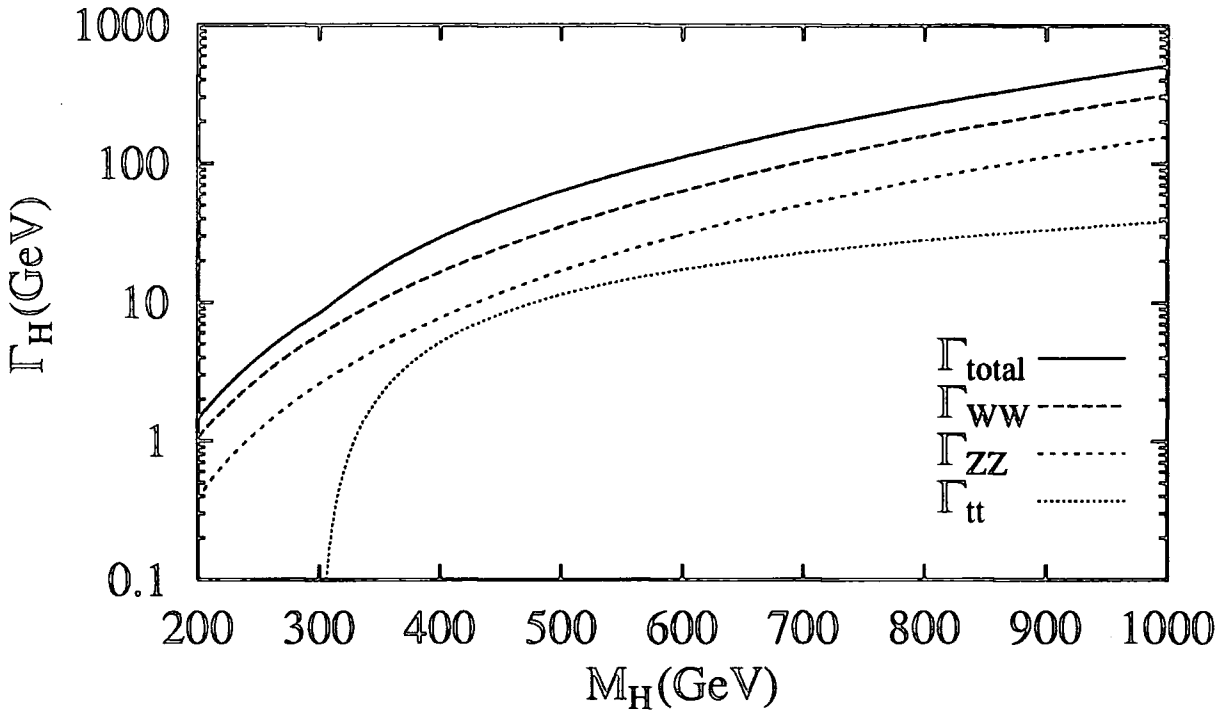


*Fig.*3.1.1 The dominant partial, and total widths for a heavy Higgs in the case when $m_t = 150\,\text{GeV}$

3.2 Heavy Higgs

By a heavy Higgs we mean a Higgs with mass larger than $M_H > 2M_Z$, a Higgs this heavy can decay to two on mass shell W's or Z's. Now the Higgs partial widths are given by [17],

$$\Gamma(H \to f\bar{f}) = \frac{n_C G_F m_f^2 M_H}{4\pi\sqrt{2}} \left(1 - \frac{4m_f^2}{M_H^2}\right)^{3/2} , \tag{3.2.1}$$

$$\Gamma(H \to W^+W^-) = \frac{G_F M_H^3}{8\pi\sqrt{2}} \left(1 - \frac{4M_W^2}{M_H^2}\right)^{1/2} \left(1 - \frac{4M_W^2}{M_H^2} + \frac{12M_W^4}{M_H^4}\right) , \tag{3.2.2}$$

$$\Gamma(H \to ZZ) = \frac{1}{2}\Gamma(H \to W^+W^-)\Big|_{M_W \leftrightarrow M_Z} , \tag{3.2.3}$$

where n_C is the usual colour factor. These widths are shown in *Fig.*3.1.1 for a heavy Higgs in the case where $m_t = 150\,\text{GeV}$. So when $M_H > 2M_Z$ the decays $H \to W^+W^-, ZZ$ dominate over Higgs decays to fermions, including the decay $H \to t\bar{t}$, which for $m_t > M_Z$ never has a branching ratio larger than 25%. As the top quark is turning out to be heavy ($m_t > 108\,\text{GeV}$[18]) the dominant production mechanism of the Higgs boson is via gluon gluon fusion shown in *Fig.*3.2.1. Although this is a loop process it dominates over tree level processes due to the large luminosity of gluons inside the proton, and also due to a large

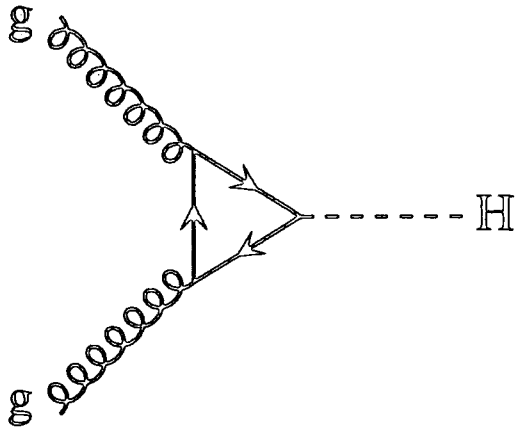


Fig.3.2.1 The Feynman diagram for the gluon fusion process to form a Higgs, the fermion in the loop tends to be the heaviest fermion, *i.e.*, the top quark.

contribution coming from a top quark loop, which has a very strong coupling to the Higgs due to its large mass.

In *Fig*.3.2.2 we show the cross-section for the process, *Fig*.3.2.1,

$$pp \to ggX \to HX \quad , \tag{3.2.4}$$

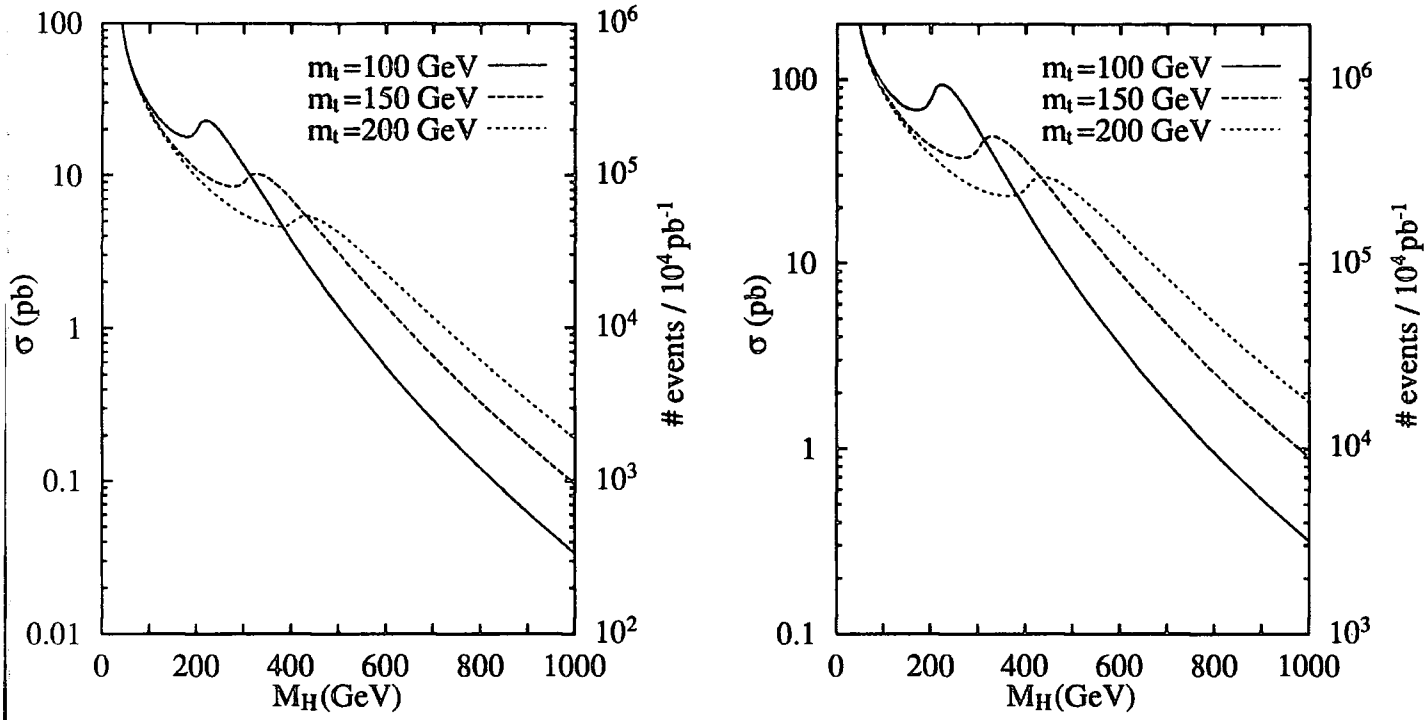
both at the LHC, which has $\sqrt{s} = 16\,\text{TeV}$, and at the SSC, which has $\sqrt{s} = 40\,\text{TeV}$. The SSC cross-section are larger because for a set Higgs mass the $\sqrt{s}$ being larger than at the LHC mean that we are at smaller Bjorken x, and the gluon parton distribution of the protons grows rapidly at small x. Also the cross-section does not drop as rapidly at the SSC as the LHC for increasing M_H due to the larger $\sqrt{s}$. We also show how many events these cross-sections correspond to for an integrated luminosity of $\mathcal{L} = 10^4\,\text{pb}^{-1}$ which is the canonical luminosity for both the LHC and the SSC.

In evaluating *Fig*.3.2.2 we have made use of the narrow width approximation for the Higgs, for this we require $\Gamma_H \ll M_H$, as *Fig*.3.1.1 shows this is not true for Higgs with $M_H > \mathcal{O}(500\,\text{GeV})$. In this case we should take *Fig*.3.2.2 with a pinch of salt - we can no longer talk about the production of a Higgs as an isolated particle, it plays a part in the physics of symmetry breaking as a whole. One of the roles that the Higgs particle plays in the Standard Model is to tame the bad high energy behaviour of processes like,

$$VV \to VV \tag{3.2.5}$$

$$VV \leftrightarrow t\bar{t} \quad , \tag{3.2.6}$$

where $V = Z, W^+, W^-$, now if we take the limit $M_H \to \infty$ then the Higgs part of the Standard Model decouples, and (3.2.5,3.2.6) regain their bad high energy behavior. This means that if we take M_H large, then the processes at low energies in comparison to the



a) at the LHC with $\sqrt{s} = 16\,\text{TeV}$. b) at the SSC with $\sqrt{s} = 40\,\text{TeV}$.

*Fig.*3.2.2 The cross-section for the process $pp \to ggX \to HX$ at the LHC and SSC. The dominant contribution comes from top quark loops, and so we are quite sensitive to the top quark mass, m_t.

Higgs mass look like they have this bad high energy behaviour; its only as we reach energies of about the Higgs mass that the cross-sections start behaving themselves. These large cross-sections are of the same overall size as the "resonant" Higgs production, and have the same initial and final states as Higgs production; and so we must include them in any calculation that we do. For this case where the Higgs is heavy we need to make a very careful study of the signals that we look for, and ask questions about the nature of the physics that we are trying to observe; this goes beyond the realms of this thesis – and I shall say no more on the subject.

As long as the Higgs is light enough that its width is narrow enough so we can talk about Higgs production in isolation then from *Fig.*3.2.2 it is clear that we have numerous events at both the LHC and SSC. Now approximately one third of these heavy Higgs decay to two Z's, and these Z's we can detect with virtually no background in their $Z \to e^+e^-, \mu^+\mu^-$ mode. This decay mode of the Z has a branching ratio of about 7%, and so if we insist that both Z's decay to electrons or muons we get a branching ratio,

$$\text{Br}(H \to ZZ \to 4l) = 0.15\% \quad . \tag{3.2.7}$$

So the signal we look for is for 4 leptons with 2 pairs that reconstruct the Z mass, and whose

total invariant mass show a peak at the Higgs mass. Now this signal has backgrounds from resonant processes, where two pairs of leptons each reconstruct the Z mass, like,

$$pp \to Xq\bar{q} \to XZZ \quad , \tag{3.2.8}$$

and also from non resonant processes like,

$$pp \to t\bar{t} \to b\bar{b}W^+\bar{W}^- \to 4l \quad , \tag{3.2.9}$$

where two of the leptons come from b decays and so are not isolated. In practice none of these backgrounds are serious[19] and it is only a question of rate. For a branching ratio $\mathrm{Br}(H \to ZZ \to 4l) = 0.15$ if we require 10 events for detection of the Higgs then we require $\mathcal{O}(10^4)$ Higgs events. From *Fig.*3.2.2 it is clear that the SSC will be able to detect the Higgs in this channel right up to where we can no longer talk about Higgs production in isolation; whereas the LHC we run out of events by $M_H = \mathcal{O}(500\,\mathrm{GeV})$ for the standard luminosity of $\mathcal{L} = 10^4\,\mathrm{pb}^{-1}$ and to probe to any higher mass we would require the high luminosity option for the LHC of $\mathcal{L} = 10^5\,\mathrm{pb}^{-1}$. More detailed studies confirm that to probe the very heavy Higgs sector in the 4 lepton decay channel at the LHC we require the high luminosity option[19].

Other decay channels of a heavy Higgs are not nearly so clean and easy to detect as its decay $H \to ZZ \to 4l$, but offer some hope of extending the upper Higgs limit to beyond 500 GeV at the standard luminosity LHC[19].

3.3 The Intermediate Mass Higgs

If the Higgs boson happens to be lighter than $2M_Z$ then it can still have a significant branching ratio into $2Z$'s where one Z is off mass shell, this is shown in *Fig.*3.2.3 where we can see that the Higgs branching ratio into $2Z$'s remains above 1% for Higgs masses greater than 120 GeV, the dip at $M_H = 160\,\mathrm{GeV}$ is because the decay channel $H \to WW$ has just opened up, and so $\Gamma(H \to WW)$ is increasing rapidly, this suppresses the branching ratio $\mathrm{Br}(H \to ZZ)$. As we approach the $H \to ZZ$ threshold the width $\Gamma(H \to ZZ)$ also starts increasing rapidly, and tends towards $\mathrm{Br}(H \to ZZ) = \frac{1}{3}$, its asymptotic value.

This means that we can still look for the Higgs in its decay,

$$H \to ZZ^* \to 4l \quad . \tag{3.3.1}$$

However because one Z is off mass shell we can no longer look for 2 pairs of leptons reconstructing the Z mass, for Higgs masses larger than M_Z though, one Z is effectively on mass shell; and so we can still insist that 1 pair of leptons reconstruct the Z mass.

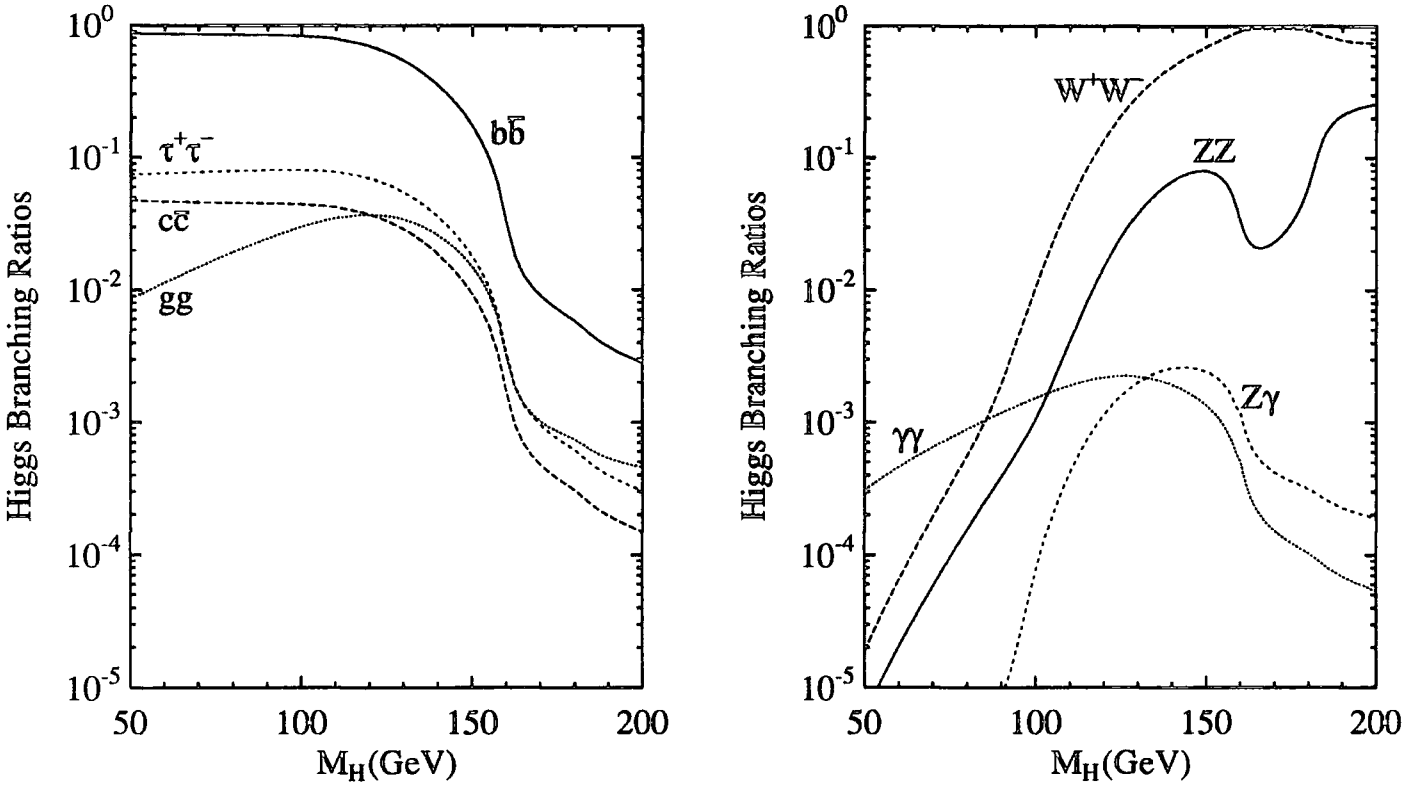


*Fig.*3.2.3 The main branching ratios of a Standard Model Higgs boson. The $Z\gamma$ width is obtained in Appendix G, and the other widths are from Ref. 20.

We now get irreducible backgrounds from two sources,

$$pp \to ZZ^*, Z\gamma^* \to 4l \quad . \tag{3.3.2}$$

The leptons coming from the γ^* decay peak at low invariant mass, and so we can largely eliminate the $Z\gamma^*$ background by insisting that the 2 leptons that don't reconstruct the Z mass must have invariant mass greater than some cut off. If we choose this cut off to be,

$$M_{ll} > 12\,\text{GeV} \quad , \tag{3.3.3}$$

then with reasonable experimental cuts[21],

$$p_\perp > 20\,\text{GeV} \qquad |\eta| < 3 \tag{3.3.4}$$

$$|M_Z - M_{ll}| < 10\,\text{GeV} \quad , \tag{3.3.5}$$

and for $l = e, \mu$ we find the number of events shown in *Table* 3.3.1 at the high luminosity LHC ($\mathcal{L} = 10^5\,\text{pb}^{-1}$) with a similar story at the standard luminosity SSC ($\mathcal{L} = 10^4\,\text{pb}^{-1}$), where cross-sections are about 4 times larger than at the LHC. It is clear from *Table* 3.3.1 that with the high luminosity LHC we have reasonable numbers of events for $M_H > 130\,\text{GeV}$,

m_H (GeV)	120	130	140	150	160	170	180
$\#H \to ZZ^* \to 4l$	13	68	161	229	114	75	224
$\#q\bar{q}, gg \to ZZ^*, Z\gamma^* \to 4l$	9	19	26	30	36	39	38

Table 3.3.1 The expected number of $pp \to H \to ZZ^* \to 4l$ signal events and $pp \to gg, q\bar{q} \to ZZ^*, Z\gamma^* \to 4l$ background events at the high luminosity LHC ($\mathcal{L} = 10^5\,\mathrm{pb}^{-1}$) for various mass values of the Standard Model Higgs boson. Taken from Ref. 21

but that with the standard luminosity there is only a small window about $M_H = 150\,\mathrm{GeV}$ where we have enough events to have any chance of detecting the Higgs.

As well as the ZZ^* and $Z\gamma^*$ irreducible backgrounds we also have reducible backgrounds from the processes,

$$pp \to t\bar{t}X \to WWb\bar{b}X \to 4lX \tag{3.3.6}$$

$$pp \to b\bar{b}ZX \to 4lX \quad , \tag{3.3.7}$$

and these processes have massive cross-sections, shown in *Table* 3.3.2, which are more than an order of magnitude larger than the $H \to ZZ^* \to 4l$ signal.

m_H (GeV)	120	130	140	150	160	170	180
$\#H \to ZZ^* \to 4l$	13	68	161	229	114	75	224
$\#t\bar{t} \to 4l$	695	1578	1885	1677	1448	1033	654
$\#Zb\bar{b} \to 4l$	153	309	448	535	712	547	777

Table 3.3.2 The expected number of $pp \to H \to ZZ^* \to 4l$ signal events and $pp \to t\bar{t}, Zb\bar{b} \to 4l$ reducible background events at the high luminosity LHC ($\mathcal{L} = 10^5\,\mathrm{pb}^{-1}$) for various mass values of the Standard Model Higgs boson before isolation cuts are applied; the top mass has been set to $m_t = 130\,\mathrm{GeV}$. Taken from Ref. 21

However in each of these backgrounds we have a b and a $\bar{b}$ quark which decay into leptons. As these b quarks are fairly energetic (because we insist that the leptons that they decay into are energetic) the remnants of hadronization and of the B hadron decay tend to be collinear with the lepton direction. This means that these leptons are not normally isolated but have hadronic activity associated with them. So if we veto events where the lepton has this hadronic activity associated with it we can largely eliminate this background. Conservatively we expect that only 1 in 7 leptonic b decays will *look* isolated; and so if we insist that all leptons are isolated we get a suppression by 7^2, as both (3.3.6,3.3.7) contain two leptons that

m_H (GeV)	120	130	140	150	160	170	180
#Signal $(H \to ZZ^*) \to 4l$	9	48	113	161	80	53	171
#Background (irreducible + reducible)	23	51	32	65	69	59	28
Significance $= S/\sqrt{S+B}$	1.6	4.8	9.4	10.7	6.6	5.0	12.1

Table 3.3.3 The expected number of $pp \to H \to ZZ^* \to 4l$ signal events and total number of background events at the high luminosity LHC ($\mathcal{L} = 10^5\,\mathrm{pb}^{-1}$) for various mass values of the Standard Model Higgs boson with lepton isolation cuts; the top mass has been set to $m_t = 130\,\mathrm{GeV}$. We also give the significance in s.d. that this signal would be if observed. Taken from Ref. 21

come from b decay[22]. If we apply this cut we find the numbers of events at the LHC shown in *Table* 3.3.3.

For $M_H > 130\,\mathrm{GeV}$ we have a healthy signal with signal to background ratios larger than about 1, this gives a high significance greater than about 5 s.d.. However for $M_H < 130\,\mathrm{GeV}$ clearly the signal rate is dropping very rapidly, this is because the branching ratio $\mathrm{Br}(H \to ZZ^* \to 4l)$ is dropping very rapidly as can be seen from *Fig*.3.2.3. Clearly whatever strings we pull we can not overcome this rapidly falling branching ratio and it will be *very hard* to use this channel to look for Higgs with $M_H \lesssim 130\,\mathrm{GeV}$.

At the SSC the same conclusions are true with the standard luminosity $\mathcal{L} = 10^4\,\mathrm{pb}^{-1}$, we can detect a Higgs with mass $M_H \gtrsim 130\,\mathrm{GeV}$ but can go no lower in M_H in this mode. In terms of the luminosity required at the LHC and SSC it is clear that the SSC benefits from a larger $\sqrt{s}$, which means that we make use of the higher gluon luminosity at small Bjorken x. Another advantage of being at smaller Bjorken x is that we improve our signal to background ratio, because as x decreases the gluon luminosity grows faster than the quark luminosity as we get decreasing contribution from valence quarks; now our signal is almost totally produced from the gg initial state, where as the irreducible backgrounds (3.3.2) are about 50% from the gg initial state and 50% from the $q\bar{q}$ initial state, clearly increasing the gluon luminosity relative to the quark luminosity increases the signal to background ratio.

3.3.1 Two photon decays of the Higgs boson

For $M_H < 130\,\mathrm{GeV}$, where we can't use the $H \to ZZ^* \to 4l$ decay mode to detect the Higgs, most other decay channels of the Higgs are overwhelmed by huge QCD backgrounds, and are totally unobservable. The only other decay channel that has remotely manageable

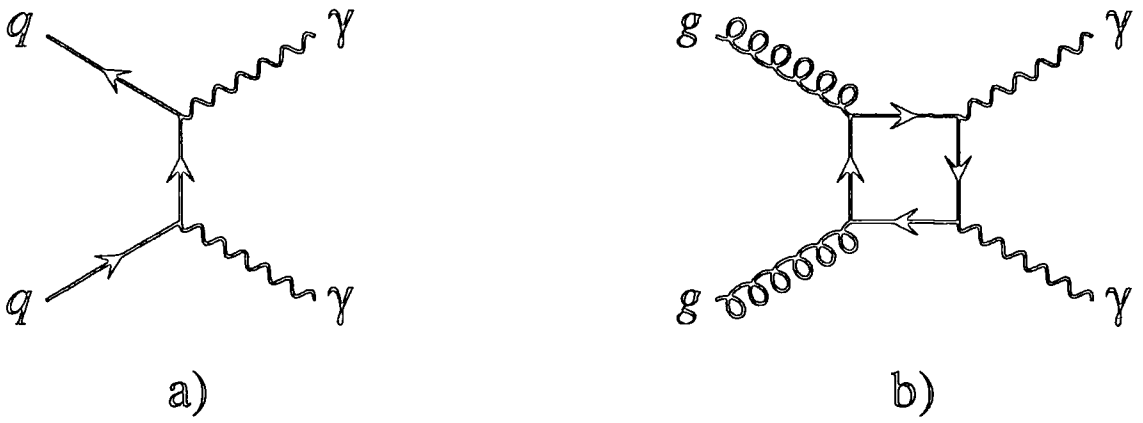


*Fig.*3.3.1 Feynman diagrams for the process $pp \rightarrow gg, qq \rightarrow \gamma\gamma$.

backgrounds is the decay,

$$H \rightarrow \gamma\gamma \quad . \tag{3.3.8}$$

This has a branching ratio of only about $\mathrm{Br}(H \rightarrow \gamma\gamma) = 10^{-3}$, see *Fig.*3.2.3, but we do still have numerous Higgs events, $\mathcal{O}(10^6)$, see *Fig.*3.2.2, at both the LHC and SSC, and so we should have $\mathcal{O}(10^3)$ Higgs decaying to 2 photons.

Backgrounds for this process come from two main sources, we have the huge irreducible continuum background from prompt photon production show in *Fig.*3.3.1; and the reducible background from jets that fragment in such a way that they look like photons. The latter may not sound particularly likely – but we should not forget that QCD jet production is about 7 orders of magnitude larger than prompt photon production, and that occasionally a jet will consist of a hard π^0 with the only other hadronic activity being soft. Now a π^0 decays about 99% of the time into two photons, and if the π^0 is energetic enough these γ have almost no separation and so look like a single photon. Now we can separate out the QCD reducible background because the produced π^0's normally have hadronic activity associated with them, and so if we insist that our photons are isolated we can largely eliminate the background; also for low energy π^0's, which dominate the cross-section, we can separate out the two photons and tell that the particle is a π^0. Non the less this calculation has great uncertainties:-

- We don't know the overall jet cross-section to better than a factor of 3.
- We don't know accurately exactly how often most of the energy of a jet is carried by a single π^0, this is right in the tail of a distribution and so hard to measure.
- Exactly how well we can differentiate photons from π^0's is unknown.

Because we can't accurately model this QCD background we must aim to make it as small as possible, and certainly far smaller than the prompt photon background. As the QCD background is about 7 orders of magnitude larger than the prompt photon production we

m_H (GeV)	Δm	$N_{\rm Sig}$	$N_{\rm Back}$	Significance $= \frac{N_{\rm Sig}}{\sqrt{N_{\rm Back}}}$
80	1.0	560	15400	4.5
100	1.5	1110	17300	8.4
150	2.0	880	6800	10.7

Table 3.3.4 The expected number of $pp \to H \to \gamma\gamma$ signal events and total number of background events at the high luminosity LHC ($\mathcal{L} = 10^5\,{\rm pb}^{-1}$) for various mass values of the Standard Model Higgs boson assuming excellent diphoton mass resolution of about 1.5%. We also give the significance in s.d. that this signal would be if observed. Taken from Ref. 22

should aim to be able to differentiate photons from jets to 1 part in 10^4; 10^4 because we have two photons to isolate, so the combined isolation is 1 part in 10^8.

Assuming that we can achieve this level of photon isolation then the only background we need worry about is prompt photon production, *Fig*.3.3.1,

$$pp \to q\bar{q}, gg \to \gamma\gamma \quad . \tag{3.3.9}$$

Unfortunately this background is also huge, and in this case irreducible. The only way of decreasing the contribution from this prompt photon background is by being able to measure the diphoton invariant mass very accurately, the Higgs width for an intermediate mass Higgs is only $\mathcal{O}$(MeV) and so the signal diphoton invariant mass is far narrower then we will ever be able to experimentally measure; however the prompt photon background is a continuum background, as we increase the accuracy of the diphoton invariant mass we can proportionally cut away at the continuum background. Typical numbers of events that we find are shown in *Table* 3.3.4 at the high luminosity LHC. The numbers shown are for a very excellent diphoton mass resolution of 1.5%, and still the signal is an order of magnitude less than the background; however as we have such large numbers of events we are able to get a high significance.

At the standard luminosity SSC again the results are similar, with a good veto for jets looking like isolated photons, and a very excellent diphoton mass resolution we can get a significant effect. Clearly the high luminosity is required at the LHC in order to probe Higgs masses down to the LEP II limit of $M_H = 80\,{\rm GeV}$, however this brings in its own complications. At the high luminosity LHC we will have about 10 individual events per bunch crossing – in such an environment it is not clear that it is sensible to talk about isolation as a meaningful concept. The standard luminosity SSC does not suffer nearly as badly as the LHC from multiple events per bunch crossing, there we only expect 4 such events.

3.4 Associate Production

The previous sections have shown that it is very hard to detect a light 'intermediate mass' Higgs boson at the LHC and SSC, at least if we just try to detect the Higgs in isolation. The Higgs boson only couples strongly to massive particles, this in practice means either W's, Z's, or t quarks. Now if we look at an ensemble of events that contain a heavy particle then we expect this will contain a Higgs boson a greater percentage of the time than events that contain no heavy particles. W's, Z's, and t quarks can be fairly easily tagged in their leptonic decay modes, and so we can ask if we can see the Higgs in association with this heavy particle. Of course by using such a strategy we expect a lower signal rate, but hope for a vastly increased signal to background ratio – we should not forget that in the previous sections it was not for lack of signal events that we had difficulty detecting the Higgs, but that the backgrounds were huge and overwhelmed our signal.

Again the dominant branching ratios of the Higgs are overwhelmed by huge QCD backgrounds, and we are forced to look at the rare branching ratios of the Higgs. For a light 'intermediate mass' Higgs this means looking at the two photon decay of the Higgs, $H \to \gamma\gamma$.

There are 3 main cases, for the 3 heavy particles, W's, Z's and t quarks, and one less important case for t quarks.

3.5 ZH Associate Production

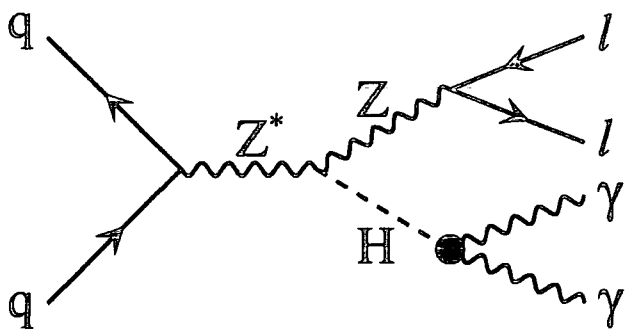


*Fig.*3.5.1 The Feynman diagram for ZH production at hadron colliders. We detect the Higgs in its two photon decay mode, $H \to \gamma\gamma$, and the Z in its leptonic decay mode, $Z \to l^+l^-$ $(l = e, \mu)$.

This process is show in *Fig.*3.5.1, the Z boson can be cleanly tagged in its decay $Z \to l^+l^-$ $(l = e, \mu)$ unfortunately this process only has a branching ratio of $\mathrm{Br}(H \to l^+l^-) = 7\%$ and this means that there are too few events to observe at even a high luminosity SSC [23].

3.6 WH Associate Production

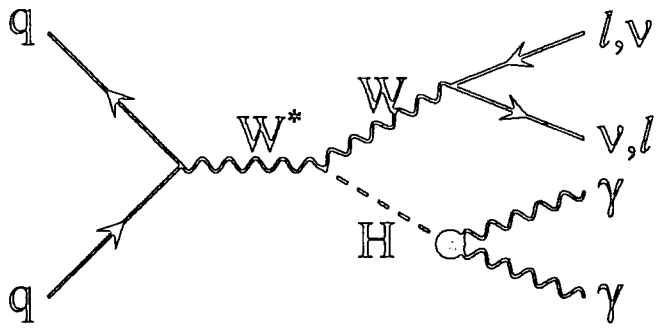


Fig.3.6.1 The Feynman diagram for WH production at hadron colliders.
We detect the Higgs in its two photon decay mode, $H \to \gamma\gamma$,
and the W in its leptonic decay mode, $W \to l^+\nu, l^-\bar{\nu}$ $(l = e, \mu)$.

This process is shown in *Fig*.3.6.1, again the W can be tagged in its leptonic decay mode, although in this case we only get a single lepton. Although this process is similar to ZH production the rate is about 6 times larger than ZH production[23] due to a larger leptonic branching ratio, $\mathrm{Br}(W \to l\bar{\nu}) = 20\%$ (c.f. $\mathrm{Br}(Z \to l^+l^-) = 7\%$) and also because the $V-A$ coupling of the W to the initial state quark are larger than the corresponding Z couplings.

If we concentrate on the two photon decay of the Higgs, then the signal we look for is a lepton from the W decay and two photons from the H decay, in order that we see these particles they must fall inside the detector, this means that they pass cuts like,

$$p_\perp > 20\,\mathrm{GeV} \qquad |\eta| < 2.5 \quad , \tag{3.6.1}$$

where $p_\perp$ is the momentum transverse to the beam pipe, and η is pseudorapidity, defined by,

$$\eta = -\ln(\tan(\theta/2)) \qquad \left(= \frac{1}{2}\ln\left(\frac{E+p_L}{E-p_L}\right) \text{ for massless particles.}\right) \tag{3.6.2}$$

Pseudorapidity is a more natural parameter to describe how forward or backward a particle is at hadron colliders than say θ or $\cos\theta$, because pseudorapidity has the advantage that differences in pseudorapidities are invariant under longitudinal Lorentz boosts for massless particles, and in a hadron collider the rest frame of the colliding partons tends to be boosted along the beam direction; and so if we use pseudorapidities the difference between the outgoing particles is independent of this boost.

Now we must also insist that both the photons and lepton are isolated, as otherwise we have large backgrounds from semileptonic heavy quark (b, c) decay faking leptons, and π^0's

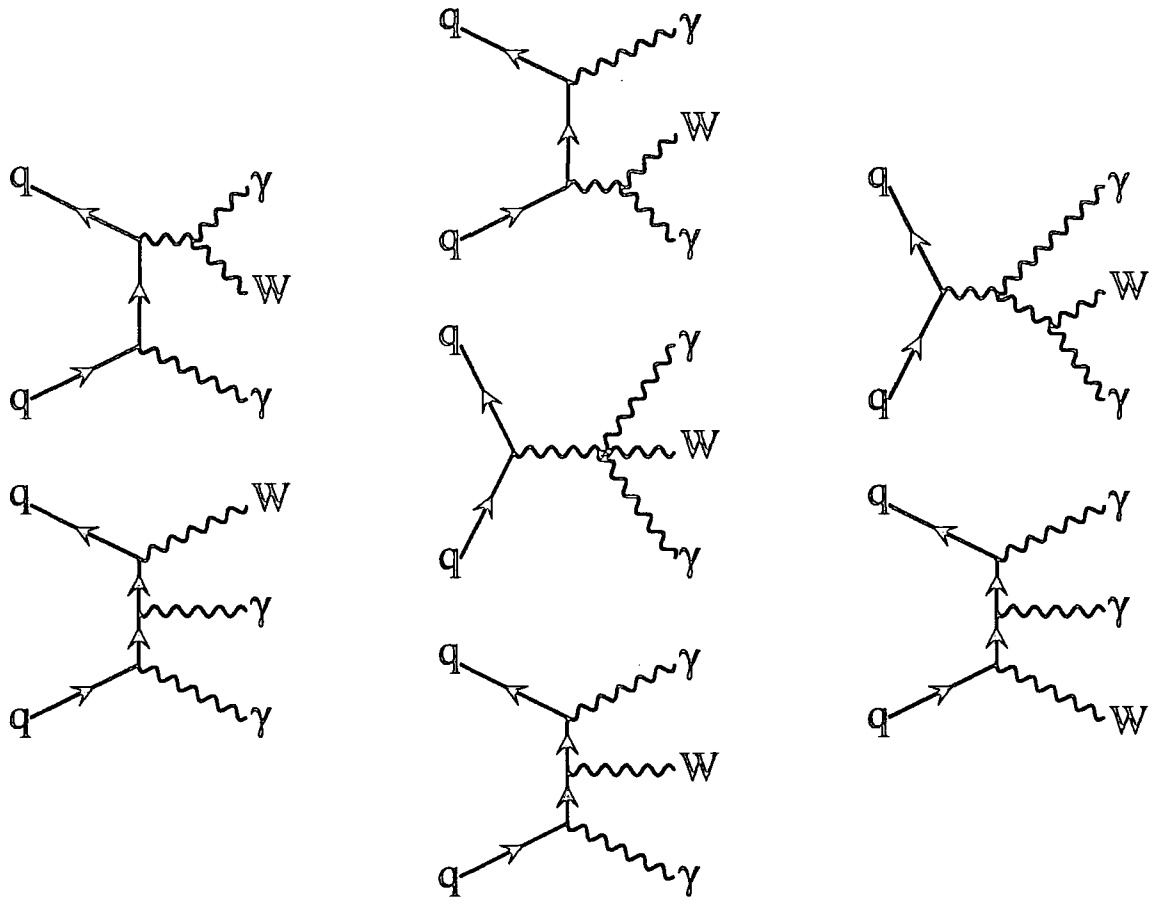


*Fig.*3.6.2 The Feynman diagrams for the process $q\bar{q} \rightarrow W\gamma\gamma$, for all the diagrams, except the central one, we must also add the diagrams where the two photon legs are exchanged.

decaying faking photons. This we implement by insisting that no other particles fall in a cone surrounding the leptons and photons, this cone we define by ΔR cuts,

$$\Delta R > 0.4 \qquad \text{where } \Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2} \quad . \tag{3.6.3}$$

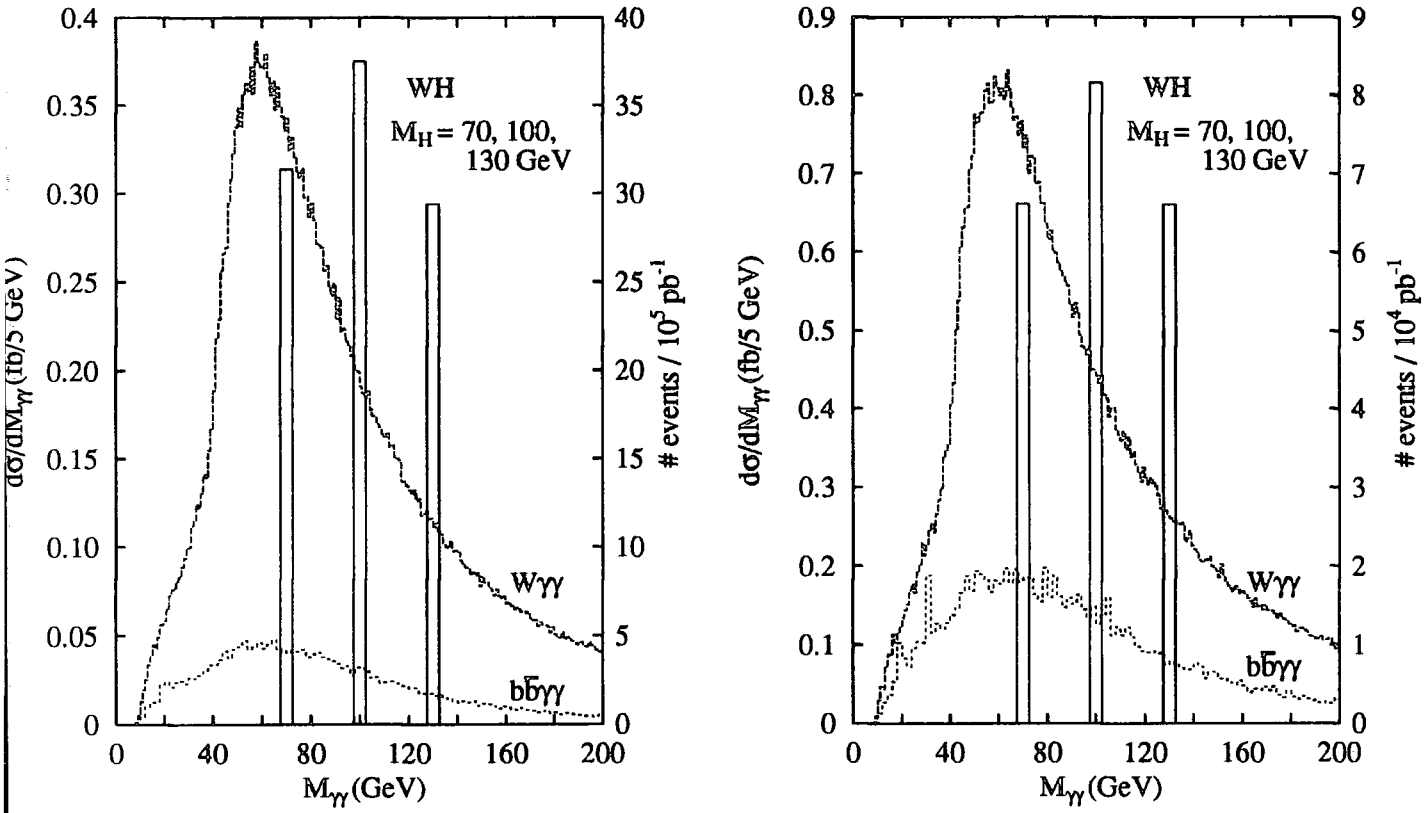
This cuts out a circular cone about photons or leptons when the cone is in the central part of the detector, with $\eta \approx 0$, away from the centre of the detector these cones become distorted, and turn into elliptical cones.

If we do this then the main background is the irreducible background,

$$pp \rightarrow q\bar{q}' \rightarrow W\gamma\gamma \quad . \tag{3.6.4}$$

At tree level this has the Feynman diagrams shown in *Fig.*3.6.2. The cross-sections and event rates for this signal and background are shown in *Fig.*3.6.3[23].

If we can achieve a rejection of jets faking photons to 1 part in 10^4, *i.e.*, the same rejection factor as used to detect the Higgs in its two photon decay, then we can eliminate to



a) for $\sqrt{s} = 16\,\text{TeV}$ b) for $\sqrt{s} = 40\,\text{TeV}$

*Fig.*3.6.3 The signal $pp \to WH \to l\gamma\gamma$ and the main irreducible background $pp \to W\gamma\gamma \to l\gamma\gamma$, and reducible background $pp \to b\bar{b}\gamma\gamma \to l\gamma\gamma$. Shown for LHC energies of $\sqrt{s} = 16\,\text{TeV}$ with integrated luminosity $\mathcal{L} = 10^5\,\text{pb}^{-1}$, and SSC energies of $\sqrt{s} = 40\,\text{TeV}$ with integrated luminosity $\mathcal{L} = 10^4\,\text{pb}^{-1}$, with the cuts (3.6.1,3.6.3) shown in the text.

an insignificant level the background from[24],

$$\begin{aligned} pp &\to W jet\, jet \to W\gamma\gamma \\ pp &\to W\gamma\, jet \to W\gamma\gamma \quad . \end{aligned} \tag{3.6.5}$$

The background from heavy quarks faking leptons comes from processes like,

$$pp \to b\bar{b}\gamma\gamma \to Xl\gamma\gamma \quad , \tag{3.6.6}$$

the question here is how well can we distinguish a semileptonically decaying b quark from a genuine isolated lepton, conservatively we expect that only $1/7$ semileptonic b decays will pass our isolation cut (3.6.3)[22]. Using this value for the rejection factor coming from isolation, and calculating just the tree level gluon fusion subprocess,

$$pp \to gg \to b\bar{b}\gamma\gamma \to Xl\gamma\gamma \quad , \tag{3.6.7}$$

which we expect to dominate at both the LHC and SSC then this process does not present a serious background[23, 14, 25]. The rate is shown in *Fig*.3.6.3.

Looking at *Fig*.3.6.3 it should be realized that we have binned the photon–photon invariant mass in 5 GeV bins, in practice we expect to be able to achieve a far better experimental resolution than this. So binning the cross-section in narrower bins will cause the continuum backgrounds to be lowered, without lowering the signal at all (remember that for a 'intermediate mass' Higgs its width is $\mathcal{O}(\text{MeV})$ and so is far smaller than any experimental resolution). It is clear from *Fig*.3.6.3 that we have good signal to background ratios, far better than the signal to background ratios for the plain $H \to \gamma\gamma$ detection mode for the Higgs. However our signal rate is low, we have only a handful of events; this means that we will require the highest luminosity possible at both the LHC and the SSC. In *Table* 3.6.1 we show the numbers of signal events expected at the high luminosity LHC and SSC (with $\mathcal{L} = 10^5\,\text{pb}^{-1}$) we also show the numbers of background events for 3% $M_{\gamma\gamma}$ resolution, and give the significance of this signal.

3.6.1 Difficulties with Next–to–leading order?

Normally in any process we expect the next–to–leading order (NLO) corrections to be $\mathcal{O}(\alpha)$, and so for QCD processes at high energies to only be of $\mathcal{O}(10\%)$, this means that we normally have confidence that NLO corrections will not change the overall picture of a process that we have, but just change the fine details. Indeed if we look at the NLO corrections to WH production[26],

$$\begin{aligned} pp &\to q\bar{q}' \to WHg \\ pp &\to g\bar{q}' \to WH\bar{q} \\ pp &\to qg \to WHq' \quad , \end{aligned} \tag{3.6.8}$$

then the corrections increase the cross-section by about 10% [26].

However if we consider at the process[27],

$$pp \to q\bar{q}' \to W\gamma \quad , \tag{3.6.9}$$

then the NLO correction to this process,

$$\begin{aligned} pp &\to q\bar{q}' \to W\gamma g \\ pp &\to g\bar{q}' \to W\gamma\bar{q} \\ pp &\to qg \to W\gamma q' \quad , \end{aligned} \tag{3.6.10}$$

turn out to enhance the cross-section by a factor of about 3. Clearly something odd is going on. In this case it is understood what is going on, in the LO process (3.6.9) then when the photon makes a certain angle with respect to the $q\bar{q}'$ pair then there is a *radiation zero*, where

$\sqrt{s}$ (TeV)	m_H (GeV)	Δm (GeV)	WH # events	$W\gamma\gamma$ # events	$b\bar{b}\gamma\gamma$ # events	Significance $= \frac{N_{\rm Sig}}{\sqrt{N_{\rm Back}}}$
16	60	1.8	27.6	13.4	1.6	7.1
	70	2.1	32.9	14.1	1.7	8.3
	80	2.4	36.2	14.0	1.8	9.1
	90	2.7	38.2	12.8	1.9	10.0
	100	3.0	39.2	11.7	1.9	10.6
	110	3.3	38.8	10.9	1.6	11.0
	120	3.6	35.7	9.6	1.5	10.7
	130	3.9	29.4	8.9	1.3	9.2
	140	4.2	20.8	8.1	1.0	6.9
	150	4.5	12.0	7.3	1.0	4.2
	160	4.8	3.8	6.7	0.9	1.4
40	60	1.8	57.8	29.2	6.4	9.7
	70	2.1	69.3	31.4	7.6	11.1
	80	2.4	77.0	30.4	8.0	12.4
	90	2.7	82.1	28.9	8.4	13.4
	100	3.0	85.3	26.6	8.0	14.5
	110	3.3	85.3	24.2	7.6	15.1
	120	3.6	79.5	22.3	6.5	14.8
	130	3.9	66.1	20.5	5.9	12.9
	140	4.2	47.2	18.7	6.0	9.5
	150	4.5	27.7	17.2	5.4	5.8
	160	4.8	8.8	16.0	4.8	1.9

Table 3.6.1 The expected number of $pp \to WH \to l\gamma\gamma$ signal events and number of background events at the high luminosity LHC and SSC ($\mathcal{L} = 10^5\,\mathrm{pb}^{-1}$) for various mass values of the Standard Model Higgs boson assuming reasonable diphoton mass resolution of about 3%. We also give the significance in s.d. that this signal would be if observed.

the cross-section vanishes identically; this is due to destructive interference between photons radiated off the W and quark legs. Now this radiation zero dominates the whole production pattern of the $W\gamma$ pair, and effectively suppresses radiation over a large range of angles; and this lowers the total cross-section considerably over what you would naively expect. Now when we move to NLO then this radiation zero for certain photon angles disappears (this should not surprise us, as the flow of charge into and out of the NLO processes (3.6.10) is completely different from the LO process (3.6.9)), and so the NLO corrections are $\mathcal{O}(\alpha_s)$ of the naive expectation for $W\gamma$ production rather the $\mathcal{O}(\alpha_s)$ of the radiation zero dominated cross-section. This does not necessarily mean that our perturbation theory is out of control, as we more to higher orders in α_s we do not expect similar enhancements; we should look at this result as saying that the LO cross-section is abnormally small, rather than the NLO cross-section being abnormally large.

However this raises the question as to how large are the NLO corrections to $W\gamma\gamma$ pro-

duction. Unfortunately these have not yet been calculated, and indeed the NLO corrections to this process would stretch current technology on loop calculations to the limit due to the presence of pentagon diagrams like *Fig*.3.6.4.

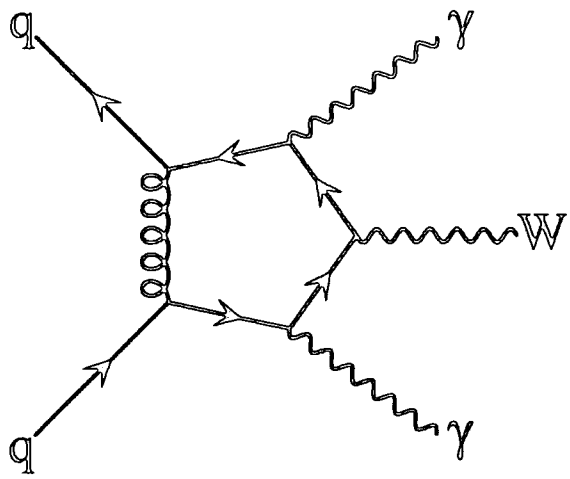


Fig.3.6.4 One of the pentagon Feynman diagrams that contributes to the NLO calculation of $p\bar{p} \rightarrow W\gamma\gamma$.

What we can do is to ask whether we expect large NLO enhancements in analogy with $W\gamma$ production. Now the matrix element for the LO $q\bar{q}' \rightarrow W\gamma\gamma$ process (3.6.4) is not the simplest of matrix elements, and finding zeros of it is far from trivial. However if we use a purely numerical approach, and compare the numerical size of the matrix element squared to the matrix element squared for individual Feynman diagrams it is clear that there are large areas of phase space where the matrix element is strongly suppressed; and so even if exact radiation zeros are not present, the large suppression of the matrix elements means that we expect the LO $q\bar{q}' \rightarrow W\gamma\gamma$ to be abnormally small.

Recently the real contributions to the NLO calculation of $pp \rightarrow W\gamma\gamma$,

$$\begin{aligned} pp &\rightarrow q\bar{q}' \rightarrow W\gamma\gamma g \\ pp &\rightarrow g\bar{q}' \rightarrow W\gamma\gamma\bar{q} \\ pp &\rightarrow qg \rightarrow W\gamma\gamma q' \quad , \end{aligned} \tag{3.6.11}$$

have been calculated [28], and here they find that LO $W\gamma\gamma + 1jet$ cross-section is about 3.5 times larger than the LO $W\gamma\gamma$ cross-section. This result should be taken with a pinch of salt, the LO calculation for $W\gamma\gamma+1jet$ production is formally divergent when the final jet becomes collinear to the incoming partons, or becomes soft; now this divergence cancels with the 1 loop contribution to $W\gamma\gamma$ production, because when the final jet is either collinear of soft it can not be experimentally differentiated from $W\gamma\gamma$ production. This means that until these virtual 1 loop contributions have been calculated we can not know how much of the divergent

result cancels. In practice what is done is that a minimum $p_\perp$ cut is imposed on the jet, that at least in principle keeps the jet experimentally observable, and so separates the $W\gamma\gamma + 1jet$ process from the $W\gamma\gamma$ process. If this cut is chosen too small then we *feel* too much of the divergent cross-section, unfortunately we can not know how small too small is without doing the complete NLO corrections. Hence we should be wary of the results of Ref. 28. Having said all this though Ref. 28 finds a considerable contribution from $W\gamma\gamma + 1jet$ even with large $p_\perp$ cuts (*e.g.*, $p_\perp > 50\,\mathrm{GeV}$); and this, in conjunction to the apparent large suppression in the LO $W\gamma\gamma$ production *should* fill us with worry.

3.6.2 What if the next–to–leading order corrections are large?

If the NLO corrections do turn out to be very large for $W\gamma\gamma$ production then at first sight it seems that the significance of our WH signal will be drastically decreased. If the K factor (that is the NLO cross-section divided by the LO cross-section) is, for example, 3, as with $W\gamma$ production, then our significance will be decreased by a factor of $\sqrt{3}$ which is clearly a large price to pay. *Is there anything that we can do to resurrect this signal?* If we return to $W\gamma$ production then although the total cross-section has very large NLO correction if we look at the more exclusive channel,

$$pp \to W\gamma + 0\ jets \quad , \tag{3.6.12}$$

(*i.e.*, we veto on the presence of any observed jets), then this process has far smaller NLO corrections, as the bulk of the NLO cross-section is coming from $pp \to W\gamma + 1\ jets$. Presumably we will be able to play the same game if $W\gamma\gamma$ production has large NLO corrections, and hopefully in this way be able to control the $W\gamma\gamma$ background. Clearly this method has a caveat, *What is a jet?* Or how hard does a jet have to become before we can experimentally call it a jet? We must not forget that each event will have an underlying event from the break up of the proton remnants, these underlying particles are not part of the hard scattering and so we do not wish to call this part of the jet. Also for low $p_\perp$ jets our fixed order perturbation theory no longer describes experiment at all well, this becomes the realm of exponentiation and of Monte Carlo models of hadronization; this by its very nature is an inexact science, and as such we wish to keep as far from this situation as feasible. In practice this means that we only observe jets with $p_\perp$ larger than some experimental cut off; if this cut off values turns out to be large, say 100 GeV, which seems quite likely at the hadron super colliders the LHC and the SSC, then we will include much of NLO correction cross-section in measuring the process (3.6.12), and so will not escape from large enhancements in the background process to WH production.

Also, as we have seen in *Table* 3.6.3, we only have reasonable numbers of signal events at the LHC and SSC if we go for the highest integrated luminosity available of $\mathcal{L} = 10^5\,\mathrm{pb}^{-1}$,

and at these high luminosity both the LHC and SSC will have multiple hard scattering per bunch crossing. In such an environment it is not clear what it means to ask to see a process with 0 jets (3.6.12).

There are many questions left unanswered in this area of looking for a light 'intermediate mass' Higgs, in association with a W boson, in the isolated lepton and two photon channel. Much of the early enthusiasm in the significance levels shown in *Table* 3.6.3 *must* be held back until a far more complete study of this process has been carried out; both on the part of theoreticians, in calculating the NLO corrections to $W\gamma\gamma$ production; and on the part of experimentalists, in carrying out detailed Monte Carlo simulations of how these events will actually look in our detectors.

3.7 $t\bar{t}H$ Associate Production

The Feynman diagrams for $t\bar{t}H$ production,

$$pp \rightarrow gg, q\bar{q} \rightarrow t\bar{t}H \quad , \tag{3.7.1}$$

are shown in *Fig*.3.7.1. At the LHC and SSC hadron super colliders the gluon fusion process dominates, due to the large gluon luminosity, and the colour structure of the matrix elements; however the quark–antiquark fusion still contributes about 10% of the gluon contribution at the SSC and 25% at the LHC, and so should not be neglected[15]. Now as $m_t > M_W$ [18, 29], within the Standard Model at least, 100% of t quarks decay to on mass shell W's and to tag the heavy particle all we need to do is to tag the W, which means the semileptonic W decay $W \rightarrow l$ with $l = \mu, e$. Also if we only tag on a single lepton, then we gain a combinatoric factor of 2 from the two t quark decays. Again we insist that the Higgs decays to two photons, being the only clean channel. Also we must insist that both photons and the lepton are isolated, in order to escape backgrounds from π^0 decay, and heavy quark decay. This leads to the set of cuts,

$$\begin{aligned} p_\perp > 20\,\mathrm{GeV} \qquad & |\eta| < 2.5 \\ \Delta R > 0.4 \qquad & \text{where } \Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2} \end{aligned} \quad . \tag{3.7.2}$$

Although we are tagging on the same signal as for WH production, these $t\bar{t}H$ events will look considerably different to the WH events. In the WH events the hard scattering just produces two isolated photons, an isolated lepton, and unobserved neutrino. However the $t\bar{t}H$ final state is far more active, from one t quark decay as well as getting a lepton, we also have a produced b quark, giving a jet; now the other t quark will typically decay hadronically to produce 3 further jets (one of which is a b jet): and so in total we will typically have 4

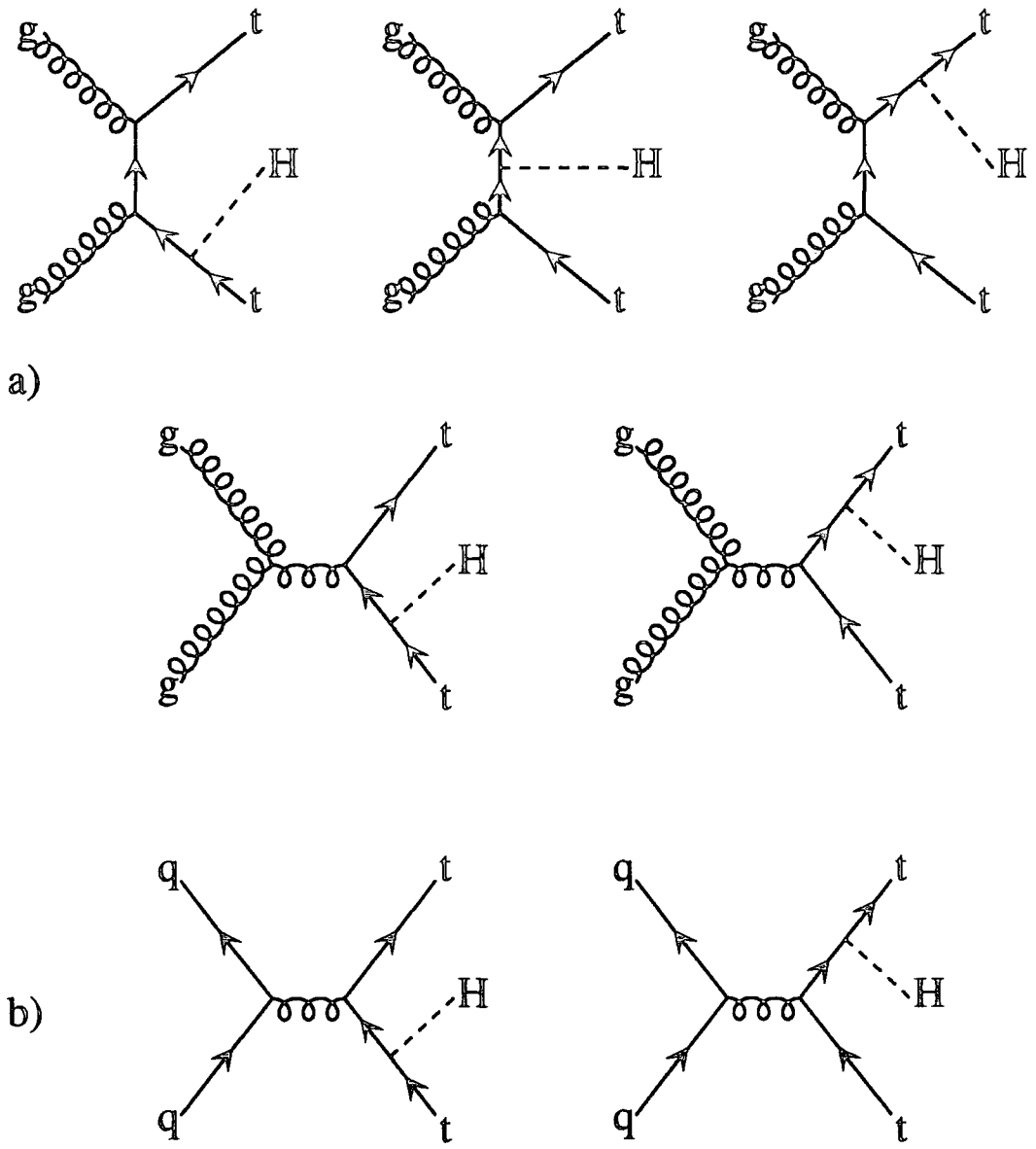


*Fig.*3.7.1 The Feynman diagrams for the processes, a) $gg \rightarrow t\bar{t}H$ b) $q\bar{q} \rightarrow t\bar{t}H$.

jets in addition to the isolated lepton and two photons, this is clearly a far more active final state than the WH final state. This means that we can differentiate this final state from the WH final state, and this in turn means that this process has different backgrounds from the WH signal.

The main background for this process is the irreducible,

$$pp \rightarrow gg, q\bar{q} \rightarrow t\bar{t}\gamma\gamma \quad , \tag{3.7.3}$$

both the gg fusion and $q\bar{q}$ fusion have numerous Feynman diagrams, too numerous to list here; however the basic structure is that we take the bare Feynman diagrams for $gg \rightarrow t\bar{t}$ and $q\bar{q} \rightarrow t\bar{t}$, and then tack a photon onto every charged line, then do the same again. This means that we end up with 30 Feynman diagrams for $gg \rightarrow t\bar{t}\gamma\gamma$, and 20 Feynman diagrams

for $q\bar{q} \to t\bar{t}\gamma\gamma$. Using the cuts (3.7.2) we end up with the cross-sections shown in *Fig*.3.7.2. It is clear that for both the signal and background the effect of the $q\bar{q}$ initial state is far larger at the LHC than at the SSC. This is because we are at smaller $\sqrt{s}$ at the LHC and so for the same $\sqrt{\hat{s}}$ we require larger Bjorken x's and so we feel the effect of the valence quarks in the proton more. For the signals the $q\bar{q}$ initial state is about 10% at the SSC, and 25% at the LHC, of the gg initial state; this is about double the enhancement that we get for just $t\bar{t}$ production without the H, this is because the extra particle in the final state (the Higgs) forces the $\sqrt{\hat{s}}$ to be larger and so forces us to larger Bjorken x where we again feel the effect of the valence quarks more. At the SSC the $q\bar{q}$ background is 20% of the gg background for $m_t = 100\,\text{GeV}$ increasing to 35% for $m_t = 180\,\text{GeV}$. At the LHC the $q\bar{q}$ background is 50% of the gg background for $m_t = 100\,\text{GeV}$ increasing to 120% for $m_t = 180\,\text{GeV}$. It is clear that the $q\bar{q}$ initial state is a very important source of background events, especially for larger m_t, although fortunately for larger m_t the $t\bar{t}\gamma\gamma$ background is far smaller than the $t\bar{t}H$ signal. For lighter m_t where the $t\bar{t}\gamma\gamma$ background is far more severe the $q\bar{q}$ is of lesser importance – but still significantly enhances the background. The large contribution of the $q\bar{q}$ initial state to the background relative to the signal is because for the signal the Higgs can only couple to the final state t quarks for both the gg and the $q\bar{q}$ initiated process; whereas for the background with the $q\bar{q}$ initial state the photons can also couple to the initial state quarks, and this interferes constructively with the final state radiation, also initial state radiation lowers the Q^2 of the exchanged gluon, and so puts it closer to its mass shell.

The background being enhanced by such a large amount from initial state radiation begs the question as to whether or not it can be reduced by careful cuts. At first glance this seems impossible as the background is an irreducible background; however we should not forget that the production mechanism is considerably different from the $t\bar{t}H$ signals and the gg initiated $t\bar{t}\gamma\gamma$ background. Now the $q\bar{q} \to t\bar{t}\gamma\gamma$ process has radiation from the massless initial state, this means that we have collinear singularities whenever the photons become collinear to the beam pipe, and so will get a large contribution from this configuration. The process without initial state radiation do not have these singularities when the photons come close to the beam pipe. Now for the $q\bar{q} \to t\bar{t}\gamma\gamma$ background these collinear singularities are regulated by the pseudorapidity cut $|\eta| < 2.5$, and to a lesser extent by the $p_\perp$ cut. This means that we may be sensitive to the $q\bar{q} \to t\bar{t}\gamma\gamma$ background by varying these cuts. In practice the $p_\perp$ distributions for all signal and background processes are very similar and we do not enhance signal to background ratios by varying the $p_\perp$ cut. However the rapidity distributions do show differences, we show this in *Fig*.3.7.3 where we have relaxed the $|\eta|$ cut. Unfortunately the differences are not huge, and, given the loss in events, we do not do any better by imposing a more restrictive rapidity cut than $|\eta| < 2.5$.

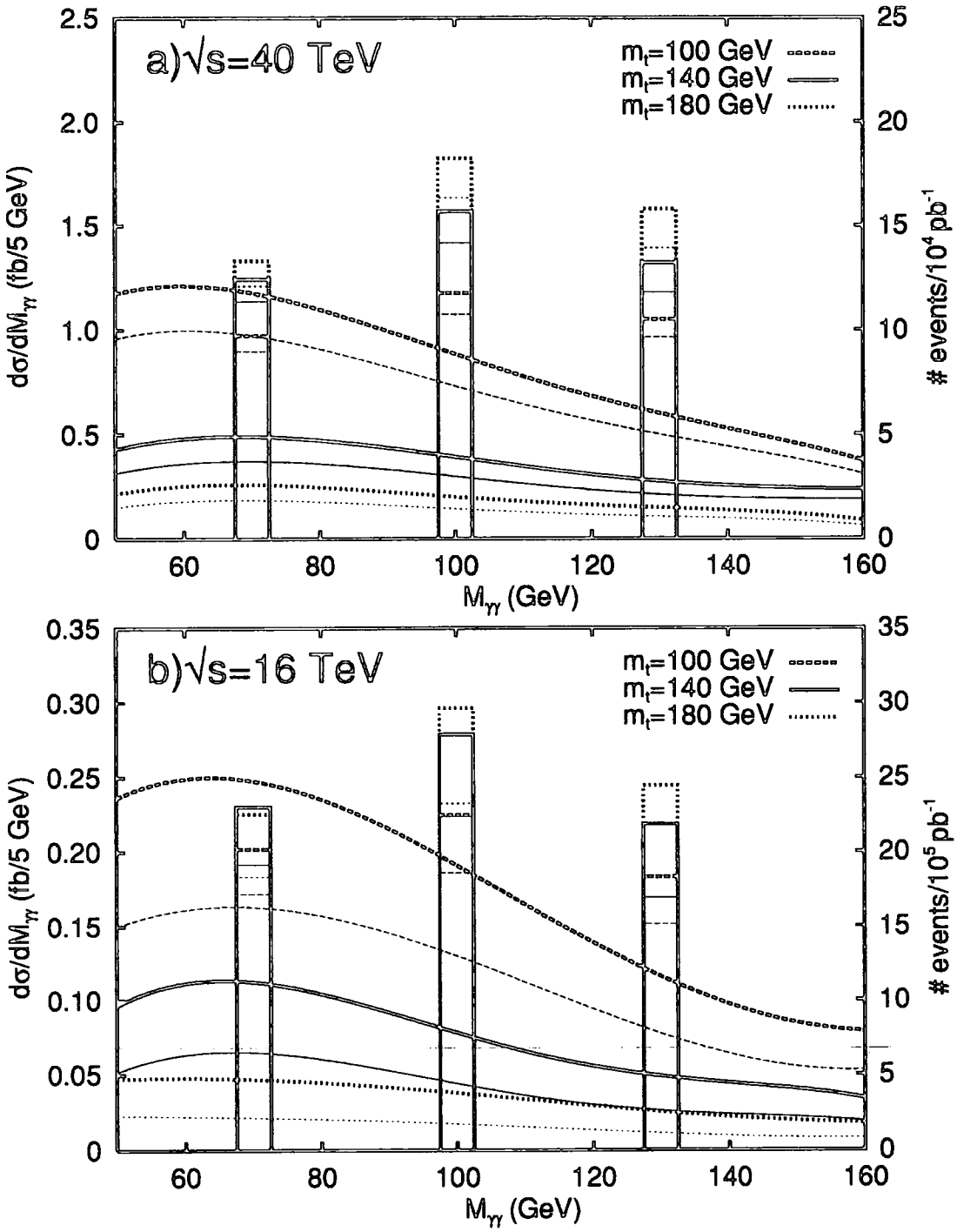


*Fig.*3.7.2 The differential cross sections ($d\sigma/dM_{\gamma\gamma}$ (fb/5 GeV)) for the process $gg \to t\bar{t}\gamma\gamma$, shown with light lines, and for $gg + q\bar{q} \to t\bar{t}\gamma\gamma$, shown with thick lines at the a) the SSC with $\sqrt{s} = 40$ TeV and b) the LHC with $\sqrt{s} = 16$ TeV. Shown for three values of $m_t = 100, 140, 180$ GeV. Also shown are the expected numbers of events for the standard luminosity SSC ($\mathcal{L} = 10^4\,\mathrm{pb}^{-1}$) and the high luminosity LHC ($\mathcal{L} = 10^5\,\mathrm{pb}^{-1}$). Superimposed are the cross sections for the processes $gg \to t\bar{t}H$ (light lines) and $gg + q\bar{q} \to t\bar{t}H$ (thick lines) for three values of $M_H = 70, 100, 130$ GeV.

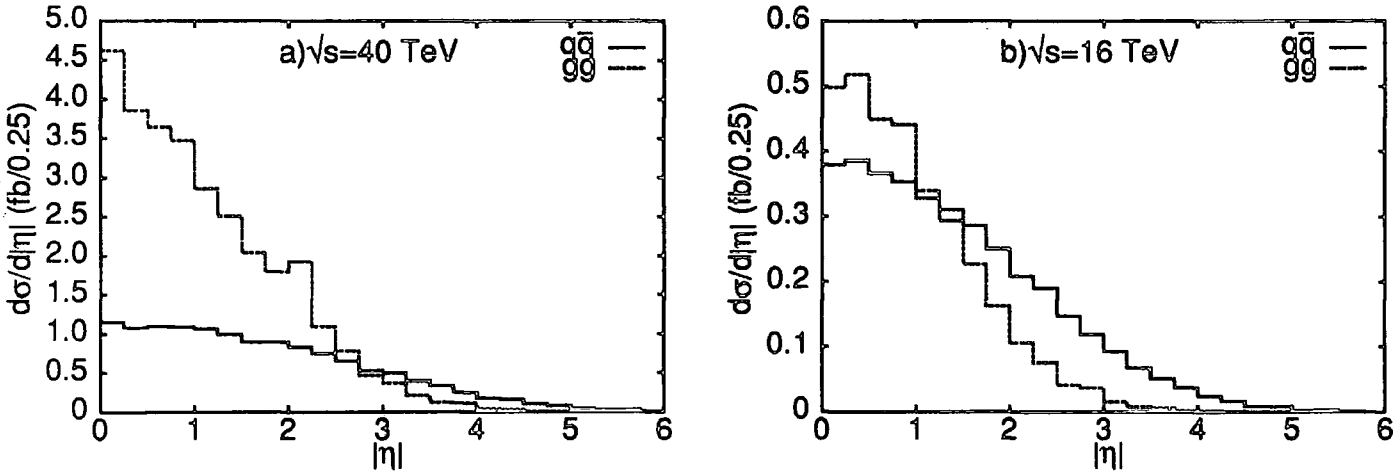


*Fig.*3.7.3 The differential cross sections $(d\sigma/d|\eta|(\,\mathrm{fb}/0.25))$ for the processes $q\bar{q} \to t\bar{t}\gamma\gamma$ (solid) and $gg \to t\bar{t}\gamma\gamma$ (dashed) for $m_t = 140\,\mathrm{GeV}$.

m_H	$m_t = 100\,\mathrm{GeV}$			$m_t = 140\,\mathrm{GeV}$			$m_t = 180\,\mathrm{GeV}$		
(GeV)	$t\bar{t}H$	$t\bar{t}\gamma\gamma$	$\frac{N_{t\bar{t}H}}{\sqrt{N_{t\bar{t}\gamma\gamma}}}$	$t\bar{t}H$	$t\bar{t}\gamma\gamma$	$\frac{N_{t\bar{t}H}}{\sqrt{N_{t\bar{t}\gamma\gamma}}}$	$t\bar{t}H$	$t\bar{t}\gamma\gamma$	$\frac{N_{t\bar{t}H}}{\sqrt{N_{t\bar{t}\gamma\gamma}}}$
	Number of events for $\sqrt{s} = 16\,\mathrm{TeV}$ with $\mathcal{L} = 10^5\,\mathrm{pb}^{-1}$								
60	18.9	9.0	6.3	20.9	4.0	10.5	19.8	1.7	15.2
70	21.3	10.4	6.6	24.7	4.7	11.4	24.2	2.0	17.1
80	22.6	11.3	6.7	27.1	5.0	12.1	27.4	2.1	18.9
90	23.2	11.6	6.8	29.0	5.0	13.0	30.3	2.2	20.4
100	23.6	11.5	7.0	29.9	4.7	13.8	32.1	2.3	21.2
110	23.3	10.9	7.1	29.8	4.3	14.4	32.7	2.2	22.0
120	21.8	10.0	6.9	27.7	4.0	13.9	30.9	2.1	21.3
130	18.4	9.0	6.1	22.9	3.8	11.7	26.0	2.0	18.4
140	13.5	8.2	4.7	16.3	3.7	8.4	18.7	1.8	13.9
150	8.2	7.6	3.0	10.0	3.6	5.3	10.9	1.8	8.1
160	2.8	7.6	1.0	3.1	3.3	1.7	3.5	1.7	2.8
	Number of events for $\sqrt{s} = 40\,\mathrm{TeV}$ with $\mathcal{L} = 10^4\,\mathrm{pb}^{-1}$								
60	8.9	4.4	4.2	11.0	1.7	8.4	11.5	0.9	12.1
70	10.3	4.9	4.7	13.2	2.1	9.1	14.3	1.1	13.6
80	11.1	5.3	4.8	14.8	2.3	9.8	16.4	1.2	15.0
90	11.8	5.4	5.1	16.1	2.4	10.4	18.4	1.2	16.8
100	12.3	5.3	5.3	16.8	2.4	10.8	19.7	1.2	18.0
110	12.6	5.1	5.6	17.2	2.3	11.3	20.5	1.2	18.7
120	12.1	4.9	5.5	16.2	2.2	10.9	19.7	1.2	18.0
130	10.6	4.7	4.9	13.9	2.1	9.6	16.8	1.2	15.3
140	7.9	4.4	3.8	10.1	2.1	7.0	12.3	1.1	11.7
150	4.9	4.1	2.4	6.1	2.2	4.1	7.4	1.1	7.1
160	1.7	3.6	0.9	2.0	2.3	1.3	2.4	0.8	2.9

Table 3.7.1 The numbers of events for the $t\bar{t}H$ signal, and $t\bar{t}\gamma\gamma$ background, assuming a $M_{\gamma\gamma}$ resolution of 3%; at the high luminosity LHC ($\mathcal{L} = 10^5\,\mathrm{pb}^{-1}$) and the standard luminosity SSC ($\mathcal{L} = 10^4\,\mathrm{pb}^{-1}$). We also give the significance, $\frac{N_{t\bar{t}H}}{\sqrt{N_{t\bar{t}\gamma\gamma}}}$, that this signal would have if it were observed.

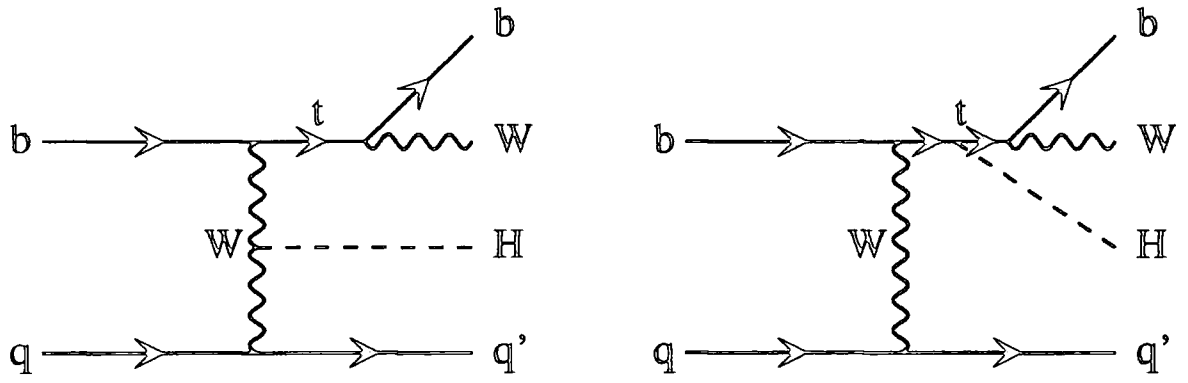


*Fig.*3.8.1 The Feynman diagrams for $qb \to tHq' \to bWHq'$.

In *Table* 3.7.3 I show the numbers of events that we expect to see at the high luminosity LHC ($\mathcal{L} = 10^5\,\mathrm{pb}^{-1}$), and the standard luminosity SSC ($\mathcal{L} = 10^4\,\mathrm{pb}^{-1}$). In order to have a reasonable number of events it is clear that the standard luminosity SSC is sufficient, whereas the LHC needs a larger luminosity, $\mathcal{L} = 5 \times 10^5\,\mathrm{pb}^{-1}$ will be sufficient. This is in contrast to WH associate production where both the LHC and the SSC required high luminosity to see the signal; the difference is that $t\bar{t}H$ production is dominated by the gluon-gluon initial state, where as WH production is dominated by the quark-antiquark initial state. Now because the SSC has larger $\sqrt{s}$ energy than the LHC to probe a certain $\sqrt{\hat{s}}$ we go to smaller Bjorken x at the SSC, and as we decrease x the gluon distribution of the proton grows quicker than the quark distribution.

3.8 tH Associate Production

High energy hadron colliders as well as producing numerous $t\bar{t}$ pairs also produce single t quarks at a smaller rate (for $m_t < 300\,\mathrm{GeV}$), due to Wb fusion. Now we can also produce a Higgs in association with this final state, the Feynman diagrams are shown in *Fig.*3.8.1. The t quark (or indeed the $\bar{t}$ antiquark) we can again detect in its semileptonic decay $t \to b\bar{\nu}e, \mu$ and the Higgs in its relatively clean two photon decay, $H \to \gamma\gamma$. Also we must apply similar cuts to $t\bar{t}H$ and WH production, *i.e.*, ,

$$\begin{aligned} p_\perp > 20\,\mathrm{GeV} \qquad & |\eta| < 2.5 \\ \Delta R > 0.4 \qquad & \text{where } \Delta R = \sqrt{\Delta\phi^2 + \Delta\eta^2} \quad . \end{aligned} \tag{3.8.1}$$

With these cuts we get the numbers of events show in *Table* 3.8.1, where we also show the numbers of events from $t\bar{t}H$ production and also WH production; clearly there are too few events to detect by themselves, although they do produce an extra *few* events that enhance the Higgs signal.

m_H	$m_t = 100\,\text{GeV}$			$m_t = 140\,\text{GeV}$			$m_t = 180\,\text{GeV}$		
(GeV)	$t\bar{t}H$	WH	$^{(-)}_{t}H$	$t\bar{t}H$	WH	$^{(-)}_{t}H$	$t\bar{t}H$	WH	$^{(-)}_{t}H$
	Number of events for $\sqrt{s} = 16\,\text{TeV}$ with $\mathcal{L} = 10^5\,\text{pb}^{-1}$								
60	18.9	27.6	0.4	20.9	27.6	0.9	19.8	27.6	1.3
70	21.3	32.9	0.6	24.7	32.9	1.3	24.2	32.9	1.7
80	22.6	36.2	0.9	27.1	36.2	1.7	27.4	36.2	2.0
90	23.2	38.2	1.2	29.0	38.2	2.2	30.3	38.2	2.3
100	23.6	39.2	1.5	29.9	39.2	2.7	32.1	39.2	2.7
110	23.3	38.8	1.8	29.8	38.8	3.3	32.7	38.8	3.2
120	21.8	35.7	2.2	27.7	35.7	3.8	30.9	35.7	3.4
130	18.4	29.4	2.1	22.9	29.4	3.7	26.0	29.4	3.1
140	13.5	20.8	1.7	16.3	20.8	3.2	18.7	20.8	2.6
150	8.2	12.0	1.2	10.0	12.0	2.2	10.9	12.0	1.7
160	2.8	3.8	0.5	3.1	3.8	0.8	3.5	3.8	0.7
	Number of events for $\sqrt{s} = 40\,\text{TeV}$ with $\mathcal{L} = 10^4\,\text{pb}^{-1}$								
60	8.9	5.8	0.3	11.0	5.8	0.5	11.5	5.8	0.7
70	10.3	6.9	0.4	13.2	6.9	0.9	14.3	6.9	1.0
80	11.1	7.7	0.5	14.8	7.7	1.1	16.4	7.7	1.2
90	11.8	8.2	0.8	16.1	8.2	1.4	18.4	8.2	1.5
100	12.3	8.5	1.0	16.8	8.5	1.8	19.7	8.5	1.8
110	12.6	8.5	1.2	17.2	8.5	2.2	20.5	8.5	2.2
120	12.1	8.0	1.4	16.2	8.0	2.6	19.7	8.0	2.5
130	10.6	6.6	1.4	13.9	6.6	2.6	16.8	6.6	2.3
140	7.9	4.7	1.3	10.1	4.7	2.2	12.3	4.7	2.0
150	4.9	2.8	0.8	6.1	2.8	1.6	7.4	2.8	1.4
160	1.7	0.9	0.3	2.0	0.9	0.6	2.4	0.9	0.5

Table 3.8.1 The numbers of events for the processes $pp \to t\bar{t}H$, $pp \to WH$, and $pp \to {}^{(-)}_{t}H$, detected in the isolated lepton and two photon mode, for the high luminosity LHC ($\mathcal{L} = 10^5\,\text{pb}^{-1}$) and the standard luminosity SSC ($\mathcal{L} = 10^4\,\text{pb}^{-1}$).

3.9 Conclusions

Higgs searches at Hadron colliders are vastly different to Higgs searches at e^+e^- colliders. This is due to far more difficult QCD dominated backgrounds. As such we are forced to look in far rarer decay channels that we would use at e^+e^- colliders. However hadron colliders typically have far larger $\sqrt{s}$ energies than e^+e^- machines and this means that far heavier Higgs can be produced than at e^+e^- machines.

At hadron colliders we have large numbers of Higgs produced through gluon fusion via a t quark loop,

$$gg \to H \quad . \tag{3.9.1}$$

For Higgs masses greater than $M_H > 2M_Z$ we have a clean signal in the,

$$H \to ZZ \to 4l \quad , \tag{3.9.2}$$

decay mode. For the standard luminosity SSC ($\mathcal{L} = 10^4\,\mathrm{pb}^{-1}$) and high luminosity LHC ($\mathcal{L} = 10^5\,\mathrm{pb}^{-1}$) we have numerous detectable events for Higgs masses up to where the Higgs width becomes so large that we can no longer talk about the Higgs boson as a *pseudo* stable particle.

For a Standard Model Higgs boson with $M_H < 2M_Z$ the experimental searches become harder. For the mass range,

$$130\,\mathrm{GeV} \lesssim M_H \lesssim 2M_Z \quad , \tag{3.9.3}$$

we can still look for the Higgs in its two Z to four lepton decay, however we now have one Z off mass shell,

$$H \rightarrow ZZ^* \rightarrow 4l \quad . \tag{3.9.4}$$

Again we require the high luminosity LHC with $\mathcal{L} = 10^5\,\mathrm{pb}^{-1}$, or the standard luminosity SSC with $\mathcal{L} = 10^4\,\mathrm{pb}^{-1}$. For Higgs masses lower than $M_H < 130\,\mathrm{GeV}$ the branching ratio $\mathrm{Br}(H \rightarrow Z^*Z^*)$ falls very quickly and we have no four lepton events to detect.

For such a Higgs with $M_H < 130\,\mathrm{GeV}$ then only other relatively clean decay channel is the decay,

$$H \rightarrow \gamma\gamma \quad , \tag{3.9.5}$$

although this decay gives us numerous events we have a large continuum background from two photon production,

$$gg, q\bar{q} \rightarrow \gamma\gamma \quad . \tag{3.9.6}$$

This places very strong constrains upon our experimental apparatus, in particular we need to be able to measure the $M_{\gamma\gamma}$ invariant mass to about 1% accuracy.

An alternative search strategy is to look for the Higgs boson produced in association with another heavy particle. This lowers the production rate by over an order of magnitude; however we now can tag on the presence of that heavy particle. This changes the character of the backgrounds tremendously and in particular lowers them considerably. This gives us far better signal to background ratios than we have for the plain $H \rightarrow \gamma\gamma$ detection mode, typically being larger than 1.

Two cases appear on first sight to give an observable signal,

$$q\bar{q}' \rightarrow WH \rightarrow l\nu\gamma\gamma \tag{3.9.7}$$

$$gg, q\bar{q} \rightarrow t\bar{t}H \rightarrow l\gamma\gamma X \quad , \tag{3.9.8}$$

in both we detect the heavy particle in its decay to a high $p_\perp$ lepton, and the Higgs in its $H \rightarrow \gamma\gamma$ decay mode.

This gives an observable signal at the LHC with $\mathcal{L} = 5 \times 10^4\,\mathrm{pb}^{-1}$ and the SSC with $\mathcal{L} = 10^4\,\mathrm{pb}^{-1}$ for a Higgs mass in the range,

$$60\,\mathrm{GeV} \lesssim M_H \lesssim 150\,\mathrm{GeV} \quad . \tag{3.9.9}$$

The work on these associate production mode of the Higgs is still at a very early stage, few of the processes have been put through full event simulations. Also in only a few cases have the NLO corrections been calculated. In particular the NLO contribution to,

$$pp \to W\gamma\gamma X \quad , \tag{3.9.10}$$

has not yet been calculated although there are strong hints that we expect the corrections to be very large.

This means that at the moment we should be cautious in our optimism in these associate production modes. There remain many questions to be answered both on the theoretical and experimental front.

Chapter 4

Conclusions

Before the advent of LEP I our bounds on the Standard Model Higgs boson were remarkably weak. Theoretically we knew that the spontaneously broken symmetry was probably the global minimum of the Higgs potential (there is a possibility that it is only a local minima – however when the universe was far younger and hotter than it is today we would almost certainly have been moved out of any local minima of the potential to the global minima) this means that the Higgs mass must be larger than some limit[30], M_{LW}. In the limit where we can ignore fermion masses this gives,

$$M_{LW}^2 = \frac{3\alpha M_W^2}{8\pi \sin^2\theta_W}(1+\frac{1}{2}\sec^4\theta_W) \simeq (7\,\mathrm{GeV})^2 \quad . \tag{4.1}$$

However if we include fermion masses then if $m_t = 78\,\mathrm{GeV}$ then this lower bound goes away and we can say nothing about the lower bound for the Higgs mass. For a heavier t quark then we can derive a similar bound for the vacuum to be stable [31] that is given approximately by,

$$M_H^2 > M_W^2 + \frac{1}{2}M_Z^2 - m_t^2 + \sqrt{\left(M_W^2 + \frac{1}{2}M_Z^2 - m_t^2\right)^2 + 4m_t^2 - 2M_W^4 - M_Z^4} \quad . \tag{4.2}$$

These bounds meant that theoretically a light Higgs boson was unfavoured, although there was till a window about $m_t = 80\,\mathrm{GeV}$ where the Higgs could be massless.

Experimentally we had some weak bounds from, for example, Upsilon decays,

$$\Upsilon, \Upsilon'' \to H\gamma \quad , \tag{4.3}$$

however these limits had very large theoretical uncertainties and also only gave limits comparable to the above theoretical M_{LW} bound.

However since LEP I came on line we have made great strides in our experimental limits on the Higgs mass. Mainly through non observation of the Bjorken process [8],

$$Z \rightarrow \mu^+\mu^- H \quad , \tag{4.4}$$

we have the current lower bound on the Higgs mass of [6]

$$M_H > 62.5\,\mathrm{GeV} \qquad \text{at the 95\% confidence level.} \tag{4.5}$$

This bound makes remarkably few demands on theoretical calculations, and is *largely* model independent. However the cross-section for this Bjorken process to occur drops very rapidly for $M_H > 60\,\mathrm{GeV}$ and this means that we will be unable to extend the Higgs search much further at LEP I.

At LEP II we still produce a reasonable number of events via ZH production,

$$e^+e^- \rightarrow ZH \quad , \tag{4.6}$$

up to the limit imposed by having any available phase space,

$$\sqrt{s} > M_H + 100\,\mathrm{GeV} \quad . \tag{4.7}$$

Now as the maximum $\sqrt{s}$ energy of LEP II is unlikely to be larger than 180 GeV we are only sensitive to a Higgs boson with mass less than,

$$M_H < 80\,\mathrm{GeV} \quad . \tag{4.8}$$

For a Higgs mass in this range we have a clean signal in the channel,

$$e^+e^- \rightarrow ZH \rightarrow l^+l^-, \nu\bar{\nu} + \text{jets} \quad . \tag{4.9}$$

For heavier Higgs masses we must, as long as no higher energy e^+e^- machine is built, rely upon the hadron super colliders: the LHC and the SSC. Both these machines produce numerous Higgs bosons through the process of gluon–gluon fusion,

$$pp \rightarrow ggX \rightarrow HX \quad . \tag{4.10}$$

for heavier Higgs with $M_H > 2M_Z$ we have a clean signal in the decay to four leptons via two Z's,

$$pp \rightarrow ggX \rightarrow HX \rightarrow ZZX \rightarrow 4lX \quad . \tag{4.11}$$

This process has both irreducible continuum backgrounds from,

$$pp \rightarrow gg, q\bar{q}X \rightarrow ZZX \rightarrow 4lX \quad , \tag{4.12}$$

and reducible backgrounds from processes like,

$$pp \to gg, q\bar{q}X \to t\bar{t}X \to W^+W^-b\bar{b}X \to 4lX \quad . \tag{4.13}$$

However with reasonable cuts these backgrounds can be easily overcome.

For a lighter Higgs with $2M_Z > M_H > 130\,\text{GeV}$ we can still look for the Higgs in its decay to two Z's, in this case with one Z virtual,

$$pp \to ggX \to HX \to ZZ^*X \to 4lX \quad . \tag{4.14}$$

this gives us the same backgrounds as before, and also

$$pp \to gg, q\bar{q}X \to Z\gamma^*X \to 4lX \quad . \tag{4.15}$$

The backgrounds are now harder to deal with due to the rapidly falling branching ratios $\text{Br}(H \to Z^*Z^*)$; however these can still be overcome with careful cuts.

The hardest Standard Model Higgs boson to detect is the so called light 'intermediate' mass Higgs with,

$$80\,\text{GeV} \lesssim M_H \lesssim 140\,\text{GeV} \quad , \tag{4.16}$$

being too heavy to be produced at LEP I but too light to be detected in the $H \to ZZ^*$ decay mode.

The best mode to detect the Higgs in isolation is in the decay,

$$H \to \gamma\gamma \quad . \tag{4.17}$$

However this has a tremendous continuum background from the processes,

$$pp \to gg, q\bar{q}X \to \gamma\gamma X \quad . \tag{4.18}$$

To overcome this background we need to measure the diphoton invariant mass to about the 1% level of accuracy, which will be very hard to achieve experimentally.

Another approach, that is proposed in this thesis, is to look for the Higgs in association with either a W boson or a t quark. This decreases our event rate, but vastly increases our signal to background ratio. In practice the easiest decay modes to detect are,

$$pp \to q\bar{q}X \to WHX \to l\nu\gamma\gamma X \tag{4.19}$$

$$pp \to gg, q\bar{q}X \to t\bar{t}HX \to l\gamma\gamma X \quad . \tag{4.20}$$

These detection modes put far less constraint on our diphoton mass measurement, 3% is enough to give an observable signal.

The study of these associate production modes is still in its infancy, and there still remains much work to be done; both on the part of theoreticians, in calculating all the backgrounds and calculating higher order corrections; and on the part of experimentalists, in asking how these events will actually look in our detectors.

All the above mentioned Higgs searches are based upon a Standard Model Higgs boson, if we move beyond the Standard Model then detection of Higgs bosons (that is particles responsible for generating vector boson and fermion masses) typically becomes far harder.

At LEP we are only sensitive to Higgs bosons that generate the mass of the Z boson (which implies that there is a ZZH coupling). At hadron colliders the story is far more complicated. As well as the gluon fusion via a t quark loop production mechanism we also can produce Higgs at a lower rate from the mechanisms,

$$W^+W^-, ZZ \rightarrow H \quad . \tag{4.21}$$

This means that we will produce Higgs bosons whether they couple to vector bosons or fermions. However at hadron colliders we must also consider the detection mechanism for these Higgs. The decay to ZZ only occurs for Higgs that generate vector boson mass, similarly the decay $H \rightarrow \gamma\gamma$ is vastly suppressed for Higgs that only couple to fermions. This means that Higgs that predominantly generate fermion masses will be very hard to detect at hadron colliders, even though the machines will produce large numbers of them.

Appendix A

From Classical Mechanics to Quantum Field Theory

Pooh Looked after her as she went.

'I wish I could jump like that,' he thought. 'Some can and some can't. That's how it is.'

But there were moments when Piglet wished that Kanga couldn't. Often, when he had had a long walk home through the Forest, he had wished that he were a bird; but now he thought jerkily to himself at the bottom of Kanga's pocket,

'If this is shall really take flying I never to it.'

Winnie–the–Pooh, A.A.Milne

A.1 Introduction

In this appendix I review the progression from classical mechanics to quantum mechanics, and from classical field theory to quantum field theory. This is heavily based upon material from the books [32, 3].

A.2 Classical Mechanics

If we consider a classical system of particles that are described by n generalized coordinates, q_i, (*e.g.*, x, y and z coordinates, or angular coordinates, or indeed any other set of

position variables that we choose) then the equations of motion can be conveniently written down in Lagrange's equation,

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{q}_i}\right) - \frac{\partial L}{\partial q_i} = 0 \quad , \tag{A.2.1}$$

where $\dot{q}_i = \frac{dq}{dt}$ and the Lagrangian, L, is defined as the difference in kinetic and potential energies,

$$L = T - V \quad . \tag{A.2.2}$$

An alternative formalism, that is equivalent, is Hamilton's principle, there the action, I, is defined as,

$$I = \int_{t_1}^{t_2} L dt \quad , \tag{A.2.3}$$

and then the equations of motion can be stated as,

$$\delta I = 0 \quad . \tag{A.2.4}$$

Now we can also define conjugate momenta for the q_i coordinates,

$$p_i = \frac{\partial L}{\partial \dot{q}_i} \quad , \tag{A.2.5}$$

and from this form the Hamiltonian,

$$H(p,q) = \dot{q}_i p_i - L \quad , \tag{A.2.6}$$

and in this form the Hamiltonian equations of motion take the form,

$$\dot{q}_i = \frac{\partial H}{\partial p_i} \qquad \dot{p}_i = -\frac{\partial H}{\partial q_i} \quad , \tag{A.2.7}$$

which is puts both the coordinates and momenta on an equal footing. Now for a general function f we can find its time derivative from,

$$\frac{df}{dt} = \frac{\partial f}{\partial t} + \sum_i \frac{\partial H}{\partial p_i}\frac{\partial f}{\partial q_i} - \frac{\partial H}{\partial q_i}\frac{\partial f}{\partial p_i} = \frac{\partial f}{\partial t} + \{H, f\} \quad , \tag{A.2.8}$$

where we have introduced the Poisson bracket notation,

$$\{u, v\} = \sum_i \frac{\partial u}{\partial p_i}\frac{\partial v}{\partial q_i} - \frac{\partial u}{\partial q_i}\frac{\partial v}{\partial p_i} \quad . \tag{A.2.9}$$

The Possion brackets of the generalized coordinates and momenta are trivially,

$$\begin{aligned} \{q_j, q_k\} &= 0 = \{p_j, p_k\} \\ \{p_j, q_k\} &= \delta_{jk} = -\{q_j, p_k\} \end{aligned} \quad . \tag{A.2.10}$$

Also the Hamiltonian equations of motion can be written as,

$$\dot{q}_i = \{q_i, H\} \qquad \dot{p}_i = \{p_i, H\} \quad . \tag{A.2.11}$$

If the Lagrangian is invariant under some symmetry group then Noether's theorem tells us that this system has some conserved quantity. For example if the Lagrangian, L, is invariant under,

$$\phi \rightarrow \phi' = \phi + \delta\phi \quad , \tag{A.2.12}$$

then under this transformation the Lagrangian is modified by,

$$\delta L = \frac{\delta L}{\delta \phi}\delta\phi + \frac{\delta L}{\delta(\partial_\mu \phi)}\delta(\partial_\mu \phi) \quad , \tag{A.2.13}$$

and after a bit of manipulation we get,

$$\delta L = \partial_\mu \left[\frac{\delta L}{\delta(\partial_\mu \phi)}\delta\phi\right] \quad , \tag{A.2.14}$$

and if L is unchanged under this transformation we have $\delta L = 0$, and so we have,

$$\partial^\mu J_\mu = 0 \quad , \tag{A.2.15}$$

where,

$$J_\mu = \frac{\delta L}{\delta(\partial_\mu \phi)}\phi \quad , \tag{A.2.16}$$

and we have a conserved charge Q given by,

$$Q = \int d^3x J_0 \quad . \tag{A.2.17}$$

A.3 Quantum Mechanics

If we now wish to quantize this classical system of particles then first we interpret the generalized coordinates and momenta as operators, rather than numbers, and then apply the following substitution,

$$\{A, B\} \rightarrow \frac{1}{i\hbar}[A, B] \quad , \tag{A.3.1}$$

where $[A, B] = AB - BA$ is the commutator of the operators A and B this means that the Possion brackets of momenta and coordinates (A.2.10) becomes,

$$\begin{aligned} [q_j, q_k] &= 0 = [p_j, p_k] \\ [p_j, q_k] &= \delta_{jk} i\hbar = -[q_j, p_k] \end{aligned} \quad . \tag{A.3.2}$$

Now if the generalized coordinates and momenta are operators then they must act on some function, $|\phi\rangle$, this function that they act on is the wave function, or the state vector of the quantum state. Now we define the conjugate wavefunction, $\langle\phi|$ as,

$$\langle\phi| = |\phi\rangle^\dagger \quad , \tag{A.3.3}$$

and insist that the normalisation of this state vector is,

$$\int_{\mathbb{R}^3} \langle\phi|\phi\rangle dx^3 = 1 \quad . \tag{A.3.4}$$

Also we must ask what it means to experimentally measure the value of a coordinate or momenta, it is usually taken to have the *mean* value of the operator, defined by,

$$\bar{f} = \int_{\mathbb{R}^3} \langle\phi|f|\phi\rangle dx^3 \quad . \tag{A.3.5}$$

Now the evolution of this quantum state with time is, in the same way as the classical theory, given by,

$$i\hbar \frac{d\langle\phi|A|\phi\rangle}{dt} = \frac{\partial\langle\phi|A|\phi\rangle}{\partial t} + \langle\phi|[H, A]|\phi\rangle \quad . \tag{A.3.6}$$

Now it is clear that any physical measurement (A.3.5) is invariant under the unitary transformation,

$$\begin{array}{c} |\phi\rangle \rightarrow U|\phi\rangle \\ \langle\phi| \rightarrow \langle\phi|U^\dagger \\ A \rightarrow UAU^\dagger \end{array} \quad , \tag{A.3.7}$$

where U is any unitary operator (*i.e.*, $UU^\dagger = 1$). This means that we can choose whether the time dependence of our quantum state resides in the state vectors or the quantum operators. If we choose all the time dependence to lie in the quantum operators then the operator evolution is given by,

$$i\hbar \frac{dA}{dt} = \frac{\partial A}{\partial t} + [H, A] \quad . \tag{A.3.8}$$

This is called the Heisenberg picture. Alternately we can choose all the time dependence of the quantum state to resides in the state vectors, $|\phi\rangle_S$, and this is called the Schrodinger picture. So state vectors in the Heisenberg and Schrodinger pictures are related by,

$$|\phi\rangle_H = e^{iHt/\hbar}|\phi\rangle_S \quad . \tag{A.3.9}$$

Often we find the the Hamiltonian for a system can be written as a sum of two parts,

$$H = H_0 + H_I \quad , \tag{A.3.10}$$

where the equations of motion can be solved exactly for $H = H_0$ and H_I gives a small perturbation from this exactly solvable state. In this case it is convenient to work in the Dirac Interaction picture defined by,

$$|\phi\rangle_D = e^{iH_0t/\hbar}|\phi\rangle_S \quad , \tag{A.3.11}$$

where we have some time dependence in the state vectors, and some in the quantum operators. This picture is useful if we look at the evolution of a quantum state from time t_0 to time t, we find,

$$|\phi, t\rangle = \frac{1}{i\hbar}\int_{t_0}^{t} dt' H_I|\phi, t'\rangle + |\phi, t_0\rangle \quad . \tag{A.3.12}$$

Now this equation can be iteratively substituted into itself to give,

$$\begin{aligned} |\phi, t\rangle &= S(t, t_0)|\phi, t_0\rangle \\ &= \left(1 + \frac{1}{i\hbar}\int_{t_0}^{t} dt' H_I|\phi, t'\rangle + \frac{1}{(i\hbar)^2}\int_{t_0}^{t} dt' \int_{t_0}^{t'} dt'' H_I|\phi, t'\rangle H_I|\phi, t''\rangle + \ldots\right)|\phi, t_0\rangle \ , \end{aligned} \tag{A.3.13}$$

and in this form we can find the scattering matrix S to go from some initial state $|\phi, t_0\rangle$ to some final state $|\phi, t\rangle$ in perturbative expansion in terms of H_I.

Usually, due to the number of c's and $\hbar$'s that appear in equations, it it easiest to work in natural units where we define,

$$c = 1 = \hbar \quad , \tag{A.3.14}$$

and this is what I shall do for the rest of this thesis.

A.4 Field Theory

If we now consider a classical field theory we can again form the action,

$$S = \int d^4x \mathcal{L}(\phi_i, \partial_\mu \phi_i) \quad , \tag{A.4.1}$$

where $\mathcal{L}$ is the Lagrangian density, the kinetic minus the potential energy of the field per unit volume, and again the dynamics of this field is given by,

$$\delta S = 0 \quad . \tag{A.4.2}$$

Again this leads to the Euler–Lagrange equation of motion,

$$\partial_\mu \frac{\delta \mathcal{L}}{\delta(\partial_\mu \phi_i)} - \frac{\delta \mathcal{L}}{\delta \phi_i} = 0 \quad . \tag{A.4.3}$$

As with classical mechanics we can define conjugate momenta by,

$$\pi_i = \frac{\delta \mathcal{L}}{\delta(\partial_0 \phi_i)} \quad . \tag{A.4.4}$$

We quantize this system using the equal time canonical commutation relations,

$$\begin{aligned} [\phi_i(x, t), \pi_j(x', t)] &= i\delta_{ij}\delta^3(x - x') \\ [\phi_i(x, t), \phi_j(x', t)] &= 0 = [\pi_i(x, t), \pi_j(x', t)] \quad . \end{aligned} \tag{A.4.5}$$

As before the Hamiltonian,

$$H = \int d^3x[\pi(x)\partial_o\phi(x) - \mathcal{L}(x)] \quad , \tag{A.4.6}$$

governs the dynamics of this quantum field theory through,

$$\begin{aligned} \partial_0\phi_i &= i[H, \phi_i] \\ \partial_0\pi_i &= i[H, \pi_i] \end{aligned} \quad . \tag{A.4.7}$$

If we now consider the example Lagrangian density,

$$\mathcal{L} = \frac{1}{2}(\partial_\lambda\phi\partial^\lambda\phi - m^2\phi^2) \quad , \tag{A.4.8}$$

this leads, via the Euler–Lagrange equation to the Klein–Gordon equation of motion,

$$(\partial^2 + m^2)\phi = 0 \quad . \tag{A.4.9}$$

Quantizing this field we obtain,

$$\begin{aligned} \phi &= \phi^+ + \phi^- \\ &= \int \frac{d^3k}{((2\pi)^3 2\omega_k)^{1/2}}[a(k)e^{i(k.x-\omega_k t)} + a^\dagger(k)e^{-i(k.x-\omega_k t)}] \end{aligned} \quad , \tag{A.4.10}$$

where $\omega_k = (k^2 + m^2)^{1/2}$. The values of $a(k)$ and $a^\dagger(k)$ are given by,

$$\begin{aligned} [a(k), a^\dagger(k')] &= \delta^3(k - k') \\ [a(k), a(k')] = 0 &= [a^\dagger(k), a^\dagger(k')] \end{aligned} \quad , \tag{A.4.11}$$

and the Hamiltonian is given by,

$$\begin{aligned} H &= \int d^3k\omega_k \frac{1}{2}(a^\dagger(k)a(k) + a(k)a^\dagger(k)) \\ &= \int d^3k\omega_k(a^\dagger(k)a(k) + \frac{1}{2}) \end{aligned} \quad . \tag{A.4.12}$$

Now this factor of $\frac{1}{2}$ in the last equation is somewhat of a problem, because integrated over k space this gives an infinite contribution, however it doesn't affect any physics (A.4.7), as this value of $\frac{1}{2}$ commutes with everything. As a result this factor is discarded, this is imposed by using normal ordering where we always write the $a(k)$ terms to the right of $a^\dagger(k)$ terms, we denote this using :: notation, that is $: A :$ is the operator A normal ordered so the $a(k)$ terms are to the right of $a^\dagger(k)$ terms.

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

$$a(k)|0\rangle = 0 \quad , \tag{A.4.13}$$

then we can interpret $|1\rangle = a^\dagger(k)|0\rangle$ as being the state with a single particle of momenta k, and $|2\rangle = \frac{1}{\sqrt{2}}[a^\dagger(k)]^2|0\rangle$ being the state with two particles of momenta k in it, *etc.*. With this we find that,

$$\langle n|a^\dagger a|n\rangle = n\langle n|n\rangle \quad , \tag{A.4.14}$$

and so $a^\dagger(k)a(k)$ counts the number of quanta with momenta k. This also means that the normal ordered Hamiltonian gives us the total energy of the system as we expect.

If we now consider the *time-ordered* product of two fields,

$$T\{\phi(x)\phi(x')\} = \begin{cases} \phi(x)\phi(x') & \text{if } t > t' \\ \phi(x')\phi(x) & \text{if } t' > t \end{cases} \quad , \tag{A.4.15}$$

then the object

$$\langle 0|T\{\phi(x)\phi(x')\}|0\rangle \quad , \tag{A.4.16}$$

represents a particle being created at x and then propagating to x' where it is annihilated, that is if $t > t'$, and for $t' > t$ a particle propagating in the opposite direction. If we evaluate this we find,

$$\begin{aligned} \langle 0|T\{\phi(x)\phi(x')\}|0\rangle &= i\Delta_F(x - x') \\ &= \frac{i}{(2\pi)^4}\int \frac{d^4 e^{-ikx}}{k^2 - m^2} \quad . \end{aligned} \tag{A.4.17}$$

If rather than looking at this propagator in coordinate space we look at this in momentum space, we find,

$$i\Delta_F(k) = \frac{i}{k^2 - m^2} \quad , \tag{A.4.18}$$

and the propagator takes a particularly easy form.

If we move on to consider the Dirac equation,

$$i\gamma^\mu \frac{\partial \psi}{\partial x^\mu} - m\psi = 0 \quad , \tag{A.4.19}$$

for this to agree with the Klein–Gordon equation we require that,

$$[\gamma^\mu, \gamma^\nu]_+ = 2g^{\mu\nu} \quad , \tag{A.4.20}$$

where $[A, B]_+ = AB + BA$. If we also require Hermiticity we find,

$$\gamma^{\mu\dagger} = \gamma^0\gamma^\mu\gamma^0 \quad . \tag{A.4.21}$$

Defining $\bar{\psi} = \psi^\dagger\gamma^0$ we find the Lagrangian density,

$$\mathcal{L} = \bar{\psi}\left(i\gamma^\mu \frac{\partial}{\partial x^\mu} - m\right)\psi \quad , \tag{A.4.22}$$

unfortunately this does not make clear that ψ and $\bar{\psi}$ are on an equal footing; to do this we rewrite the Lagrangian density as,

$$\mathcal{L} = \bar{\psi}(i\gamma^\mu \overleftrightarrow{\partial}_\mu - m)\psi \quad , \tag{A.4.23}$$

where,

$$\bar{\psi}\overleftrightarrow{\partial}_\mu\psi = \frac{1}{2}\left(\bar{\psi}\frac{\partial\psi}{\partial x^\mu} - \frac{\partial\bar{\psi}}{\partial x^\mu}\psi\right) \quad . \tag{A.4.24}$$

If we now go on to quantize this in the usual manner we find that the Hamiltonian is given by,

$$H = \sum E(N(p) - \bar{N}(p)) \quad , \tag{A.4.25}$$

where $N(p)$ gives the number of ψ quanta with momenta p, and $\bar{N}(p)$ gives the number of $\bar{\psi}$ quanta with momenta p. Unfortunately this Hamiltonian is not bounded below as $\bar{N}(p)$ can become arbitrarily large, and this is physically unreasonable.

Instead we quantize the Dirac equation according to anticommutation relations, rather than commutation relations, where we replace commutators $[\,,\,]$ with anticommutators $[\,,\,]_+$. This leads to a solution of the form,

$$\begin{aligned}\psi &= \psi^+ + \psi^- \\ &= \sum \frac{1}{E_p^{1/2}}\left(c_r(p)u_r(p)e^{-ipx} + d_r^\dagger(p)v_r(p)e^{ipx}\right) \quad ,\end{aligned} \tag{A.4.26}$$

and

$$\begin{aligned}\bar{\psi} &= \bar{\psi}^+ + \bar{\psi}^- \\ &= \sum \frac{1}{E_p^{1/2}}\left(d_r(p)\bar{v}_r(p)e^{-ipx} + c_r^\dagger(p)\bar{u}_r(p)e^{ipx}\right) \quad ,\end{aligned} \tag{A.4.27}$$

where we have,

$$[c_r(p), c_s^\dagger(p')]_+ = \delta_{rs}\delta_{pp'} = [d_r(p), d_s^\dagger(p')]_+ \quad , \tag{A.4.28}$$

and all other anticommutators of c and d vanish. With these plane wave solutions the Dirac equation becomes,

$$\begin{aligned}(\not{p} - m)u(p) &= 0 \qquad (\not{p} + m)v(p) = 0 \\ \bar{u}(p)(\not{p} - m) &= 0 \qquad \bar{v}(p)(\not{p} + m) = 0 \quad .\end{aligned} \tag{A.4.29}$$

Proceeding as before we find the propagator for this Dirac field to be,

$$iS_F(p) = \frac{i(\not{p} + m)}{p^2 - m^2} \quad . \tag{A.4.30}$$

Here we have made use of the *slash* notation where we define,

$$\not{p} = \gamma^\mu p_\mu \quad . \tag{A.4.31}$$

Appendix B

Gauge Theories and Spontaneous Symmetry Breaking

Pooh knew what he meant, but being a Bear of Very Little Brain, couldn't think of the words.

The House At Pooh Corner, A.A.Milne

B.1 Introduction

In this appendix I extend the ideas of the previous appendix and introduce the important ideas of gauge theories, and spontaneous symmetry breaking. This is very strongly based upon the books listed in Ref. 3.

B.2 Gauge Theories

In the previous appendix we saw that the Lagrangian density for the Dirac equation, which describes electrons, is given by,

$$\mathcal{L} = \bar{\psi}(x)(i\partial\!\!\!/ - m)\psi(x) \quad . \tag{B.2.1}$$

Now this Lagrangian is invariant under the $U(1)$ symmetry,

$$\begin{aligned} \psi(x) &\rightarrow \psi'(x) = e^{-i\alpha}\psi(x) \\ \bar{\psi}(x) &\rightarrow \bar{\psi}'(x) = e^{i\alpha}\bar{\psi}(x) \quad , \end{aligned} \tag{B.2.2}$$

and because of Noether's theorem this means that we have a conserved quantity, in this case electric charge.

We now ask what happens if we change this symmetry from a global symmetry to a local symmetry.

$$\begin{aligned} \psi(x) &\rightarrow \psi'(x) = e^{-i\alpha(x)}\psi(x) \\ \bar{\psi}(x) &\rightarrow \bar{\psi}'(x) = e^{i\alpha(x)}\bar{\psi}(x) \end{aligned} \quad . \tag{B.2.3}$$

Now the $m\bar{\psi}\psi$ term is invariant under this local symmetry however the derivative term does not transform so simply, we have

$$\begin{aligned} \bar{\psi}\partial_\mu\psi \rightarrow \bar{\psi}'\partial_\mu\psi' &= \bar{\psi}e^{i\alpha}\partial_\mu(e^{-i\alpha}\psi) \\ &= \bar{\psi}\partial_\mu\psi - i\bar{\psi}\partial_\mu\alpha\psi \end{aligned} \quad , \tag{B.2.4}$$

and this spoils the symmetry. However if we replace ∂_μ with D_μ, where,

$$D_\mu = \partial_\mu + ieA_\mu \quad , \tag{B.2.5}$$

then $D_\mu\psi$ transforms as,

$$D_\mu\psi \rightarrow (D_\mu\psi)' = e^{-i\alpha(x)}D_\mu\psi(x) \quad , \tag{B.2.6}$$

as long as $A_\mu(x)$ transforms as,

$$A_\mu(x) \rightarrow A'_\mu(x) = A_\mu(x) + \frac{1}{e}\partial_\mu\alpha(x) \quad . \tag{B.2.7}$$

At this stage our Lagrangian has become,

$$\mathcal{L} = \bar{\psi}(i\gamma^\mu(\partial_\mu + ieA_\mu) - m)\psi \quad , \tag{B.2.8}$$

and we seem to have some interaction term between the ψ,$\bar{\psi}$ fields, and some A_μ field. However at the moment we have no kinetic term for the A_μ field. We wish to add a kinetic term that is also gauge invariant, that is invariant under the change of the local symmetry. The easiest way to do this is to consider,

$$(D_\mu D_\nu - D_\nu D_\mu)\psi = ieF_{\mu\nu}\psi \quad , \tag{B.2.9}$$

where $F_{\mu\nu}$ is invariant under the gauge transformation, and is given by,

$$F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu \quad . \tag{B.2.10}$$

So we construct a gauge invariant kinetic term for the A_μ field as $\mathcal{L}_A = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu}$, and the full Lagrangian becomes,

$$\mathcal{L} = \bar{\psi}(i\gamma^\mu D_\mu - m)\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} \quad . \tag{B.2.11}$$

Notice that the A^μ field is massless because there is no $A_\mu A^\mu$ term, and indeed this field is forced to be massless because such a $A_\mu A^\mu$ term is not gauge invariant.

Unfortunately the A^μ field in (B.2.11) can not be quantized, because its conjugate momenta vanishes identically, and so the equal time canonical commutation relations can not be applied. This is solved by adding a *gauge fixing* term to the Lagrangian, *e.g.*, $-\lambda/2(\partial_\mu A^\mu)^2$, and so we end up with,

$$\mathcal{L} = \bar{\psi}(i\gamma^\mu D_\mu - m)\psi - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} - \lambda/2(\partial_\mu A^\mu)^2 \quad , \tag{B.2.12}$$

and this Lagrangian can be quantized. Now we should ask how this extra gauge fixing term changes our theory ? Well in the original Lagrangian we have a gauge symmetry under which the Lagrangian, and hence the physics, is unchanged; now if we fix this gauge and any calculation that do is unaltered by this, in particular if we choose a gauge under which the gauge fixing term is identically zero then this gauge fixing term can play no part in the physics of the Lagrangian. This means that we need to choose,

$$\partial_\mu A^\mu = 0 \quad . \tag{B.2.13}$$

However again we can not do this because the equal time commutation relations tell us that this is not zero in the quantum theory. The solution to this quandary was put forward by Gupta and Bleuler; they noted that all we require is that for all physical observable $\partial_\mu A^\mu = 0$; or in particular that,

$$\langle\psi|\partial_\mu A^\mu|\psi\rangle = 0 \quad , \tag{B.2.14}$$

and we can arrange this if we insist that,

$$\partial_\mu A^{\mu+}|\psi\rangle = 0 \quad , \tag{B.2.15}$$

or that all our $a(k)$ terms annihilate the vacuum.

However in applying a gauge fixing term we have introduced an extra parameter into the theory in the form of λ, now λ has no physical significance, and so should not change any physical quantities that we calculate, and this provides us with a very strong test of any calculation that we do – that the answer should be independent of λ.

B.2.1 Larger Gauge Groups

If the fermion field associated with the Dirac equation comes in several different versions, *i.e.*, the fermion field has several components,

$$\Psi = \begin{pmatrix} \psi_1 \\ \vdots \\ \psi_n \end{pmatrix} \quad , \tag{B.2.16}$$

then we can transform this field under an $SU(n)$ transformation,

$$\Psi(x) \rightarrow \Psi'(x) = e^{-iT^a\theta_a}\Psi(x) \quad , \tag{B.2.17}$$

where θ_a and the Euler angles that parametrise the $SU(n)$ rotation and the T_a are the basis elements of $SU(n)$ under the representation of Ψ. Then in order for the Dirac equation to be invariant under a local transformation we rewrite ∂_μ as D_μ where,

$$D_\mu\Psi = (\partial_\mu - igT^aA^a_\mu)\Psi \quad . \tag{B.2.18}$$

Now the representation matrices T^a obey the algebra,

$$[T^a, T^b] = if^{ab}_c T^c \quad . \tag{B.2.19}$$

The second-rank tensor for gauge fields is given by,

$$F^a_{\mu\nu} = \partial_\mu A^a_\nu - \partial_\nu A^a_\mu + gf^a_{bc}A^b_\mu A^c_\nu \quad , \tag{B.2.20}$$

and under an infinitesimal variation, $A^a_\mu(x)$ transforms as,

$$A^a_\mu(x) \rightarrow A^{a\prime}_\mu(x) = A^a_\mu + f^a_{bc}\theta^b(x)A^c_\mu(x) - \frac{1}{g}\partial_\mu\theta^a(x) \quad . \tag{B.2.21}$$

When the underlying group is larger than $U(1)$ then $f^a_{bc} \not\equiv 0$ and so the gauge field kinetic term contains A^3 and A^4 terms. This means that there are both 3 and 4 point vertices, and that the gauge field is itself charged and interacts with itself.

B.3 Spontaneous Symmetry breaking

For a theory with a gauge symmetry the gauge field is forced to be massless because a mass term of the form $mA_\mu A^\mu$ is not gauge invariant, however there is a clever mechanism, called the Higgs mechanism, whereby we can generate a mass term for the gauge field, while still retaining the nice properties which gauge theories have. If we consider an Abelian $U(1)$ gauge theory with a complex scalar field we have the Lagrangian,

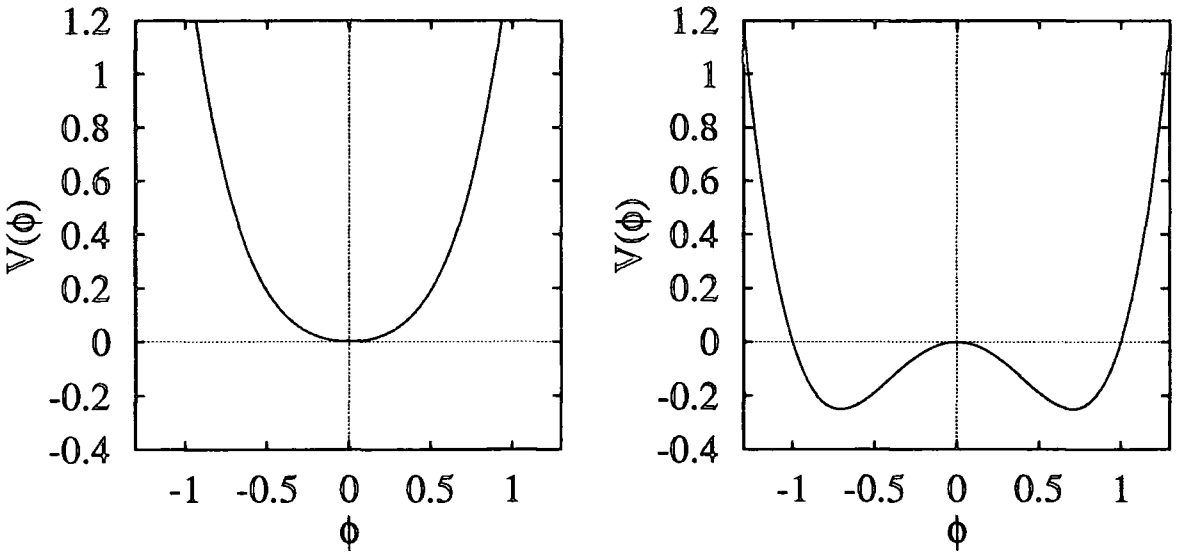
$$\mathcal{L} = (D_\mu\phi)^\dagger(D^\mu\phi) + \mu^2\phi^\dagger\phi - \lambda(\phi^\dagger\phi)^2 - \frac{1}{4}F_{\mu\nu}F^{\mu\nu} \quad , \tag{B.3.1}$$

where,

$$\begin{aligned} D_\mu\phi &= (\partial_\mu - igA_\mu)\phi \\ F_{\mu\nu} &= \partial_\mu A_\nu - \partial_\nu A_\mu \quad . \end{aligned} \tag{B.3.2}$$

Now we are used to having $\mu^2 < 0$ and so the minimum of the potential,

$$V(\phi) = -\mu^2\phi^\dagger\phi + \lambda(\phi^\dagger\phi)^2 \quad , \tag{B.3.3}$$



a) with $\mu^2 < 0$ b) with $\mu^2 > 0$

Fig.B.3.1 The potential $V(\phi) = -\mu^2\phi^\dagger\phi + \lambda(\phi^\dagger\phi)^2$ for the two cases a) $\mu^2 < 0$ where the minimum occurs at $\phi = 0$ and b) $\mu^2 > 0$ where the minimum occurs at $\phi = v/\sqrt{2}$ for $v = (\mu^2/\lambda)^{1/2}$

is at $\phi = 0$, this is shown in *Fig*.B.3.1.a, in this case the vacuum state has $\phi = 0$. However if we choose $\mu^2 > 0$ then the minimum of the potential occurs at $|\phi| = v/\sqrt{2}$ where $v = (\mu^2/\lambda)^{1/2}$ and this means that the vacuum has a non zero expectation value, *i.e.*,

$$|\langle 0|\phi|0\rangle| = v/\sqrt{2} \quad . \tag{B.3.4}$$

This is shown in *Fig*.B.3.1.b.

Now there are a ring of minima for the potential, all at $|\phi| = v/\sqrt{2}$. *How the the vacuum decide in which minima to lie ?* well clearly it doesn't, it just lies in a minima at random. So although the potential is invariant under a change of phase in ϕ the vacuum state is clearly not. This is called a spontaneous breaking of the symmetry. If we choose the vacuum state to be purely real, and write ϕ in terms of real fields ϕ_1 and ϕ_2,

$$\phi = \frac{1}{\sqrt{2}}(\phi_1 + i\phi_2) \quad , \tag{B.3.5}$$

with,

$$\langle 0|\phi_1|0\rangle = v \qquad \text{and} \qquad \langle 0|\phi_2|0\rangle = 0 \quad . \tag{B.3.6}$$

Now we wish to expand about the vacuum, and so use the shifted fields,

$$\phi_1' = \phi_1 - v \qquad \text{and} \qquad \phi_2' = \phi_2 \quad . \tag{B.3.7}$$

In terms of these transformed fields the kinetic ϕ term becomes

$$(D_\mu\phi)^\dagger D^\mu\phi = \frac{1}{2}(\partial_\mu\phi_1' + gA_\mu\phi_2')^2 + \frac{1}{2}(\partial_\mu\phi_2' - gA_\mu\phi_1')^2 - gvA^\mu(\partial_\mu\phi_2' + \partial_\mu\phi_1') + \frac{g^2v^2}{2}A^\mu A_\mu \, , \tag{B.3.8}$$

and this last term, $\frac{1}{2}g^2v^2A^\mu A_\mu$, can be interpreted as a mass term for the quanta of the A^μ field, and these quanta will have mass $M = gv$.

Appendix C

The Matrix Element

Christopher Robin came down from the Forest to the bridge, feeling all sunny and careless, and just as if twice nineteen didn't matter a bit, as it didn't on such a happy afternoon, and he though that if he stood an the bottom rail of the bridge, and leant over, and watch the river slipping slowly away beneath him, then he would suddenly know everything there was to know, and he would be able to tell Pooh, who wasn't quite sure about some of it. But when he got to the bridge and saw all the animals there, then he knew that it wasn't that kind of afternoon, but the other kind, when you wanted to do something.

The House At Pooh Corner, A.A.Milne

C.1 Introduction

In this appendix I will go over the methods used for calculating the matrix element squared for any particular process, that is the probability to get a particular final state from a particular initial state, per unit volume of the final state particles. This appendix is largely based upon the methods of Ref. 33,34, the analytic continuation to negative energy spinors and the matrix notation are my own work.

If we consider a process, say $e^+e^- \rightarrow Z \rightarrow \mu^+\mu^-$ the first step is to write down all Feynman diagrams that contribute to this process, for this process at lowest order there is only the 1 diagram.

Then we use Feynman rules to write down the matrix element, for the electron current we get,

$$J_e^\alpha = \bar{v}(2) g_Z \gamma^\alpha (g_{Ve} - g_{Ae}\gamma_5) u(1) \quad , \tag{C.1.1}$$

where $g_Z = \frac{ig}{4\cos\theta_W}$; and for the muon current,

$$J_\mu^\nu = \bar{u}(4) g_Z \gamma^\nu (g_{V\mu} - g_{A\mu}\gamma_5) v(3) \quad , \tag{C.1.2}$$

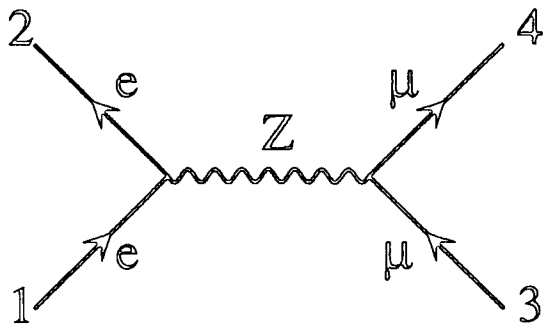


*Fig.*C.1.1 The process $e^+e^- \rightarrow Z \rightarrow \mu^+\mu^-$

and for the Z propagator,

$$iD_{\alpha\nu} = \frac{i(-g_{\alpha\nu} + k_\alpha k_\nu/m_Z^2)}{k^2 - m_Z^2 + im_Z\Gamma_Z} \quad , \tag{C.1.3}$$

giving the matrix element,

$$\mathcal{M} = J_e^\alpha iD_{\alpha\nu}J_\mu^\nu \quad . \tag{C.1.4}$$

Now in this form the matrix element is not usable, traditionally it has been converted to a usable form via Traceology.

C.2 Traceology

In the end we are always going to have to square the matrix element, in Traceology we do this at the outset, this gives,

$$|\mathcal{M}|^2 = J_e^\alpha iD_{\alpha\nu}J_\mu^\nu J_e^{*\beta}(-i)D_{\beta\mu}^* J_\mu^{*\mu} \quad . \tag{C.2.1}$$

Now concentrating on the $J_\mu^\nu J_\mu^{*\mu}$ term,

$$J_\mu^\nu J_\mu^{*\mu} = J_\mu^\nu J_\mu^{\dagger\mu} \tag{C.2.2}$$

$$= |g_Z|^2\bar{u}(4)\gamma^\nu(g_{V\mu} - g_{A\mu}\gamma_5)v(3)\bar{v}(3)\gamma^\nu(g_{V\mu} - g_{A\mu}\gamma_5)u(4) \tag{C.2.3}$$

$$= |g_Z|^2\mathrm{Tr}\left[u(4)\bar{u}(4)\gamma^\nu(g_{V\mu} - g_{A\mu}\gamma_5)v(3)\bar{v}(3)\gamma^\nu(g_{V\mu} - g_{A\mu}\gamma_5)\right] . \tag{C.2.4}$$

Now summing over the chiralities of the fermion lines,

$$\sum J_\mu^\nu J_\mu^{*\mu} = |g_Z|^2\mathrm{Tr}\left[(\not{p}_4 + m)\gamma^\nu(g_{V\mu} - g_{A\mu}\gamma_5)(\not{p}_3 - m)\gamma^\mu(g_{V\mu} - g_{A\mu}\gamma_5)\right] \tag{C.2.5}$$

$$= |g_Z|^2\mathrm{Tr}\Big[(g_{V\mu} - g_{A\mu}\gamma_5)((g_{V\mu} - g_{A\mu}\gamma_5)\not{p}_4 + (g_{V\mu} + g_{A\mu}\gamma_5)m) \gamma^\nu(\not{p}_3 - m)\gamma^\mu] \tag{C.2.6}$$

$$= |g_Z|^2\Big(\mathrm{Tr}\left[(g_{V\mu}^2 + g_{A\mu}^2)\not{p}_4\gamma^\nu(\not{p}_3 - m)\gamma^\mu\right]$$

$$+ \mathrm{Tr}\left[-2g_{V\mu}g_{A\mu}\gamma_5 \not{p}_4 \gamma^\nu (\not{p}_3 - m)\gamma^\mu\right] \tag{C.2.7}$$
$$+ \mathrm{Tr}\left[(g_{V\mu}^2 - g_{A\mu}^2)\, m\gamma^\nu (\not{p}_3 - m)\gamma^\mu\right]\Big)$$
$$= |g_Z|^2\Big((g_{V\mu}^2 + g_{A\mu}^2)\mathrm{Tr}\left[\not{p}_4\gamma^\nu \not{p}_3\gamma^\mu\right]$$
$$- 2g_{V\mu}g_{A\mu}\mathrm{Tr}\left[\gamma_5 \not{p}_4\gamma^\nu(\not{p}_3)\gamma^\mu\right] \tag{C.2.8}$$
$$- (g_{V\mu}^2 - g_{A\mu}^2)\, m^2 \mathrm{Tr}\left[\gamma^\nu\gamma^\mu\right]\Big)$$
$$= |g_Z|^2\Big((g_{V\mu}^2 + g_{A\mu}^2)4(p_4^\nu p_3^\mu + p_4^\mu p_3^\nu - g^{\mu\nu}\, p_3\cdot p_4\,)$$
$$- 2g_{V\mu}g_{A\mu}(-4i\epsilon^{\alpha\nu\beta\mu}p_{4\alpha}p_{3\beta}) \tag{C.2.9}$$
$$- (g_{V\mu}^2 - g_{A\mu}^2)\, m^2(4g^{\mu\nu})\Big) \quad .$$

We can contract the electron terms similarly to get,

$$J_e^\nu J_e^{*\mu} = |g_Z|^2\Big((g_{Ve}^2 + g_{Ae}^2)4(p_2^\nu p_1^\mu + p_2^\mu p_1^\nu - g^{\mu\nu}\, p_1\cdot p_2\,)$$
$$- 2g_{Ve}g_{Ae}(-4i\epsilon^{\alpha\nu\beta\mu}p_{2\alpha}p_{1\beta})\Big) \quad . \tag{C.2.10}$$

However before we do this it is worth considering the $k_\alpha k_\nu$ term from the Z propagator $iD_{\alpha\nu}$ contracted on the electron current. We have

$$J_e^\nu k_\nu = \bar{v}(2)g_Z \not{k}(g_{Ve} - g_{Ae}\gamma_5)u(1) \tag{C.2.11}$$
$$= g_Z\bar{v}(2)(\not{p}_1 + \not{p}_2)(g_{Ve} - g_{Ae}\gamma_5)u(1) \quad , \tag{C.2.12}$$

now if the fermion line is massless then the $\not{p}_2$ term cancels on the $\bar{v}(2)$ spinor on the left by the Dirac equation, and the $\not{p}_1$ term cancels on the $u(2)$ spinor on the right, also by the Dirac equation. This means that we only need consider the $-g_{\mu\nu}$ term in the Z propagator as long as we treat at least one fermion (in this case the electron) as massless. As such equation (C.2.1) becomes,

$$\sum |\mathcal{M}|^2 = \frac{J_{e\nu}J_\mu^\nu J_{e\mu}^* J_\mu^{*\mu}}{(k^2 - m_Z^2)^2 + m_Z^2\Gamma_Z^2} \quad . \tag{C.2.13}$$

Now substituting in (C.2.9,C.2.10) we get,

$$J_{e\nu}J_{e\mu}^* J_\mu^\nu J_\mu^{*\mu} = |g_Z|^4\Big[(g_{V\mu}^2 + g_{A\mu}^2)4(p_4^\nu p_3^\mu + p_4^\mu p_3^\nu - g^{\mu\nu}\, p_3\cdot p_4\,) - (g_{V\mu}^2 - g_{A\mu}^2)\, m^2(4g^{\mu\nu})$$
$$- 2g_{V\mu}g_{A\mu}(-4i\epsilon^{\alpha\nu\beta\mu}p_{4\alpha}p_{3\beta})\Big]$$
$$\times\Big[(g_{Ve}^2 + g_{Ae}^2)4(p_2^\nu p_1^\mu + p_2^\mu p_1^\nu - g^{\mu\nu}\, p_1\cdot p_2\,)$$
$$- 2g_{Ve}g_{Ae}(-4i\epsilon^{\alpha\nu\beta\mu}p_{2\alpha}p_{1\beta})\Big] \quad . \tag{C.2.14}$$

Now the terms on lines 1 and 3 of this equation are symmetric under the interchanges $\mu \leftrightarrow \nu$, and the terms on lines 2 and 4 are antisymmetric; now when we contract a symmetric line

with an antisymmetric line we get 0. So we get,

$$\sum J_{e\nu} J^*_{e\mu} J^\nu_\mu J^{*\mu}_\mu = |g_Z|^4 \Big[\Big[(g_{V\mu}^2 + g_{A\mu}^2) 4(p_4^\nu p_3^\mu + p_4^\mu p_3^\nu - g^{\mu\nu}\, p_3 \cdot p_4) - (g_{V\mu}^2 - g_{A\mu}^2)\, m^2 (4 g^{\mu\nu}) \Big]$$
$$\times \Big[(g_{Ve}^2 + g_{Ae}^2) 4(p_2^\nu p_1^\mu + p_2^\mu p_1^\nu - g^{\mu\nu}\, p_1 \cdot p_2) \Big]$$
$$- 64 g_{V\mu} g_{A\mu} g_{Ve} g_{Ae} \epsilon^{\alpha\nu\beta\mu} \epsilon_{\gamma\nu\delta\mu} p_{4\alpha} p_{3\beta} p_2^\gamma p_1^\delta \Big] \quad . \qquad \text{(C.2.15)}$$

Now contracting the first term and using the fact that $\epsilon^{\alpha\nu\beta\mu}\epsilon_{\gamma\nu\delta\mu} = -2(g^\alpha_\gamma g^\beta_\delta - g^\alpha_\delta g^\beta_\gamma)$ we get

$$J_{e\nu} J^*_{e\mu} J^\nu_\mu J^{*\mu}_\mu = |g_Z|^4 \Big[16 (g_{Ve}^2 + g_{Ae}^2)(g_{V\mu}^2 + g_{A\mu}^2)(2\, p_4 \cdot p_2\; p_3 \cdot p_1 + 2\, p_4 \cdot p_1\; p_3 \cdot p_2)$$
$$- 16 m^2 (g_{Ve}^2 + g_{Ae}^2)(g_{V\mu}^2 - g_{A\mu}^2)(-2\, p_1 \cdot p_2)$$
$$- 64 g_{V\mu} g_{A\mu} g_{Ve} g_{Ae} (-2(\, p_4 \cdot p_2\; p_3 \cdot p_1 \; - \; p_4 \cdot p_1\; p_3 \cdot p_2\,)) \Big] \qquad \text{(C.2.16)}$$
$$= |g_Z|^4 \Big[32 (g_{Ve}^2 + g_{Ae}^2)(g_{V\mu}^2 + g_{A\mu}^2)(\, p_4 \cdot p_2\; p_3 \cdot p_1 \; + \; p_4 \cdot p_1\; p_3 \cdot p_2\,)$$
$$+ 32 m^2 (g_{Ve}^2 + g_{Ae}^2)(g_{V\mu}^2 - g_{A\mu}^2)\, p_1 \cdot p_2$$
$$+ 128 g_{V\mu} g_{A\mu} g_{Ve} g_{Ae} (\, p_4 \cdot p_2\; p_3 \cdot p_1 \; - \; p_4 \cdot p_1\; p_3 \cdot p_2\,) \Big] \quad , \qquad \text{(C.2.17)}$$

and in this form we can easily evaluate the summed and averaged matrix element.

C.3 The $k^\mu k^\nu$ term in the Z propagator and massless fermions

The fact that the $k^\mu k^\nu$ term in the Z propagator in the previous section gave no contribution is a special case of a more general theorem that says that in any process if a Z (or W) couples to a massless fermion current then the $k_\mu k_\nu$ term in the propagator is always identically zero by gauge invariance, even if the fermion line is off shell. To see this consider doing the calculation in the $\overline{R_\xi}$ gauge. Then as well as the coupling of the Z to the fermion line, *Fig*.C.3.1.a, we also have a coupling of the Goldstone boson associated with the Z to the fermion line, *Fig*.C.3.1.b,

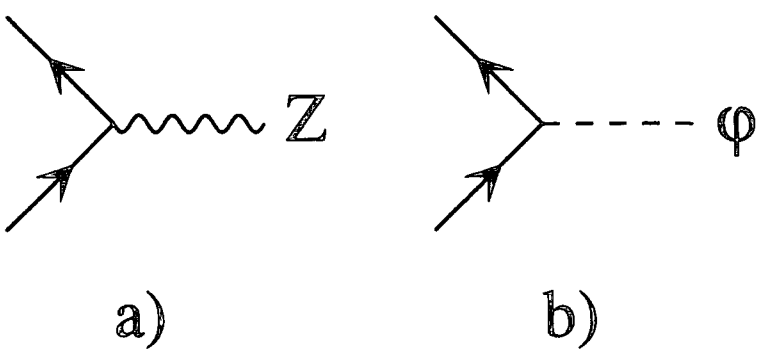


Fig.C.3.1 The process $Z, \phi \rightarrow \mu^+\mu^-$

If we calculate with $\xi = \infty$ then we just get diagram *Fig*.C.3.1.a and the Z propagator is the same as in the unitary gauge (in fact the R_ξ gauge with $\xi = \infty$, despite having a different gauge fixing term to the unitary gauge, is identical to the unitary gauge). Now calculating the $\xi = 1$ 't Hooft-Feynman gauge *Fig*.C.3.1.a has the standard Z propagator *without* the $k_\mu k_\nu$ term – in general this gives a different answer to the calculation in the $\xi = \infty$ unitary gauge; however this gauge dependence is canceled by diagram *Fig*.C.3.1.b. Now if the fermion is massless then diagram *Fig*.C.3.1.b is identically zero – because Goldstone bosons don't couple to massless fermions; this means that if the sum of diagrams *Fig*.C.3.1.a,b is to be gauge independent (as it must be) then diagram *Fig*.C.3.1.a had better be gauge independent by itself, and hence the $k^\mu k^\nu$ had better be zero. Notice that this is only a statement about gauge invariance – and as such it is only gauge invariant quantities that are independent of this $k^\mu k^\nu$ term, so in the case where say an arbitrary number of photons are radiated from the fermion leg each individual Feynman diagram does have a contribution from the $k^\mu k^\nu$ term as these are not gauge invariant quantities ... it is only in the sum over Feynman diagrams, which is gauge invariant, that the dependence on the $k^\mu k^\nu$ term drops out. This makes a useful test of the gauge invariance of internal Z or W bosons – and is invaluable as a cross-check of calculated matrix elements.

C.4 Spinor Techniques

Although the method of finding a usable form for the matrix element via traceology works well for process with only one or two Feynman diagrams as the number of Feynman diagrams increases the process becomes increasingly complicated, for example if we are calculating a process that is the sum of say 3 Feynman diagrams then we get,

$$\mathcal{M} = \mathcal{M}_1 + \mathcal{M}_2 + \mathcal{M}_3 \tag{C.4.1}$$

$$|\mathcal{M}|^2 = \mathcal{M}\mathcal{M}^* \tag{C.4.2}$$

$$= (\mathcal{M}_1 + \mathcal{M}_2 + \mathcal{M}_3)(\mathcal{M}_1 + \mathcal{M}_2 + \mathcal{M}_3)^* \tag{C.4.3}$$

$$= |\mathcal{M}_1|^2 + |\mathcal{M}_2|^2 + |\mathcal{M}_3|^2 + 2\Re(\mathcal{M}_1\mathcal{M}_2^*) + 2\Re(\mathcal{M}_1\mathcal{M}_3^*) + 2\Re(\mathcal{M}_2\mathcal{M}_3^*) \quad , \tag{C.4.4}$$

and already we have 6 terms that we need to evaluate via traceology. Clearly if we have n Feynman diagrams then we will have $n(n+1)/2$ terms to evaluate via traceology – and this rapidly becomes intractable. Now for spinor techniques we try and calculate the value of the matrix element at the matrix element level, then to get the matrix element squared we just sum over the (complex) values for the various matrix element associated with each

Feynman diagram. While we still have to sum over the different spin states in the problem as the number of Feynman diagrams becomes larger this method will become superior.

The spinor techniques that I present here are strongly based upon the spinor techniques of Kleiss and Stirling[33, 34].

We start by choosing an explicit representation for our fermion spinors. We require our spinors to be solutions to the Dirac equation,

$$\begin{aligned} (\not{p} - m)u_\lambda(p) = 0 \quad &, \quad (\not{p} + m)v_\lambda(p) = 0 \\ \bar{u}_\lambda(p)\,(\not{p} - m) = 0 \quad &, \quad \bar{v}_\lambda(p)\,(\not{p} + m) = 0 \quad , \end{aligned} \tag{C.4.5}$$

and choose there normalisation † as,

$$\begin{aligned} \bar{u}_{\lambda_1}(p)u_{\lambda_2}(p) &= 2m\,\delta_{\lambda_1\lambda_2} \\ \bar{v}_{\lambda_1}(p)v_{\lambda_2}(p) &= -2m\,\delta_{\lambda_1\lambda_2} \quad , \end{aligned} \tag{C.4.6}$$

or equivalently as,

$$\begin{aligned} \sum_\lambda u_\lambda(p)\bar{u}_\lambda(p) &= \not{p} + m \\ \sum_\lambda v_\lambda(p)\bar{v}_\lambda(p) &= \not{p} - m \quad , \end{aligned} \tag{C.4.7}$$

which has the added convenience of being meaningful when $m = 0$. Now these equations have the symmetry,

$$u \leftrightarrow v \quad , \quad m \leftrightarrow -m \quad , \tag{C.4.8}$$

this means that we can usually treat a v spinor as a u spinor with negative mass. Henceforth, unless otherwise stated, I will just use u spinors, the v spinors can be obtained by reversing the sign of the mass. In particular for massless spinors v spinors are equivalent to u spinors and I will always use u spinors.

To obtain an explicit representation for the spinors we first choose two four-vectors such that,

$$k_0 \cdot k_0 = 0 \quad , \quad k_1 \cdot k_1 = -1 \quad , \quad k_0 \cdot k_1 = 0 \quad . \tag{C.4.9}$$

We have one spinor, $u_-(k_0)$, which is used to define all other spinors, such that,

$$u_-(k_0)\bar{u}_-(k_0) = \omega_- \not{k}_0 \quad , \tag{C.4.10}$$

† We choose this normalization rather than,

$$\begin{aligned} \bar{u}_{\lambda_1}(p)u_{\lambda_2}(p) &= \delta_{\lambda_1\lambda_2} \\ \bar{v}_{\lambda_1}(p)v_{\lambda_2}(p) &= -\delta_{\lambda_1\lambda_2} \quad , \end{aligned}$$

as this avoids having factors of $2m$ in the phase space, and means that we treat both bosons and fermions on an equal footing in the phase space. It also saves us doing all calculations with a finite fermion mass and then only in the final answer taking the fermion mass to zero when we multiply the matrix element by the phase space

where,

$$\omega_{\pm} = \frac{1}{2}(1 \pm \gamma_5) \quad . \tag{C.4.11}$$

This spinor represents the negative chirality state of a massless fermion with momentum k_0. We define the positive chirality spinor as,

$$u_+(k_0) = \not{k}_1 u_-(k_0) \quad . \tag{C.4.12}$$

Now we define a general spinor as,

$$u_\lambda(p) = \frac{(\not{p} + m)u_{-\lambda}(k_0)}{\sqrt{2\,p\cdot k_0}} \quad , \tag{C.4.13}$$

which satisfies equations (C.4.5,C.4.6,C.4.7) and hence makes a satisfactory definition of a spinor.

For massless spinors $\omega_\pm$ act as projection operators for the different chirality states and we have,

$$\begin{aligned} \omega_\pm u_\pm(p) &= u_\pm(p) \\ \omega_\mp u_\pm(p) &= 0 \end{aligned} \qquad \text{for } m = 0 \quad . \tag{C.4.14}$$

The beauty of defining the spinors in this way is that we are able to get an explicit value for terms like $\bar{u}_{\lambda_1}(p_1)u_{\lambda_2}(p_2)$, for example we find

$$\begin{aligned} \bar{u}_+(p_1)u_-(p_2) &= \frac{\bar{u}_-(k_0)(\not{p}_1 + m_1)(\not{p}_2 + m_2)u_+(k_0)}{\sqrt{4\,p_1\cdot k_0\; p_2\cdot k_0}} \\ &= \frac{\mathrm{Tr}\,[\omega_- \not{k}_0(\not{p}_1 + m_1)(\not{p}_2 + m_2)\not{k}_1]}{\sqrt{4\,p_1\cdot k_0\; p_2\cdot k_0}} \\ &= \frac{\frac{1}{2}\mathrm{Tr}\,[\not{k}_0\not{p}_1\not{p}_2\not{k}_1] - \frac{1}{2}\mathrm{Tr}\left[\gamma_5\not{k}_0\not{p}_1\not{p}_2\not{k}_1\right]}{\sqrt{4\,p_1\cdot k_0\; p_2\cdot k_0}} \\ &= \frac{p_1\cdot k_0\; p_2\cdot k_1 \;-\; p_1\cdot k_1\; p_2\cdot k_0 \;- i\epsilon_{\mu\nu\rho\sigma}k_0^\mu k_1^\nu p_1^\rho p_2^\sigma}{\sqrt{p_1\cdot k_0\; p_2\cdot k_0}} \quad . \end{aligned} \tag{C.4.15}$$

If we now choose explicit vectors for k_0 and k_1 then this last expression becomes particularly simple. For example with,

$$\begin{aligned} p_i^\mu &= (p_i^0, p_i^x, p_i^y, p_i^z) \\ k_0^\mu &= (1,1,0,0) \\ k_1^\mu &= (0,0,1,0) \quad , \end{aligned} \tag{C.4.16}$$

we find,

$$\bar{u}_+(p_1)u_-(p_2) = (p_1^y + ip_1^z)\left[\frac{p_2^0 - p_2^x}{p_1^0 - p_1^x}\right]^{1/2} - (p_2^y + ip_2^z)\left[\frac{p_1^0 - p_1^x}{p_2^0 - p_2^x}\right]^{1/2} \quad . \tag{C.4.17}$$

As spinor products like these occur particularly frequently it proves convenient to introduced a more compact s notation. We use,

$$s_{+-}(1,2) = s_{+-}(p_1,p_2) \equiv \bar{u}_+(p_1)u_-(p_2) = -s_{+-}(2,1) \tag{C.4.18}$$

$$s_{-+}(1,2) = s_{-+}(p_1,p_2) \equiv \bar{u}_-(p_1)u_+(p_2) = [s_{+-}(2,1)]^* \quad . \tag{C.4.19}$$

Now for massless momenta the other spinor contractions vanish, however for massive momenta we also need to define two more spinor contractions,

$$\begin{aligned} s_{++}(1,2) = s_{++}(p_1,p_2) &\equiv \bar{u}_+(p_1)u_+(p_2) \\ &= m_1 \left[\frac{p_2^0 - p_2^x}{p_1^0 - p_1^x}\right]^{1/2} + m_2 \left[\frac{p_1^0 - p_1^x}{p_2^0 - p_2^x}\right]^{1/2} \end{aligned} \tag{C.4.20}$$

$$s_{--}(1,2) = s_{--}(p_1,p_2) \equiv \bar{u}_-(p_1)u_-(p_2) = s_{++}(1,2) \quad . \tag{C.4.21}$$

Usually in traceology we seek to rewrite all spinor products in terms of gamma matrices and then evaluate the trace, however in spinor techniques we do exactly the opposite and convert all gamma matrices into sums over spinors; and then evaluate these spinors via equations (C.4.17,C.4.18,C.4.19,C.4.20,C.4.21).

C.4.1 Useful identities for manipulating spinors and gamma matrices

Now gamma matrices can only occur in a few specific ways in the matrix element, we use the following rules to reduce them to spinors.

- Gamma matrices contracted on other gamma matrices.
  For massless momenta p_1 and p_2 we have the Chisholm identity,

$$\{\bar{u}_\lambda(p_1)\gamma^\mu u_\lambda(p_2)\}\gamma_\mu = 2u_\lambda(p_2)\bar{u}_\lambda(p_1) + 2u_{-\lambda}(p_1)\bar{u}_{-\lambda}(p_2) \quad . \tag{C.4.22}$$

- Line reversal trick.
  For massless momenta p_1 and p_2, and Γ an arbitrary string of γ matrices, we have,

$$\bar{u}_{\lambda_1}(p_1)\Gamma u_{\lambda_2}(p_2) = \lambda_1\lambda_2\bar{u}_{-\lambda_2}(p_2)\Gamma^R u_{-\lambda_1}(p_1) \quad , \tag{C.4.23}$$

  where Γ^R is the string of γ matrices, Γ, reversed.

- Gamma matrices contracted with vectors.
  For a general term $\not{p}$ we first write p as a sum of two massless momenta, for a general massless momenta p_2 we write,

$$p = p_1 + a\,p_2 \quad , \tag{C.4.24}$$

then choosing,

$$a = \frac{p^2}{2\,p\cdot p_2} \quad , \tag{C.4.25}$$

forces $p_1^2 = 0$. We can now write,

$$\begin{aligned} \not{p} &= \not{p}_1 + a\,\not{p}_2 \\ &= \sum_{\lambda=\pm} u_\lambda(p_1)\bar{u}_\lambda(p_1) + a \sum_{\lambda=\pm} u_\lambda(p_2)\bar{u}_\lambda(p_2) \end{aligned} \quad , \tag{C.4.26}$$

from the completeness relation, equation (C.4.7). Notice that p_2 is a free choice, in many cases by a careful choice of p_2 we can arrange for the second set of spinors to cancel on other factors which leaves only the p_1 set of spinors.

We can also play a similar game in cases where we have an extra added scalar term as in a heavy quark propagator. Starting from a term $\not{p} + m$ we write,

$$p = p_1 + ap_2 \qquad \text{where } p_2^2 = 0 \quad , \tag{C.4.27}$$

then choosing,

$$a = \frac{p^2 - m^2}{2\,p\cdot p_2} \quad , \tag{C.4.28}$$

forces $p_1^2 = m^2$. And now as before we get,

$$\begin{aligned} \not{p} + m &= \not{p}_1 + m + \not{p}_2 \\ &= \sum_{\lambda=\pm} u_\lambda(p_1)\bar{u}_\lambda(p_1) + a \sum_{\lambda=\pm} u_\lambda(p_2)\bar{u}_\lambda(p_2) \end{aligned} \quad . \tag{C.4.29}$$

- γ_5

If we have a term, $g_V - g_A\gamma_5$, that contains a γ_5 term and if there is an adjacent massless spinor on the right (or antispinor on the left) we can use the following trick,

$$\begin{aligned} (g_V - g_A\gamma_5)u_\lambda(p) &= ((g_V - g_A)\omega_+ + (g_V + g_A)\omega_-)u_\lambda(p) \\ &= (g_V - \lambda g_A)u_\lambda(p) \quad . \end{aligned} \tag{C.4.30}$$

C.4.2 Polarization vectors for massless gauge bosons

Although we can use this second trick to take care of all polarization vectors of vector bosons this is not normally a good way to proceed. For a massless vector boson with a definite helicity and momentum k we require the polarization vectors to satisfy,

$$\begin{aligned} \epsilon_\lambda\cdot k = 0 \quad &, \quad \epsilon_\lambda\cdot\epsilon_\lambda = 0 \\ \epsilon^\mu_{-\lambda} = (\epsilon^\mu_\lambda)^* \quad &, \quad \epsilon_\lambda\cdot\epsilon_{-\lambda} = -1 \quad . \end{aligned} \tag{C.4.31}$$

Now the object,

$$\epsilon_\lambda^\mu = \frac{\bar{u}_\lambda(k)\gamma^\mu u_\lambda(p)}{\sqrt{4\,p\cdot k}} \quad , \tag{C.4.32}$$

where p is an arbitrary lightlike vector not collinear to k, satisfies these requirements and hence makes a good choice of polarization vector. Notice again that p is an arbitrary vector, we can usually choose it in such a way as to make many spinor products involving this polarization vector zero.

C.4.3 Polarization vectors for massive gauge bosons

Most massive vector bosons so far discovered decay by their very nature, this decay is easily included using spinor techniques, and usually this is the easiest solution for external massive vector bosons. If the matrix element for specific polarizations is required then the easiest solution is to explicitly construct the polarization vectors and then remove them by converting them to a sum over spinors as above.

C.4.4 Spinors for massive fermions

Many of the previous rules only work for massless fermion spinors – and we have no simplifying rules for massive fermion spinors. Usually this does not cause difficulties because after we have manipulated the massless spinors we no longer have any troublesome terms. For example if we consider the decay $t \to bl^+\nu$,

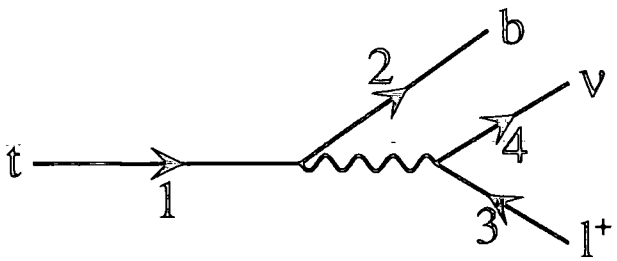


*Fig.*C.4.1 The Feynman diagram for $t \to bl^+\nu$

including the mass of the t and b quarks while keeping both leptons massless, we have the matrix element,

$$\begin{aligned}\mathcal{M} = {} & \bar{u}_{\lambda_b}(2) g_W \gamma^\mu (1-\gamma_5) u_{\lambda_t}(1) \\ & \times \frac{i(-g_{\mu\nu} + k_\mu k_\nu / m_W^2)}{k^2 - m_w^2 + i m_w \Gamma_w} \\ & \times \bar{u}_{\lambda_\nu}(4) g_W \gamma^\nu (1-\gamma_5) u_{\lambda_l}(3) \quad ,\end{aligned} \tag{C.4.33}$$

now the $(1-\gamma_5) = 2\omega_-$ and so projects out the $\lambda_\nu = - = \lambda_l$ state, and also the $k_\mu k_\nu$ term

in the W propagator cancels on the massless lepton current. This gives

$$\mathcal{M} = \frac{-i}{k^2 - m_w^2 + i m_w \Gamma_w} \bar{u}_{\lambda_b}(2) g_W \gamma^\mu 2\omega_- u_{\lambda_t}(1)\ \bar{u}_-(4) g_W \gamma_\mu 2 u_-(3) \quad . \tag{C.4.34}$$

The term $\bar{u}_-(4) g_W \gamma_\mu 2 u_-(3)$ is a scalar and so we can move it where we like,

$$\mathcal{M} = \frac{-4ig_W^2}{k^2 - m_w^2 + i m_w \Gamma_w} \bar{u}_{\lambda_b}(2) \{\bar{u}_-(4)\gamma_\mu u_-(3)\} \gamma^\mu \omega_- u_{\lambda_t}(1) \quad , \tag{C.4.35}$$

now using the Chisholm identity (C.4.22) on the massless lepton current,

$$\mathcal{M} = \frac{-4ig_W^2}{k^2 - m_w^2 + i m_w \Gamma_w} \bar{u}_{\lambda_b}(2)(2u_-(3)\bar{u}_-(4) + 2u_+(4)\bar{u}_+(3))\omega_- u_{\lambda_t}(1) \quad , \tag{C.4.36}$$

and we can now cancel the remaining $(1 - \gamma_5)$ on the massless lepton spinors to give,

$$\mathcal{M} = \frac{-8ig_W^2}{k^2 - m_w^2 + i m_w \Gamma_w} \bar{u}_{\lambda_b}(2) u_-(3)\ \bar{u}_-(4) u_{\lambda_t}(1) \tag{C.4.37}$$

$$= \frac{-8ig_W^2}{k^2 - m_w^2 + i m_w \Gamma_w} s_{\lambda_b -}(2,3) s_{-\lambda_t}(4,1) \quad . \tag{C.4.38}$$

However if we wish to keep the mass dependence of the lepton line we clearly cannot proceed in the way we have. For cases like this it turns out to be useful to redefine massive spinors as sums over massless spinors by the following method.

$$u_{+s}(q) = \frac{(\not{q} + m) u_-(k)}{\sqrt{2\, q \cdot k}} \quad , \tag{C.4.39}$$

is a good spinor for a massive fermion with momentum q and mass m. Now if we choose two lightlight momenta that make up q,

$$p_1^2 = 0 = p_2^2 \quad , \quad q^\mu = p_1^\mu + p_2^\mu \quad , \tag{C.4.40}$$

and choose $k = p_2$ we get,

$$u_{+s}(q) = \frac{1}{m}(\not{p}_1 + \not{p}_2 + m) u_-(p_2) \tag{C.4.41}$$

$$= \frac{s_{+-}(p_1, p_2)}{m} u_+(p_1) + u_-(p_2) \quad . \tag{C.4.42}$$

Since $|s_{+-}(p_1, p_2)| = m$ the coefficient of $u_+(p_1)$ is just a complex phase. The direction of spin of this spinor is,

$$s^\mu = \frac{p_1^\mu - p_2^\mu}{m} \quad , \tag{C.4.43}$$

which follows from,

$$\begin{aligned} u_{+s}(q)\bar{u}_{+s}(q) &= \frac{(\not{q} + m)\omega_- \not{k} (\not{q} + m)}{2\, q \cdot k} \\ &= \frac{[(\not{q} + m)\not{k} + \gamma_5 (\not{q} - m)\not{k}](\not{q} + m)}{4\, q \cdot k} \\ &= \frac{[2\, q \cdot k\ + 2\, q \cdot k\, \gamma_5 - 2m\gamma_5 \not{k}](\not{q} + m)}{4\, q \cdot k} \\ &= \frac{1}{2}(1 + \gamma_5 \not{s})(\not{q} + m) \quad , \end{aligned} \tag{C.4.44}$$

and $s \cdot s = -1$ and $q \cdot s = 0$ and so s is a pure spin state. We can similarly obtain the other spinor products,

$$u_{+s}(q) = \frac{s_{+-}(p_1, p_2)}{m} u_+(p_1) + u_-(p_2) \tag{C.4.45}$$

$$u_{-s}(q) = \frac{s_{-+}(p_1, p_2)}{m} u_-(p_1) + u_+(p_2) \quad . \tag{C.4.46}$$

C.4.5 An analytic continuation to negative energy spinors

If we consider the process $\gamma\gamma \to t\bar{t}$,

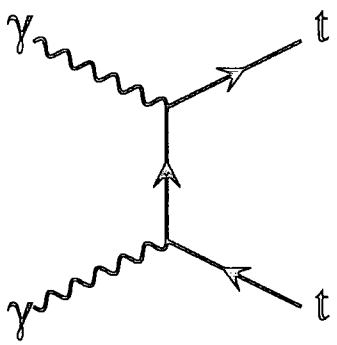


*Fig.*C.4.2 One of the Feynman diagrams for $\gamma\gamma \to t\bar{t}$

then the momentum of the internal quark line may either have positive energy or negative energy, now the spinors defined so far only describe positive energy fermions. To see this consider the spinor products,

$$\bar{u}_+(q) u_-(p) \bar{u}_-(p) u_+(q) = s_{+-}(q,p) s_{-+}(p,q) = 2p.q \quad , \tag{C.4.47}$$

however $s_{-+}(p,q) = (s_{+-}(q,p))^*$ and so we get,

$$2p.q = s_{+-}(q,p) s_{-+}(p,q) > 0 \quad , \tag{C.4.48}$$

which in turn implies that the signs of the energies of p and q must be equal. This must be true for all pairs of momenta, and so the energies of all spinors must have the same sign. Now this does not cause any difficulties, as if we have any momenta with negative energy we can always write them as minus momenta with positive energy. So for example for the process *Fig.*C.4.2 the internal quark propagator has the form,

$$(\not{p} + m) = (\not{p}_1 + \not{p}_2 + m) \quad , \tag{C.4.49}$$

where,

$$p_1^2 = 0 \quad , \quad p_2^2 = m^2 \quad , \tag{C.4.50}$$

considering just the $\not{p}_2 + m$ term we write,

$$\not{p}_2 + m = \begin{cases} \sum_\lambda u_\lambda(p_2)\bar{u}_\lambda(p_2) & \text{if } p_2^0 > 0 \\ -\sum_\lambda v_\lambda(-p_2)\bar{v}_\lambda(-p_2) & \text{if } p_2^0 < 0 \end{cases} \quad , \tag{C.4.51}$$

and so our spinors are always spinors of positive energy momenta. However this means that we need to set up special cases if any momenta go negative, and this can greatly complicate the calculation of any process. Clearly it would be more convenient to give a definition of our spinors valid for negative energy momenta, that work for all processes. From equation (C.4.51) it is clear that we require,

$$\sum_\lambda u_\lambda(p_2)\bar{u}_\lambda(p_2) = -\sum_\lambda v_\lambda(-p_2)\bar{v}_\lambda(-p_2) \quad , \tag{C.4.52}$$

a convenient solution for all cases is,

$$u_\lambda(-p) = iv_\lambda(p) \tag{C.4.53a}$$
$$v_\lambda(-p) = iu_\lambda(p) \tag{C.4.53b}$$
$$\bar{u}_\lambda(-p) = i\bar{v}_\lambda(p) \tag{C.4.53c}$$
$$\bar{v}_\lambda(-p) = i\bar{u}_\lambda(p) \quad , \tag{C.4.53d}$$

where p is a momenta with positive energy. Of course this analytic continuation destroys the property $s_{-+}(1,2) = [s_{+-}(2,1)]^*$ (equation (C.4.19)), as we know it must from equation (C.4.48).

This analytic continuation has also been obtained independently by Gunion and Kunszt[35].

We should check that this analytic continuation does not cause any difficulties in any calculation, however this is clear from

- Internal Fermions.
  This is the case just considered, we require an analytic continuation along the lines of (C.4.53).
- Internal Bosons.
  Internal boson propagators contain no spinors or gamma matrices and so are unaffected by any definition of those spinor products.
- External Particles.
  The factor of i that we have introduced in equation (C.4.53) does not change the modulus of any matrix element, but only changes its phase. As such it can have no effect in cases where we sum amplitudes at the matrix element squared level, and can only have any effect in the interference terms of sums of amplitudes at the matrix element level; however

we only sum diagrams at the matrix element level when they contain identical external particles; and this means that each diagram that we sum over contains an identical set of i's that can be taken as a common phase out front. Then when we square, this overall phase is unobserved.

It turns out to be particularly convenient in the calculation of any process to define all final state particles as having positive energy and all initial state particles as having negative energy, as this means that the momenta on any internal particles are easily reconstructed from *sums* over the external particle momenta; and with the above analytic continuation even if this internal momenta has negative energy we encounter no difficulties. This means that we end up with a matrix element that is manifestly crossing symmetric.

C.5 A convenient Matrix notation and $q\bar{q} \to t\bar{t}H$

In this section I will introduce a convenient matrix notation for calculating the matrix elements for a process in which we calculate all the chiralities of the fermion lines simultaneously. As well as being a compact way of doing that calculation, it speeds the numerical calculation of the matrix element up because many of the numerical factors used are identical for different chirality fermions; if we were to calculate each chirality separately we would have to calculate these factors several times – however if we calculate all the chiralities simultaneously we only need calculate these factors once. I will show these methods *in use* while calculating the process $q\bar{q} \to t\bar{t}H$, shown in *Fig*.C.5.1,

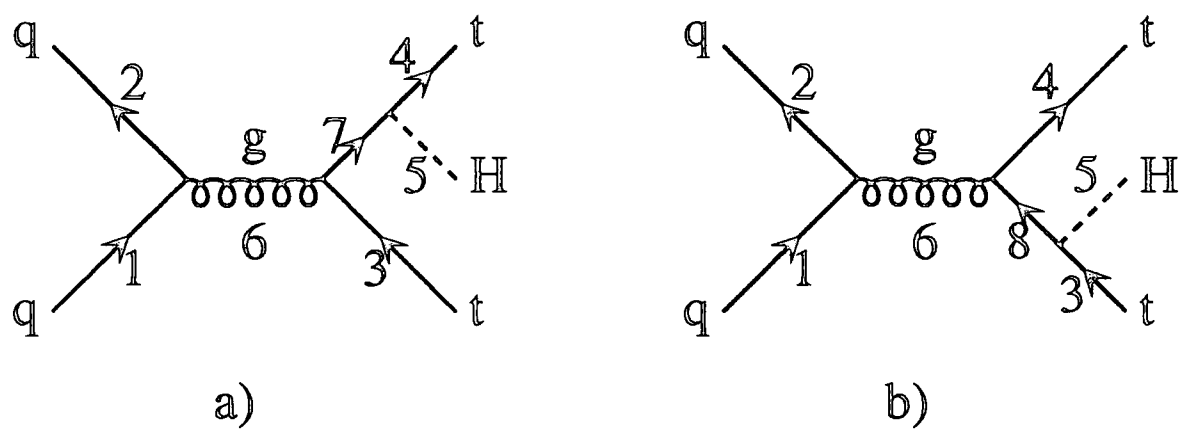


Fig.C.5.1 The Feynman diagrams for $q\bar{q} \to t\bar{t}H$

If we just consider the Feynman diagram *Fig*.C.5.1.a we have the matrix element,

$$\mathcal{M}^a = \bar{u}_{\lambda_4}(4)\left(-i\frac{m_t}{v}\right)\left(\frac{\not{p}_7 + m_t}{p_7^2 - m_t^2}\right)(ig_S\gamma_\mu)\,u_{\lambda_3}(3)$$
$$\times\left(i\frac{-g^{\mu\nu}}{p_6^2}\right)\quad \bar{u}_{\lambda_2}(2)\,(ig_S\gamma_\nu)\,u_{\lambda_1}(1) \tag{C.5.1}$$

$$= -\frac{m_t g_S^2}{v}\frac{1}{p_6^2}\frac{1}{p_7^2 - m_t^2}\quad \bar{u}_{\lambda_4}(4)(\not{p}_7 + m_t)\gamma_\mu u_{\lambda_3}(3)\quad \bar{u}_{\lambda_2}(2)\gamma^\mu u_{\lambda_1}(1) \tag{C.5.2}$$

Now looking at the light quark current the spinor products vanish unless $\lambda_2 = \lambda_1$ as we can imagine $\bar{u}_{\lambda_2}(2) = \bar{u}_{\lambda_2}(2)\omega_{-\lambda_2}$ and $u_{\lambda_1}(1) = \omega_{\lambda_1}u_{\lambda_1}(1)$ and so,

$$\begin{aligned}\bar{u}_{\lambda_2}(2)\gamma^\mu u_{\lambda_1}(1) &= \bar{u}_{\lambda_2}(2)\omega_{-\lambda_2}\gamma^\mu\omega_{\lambda_1}u_{\lambda_1}(1)\\ &= \bar{u}_{\lambda_2}(2)\gamma^\mu\omega_{\lambda_2}\omega_{\lambda_1}u_{\lambda_1}(1)\\ &= 0 \qquad\qquad \text{if } \lambda_1 \neq \lambda_2 \quad .\end{aligned} \tag{C.5.3}$$

Now contracting this current on the γ_μ term of the massive propagator we get,

$$\gamma_\mu\{\bar{u}_{\lambda_1}(2)\gamma^\mu u_{\lambda_1}(1)\} = 2u_{\lambda_1}(2)\bar{u}_{\lambda_1}(1) + 2u_{\lambda_1}(1)\bar{u}_{\lambda_1}(2) \tag{C.5.4}$$
$$= 2u_+(a)\bar{u}_+(b) + 2u_-(b)\bar{u}_-(b) \tag{C.5.5}$$
$$= 2\,(\,u_+(a)\quad u_-(b)\,)\begin{pmatrix}\bar{u}_+(b)\\ \bar{u}_-(a)\end{pmatrix}\quad , \tag{C.5.6}$$

where we define,

$$\left.\begin{matrix}a=2\\ b=1\end{matrix}\right\}\text{ If }\lambda_1 = +\qquad \left.\begin{matrix}a=1\\ b=2\end{matrix}\right\}\text{ If }\lambda_1 = -\quad . \tag{C.5.7}$$

If we now consider the $\not{p}_7 + m_t$ term on the massive line, we can write,

$$(\not{p}_7 + m_t) = (\not{p}_9 + \not{p}_{10} + m_t)\quad , \tag{C.5.8}$$

where,

$$p_9^2 = 0\quad ,\quad p_{10}^2 = m_t^2\quad , \tag{C.5.9}$$

and then writing this as a sum over spinors get,

$$(\not{p}_7 + m_t) = \sum_\lambda u_\lambda(10)\bar{u}_\lambda(10) + \sum_\lambda u_\lambda(9)\bar{u}_\lambda(9) \tag{C.5.10}$$

$$= (\,u_+(10)\quad u_-(10)\quad u_+(9)\quad u_-(9)\,)\begin{pmatrix}\bar{u}_+(10)\\ \bar{u}_-(10)\\ \bar{u}_+(9)\\ \bar{u}_-(9)\end{pmatrix}\quad . \tag{C.5.11}$$

Putting all these terms together we get,

$$\begin{aligned}\begin{pmatrix} \mathcal{M}^a_{++} & \mathcal{M}^a_{+-} \\ \mathcal{M}^a_{-+} & \mathcal{M}^a_{--} \end{pmatrix} &= -\frac{2m_t g_S^2}{v}\frac{1}{p_6^2}\frac{1}{p_7^2-m_t^2} \\ &\times \begin{pmatrix} \bar{u}_+(4) \\ \bar{u}_-(4) \end{pmatrix} \begin{pmatrix} u_+(10) & u_-(10) & u_+(9) & u_-(9) \end{pmatrix} \\ &\times \begin{pmatrix} \bar{u}_+(10) \\ \bar{u}_-(10) \\ \bar{u}_+(9) \\ \bar{u}_-(9) \end{pmatrix} \begin{pmatrix} u_+(a) & u_-(b) \end{pmatrix} \\ &\times \begin{pmatrix} \bar{u}_+(b) \\ \bar{u}_-(a) \end{pmatrix} \begin{pmatrix} u_+(3) & u_-(3) \end{pmatrix} \qquad \text{(C.5.12)} \\ &= -\frac{2m_t g_S^2}{v}\frac{1}{p_6^2}\frac{1}{p_7^2-m_t^2} \\ &\times \begin{pmatrix} s_{++}(4,10) & s_{+-}(4,10) & s_{++}(4,9) & s_{+-}(4,9) \\ s_{-+}(4,10) & s_{--}(4,10) & s_{-+}(4,9) & s_{--}(4,9) \end{pmatrix} \\ &\times \begin{pmatrix} s_{++}(10,a) & s_{+-}(10,b) \\ s_{-+}(10,a) & s_{--}(10,b) \\ s_{++}(9,a) & s_{+-}(9,b) \\ s_{-+}(9,a) & s_{--}(9,b) \end{pmatrix} \\ &\times \begin{pmatrix} s_{++}(b,3) & s_{+-}(b,3) \\ s_{-+}(a,3) & s_{--}(a,3) \end{pmatrix} \quad . \qquad \text{(C.5.13)}\end{aligned}$$

Notice that in calculating these matrices many terms are repeated, this means that the matrix can be calculated far quicker than it would take to generate each entry of the matrix separately. Using the same methods we can find $\mathcal{M}^b$. Setting

$$(\not{p}_8 + m_t) = (\not{p}_{11} + \not{p}_{12} + m_t) \quad , \qquad \text{(C.5.14)}$$

where,

$$p_{11}^2 = 0 \quad , \quad p_{12}^2 = m_t^2 \quad , \qquad \text{(C.5.15)}$$

we find,

$$\begin{aligned}\begin{pmatrix} \mathcal{M}^b_{++} & \mathcal{M}^b_{+-} \\ \mathcal{M}^b_{-+} & \mathcal{M}^b_{--} \end{pmatrix} &= -\frac{2m_t g_S^2}{v}\frac{1}{p_6^2}\frac{1}{p_8^2-m_t^2} \\ &\times \begin{pmatrix} s_{++}(4,a) & s_{+-}(4,b) \\ s_{-+}(4,a) & s_{--}(4,b) \end{pmatrix} \\ &\times \begin{pmatrix} s_{++}(b,12) & s_{+-}(b,12) & s_{++}(b,11) & s_{+-}(b,11) \\ s_{-+}(a,12) & s_{--}(a,12) & s_{-+}(a,11) & s_{--}(a,11) \end{pmatrix} \\ &\times \begin{pmatrix} s_{++}(12,3) & s_{+-}(12,3) \\ s_{-+}(12,3) & s_{--}(12,3) \\ s_{++}(11,3) & s_{+-}(11,3) \\ s_{-+}(11,3) & s_{--}(11,3) \end{pmatrix} \quad . \qquad \text{(C.5.16)}\end{aligned}$$

We are now in a position to form the complete matrix element $\mathcal{M}$,

$$\mathcal{M} = \mathcal{M}^a + \mathcal{M}^b \quad , \tag{C.5.17}$$

and this matrix element we can square up numerically.

Now this process in invariant under parity transformation, and this tells us that,

$$|\mathcal{M}_{++}|^2 = |\mathcal{M}_{--}|^2 \quad , \quad |\mathcal{M}_{+-}|^2 = |\mathcal{M}_{-+}|^2 \quad , \tag{C.5.18}$$

which gives us a useful (if somewhat trivial) cross check that we have calculated the spinor matrices correctly. Also our choice of p_9, p_{10}, p_{11} and p_{12} are (to a certain extent) arbitrary (see equation (C.4.24)), and when these vectors are varied our matrix element squared should be unchanged, which gives a stronger cross-check. If we calculate a process with external photons or gluons then we have a polarization vector for these photons or gluons, and again varying this vector should not influence the matrix element squared, this test of gauge invariance is a very strong test that we have correctly programmed the matrix element.

If the top quark were massless, $m_t = 0$, then all the s_{++} and s_{--} terms in the above spinorial matrices would be zero and the matrix multiplications would become far easier, breaking down to just two strings of numbers to be multiplied together, one corresponding to the $+ve$ chirality t, and one to the $-ve$ chirality t spin state of the quark, the chirality of the $\bar{t}$ antiquark being opposite of the t quark, as the fermion line's chirality is flipped when the Higgs is emitted. Both these chirality states are again identical in magnitude due to parity conservation. This happens because a massless quark's chirality is conserved when it emits a vector boson, and so we only have to calculate the flow of either $+ve$ chirality or $-ve$ chirality without any mixing between the two. Whereas for massive quark we have a chance of chirality flip terms, other than when the Higgs is emitted, through the s_{++} and s_{--} terms, this is what makes the matrix notation so easy – it sums up all the possible chiralities of the internal quark lines without us having to think about it.

Appendix D

The Phase Space

'When you wake up in the morning, Pooh,' said piglet
at last, 'what's the first thing you say to yourself?
'What's for breakfast?' said Pooh. 'What do you say, Piglet?'
'I say, I wonder what's going to happen exciting to-day?' said Piglet.
Pooh nodded thoughtfully. 'It's the same thing,' he said.

Winnie–the–Pooh, A.A.Milne

D.1 Introduction

When calculating a process the matrix element is only half the story, it only gives the probability of producing a particle per unit of phase space. So given a matrix element we find the differential cross-section from the formula,

$$d\sigma = (2\pi)^4 \frac{1}{F} |\mathcal{M}|^2 \quad \delta^4\Big(\sum_{\substack{\text{Initial}\\ \text{particles}}} p - \sum_{\substack{\text{Final}\\ \text{particles}}} p\Big) \quad d(\text{LIPS}) \qquad . \tag{D.1.1}$$

The flux factor, F, can be thought of as the probability of the two colliding particles being in the same place at the same time, its value is given by,

$$F = 4E_1 E_2 v_{\text{rel}} \qquad , \tag{D.1.2}$$

where v_{rel} is the relative velocity of the initial particles, and E_1 and E_2 are their energies. This can be written as,

$$v_{\text{rel}} = \frac{[(p_1 \cdot p_2)^2 - m_1^2 m_2^2]^{\frac{1}{2}}}{E_1 E_2} \qquad , \tag{D.1.3}$$

and so the flux factor can be written as,

$$F = 4[(\,p_1 \cdot p_2\,)^2 - m_1^2 m_2^2]^{\frac{1}{2}} \quad , \tag{D.1.4}$$

and in this form it is manifest that this is Lorentz invariant.

The phase space, d(LIPS), can be thought of as the amount of volume that the final state particles have to be produced in, it is written as,

$$d(\text{LIPS}) = \prod_{\substack{\text{Final} \\ \text{particles}}} \frac{d^3p}{(2\pi)^3 2E} \tag{D.1.5}$$

$$= \prod_{\substack{\text{Final} \\ \text{particles}}} \frac{d^4p \quad \delta(p^2 - m^2)}{(2\pi)^3} \quad , \tag{D.1.6}$$

and in the second form this is also manifestly Lorentz invariant.

The δ^4 term just ensures that 4 momentum is conserved, that is energy and momentum conservation.

Similarly we can calculate the partial width of a particle to decay as,

$$d\Gamma = (2\pi)^4 \frac{1}{G} |\mathcal{M}|^2 \quad \delta^4\Big(\sum_{\substack{\text{Initial} \\ \text{particles}}} p - \sum_{\substack{\text{Final} \\ \text{particles}}} p\Big) \quad d(\text{LIPS}) \quad , \tag{D.1.7}$$

where here the flux factor, G, is equal to $2m$.

D.2 Monte Carlo Integration

Now for both these quantities in order to evaluate the total cross-section, σ, or the total width, Γ, we need to be able to integrate over the phase space, d(LIPS). For simple process like $e^+e^- \rightarrow \mu^+\mu^-$ this is reasonably easy to do analytically – however as the process increases in complexity the analytic calculation becomes far harder, and usually is intractable. Even if the analytic calculation were tractable, often this is not what we want; usually we want the differential cross-section with respect to only certain variables (for example missing p_T in a process which contains a W decaying leptonically) with the other variables integrated out. This greatly increases the difficulty of the calculation, and if we require several different distributions we must redo the calculation for each distribution. As a result most integrations are not done analytically but numerically on a computer using a method called Monte Carlo integration.

If we have some function f that depends on many variables $\underline{x}$ then we can set up the distribution, I_1, where,

$$I_1 = f(\underline{x}) \quad , \tag{D.2.1}$$

and we choose $\underline{x}$ uniformly in some space $\mathcal{S}$. Its mean value is given by,

$$\langle I_1 \rangle = \frac{1}{V} \int_{\mathcal{S}} f(\underline{x}) d\underline{x} \quad , \tag{D.2.2}$$

where V is the volume of the space $\mathcal{S}$ given by,

$$V = \int_{\mathcal{S}} d\underline{x} \quad , \tag{D.2.3}$$

and its variance is given by,

$$\begin{aligned} \mathrm{Var}(I_1) &= \langle (I_1 - \langle I_1 \rangle)^2 \rangle & \text{(D.2.4)} \\ &= \langle I_1^2 \rangle - \langle I_1 \rangle^2 \quad . & \text{(D.2.5)} \end{aligned}$$

Now we can look at I_1 in a different way, from equation (D.2.2) we can view $V\, I_1$ as a crude estimator of the integral $\int_{\mathcal{S}} f(\underline{x}) d\underline{x}$, and as such it is called the 1 shot estimator of $\int_{\mathcal{S}} f(\underline{x}) d\underline{x}$. Of course the 1sd error, E_1, on this estimate we expect to be very large,

$$E_1 = V \sqrt{\mathrm{Var}(I_1)} \quad . \tag{D.2.6}$$

Now we can improve upon this estimate of $\int_{\mathcal{S}} f(\underline{x}) d\underline{x}$ by rather than throwing one random point in $\mathcal{S}$, throwing n random points in $\mathcal{S}$, and then taking the average. This is called the n shot estimator of $\int_{\mathcal{S}} f(\underline{x}) d\underline{x}$,

$$I_n = \frac{1}{n} \sum_{i=1,n} f(\underline{x}_i) \quad . \tag{D.2.7}$$

The mean of I_n is given by,

$$\begin{aligned} \langle I_n \rangle &= \frac{1}{n} \sum_{i=1,n} \langle I_1 \rangle \\ &= \langle I_1 \rangle \\ &= \frac{1}{V} \int_{\mathcal{S}} f(\underline{x}) d\underline{x} \quad , \end{aligned} \tag{D.2.8}$$

and so is unchanged. Now the variance of I_n* is,

$$\begin{aligned} \mathrm{Var}(I_n) &= \frac{1}{n^2} \sum_{i=1,n} \mathrm{Var}(I_1) \\ &= \frac{1}{n} \mathrm{Var}(I_1) \quad , \end{aligned} \tag{D.2.9}$$

where the variances sum if we generate the $\underline{x}_i$ independently. It is clear that the variance decreases for the n shot estimator of $\int_S f(\underline{x}) d\underline{x}$, and the 1 sd error, E_n, decreases as,

$$\begin{aligned} E_n &= V\sqrt{\mathrm{Var}(I_n)} \\ &= \frac{1}{\sqrt{n}} E_1 \quad . \end{aligned} \tag{D.2.10}$$

Also the Central Limit Theorem tells us that if we add probability distributions together that in the limit that we add an infinite number together the probability distribution becomes normal (or Gaussian). In practice we do not need to add many probability distributions together until the sum looks very like a Gaussian curve,

$$P(x) = \frac{1}{2\pi \mathrm{Var}} \exp\left[\frac{-(x - \langle P \rangle)^2}{2\,\mathrm{Var}}\right] \quad . \tag{D.2.11}$$

This can be seen in *Fig.*D.2.1 where we show the probability distributions, P_n of,

$$I_n = \frac{1}{n} \sum_{i=1,n} x_i \quad , \tag{D.2.12}$$

where we have chosen the x's in the interval $x \in [0, 1]$ with the probability distribution,

$$P_1 = 2x \quad , \tag{D.2.13}$$

* The variance of two variables added together is given by,

$$\begin{aligned} \mathrm{Var}(X+Y) &= \langle ((X+Y) - \langle X+Y \rangle)^2 \rangle \\ &= \langle ((X - \langle X \rangle) + (Y - \langle Y \rangle))^2 \rangle \\ &= \langle (X - \langle X \rangle)^2 \rangle + \langle (Y - \langle Y \rangle)^2 \rangle + 2\langle (X - \langle X \rangle)(Y - \langle Y \rangle) \rangle \\ &= \mathrm{Var}(X) + \mathrm{Var}(Y) + 2\mathrm{Cov}(X,Y) \quad , \end{aligned}$$

where $\mathrm{Cov}(X,Y) = \langle (X - \langle X \rangle)(Y - \langle Y \rangle) \rangle$, and if X and Y are independent then $\mathrm{Cov}(X,Y) = 0$ and the variances sum. Also if we scale a distribution we find,

$$\begin{aligned} \mathrm{Var}(aX) &= \langle (aX - \langle aX \rangle)^2 \rangle \\ &= \langle a^2 (X - \langle X \rangle)^2 \rangle \\ &= a^2 \mathrm{Var}(X) \quad . \end{aligned}$$

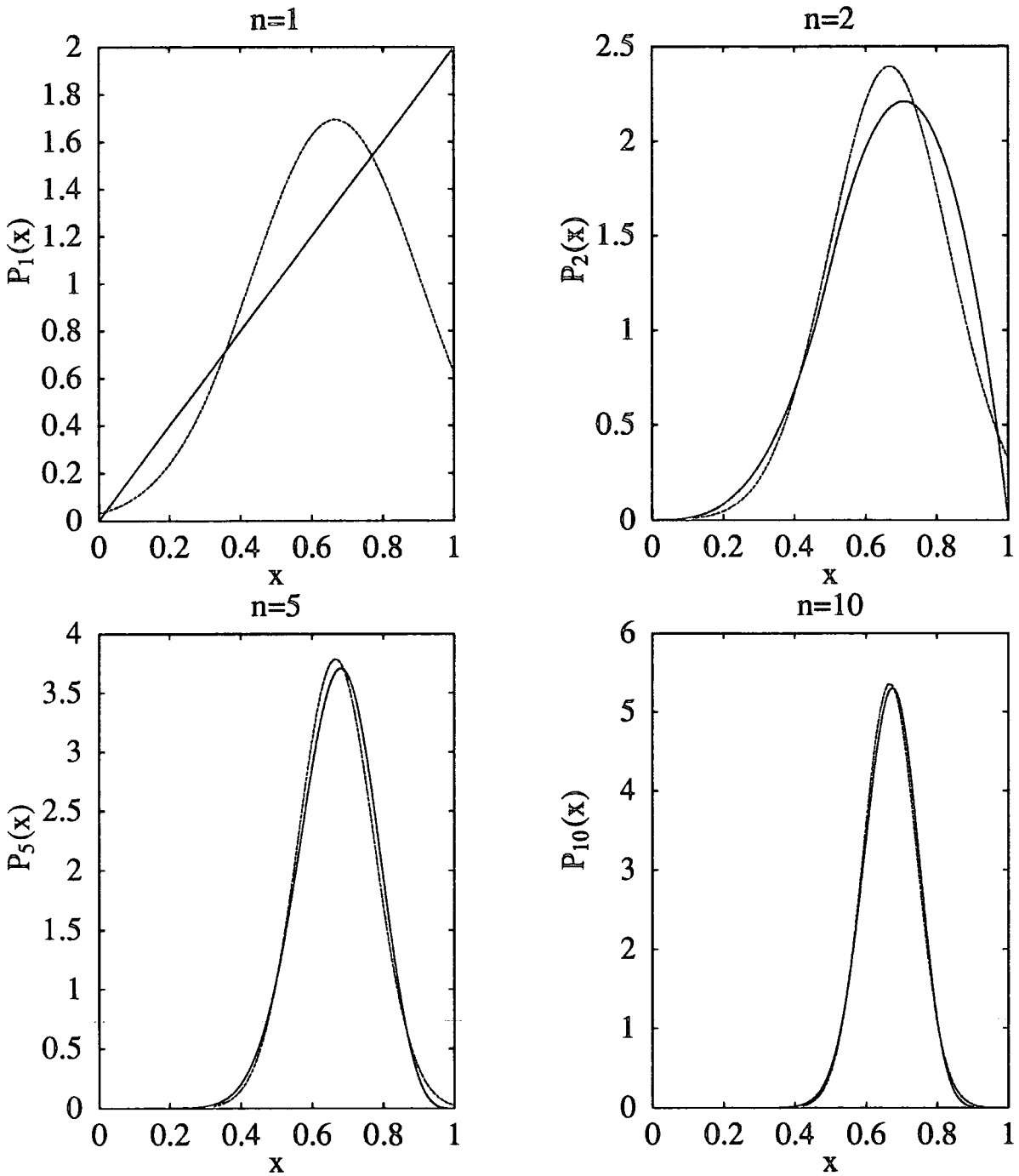


*Fig.*D.2.1 The probability distributions, P_1, P_2, P_5, P_{10} plotted with solid lines, and the Gaussian curves with the same mean and variance plotted with dashed lines

and also the Gaussian curve with the same mean and variance as the I_n, clearly the probability distributions very quickly approach the Gaussian distribution, even though the starting distribution P_1 is very different.

So if we estimate the integral $\int_{\mathcal{S}} f(\underline{x})d\underline{x}$ using the equation,

$$\int_{\mathcal{S}} f(\underline{x})d\underline{x} \simeq V\frac{1}{n}\sum_{i=1,n} f(\underline{x}_i) \quad , \tag{D.2.14}$$

we have an error term that decreases as $\frac{1}{\sqrt{n}}$ and is well behaved (*i.e.*, doesn't have a large tail). At first sight this looks like a poor rate of convergence in comparison with other methods of estimating integrals, for example if we were to use the trapezoid rule to estimate a one dimensional integral then the error term goes down as $\frac{1}{n^2}$. However where Monte Carlo integration wins out is in multi dimensional integrations – here the Monte Carlo error still decreases as $\frac{1}{\sqrt{n}}$ whereas other integration methods do less well, for example for a m dimensional integral the trapezoid rule error decreases as $\frac{1}{n^{2/m}}$ and for $m > 4$ this does worse than Monte Carlo integration. When calculating an n body phase space we work in $3n-4$ dimensions (unless we can analytically do certain integrals) and so typically we work in a large number of dimensions where Monte Carlo integrations rate of convergence is far better than other methods.

Now the Monte Carlo integration 1 sd error we know goes as equation (D.2.10), however we can also estimate $\mathrm{Var}(I_1)$ by Monte Carlo integration, from equation (D.2.5) we have,

$$\begin{aligned}\mathrm{Var}(I_1) &= \langle I_1^2\rangle - \langle I_1\rangle^2 \\ &= \frac{1}{V}\int_{\mathcal{S}}(f(\underline{x}))^2 d\underline{x} - \left(\frac{1}{V}\int_{\mathcal{S}} f(\underline{x})d\underline{x}\right)^2 \qquad \text{(D.2.15)} \\ &\simeq \frac{1}{n}\sum_{i=1,n}(f(\underline{x}_i))^2 - \left(\frac{1}{n}\sum_{i=1,n} f(\underline{x}_i)\right)^2 \quad , \qquad \text{(D.2.16)}\end{aligned}$$

and so by also forming sums of f^2 as well as f we can estimate the error on I_1 and so also the error on I_n via (D.2.10).

D.3 Importance Sampling

While the error in Monte Carlo integration decreases as $\frac{1}{\sqrt{n}}$ clearly its overall size is dictated by the value of $\mathrm{Var}(I_1)$. Any way that we can decrease this value will pay off directly in the size of our error. From equation (D.2.4) we can see the the more "flat" the value of $f(\underline{x})$ the smaller the value of $\mathrm{Var}(I_1)$. If we consider doing the integral,

$$\int_0^1 x^2 dx = \frac{1}{3} \quad , \tag{D.3.1}$$

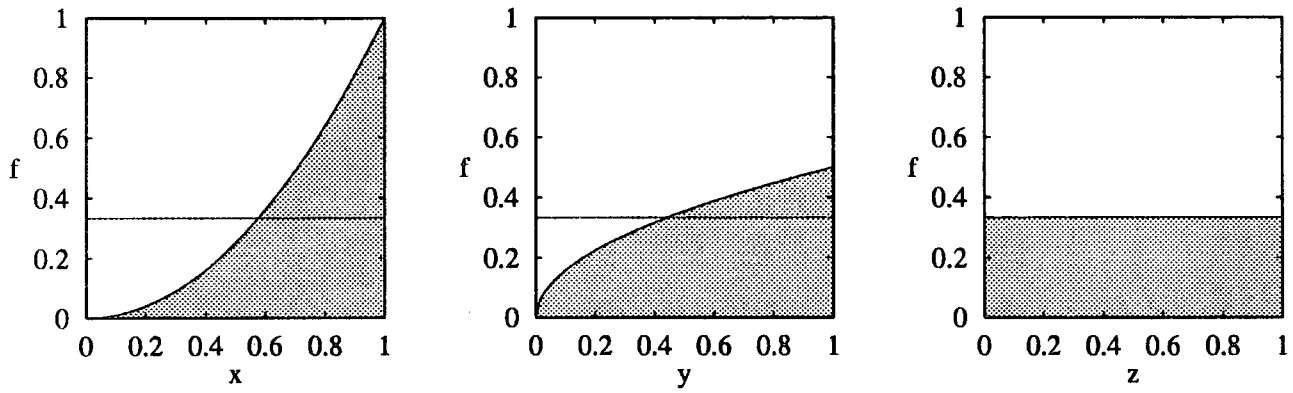


*Fig.*D.3.1 The function f, as a function of x, y, z; and the mean value of f

via Monte Carlo methods then we find that the variance is,

$$\mathrm{Var}(I_1) = \frac{4}{45} \qquad \Rightarrow \qquad E_1 = \frac{2}{3\sqrt{5}} \approx 0.298 \quad . \tag{D.3.2}$$

If we now make a change of variables $y = x^2$ we get the integral,

$$\int_0^1 \frac{1}{2} y^{1/2} dy = \frac{1}{3} \quad , \tag{D.3.3}$$

which is much "flatter" than (D.3.1) as can be seen from *Fig.*D.3.1 and calculating this by Monte Carlo integration gives the variance,

$$\mathrm{Var}(I_1) = \frac{1}{72} \qquad \Rightarrow \qquad E_1 = \frac{1}{6\sqrt{2}} \approx 0.118 \quad , \tag{D.3.4}$$

clearly this has improved the error term – how can this be ? Well in equation (D.3.1) we have chosen x uniformly in the interval $[0, 1]$; whereas in (D.3.3) we have chosen y uniformly in the interval $[0, 1]$. This is not the same thing – when we choose y uniformly in the interval $[0, 1]$ if we transform this back to x we end up choosing x values more densely as x get larger; *i.e.,* we choose more x values where the integrand is largest. This idea of choosing most of our random points where the integrand is largest is called importance sampling. In the above integral we can in fact do this importance sampling exactly, if we make the change of variables $z = x^3$ then (D.3.1) becomes,

$$\int_0^1 \frac{1}{3} dz = \frac{1}{3} \quad , \tag{D.3.5}$$

and the variance drops to zero, and so the Monte Carlo integral is exact (even if we only throw one point !). This has happened because, in this case, we can do the integral (D.3.1) exactly, however as equation (D.3.3) shows it is not necessary to be able to do the integral exactly to reduce the variance – if we can do any integral that looks even remotely similar (in this case x^2 vs x^3) then we can decrease the variance and so increase the accuracy of our Monte Carlo integral.

D.4 RAMBO

In some processes we have little or no initial idea of the structure of the matrix element – even if we are able calculate its value. As such we often do not know what changes of variables will flatten the matrix element. In such circumstances we want to map the phase space (which is universal, and so independent of the actual process) to as flat a function as possible. However this is a tricky problem. Luckily this has been done exactly for massless particles by R. Kleiss, W.J. Stirling, and S.D. Ellis, in the program called RAMBO (RAndom Momenta Beautifully Organized)[36]. This program generates events isotropically in the space defined by,

$$\delta^4\Big(\sum_{\substack{\text{Initial}\\\text{particles}}} p - \sum_{\substack{\text{Final}\\\text{particles}}} p\Big) \prod_{\substack{\text{Final}\\\text{particles}}} \frac{d^3p}{2E} \quad , \tag{D.4.1}$$

where the boundaries are defined by the δ function. The program also generates the phase space in the case where the particles are massive, however here the phase space is not mapped to an entirely flat function and there is still some small variance. In practice most processes that I've considered have $m^2 \ll s$ for at least most particles and RAMBO still does a good (if non perfect) job.

D.4.1 The Breit-Wigner line shape

Far more usual when calculating a process we have some idea of the structure of the matrix element, for example a $\mu^+\mu^-$ may come from a decaying Z^0. If this happens then we know that the matrix element contains the factor,

$$\frac{1}{((p_1+p_2)^2 - m_Z^2)^2 + m_Z^2\Gamma_Z^2} \quad , \tag{D.4.2}$$

where p_1, p_2 are the momenta of the μ^+, μ^-. In this case we can exactly remove this factor. The phase space looks like,

$$\begin{aligned}\delta^4(\textstyle\sum p - p_1 - p_2)\prod \frac{d^3p}{2E} \times \frac{d^3p_1}{2E_1}\frac{d^3p_2}{2E_2} &= \delta^4(\textstyle\sum p - p_1 - p_2)\prod \frac{d^3p}{2E} \\ &\quad \times \frac{d^3p_1}{2E_1}\frac{d^3p_2}{2E_2}\, d^4p_Z\, \delta^4(p_Z - p_1 - p_2) \qquad \text{(D.4.3)} \\ &= \left(\delta^4(\textstyle\sum p - p_Z)\prod \frac{d^3p}{2E}\frac{d^3p_Z}{2E_Z}\right) \times \left(d(p_Z^2)\right) \\ &\quad \times \left(\delta^4(p_Z - p_1 - p_2)\frac{d^3p_1}{2E_1}\frac{d^3p_2}{2E_2}\right) . \qquad \text{(D.4.4)}\end{aligned}$$

At this stage the phase space looks like the phase space to produce a Z with mass $\sqrt{p_Z^2}$ multiplied by the phase space for a Z to decay, multiplied by a factor of $d(p_Z^2)$. Now we can

use this last factor to remove the Breit-Wigner line shape, we have,

$$\frac{d(p_Z^2)}{(p_Z^2 - m_Z^2)^2 + m_Z^2\Gamma_Z^2} = \frac{d\theta}{m_Z\Gamma_Z} \quad , \tag{D.4.5}$$

where we have defined,

$$p_Z^2 - m_Z^2 = m_Z\Gamma_Z \tan\theta \quad . \tag{D.4.6}$$

So our strategy becomes first to generate θ then use equation (D.4.6) to reconstruct p_Z^2 then generate the original phase space (for example by RAMBO) but instead of generating the p_1 and p_2 momenta instead just generating *pseudo* Z momenta with mass $\sqrt{p_Z^2}$, and then decay this *pseudo* Z (again this can be done by RAMBO) into particles p_1 and p_2.

D.5 Experimental Cuts

So far I have made no mention of how to cope with the volume term V that occurs in Monte Carlo integration, the best solution is to scale all integration variables so that they have a minimum value of 0 and a maximum value of 1; then our integration space becomes a unit hypercube with volume $V = 1$. Usually, however, we can not map our integration space to a hypercube, due to complicated boundaries imposed by the experimental cuts that we apply. In this case we try to map our integration variables so that they lie *inside* a unit hypercube, and so we view our integration space as embedded in some larger space. Then we generate our random points inside this hypercube, and when they fall outside of our integration region we set the integrand to zero; and so the extra volume that we integrate over does not add any contribution to the total cross-section. As this means that at the edge of our physical parameter space we have a sharp drop off where the integrand drops to zero this makes the error term in the Monte Carlo integral worse; this should neither surprise, nor worry, us because effectively we are estimating the volume of the integration space V as well.

D.6 Differential cross-sections and a model for reality

What do we mean when we ask for the differential cross-section, $\frac{d\sigma}{dX}$ for some process? Well we mean that,

$$\sigma_{[X:X+\delta X]} = \int_X^{X+\delta X} \frac{d\sigma}{dX} dX \tag{D.6.1}$$

$$\approx \frac{d\sigma}{dX}(X)\delta X \quad , \tag{D.6.2}$$

and starting from this definition it is clear how to evaluate the differential cross-section, the cross-section on the left we can just estimate via Monte Carlo integration directly and then use the approximate equation (D.6.2) to approximate the differential cross-section. Now we can form the cross-section on the left as we form the total crosection – we just need to look at the random points that we generate in phase space and see if the X value falls in the particular bin that we are interested in; or if we are interested in the differential cross-section for all X we just bin the cross-section in the variable X. Now equation (D.6.2) only becomes exact in the limit $\delta X \to 0$ however in this limit our estimation of the $\sigma(X \to X + \delta X)$ fails because our Monte Carlo integral has very few points in it and so estimates the integral badly. Clearly we have to choose the width for our bin carefully, fine enough to give us a feel for the differential cross-section, but coarse enough to estimate the partial cross-section accurately.

Now one of the beauties of Monte Carlo integration is that we can form this differential cross-section while calculating the total cross-section, and indeed if we require two or three differential cross-sections we can form these simultaneously by just binning our random points in phase space to form the partial cross-sections. Now we seem to be able to do this until we have exhausted all the variables that the point in phase space depends on – however at this stage we can describe the point in phase space exactly, it knows about everything that an event in an actual experiment would do. As such we can take our Monte Carlo integrator as model for reality with each point in phase space being an actual event, of which we can ask question, such as, *What are particles energy? What direction do the particles travel in?*, or indeed any question which we could ask of an actual particle. These pseudo events have an obvious difference with the real world – each event in this Monte Carlo world has a weight associated with it, that depends upon the value of the matrix element and any changes of variables that we have made. As such this kind of event is called a *weighted event*, if we require events that are not weighted (as in reality) then we can always turn our weighted events into unweighted events by using the following method. First calculate the maximum weight possible, w_{max}, or an upper bound to the weights. Then as we generate our events in phase space, only accept the event if its weight, w, satisfies,

$$\frac{w}{w_{max}} > \phi \quad , \tag{D.6.3}$$

where ϕ is a random variable chosen uniformly in the interval $[0, 1]$. Then every event has an identical weight of $w_{\max}$. This method always causes a loss in accuracy – because if any event fails the above test it is totally rejected and not used in the calculation of the cross-section, such a blatant trashing of information can never help the accuracy of a calculation; however this approach is very popular with experimentalists because our unweighted events now look exactly as they would do if they were produced in reality; modulo the assumption

that the model on which we have based our Monte Carlo (be it the Standard Model, or the Minimal Supersymmetric Model; leading order, or all orders, *etc.*) is indeed correct !

Appendix E

Parton Distributions – A model of the Proton

It went on raining, and every day the water got a little higher, until now it was nearly up to Piglet's window ... and still he hadn't done anything.

'There's Pooh,' he thought to himself. 'Pooh hasn't much Brain, but he never comes to any harm. He does silly things and they turn out right. There's Owl. Owl hasn't exactly got Brain, but he Knows Things. He would know the Right Thing to Do when Surrounded by Water. There's Rabbit. He hasn't Learnt in Books, but he can always Think of a Clever Plan. There's Kanga. She isn't Clever, Kanga isn't, but she would be so anxious about Roo that she would do a Good Thing to Do without thinking about it. And then there's Eeyore. And Eeyore is so miserable anyhow that he wouldn't mind about this. But I wonder what Christopher Robin would do ?'

Winnie–the–Pooh, A.A.Milne

E.1 Introduction

So far in these appendices I have given the tools for calculating process like,

$$\begin{aligned} e^+e^- &\to q\bar{q} \\ e^-e^- &\to ZH \quad , \end{aligned} \tag{E.1.1}$$

however as well as e^+e^- colliders we also have pp and $p\bar{p}$ colliders. Indeed as pp colliders have a higher $\sqrt{s}$ energy the new physics potential is greater than at e^+e^- machines. With pp and $p\bar{p}$ collisions we need some way to model the proton; this is because the proton is not a fundamental object; but is made up of quarks and gluons, collectively called partons. It is these partons that undergo the scattering process, rather than the protons; and so we need

some way to model the partons within protons. In this appendix I review how we describe these partons in terms of parton distributions, this is based upon the books Ref. 3.

E.2 The Probabilistic Approach

If we consider a process involving protons in the initial state,

$$pe \to X \quad , \tag{E.2.1}$$

then we can split this process into two halves,

$$\begin{aligned} p &\to q + \text{debris} \\ p &\to g + \text{debris} \end{aligned} \quad , \tag{E.2.2}$$

and

$$\begin{aligned} qe &\to X \\ ge &\to X \end{aligned} \quad . \tag{E.2.3}$$

Now the rate for (E.2.3) we can calculate assuming that the quarks or gluons are the incoming particles using the standard methods of evaluating the matrix element, squaring, and integrating over the phase space. This gives us a subprocess cross-section, $\hat{s}$. We model the processes (E.2.2) through probability distributions. As we are typically at very high energies ($\mathcal{O}(\text{ TeV})$) the quarks and gluons have almost no transverse momentum ($\mathcal{O}(300\,\text{MeV})$) in comparison to longitudinal momentum. We define x to be,

$$x = \frac{p_{\text{parton}}}{p_{\text{proton}}} \quad , \tag{E.2.4}$$

then the quark and gluon distributions are,

$$\begin{aligned} g(x)\delta x &= \mathcal{P}(\text{gluon with momentum fraction } [x, x+\delta x]) \\ q(x)\delta x &= \mathcal{P}(\text{quark with momentum fraction } [x, x+\delta x]) \end{aligned} \quad . \tag{E.2.5}$$

Then the total cross-section is given by;

$$\begin{aligned} \sigma(pe \to X) = & \int dx g(x)\hat{\sigma}(ge \to X)\Big|_{p_g = x p_p} \\ & + \int dx q(x)\hat{\sigma}(qe \to X)\Big|_{p_q = x p_p} \end{aligned} \quad . \tag{E.2.6}$$

Similarly for a process with 2 protons (or a proton antiproton pair) colliding we get the total cross-section as,

$$\sigma(pp \to X) = \int dx_1 \int dx_2 g(x_1)g(x_2)\hat{\sigma}(gg \to X)\Big|_{\substack{p_{g_1} = x_1 p_{p_1} \\ p_{g_2} = x_2 p_{p_2}}} + \ldots \quad . \tag{E.2.7}$$

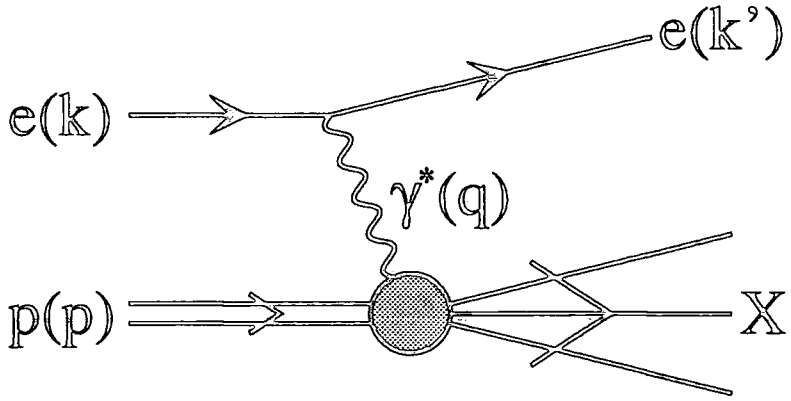


Fig.E.3.1 The Feynman diagram for *ep* scattering via an exchanged virtual photon. The particle momenta are show in brackets.

Now we have to ask *How do we find out about g and q?* Well we can measure them in one experimental process; and then apply them to another – however this begs the question what else do g and q depend upon other than x. The only thing that the proton really knows about is how hard it is hit in the collision, *i.e.*, the Q^2 of the process. *Can this* Q^2 *affect* q *and* g *?*. If the proton were just some classical bag that contained some mixture of non interacting quarks and gluons then the answer would be that the q and g don't know about Q^2, and that as Q^2 is varied that q and g remain unaltered; this is known as Bjorken scaling. Unfortunately a proton isn't just a classical bag of partons – but a more complicated quantum object. This means that this does not happen; and q and g know about Q^2.

E.3 The Altarelli–Parisi Equation

If we consider *ep* scattering, which is shown in *Fig*.E.3.1,

$$e(k) + p(p) \rightarrow e(k') + X \quad , \tag{E.3.1}$$

it is useful to define a few variables,

$$\begin{array}{ll} q = k - k' & Q^2 = -q^2 \\ x = \frac{Q^2}{2p.q} & y = \frac{p.q}{p.k} \end{array} \quad . \tag{E.3.2}$$

Now at leading order this Bjorken x corresponds to the fraction of the proton's momentum that the struck quark carries; it is useful to define it this way as it is then defined in terms of experimental observables; and so has an unambiguous definition when we go beyond leading order.

At leading order we have the subprocess shown in *Fig*.E.3.2.a; and if we calculate F_2 for

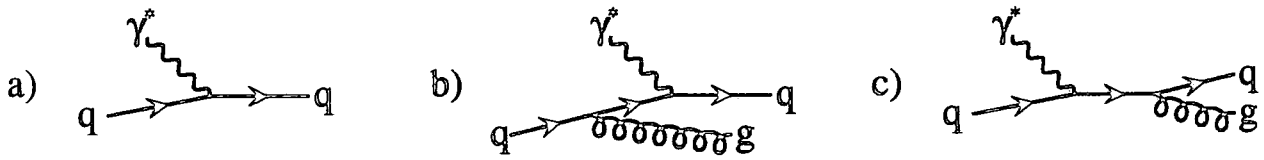


Fig.E.3.2 The Feynman diagrams for the subprocess in ep scattering, shown in a) at leading order, and b&c) when an extra gluon is also radiated

a particular x value we find;

$$\frac{F_2(x,Q^2)}{x} = \sum_i e_i^2 \int_x^1 \frac{dy}{y} q_i(y)\delta(1-x/y) = \sum_i e_i^2 q_i(x) \quad , \tag{E.3.3}$$

where we have summed over the various quarks in the proton. Reassuringly we find that at leading order we have probed the parton distribution at momentum fraction x. Also at leading order F_2 is independent of Q^2; as we expect from Bjorken scaling. If we move beyond leading order we also get a contribution where an extra gluon is radiated; the subprocess is shown in *Fig*.E.3.2.b. When the radiated gluon is approximately collinear to the initial state quark the dominant contribution comes from radiation of this initial quark shown in *Fig*.E.3.2.b. Then we can approximate the subprocess cross-section as,

$$d\hat{\sigma}(\gamma^* q \to qg) \simeq e_i^2 \hat{\sigma}_0 \frac{\alpha_S}{2\pi} P_{qq}(z) \frac{dp_\perp^2}{p_\perp^2} \quad , \tag{E.3.4}$$

where $p_\perp$ is the transverse momentum of the quark. $P_{qq}(z)$ is the quark to quark splitting function, in some sense it is proportional to the probability for a quark to fragment into a quark plus a gluon, where the final quark carries a fraction z of the momentum of the initial quark. At leading order it is given by,

$$P_{qq}(z) = \frac{4}{3}\left(\frac{1+z^2}{1-z}\right) \quad . \tag{E.3.5}$$

Now as $p_\perp \to 0$ this differential cross-section diverges; which we know physically does not happen! This is because we have missed off the virtual diagrams shown in *Fig*.E.3.3 which contributes with the same power of α and so should also be considered. In practice for this level of calculation this diagram does not need to be calculated. Already approximations have been made by using the collinear splitting function; so we do not seek an accurate answer, merely one that gets the gist of things correct. In this case we know that the final answer must be finite; so if we regulate the divergent piece and then just throw the divergence away, we make an error of $\mathcal{O}(\alpha_S)$, but correctly take care of all terms of the form $\ln(Q^2)\alpha_S$. This means that we are only working to leading log accuracy. If we regulate the divergence by having some minimum $p_\perp = \mu$ then the subprocess cross-section is given by,

$$\hat{\sigma}(\gamma^* q \to qg) \simeq e_i^2 \hat{\sigma}_0 \left(\frac{\alpha_S}{2\pi} P_{qq}(z) \ln \frac{Q^2}{\mu^2}\right) \quad . \tag{E.3.6}$$

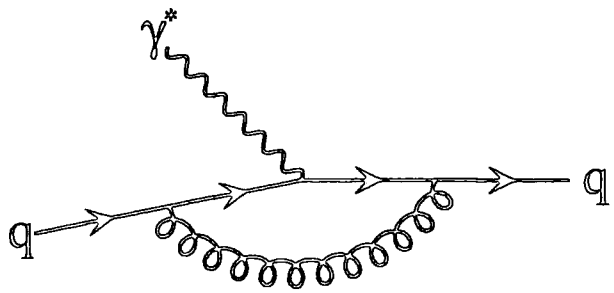


*Fig.*E.3.3 The virtual Feynman diagram that cancels the soft and collinear singularities in the tree level diagrams where a gluon is radiated.

and this gives F_2 at the leading log level as,

$$\frac{F_2(x,Q^2)}{x} = \sum_q e_q^2 \int_x^1 \frac{dy}{y} q(y) \left(\delta(1-x/y) + \frac{\alpha_S}{2\pi} P_{qq}(z) \ln \frac{Q^2}{\mu^2} \right) \quad . \tag{E.3.7}$$

This gluon radiation of the initial state quark is general to all processes which have an incoming quark, and so it makes sense to factorize this contribution off into the definition of the parton distribution. To do this we write equation (E.3.7) as,

$$\frac{F_2(x,Q^2)}{x} = \sum_q e_q^2 (q(x) + \Delta q(x,Q^2)) \quad , \tag{E.3.8}$$

where

$$\Delta q(x,Q^2) = \int_x^1 \frac{dy}{y} q(y) \left(\frac{\alpha_S}{2\pi} P_{qq}(z) \ln \frac{Q^2}{\mu^2} \right) \quad . \tag{E.3.9}$$

This means that if we look at $q(x,Q^2)$ as we vary Q^2 we find,

$$\frac{dq(x,Q^2)}{d\ln Q^2} = \frac{\alpha_S}{2\pi} \int_x^1 \frac{dy}{y} q(y,Q^2) P_{qq}\left(\frac{x}{y}\right) \quad . \tag{E.3.10}$$

This is known as one of the Altarelli–Parisi evolution equations. It tells us how the parton distribution varies as we change Q^2. Notice that the cut off parameter μ does not appear in this equation! It turns out to be convenient to define flavour singlet, and flavour non singlet distributions,

$$\begin{aligned} F^{NS} &= q_i - q_j \\ F^S &= \sum_i (q_i + \bar{q}_i) \end{aligned} \quad . \tag{E.3.11}$$

Then the full Altarelli–Parisi equations read,

$$\begin{aligned} \frac{\partial F^{NS}}{d\ln Q^2} &= \frac{\alpha_S(Q^2)}{2\pi} P_{qq} * F^{NS} \\ \frac{\partial F^S}{d\ln Q^2} &= \frac{\alpha_S(Q^2)}{2\pi} (P_{qq} * F^S + 2n_f P_{qg} * g) \\ \frac{\partial g}{d\ln Q^2} &= \frac{\alpha_S(Q^2)}{2\pi} (P_{gq} * F^S + P_{gg} * g) \end{aligned} \quad , \tag{E.3.12}$$

where n_f is the number of active flavours and $*$ denotes the convolution integral,

$$f * g = \int_x^1 \frac{dy}{y} f(y) g(x/y) \quad . \tag{E.3.13}$$

The splitting functions P_{ab} that give us the splitting of particle b into particle a are,

$$\begin{aligned} P_{qq} &= \frac{4}{3}\left(\frac{1+x^2}{1-x}\right) \\ P_{qg} &= \frac{1}{2}\left(x^2+(1-x)^2\right) \\ P_{gq} &= \frac{4}{3}\left(\frac{1+(1-x)^2}{x}\right) \\ P_{gg} &= 6\left(\frac{1-x}{x}+x(1-x)+\frac{x}{1-x}\right)-\left(\frac{1}{2}+\frac{n_f}{3}\right)\delta(1-x) \quad . \end{aligned} \tag{E.3.14}$$

There still remains one last problem with these splitting functions, many of them diverge when we do the integral over them up to $z = 1$. This is because for these splitting functions as $z \to 1$ we have a gluon whose energy goes to zero. Now cross-sections with radiation not only have singularities when the radiation becomes collinear to massless charged particles; but also have divergences when the radiation becomes soft, as is the case here. Again these soft singularities are canceled via the virtual diagram *Fig*.E.3.3; and as with the collinear singularities to leading log accuracy we can deal with these singularities just by throwing them away. In this case a convenient prescription for doing this is the *plus prescription*; where if we have the integral over some function that diverges as $x \to 1$ we use,

$$\int_0^1 dx f(x)(g(x))_+ = \int_0^1 dx (f(x) - f(1)) g(x) \quad . \tag{E.3.15}$$

This makes the convolutions over the splitting functions finite, and accurate to leading logs.

As we have factorized the initial state collinear divergence into our definition of the parton distribution we need to take care when we do calculations of a subprocess cross-section to NLO, not to double count the this divergence and include it in the subprocess cross-section.

E.4 Parton Distributions in Practice

With the Altarelli–Parisi equations we are still left with the problem that they are only a set of evolution equations. They give no firm prediction of our parton distributions, but only tell us, if we know the distribution at one Q^2, how to evolve it to another Q^2. Hence *to make any prediction about parton distributions we need to measure them !* In practice the philosophy usually used is to parametrise the parton distributions at some low $Q^2 = Q_0^2$ (typically chosen

at $Q_0^2 = 4\ \mathrm{GeV}^2$), and then evolve them up to higher Q^2 using the Altarelli–Parisi equations, and use them to make predictions for measured processes. The parametrised distributions at Q_0^2 are then varied so that the predictions at higher Q^2 agree with the experimentally measured processes. The set of distributions that result can then be used to make predictions for processes as yet unmeasured. As such we have a working model of the proton, however it is far more dependent upon experimental measurements than much of the theory of high energy physics! In practice all the better parton distributions work beyond leading log accuracy, and also include the sub leading log terms as well†. Throughout this thesis I have used one such set of parton distributions, the D0 set of MRS92′ [37]. For the processes I have considered at the LHC and SSC these parton distributions need to be evolved far from the regions where they are experimentally constrained, as such there is a fairly large uncertainty (50% or more) in all cross-sections that I have presented arising from the parton distributions. However for all processes I have used identical parton distributions, evaluated at the same scale (chosen to be $Q^2 = \hat{s}$ the subprocess energy); and so we should expect uncertainties in the cross-sections to largely cancel in signal to background ratios, as these typically will have the same initial states; and even if this is not true then the nature of the Altarelli–Parisi evolution equations is that if we increase one thing (say the gluon distribution) then we increase every distribution.

† It may be a worry that we are using NLL parton distributions, and convoluting them with LO subprocess cross-sections; and as such are mixing the order at which we are working. The view that I have taken here is that within the probabilistic approach to parton distributions the parton densities merely reflect the probability of getting a parton out of the proton. This parton emerges *on shell* (*i.e.*, its off shelledness is far smaller than the subprocess energy), and so is a physical particle. In this frame work I can ask, what is the probability of getting a parton from the proton *using* NLL evolution equations, then having obtained this parton what is the cross-section for it to interact at LO. If I were to move beyond LO subprocess cross-section then this simplistic approach can no longer held, and we must take far greater care about how we use our parton distributions in conjunction with our subprocess cross-section.

Appendix F

Measurement of Probability Distributions

Pooh began to feel a little more comfortable, because when you are a Bear of Very Little Brain, and you Think of Things, you find sometimes that a Thing which seemed very Thingish inside you is quite different when it gets out into the open and has other people looking at it.

The House At Pooh Corner, A.A.Milne

F.1 Introduction

In the previous Appendix we saw how when forming a differential cross-section we had to decide upon a bin width in which we binned our cross-section; if we choose the bin width too large we can not discern the finer details of the differential cross-section, if we choose the bin width too narrow our differential cross-section has large errors, and again does not represent the differential cross-section well. This is a problem that occurs both for theoreticians, while doing Monte Carlo integrals; and for experimentalists, while analyzing their data. In this appendix I wish to put forward an alternative approach to extracting differential cross-sections, or more generally of extracting a probability distribution from a random variable.

Usually when forming a probability distribution we *have a feeling* for the form of that probability distribution; for example that it is smooth, or that it is described by a relatively simple set of polynomials. However the random variable whose distribution we are trying to find only gives a string of values $x_1, x_2, \cdots, x_{\text{nev}}$, how do we proceed to the distribution ?

F.2 Binning and least χ^2 method

Traditionally the method is to first bin the string of data values; and then to fit a smooth distribution to the histogramed data. This is usually done by choosing possible parametrisations for the probability distribution; and then minimising the χ^2 to choose between these possible parametrisations.

If we consider a case where we are trying to form $\frac{dX}{dx}$, where we have nev data events, where we bin in nbin bins, and where x has been scaled so its largest and smallest values are $1, 0$ respectively. When we bin the data we form,

$$y_i = \frac{\text{nbin}}{\text{nev}} \sum_{x_j \text{ in } i\text{'th bin}} 1 \qquad z_i = \text{central value of } x \text{ in the bin } i \quad . \tag{F.2.1}$$

The error in n events is approximately $\sqrt{n}$ so

$$\sigma_i = \frac{\text{nbin}}{\text{nev}} \sqrt{\sum_{x_j \text{ in } i\text{'th bin}} 1} \quad . \tag{F.2.2}$$

If we now take a parametrisation of $\frac{dX}{dx}$,

$$\frac{dX}{dx} = \sum_{k=1}^{M} a_k P_k(x) \quad , \tag{F.2.3}$$

where the a_k are unknown coefficients, and the P_k predetermined functions. To get a measure of how well a particular parametrisation fits the histogrammed data we form the χ^2 as,

$$\chi^2 = \sum_{i=1}^{\text{nbin}} \left[\frac{y_i - \sum_{k=1}^{M} a_k P_k(z_i)}{\sigma_i} \right]^2 \quad . \tag{F.2.4}$$

To get the *best fit* of the parametrisation (F.2.3) to the histogram we minimise the value of this χ^2 with respect to variations in the a_k. That is we require,

$$0 = \frac{\partial \chi^2}{\partial a_l} \tag{F.2.5}$$

$$= -2 \sum_{i=1}^{\text{nbin}} \frac{\left(y_i - \sum_{k=1}^{M} a_k P_k(z_i) \right) P_l(z_i)}{\sigma_i^2} \quad . \tag{F.2.6}$$

Rearranging this expression we get the set of simultaneous equations,

$$\sum_{j=1}^{M} \alpha_{kj} a_j = \beta_k \quad , \tag{F.2.7}$$

where

$$\alpha_{kj} = \sum_{i=1}^{\text{nbin}} \frac{P_j(z_i)P_k(z_i)}{\sigma_i^2} \quad , \tag{F.2.8}$$

and

$$\beta_k = \sum_{i=1}^{\text{nbin}} \frac{y_i P_k(z_i)}{\sigma_i^2} \quad . \tag{F.2.9}$$

Now the set of M simultaneous equations (F.2.7) has M unknowns, a_j, and so can be exactly solved, by example inverting the matrix α_{kj} of values; and we thus get the best fit of the parametrisation (F.2.3) to our string of data points. How good is this fit ? Well for our fit we have a measured χ^2 from equation (F.2.4); this quantity is again a probability distribution, which (surprisingly) can be described analytically [38] in terms of just one variable, the *number of degrees of freedom*, ν, which for this example equals nbin – M (we have nbin fixed y_i values in our χ^2 minus M free varying a_j values). Although we can describe the probability rigorously in terms of the *number of degrees of freedom*, ν, this degree of rigor is not normally required – it is enough to know that its mean, $\langle\chi^2\rangle$, is ν, and its variance $\text{Var}(\chi^2)$ is 2ν (and so the 1sd error is $\sqrt{2\nu}$). Of course as the *number of degrees of freedom*, $\nu \to \infty$, the χ^2 distribution becomes normally distributed again by the Central Limit Theorem. So if after fitting our parametrised function (F.2.3) we find a χ^2 that is considerably larger than ν then we know that our function (F.2.3) is not general enough to describe the data that we have.

Also from equation (F.2.7) we can find the variance, and hence 1sd error, in the values a_i, we find,

$$\text{Var}(a_i) = (\alpha^{-1})_{ii} \quad , \tag{F.2.10}$$

where (α^{-1}) is the inverse of the matrix α.

Now this method of extracting the probability distribution has several faults, namely,

- How may bins should we bin our data in. At best there is a handwaving argument that says that about 10 events on average per bin is optimal – however this is very ad hoc, and is not always optimal (*e.g.*, see the example in section F.4)
- When we bin our data we lose knowledge about the exact value of our data, all data that falls in a bin is treated in an identical way, data that is close to falling out of the bin one one side is identical to data in the centre of the bin. This destroys some of the information content of the data – which is wasteful.

The problem seems to be that binning our data is a rather arbitrary procedure, here I suggest an alternative method that avoids binning the data altogether.

F.3 Orthogonal Polynomials and Monte Carlo Integration

In equation (F.2.3),

$$\frac{dX}{dx} = \sum_{k=1}^{M} a_k P_k(x) \quad , \tag{F.2.3}$$

we took the polynomials $P_k(x)$ to be arbitrary. Now, without loss of generality we can take x to be in the range $[0,1]$, as we can always scale the variable x so that it lies in this range. If we now take the polynomials $P_k(x)$ to be orthogonal functions on the interval $[0,1]$ with respect to some weighting function $W(x)$, that is,

$$\int_0^1 P_i(x)P_j(x)W(x)dx = \delta_{ij} l_i \quad , \tag{F.3.1}$$

then we can extract the a_i from equation (F.2.3) using these orthogonality relationships, we find,

$$a_i = \frac{1}{l_i}\int_0^1 \frac{dX}{dx} W(x)P_i(x)dx \tag{F.3.2}$$

$$= \frac{1}{l_i}\int P_i(x)W(x)dX \quad . \tag{F.3.3}$$

Also we can always write $\frac{dX}{dx}$ as a sum over orthogonal polynomials if we can write it as sum of arbitrary polynomials because we can orthogonalise these polynomials for example by Gram-Schmidt orthogonalisation[39]. If we now consider doing the integral (F.3.3) by Monte Carlo integration we need to choose points uniformly in X – however this is exactly what our string of data values $x_1, x_2, \cdots, x_{\text{nev}}$ are. As such we can estimate the a_i from,

$$a_i = \frac{1}{l_i}\langle P_i(x)W(x)\rangle \quad . \tag{F.3.4}$$

Also we know that the 1sd error on a_i is given by,

$$\Delta a_i = \frac{1}{l_i}\sqrt{\frac{\langle (P_i(x)W(x))^2\rangle - \langle P_i(x)W(x)\rangle^2}{\text{nev}}} \quad . \tag{F.3.5}$$

Using this method we can extract probability distributions without having to resort to binning the data, we just form sums of the orthogonal polynomials. Also the error term is easy to calculate.

It turns out that the relationship between the two different approaches can be made explicit.

In the traditional method we consider taking the limit as nbins $\to \infty$ then we find,

$$\frac{\alpha_{kj}}{\text{nbin}} \to \int \frac{P_j P_k}{\sigma^2} dx \quad . \tag{F.3.6}$$

On substituting this into equation (F.2.7) we get,

$$a_j \int \frac{P_j P_k}{\sigma^2} dx = \frac{1}{\text{nev}} \sum_{\text{events}} \frac{P_k(x_i)}{\sigma^2(x_i)} \quad , \tag{F.3.7}$$

and this becomes identical to equation (F.3.4) if we take $\sigma^2 = \text{constant}$ and $W(x) = 1$. In fact we need to make some assumption along these lines about the error on bins as we take the limit nbins $\rightarrow \infty$ as if this limit infinite numbers of bins have zero events in them and we can no longer use $\sqrt{n} = 0$ to estimate the error on the measurement in each bin. Indeed equations (F.2.8,F.2.9) are badly defined as the error terms go to zero.

So the new approach is equivalent to the traditional approach in the limit nbin $\rightarrow \infty$ and where the error term is taken to be a constant.

It is also worth noticing that histogramming the data is in fact a sub case of the new approach where we choose as our orthogonal polynomials,

$$P_i(x) = \begin{cases} 1 & \text{if } x \in i^{\text{th}} \text{ bin} \\ 0 & \text{if } x \notin i^{\text{th}} \text{ bin} \end{cases} \quad , \tag{F.3.8}$$

although here the histogram is the final step of our procedure, we don't go on to fit parametrised curves to it.

F.4 A Practical Example

As a practical example of these two approaches in action, if we take the probability distribution given by,

$$dX/d\theta = 1/\pi(1/2 + 1/6 \sin\theta + 1/6 \cos\theta + 1/7 \sin 7\theta + 1/8 \cos 20\theta) \quad , \tag{F.4.1}$$

on the interval $[0, 2\pi]$. For which we have a theoretical bias that distribution is of the form,

$$dX/d\theta = 1/2\pi + \sum_n a_n \cos n\theta + \sum_n b_n \sin n\theta \quad . \tag{F.4.2}$$

Now $\sin n\theta$ and $\cos n\theta$ are orthogonal polynomials on the interval $[0, 2\pi]$ with the weighting function $W(x) = 1$, and so we can find a_n and b_n from,

$$a_n = \frac{1}{\pi} \int \cos n\theta \, dX \tag{F.4.3}$$

$$b_n = \frac{1}{\pi} \int \sin n\theta \, dX \quad . \tag{F.4.4}$$

With such a distribution as (F.4.1) we are clearly going to have to consider values of n up to at least 20, here we will chose the maximum value of n to be 23,or in the traditional approach

choose $M = 46$ (23 sin and 23 cos coefficients); notice that this choice of the maximum value for n does not alter our prediction for the lower values of a_i and b_i in the new approach, whereas in the traditional approach changing the value of M changes drastically the problem that we solve, and changes *all* values of a_i and b_i. Also in the traditional approach we are going to have to choose more than 46 bins when we histogram the random variables – as otherwise we will have more free variables in equation (F.2.7) than constraints, here we will choose $nbin = 70$. If we now generate 500 random points in the interval $[0, 2\pi]$ according to the distribution (F.4.1) and extract the values of a_i and b_i via both approaches we find the results shown in *Table* F.4.1

The significance column gives the number of standard deviations that we are away from $a_n, b_n = 0$. Both methods have clearly found the non zero terms with a more than 4 sigma effect, and both have similar errors on their measurements, although the traditional approach seems to have done slightly worse for large n. In practice the traditional approach was far harder to implement, as the set of simultaneous equations (F.2.7) took a long time to solve, also the results depend quite sensitively on the number of bins chosen – if we had chosen 50 bins which the *handwaving* argument of 10 events per bin suggests then we have great difficulty extracting the $1/8 \cos 20\theta$ term, as we are fitting almost as many variables, a_n and b_n, as we have constrains, nbin. The results in the traditional method also have a slight dependence upon the number of polynomials fitted,M, for $M > 20$; this vagueness is removed in the new approach.

Finally we show in *Fig*.F.4.1 the measured probability distributions, clearly both have felt the general gist of the probability distribution but have failed to get the fine details right for this small number of events. As we increase the numbers of events the two methods still do comparably, as long as we choose the number of bins optimally, and make a reasonable choice for the order of curve to fit. The new method does not require nearly as much "fiddling" as the traditional approach, and always performs as well, if not better, than the traditional approach. To summarise,

> *The employment of a middle man* (in this case the binning of the data) *at best keeps prices static, but more usually increases prices by an unknown amount.*

F.5 High Energy Physics Applications

As has been already mentioned this new approach has applications both for theory and experiment. On the experimental side we often want to extract differential cross-sections, and the new approach can be applied directly. The only question to ask is; *What orthogonal*

The New Approach						
n	a_n	$\Delta(a_n)$	Significance	b_n	$\Delta(b_n)$	Significance
0	0.1592	0.0000	∞			
1	0.0449	0.0096	4.7	0.0504	0.0101	**5.0**
2	−0.0168	0.0096	1.7	0.0058	0.0104	0.6
3	−0.0079	0.0101	0.8	0.0174	0.0100	1.7
4	−0.0243	0.0099	2.5	0.0004	0.0102	0.0
5	−0.0181	0.0100	1.8	−0.0129	0.0101	1.3
6	0.0044	0.0099	0.4	−0.0139	0.0102	1.4
7	−0.0134	0.0099	1.4	0.0531	0.0099	**5.4**
8	−0.0085	0.0101	0.8	−0.0224	0.0100	2.2
9	0.0077	0.0100	0.8	0.0073	0.0101	0.7
10	−0.0008	0.0107	0.1	0.0051	0.0093	0.5
11	0.0128	0.0100	1.3	−0.0164	0.0101	1.6
12	−0.0075	0.0099	0.8	−0.0106	0.0102	1.0
13	−0.0161	0.0103	1.6	0.0009	0.0098	0.1
14	−0.0085	0.0100	0.8	−0.0052	0.0102	0.5
15	−0.0046	0.0100	0.5	0.0129	0.0101	1.3
16	0.0018	0.0099	0.2	0.0145	0.0102	1.4
17	0.0066	0.0099	0.7	0.0004	0.0102	0.0
18	−0.0039	0.0098	0.4	0.0009	0.0103	0.1
19	0.0029	0.0102	0.3	−0.0121	0.0099	1.2
20	0.0440	0.0096	**4.6**	0.0049	0.0103	0.5
21	0.0043	0.0098	0.4	0.0086	0.0103	0.8
22	−0.0029	0.0101	0.3	−0.0101	0.0101	1.0
23	0.0004	0.0102	0.0	−0.0137	0.0100	1.4

The Traditional Approach						
n	a_n	$\Delta(a_n)$	Significance	b_n	$\Delta(b_n)$	Significance
0	0.1592	0.0000	∞			
1	0.0426	0.0096	**4.5**	0.0536	0.0100	**5.4**
2	−0.0171	0.0096	1.8	0.0076	0.0103	0.7
3	−0.0080	0.0101	0.8	0.0208	0.0099	2.1
4	−0.0228	0.0097	2.4	0.0028	0.0102	0.3
5	−0.0148	0.0099	1.5	−0.0151	0.0100	1.5
6	0.0053	0.0099	0.5	−0.0165	0.0100	1.6
7	−0.0132	0.0100	1.3	0.0487	0.0099	**4.9**
8	−0.0125	0.0101	1.2	−0.0255	0.0101	2.5
9	0.0055	0.0100	0.5	0.0122	0.0103	1.2
10	−0.0016	0.0109	0.1	0.0084	0.0095	0.9
11	0.0130	0.0101	1.3	−0.0137	0.0105	1.3
12	−0.0076	0.0103	0.7	−0.0049	0.0105	0.5
13	−0.0152	0.0106	1.4	0.0033	0.0103	0.3
14	0.0004	0.0106	0.0	−0.0059	0.0105	0.6
15	0.0002	0.0108	0.0	0.0124	0.0105	1.2
16	−0.0003	0.0105	0.0	0.0073	0.0109	0.7
17	−0.0045	0.0108	0.4	−0.0047	0.0106	0.4
18	−0.0133	0.0108	1.2	0.0015	0.0108	0.1
19	0.0142	0.0109	1.3	−0.0092	0.0109	0.8
20	0.0450	0.0108	**4.2**	0.0111	0.0111	1.0
21	−0.0017	0.0110	0.2	0.0083	0.0113	0.7
22	−0.0101	0.0113	0.9	−0.0103	0.0114	0.9
23	0.0088	0.0114	0.8	−0.0133	0.0110	1.2

Table F.4.1 The values of a_i and b_i obtained by the traditional approach and the new approach

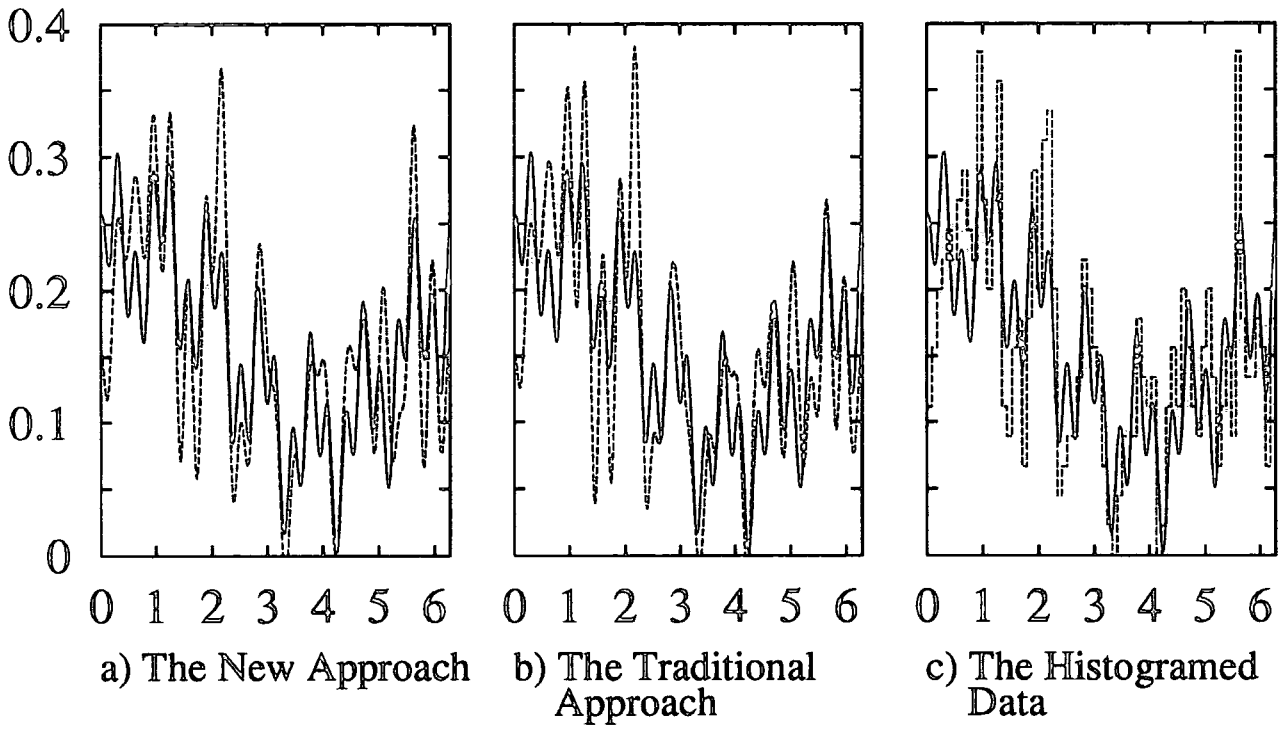


*Fig.*F.4.1 The fits to the data. The actual probability distribution is shown as a solid line and the measured probability distribution in the two approaches is shown dashed. Also show is the histogrammed data to which the traditional approach has fitted its curve.

polynomials P_n should we choose ? Clearly we should be motivated by the physics that we are searching for, if we are looking in an invariant mass distribution for a new particle that peaks at a particular mass then we need to use orthogonal polynomials that describe peaked distributions, an obvious choice is equation (F.3.8) – *i.e.*, that we just bin the data as in the traditional approach. However often, when for example doing spin physics, we will also be interested in azimuthal correlations in say a couple of decay planes. In this case we expect the differential cross-section to look like,

$$d\sigma \sim (A + B\cos\theta + C\cos^2\theta + \mathcal{O}(\cos^3\theta))\, d\cos\theta \quad , \tag{F.5.1}$$

in which case the obvious polynomials to use are Legendre polynomials, $P_n(\cos\theta)$. Similarly when looking at some p_T spectrum we should use polynomials that die away as their argument becomes large.

Theoretically we should use largely the same set of polynomials, however we can often go further because we know the process we are calculating, so for example when calculating the $pp \to t\bar{t}\gamma\gamma$ background to $pp \to t\bar{t}H$ production we know that the photon photon invariant mass, $M_{\gamma\gamma}$, has no peaks and is smooth; and so we take no risk in using Legendre polynomials

in some limited range. Clearly such a course would be fool hardy for the experimenter who is looking for a peak in this distribution as a signal for the Higgs.

One final point should be made, while extracting the coefficients to these polynomials both theoretically and experimentally may give exactly the same answer for its value; they will not give the same value for the error term. This is because our events are chosen according to different distributions, experimentally say we calculate,

$$I = \int P(x)W(x)d\sigma \quad . \tag{F.5.2}$$

Now the 1 shot variance for this is given by,

$$\mathrm{Var}_{\mathrm{exp}} = \sigma \int (P(x)W(x))^2 d\sigma - \left(\int P(x)W(x)d\sigma \right)^2 \quad . \tag{F.5.3}$$

Where as theoretically we calculate,

$$I = \int P(x)W(x)\frac{d\sigma}{dy}dy \quad , \tag{F.5.4}$$

whose 1 shot variance is given by,

$$\mathrm{Var}_{\mathrm{th}} = \int dy \int (P(x)W(x)\frac{d\sigma}{dy})^2 dy - \left(\int P(x)W(x)\frac{d\sigma}{dy}dy \right)^2 \quad . \tag{F.5.5}$$

Clearly their values are different and this means that if we can extract the value of I to a certain precision by using a certain number of theoretical Monte Carlo points that it does not follow that the same number of events in reality will give the same precision. However we can work out theoretically how many events we would require experimentally to measure I to a certain precision because we can evaluate (F.5.3) by Monte Carlo integration thus,

$$\mathrm{Var}_{\mathrm{exp}} = \sigma \int (P(x)W(x))^2 \frac{d\sigma}{dy}dy - \left(\int P(x)W(x)\frac{d\sigma}{dy}dy \right)^2 \quad . \tag{F.5.6}$$

Appendix G

Finite Width corrections to $H \to Z\gamma$

Christopher Robin came slowly down his tree.

'Silly old Bear,' he said, 'what were you doing? First you went round the spinney twice by yourself, and then Piglet ran after you and you went round again together, and then you were just going round a fourth time—'

'Wait a moment,' said Winnie–the Pooh, holding up his paw.

He sat down and thought, in the most thoughtful way he could think, then he fitted his paw into one of the Tracks ... and then he scratched his nose twice, and stood up.

'Yes,' said Winnie–the–Pooh.

'I see now,' said Winnie–the–Pooh.

'I have been Foolish and Deluded,' said he, 'and I am a Bear of no Brain at All.'

Winnie–the–Pooh, A.A.Milne

Finite width effects can be very important for Higgs decays. Because the Higgs boson couples more strongly to more massive particles we can get significant rates for a Higgs to decay to off shell massive particles, which if that particle were an asymptotic particle (*i.e.*, its width is 0 and it is stable) would be kinematically forbidden. This can be seen in the decays,

$$H \to Z^* Z^* \tag{G.1}$$

$$H \to W^* W^* \quad , \tag{G.2}$$

which are very important for Higgs masses, $H_H < 2M_Z, 2M_W$ respectively. Indeed $Br(H \to W^*W^*) > 10\%$ for $M_H \gtrsim 115\,\text{GeV}$ despite the fact that this is $45\,\text{GeV}$ beneath threshold! Now the branching ratio $Br(H \to Z\gamma)$ is at most 3×10^{-3} for $M_H \sim 150\,\text{GeV}$; and so finite width effects are going to have a minimal effect for the total Higgs width. However for

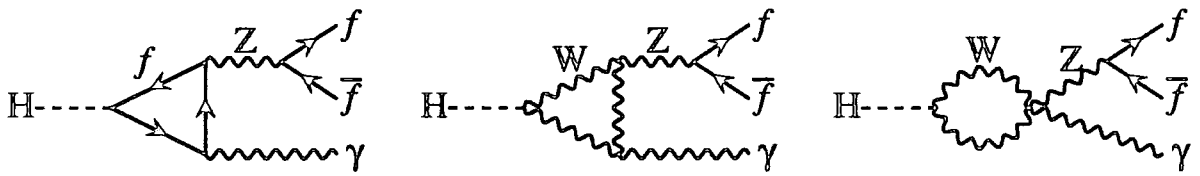


*Fig.*G.1 The Feynman diagrams for the process $H \to Z^*\gamma \to f\bar{f}\gamma$.

an intermediate mass Higgs, especially a light intermediate mass Higgs, the $H \to Z\gamma$ decay of the Higgs is one of the few clean decay channels of the Higgs. Clearly it is important to know this branching ratio as accurately as possible.

The Feynman diagrams for this decay are shown in *Fig.*G.1, it is convenient with this process to split the calculation up into two halves, into the process $H \to Z^*\gamma$ followed by the process $Z^* \to f\bar{f}$. The calculation of the process $H \to Z^*\gamma$ is identical to the calculation of $H \to Z\gamma$, which has been done many times before, see for example Ref. 9. Here I do not repeat the calculation but just quote the results. The effective coupling for $HZ\gamma$ vertex is given by,

$$\mathcal{M}^{\mu\nu}_{HZ\gamma} = A(p^\mu_Z p^\mu_\gamma - p_Z \cdot p_\gamma \, g^{\mu\nu}) \quad , \tag{G.3}$$

where A has the form [9],

$$A = \frac{\alpha g}{4\pi M_W}(A_F + A_W) \tag{G.4}$$

$$A_F = \sum_{\text{fermions}} n_c \frac{-2e_f(T^3_f - 2e_f \sin^2\theta_W)}{\sin\theta_W \cos\theta_W}[I_1(\tau_f, \lambda_f) - I_2(\tau_f, \lambda_f)] \tag{G.5}$$

$$\begin{aligned} A_W = &- \cot\theta_W \Big\{ 4(3 - \tan^2\theta_W) I_2(\tau_W, \lambda_W) \\ &+ \left[\left(1 + \frac{2}{\tau_W}\right) \tan^2\theta_W - \left(5 + \frac{2}{\tau_W}\right) \right] I_1(\tau_W, \lambda_W) \Big\} \quad , \end{aligned} \tag{G.6}$$

where,

$$\tau_f \equiv \frac{4m_f^2}{m_H^2} \qquad \lambda_f \equiv \frac{4m_f^2}{p_Z^2} \qquad \tau_W \equiv \frac{4m_W^2}{m_H^2} \qquad \lambda_W \equiv \frac{4m_W^2}{p_Z^2} \quad , \tag{G.7}$$

with,

$$I_1(a,b) = \frac{ab}{2(a-b)} + \frac{a^2b^2}{2(a-b)^2}[f(a) - f(b)] + \frac{a^2 b}{(a-b)^2}[g(a) - g(b)]$$

$$I_2(a,b) = -\frac{ab}{2(a-b)}[f(a) - f(b)]$$

$$f(\tau) = \begin{cases} \left[\sin^{-1}\left(\sqrt{1/\tau}\right)\right]^2 & \text{if } \tau \geq 1, \\ -\frac{1}{4}[\ln(\eta_+/\eta_-) - i\pi]^2 & \text{if } \tau < 1, \end{cases} \tag{G.8}$$

$$g(\tau) = \begin{cases} \sqrt{\tau - 1}\sin^{-1}(1/\sqrt{\tau}) & \text{if } \tau \geq 1, \\ \frac{1}{2}\sqrt{\tau - 1}[\ln(\eta_+/\eta_-) - i\pi] & \text{if } \tau < 1, \end{cases} \tag{G.9}$$

$$\eta_{\pm} \equiv (1 \pm \sqrt{1-\tau}) \quad . \tag{G.10}$$

Using this we find the matrix element for $H \to f\bar{f}\gamma$,

$$\mathcal{M} = \epsilon_{\gamma}^{\mu} A \left(p_Z^{\mu} p_{\gamma}^{\nu} - p_Z \cdot p_{\gamma}\, g^{\mu\nu} \right) \frac{1}{p_Z^2 - M_Z^2 + iM_Z\Gamma_Z} \bar{u}_f g_Z \gamma_{\nu} (g_V - g_A\gamma_5) u_{\bar{f}} \quad , \tag{G.11}$$

and this can be squared to give the matrix element squared as,

$$|\mathcal{M}|^2 = 8(g_V^2 + g_A^2) \frac{|A|^2 g_Z^2}{(p_Z^2 - M_Z^2)^2 + M_Z^2\Gamma_Z^2} p_f \cdot p_{\bar{f}} \left((p_{\gamma} \cdot p_f)^2 + (p_{\gamma} \cdot p_{\bar{f}})^2 \right) \quad . \tag{G.12}$$

Now the partial width is given by,

$$d\Gamma = \frac{1}{(2\pi)^5} \frac{1}{2M_H} \delta^4(p_H - p_f - p_{\bar{f}} - p_{\gamma}) |\mathcal{M}|^2 \frac{d^3p_f}{2E_f} \frac{d^3p_{\bar{f}}}{2E_{\bar{f}}} \frac{d^3p_{\gamma}}{2E_{\gamma}} \quad . \tag{G.13}$$

Most of these integrations can be done analytically; we get left with the integration over the Z line shape to get,

$$\Gamma(H \to Z^*\gamma) = \int_0^{M_H^2} dp_Z^2 \frac{\Gamma_Z}{32\pi^2 M_Z} \left(1 - \frac{p_Z^2}{M_H^2} \right)^3 \frac{A^2 p_Z^2 M_H^3}{(p_Z^2 - M_Z^2)^2 + M_Z^2\Gamma_Z^2} \quad . \tag{G.14}$$

This last integration is best done numerically. In the numerator we have collected several terms together to give the leading order Z width; this improves the accuracy of (G.14), as we know that there must be a Γ_Z in the numerator to cancel the Breit–Wigner propagator in the denominator. Having obtained this formula we are no longer constrained to use the leading order Z width; but instead can use the more realistic experimental measurement. If the Z width, Γ_Z, is narrow then we can use the narrow width approximation, where we make the approximations,

$$\int_0^{M_H^2} \frac{\Gamma_Z dp_Z^2}{(p_Z^2 - M_Z^2)^2 + M_Z^2\Gamma_Z^2} \to \begin{cases} 0 & \text{if } M_H < M_Z \\ \frac{\pi}{M_Z} & \text{if } M_H > M_Z \end{cases} \tag{G.15}$$

$$p_Z^2 \to m_Z^2 \quad , \tag{G.16}$$

then (G.14) becomes,

$$\Gamma(H \to Z\gamma) = \frac{A^2 M_H^3}{32\pi} \left(1 - \frac{M_Z^2}{M_H^2} \right)^3 \quad , \tag{G.17}$$

which, as we expect, is the width we get for $H \to Z\gamma$ where the Z is stable.

In *Fig*.G.2 we can see the numerical difference between the stable Z branching ratio, and including the Z width effects. For $M_H > 100\,\text{GeV}$ clearly both give very similar answers; this is as we would expect because the dominant width comes from producing the Z on mass

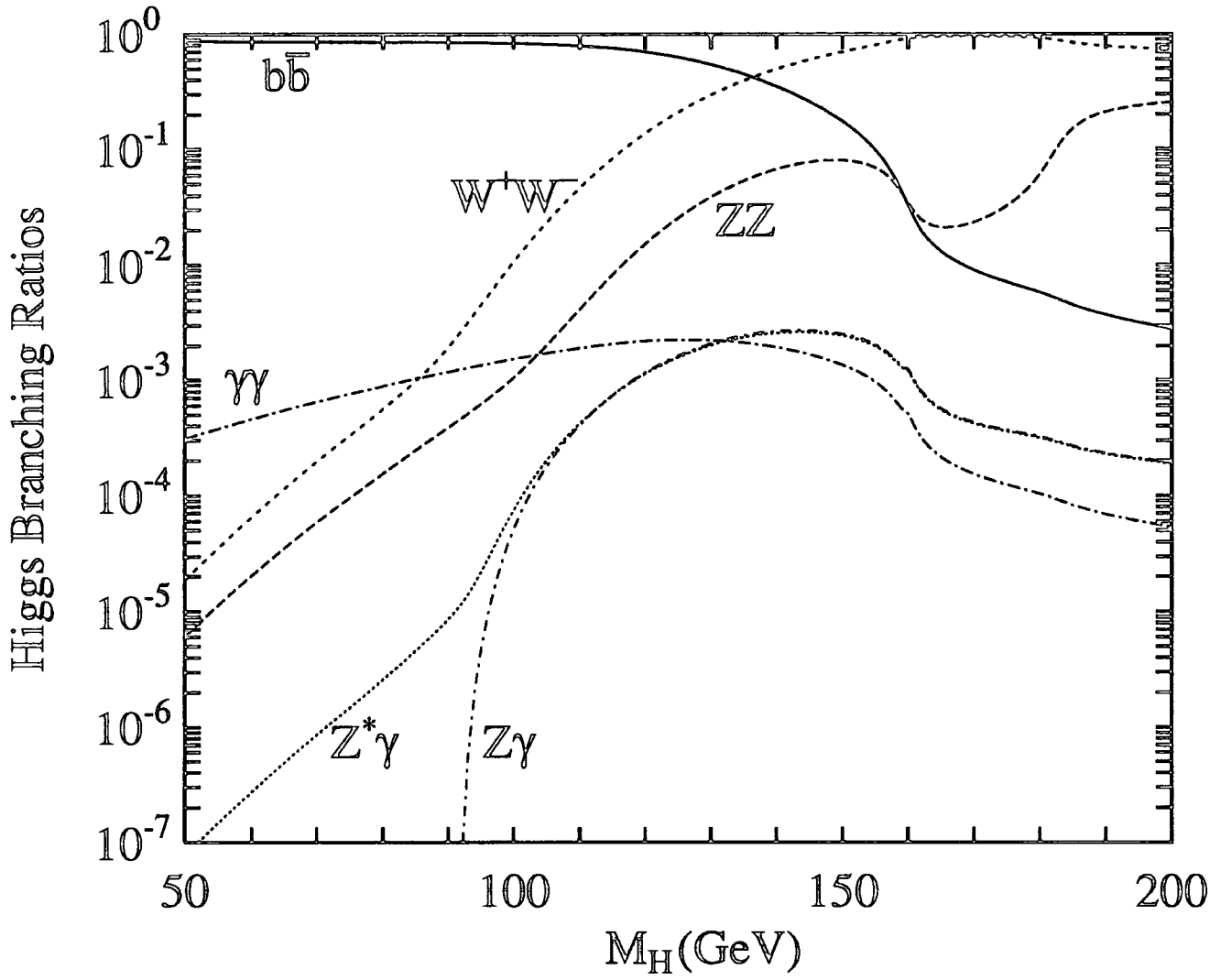


*Fig.*G.2 The main branching ratios for a Standard Model Higgs Boson, two lines are shown for the $H \to Z\gamma$ rate, $Z\gamma$ is calculated assuming a stable Z boson, whereas $Z^*\gamma$ includes the finite width of the Z boson.

shell. For $M_H < 100\,\text{GeV}$ the available phase space for $H \to Z\gamma$ causes the stable Z result to drop very rapidly. With the finite Z width we see a far milder drop off as M_H gets smaller; and still have some (small) rate for $H \to Z^*\gamma$. In conclusion we only significantly enhance the branching ratio $\text{Br}(H \to Z^*\gamma)$ for $M_H < 100\,\text{GeV}$; where the branching ratio is only a few times 10^{-5}, and so for phenomenological interest the finite widths of the Z are unimportant.

Appendix H

Measuring the Polarization of Massive Spin 1 Bosons

'Hallo, Pooh,' said Rabbit.

'Hallo, Rabbit,' said Pooh dreamily.

'Did you make that song up ?'

'Well, I sort of made it up,' said Pooh. 'It isn't Brain.' he went on humbly, 'because You Know Why, Rabbit; but it comes to me sometimes.'

The House At Pooh Corner, A.A.Milne

In many processes a spin 1 boson is produced. If we can measure the polarization of that boson we can learn much about the process involved. For example the coupling of a heavy Higgs boson to a Z_0 Boson is mainly to the longitudinal degree of freedom of the Z_0 boson; hence if in a process we detect an excess of longitudinal Z_0's this is an indication of the presence of the Higgs boson [40].

To measure the polarization of the V boson (from here on we use V to mean any massive spin 1 boson) we consider its angular decay to a fermion–anti-fermion pair. Consider the process shown in *Fig.*H.1,the matrix element for this can be written as,

$$\mathcal{M} = X_\mu \left(\frac{-g^{\mu\nu} + k^\mu k^\nu / m_V^2}{k^2 - m_V^2 + i m_V \Gamma_V} \right) \bar{u}(1) g_V \gamma_\nu (g_V - g_A \gamma_5) u(2) \quad . \tag{H.1}$$

Now we can write for $k^2 = m_V^2$

$$-g^{\mu\nu} + k^\mu k^\nu / m_V^2 = \sum_{i=L,\pm} \epsilon_i^{*\mu} \epsilon_i^\nu \quad , \tag{H.2}$$

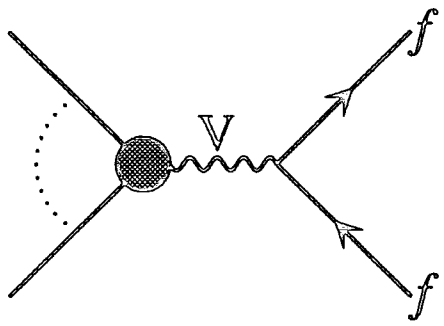


*Fig.*H.1 The general Feynman diagram for a process in which a massive vector boson is produced and decays into a fermion-antifermion pair. The blob contains the rest of the process that produced the boson.

where if $k = (k_0, 0, 0, |k|)$ we define

$$\begin{aligned} \epsilon_L &= \frac{1}{m_V}(|k|, 0, 0, k_0) \\ \epsilon_\pm &= \frac{1}{\sqrt{2}}(0, 1, \pm i, 0) \quad . \end{aligned} \tag{H.3}$$

Then equation (H.1) becomes

$$\begin{aligned} \mathcal{M} &= \sum_i \frac{g_V X.\epsilon_i^*}{k^2 - m_V^2 + i m_V \Gamma_V} \bar{u}(1) \not{\epsilon}_i (g_V - g_A \gamma_5) u(2) \\ &= \sum_i a_i \bar{u}(1) \not{\epsilon}_i (g_V - g_A \gamma_5) u(2) \quad , \end{aligned} \tag{H.4}$$

where

$$a_i = \frac{g_V X.\epsilon_i^*}{k^2 - m_V^2 + i m_V \Gamma_V} \quad . \tag{H.5}$$

To calculate the matrix elements squared we have

$$\begin{aligned} |\mathcal{M}|^2 &= \sum_{ij} a_i a_j^* \bar{u}(1) \not{\epsilon}_i (g_V - g_A \gamma_5) u(2) \bar{u}(2) \not{\epsilon}_j^* (g_V - g_A \gamma_5) u(1) \\ &= \sum_{ij} a_i a_j^* |\mathcal{M}|_{ij} \quad , \end{aligned} \tag{H.6}$$

where

$$|\mathcal{M}|_{ij} = \bar{u}(1) \not{\epsilon}_i (g_V - g_A \gamma_5) u(2) \bar{u}(2) \not{\epsilon}_j^* (g_V - g_A \gamma_5) u(1) \quad , \tag{H.7}$$

the $|\mathcal{M}|_{ij}^2$ are easily calculated using traceology. In the rest frame of the V for massless fermions with p_1 defined by

$$p_1 = \frac{m_V}{2}(1, \sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta) \quad , \tag{H.8}$$

we find

$$\begin{aligned}
|\mathcal{M}|_{LL} &= 2(g_V^2+g_A^2)m_V^2(1-\cos^2\theta) \\
|\mathcal{M}|_{\pm\pm} &= (g_V^2+g_A^2)m_V^2(1+\cos^2\theta) \pm 4g_V g_A m_V^2\cos\theta \\
|\mathcal{M}|_{L\pm} &= \sqrt{2}m_V^2 e^{\mp i\phi}(-(g_V^2+g_A^2)\cos\theta \mp 2g_V g_A)\sin\theta \\
|\mathcal{M}|_{\pm\mp} &= -(g_V^2+g_A^2)m_V^2 e^{\pm 2i\phi}(1-\cos^2\theta)
\end{aligned} \qquad \text{(F.2.7)}$$

and

$$|\mathcal{M}|_{ji} = |\mathcal{M}|^*_{ij} \quad .$$

Now to convert this into a differential cross-section (and hence a number of events) we use,

$$\frac{1}{\sigma_0}\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \sum_{ij} b_i b_j^* \frac{|\mathcal{M}|^2_{ij}}{m_V^2(g_V^2+g_A^2)} \quad , \qquad \text{(H.9)}$$

where

$$\sigma_0 = \int \mathrm{d}\sigma = \frac{1}{flux}\int |\mathcal{M}|^2 \mathrm{dLIPS} \quad . \qquad \text{(H.10)}$$

It should be noted that equation (H.9) predicts that there is a differential azimuthal decay rate in the interference terms between the different polarizations - even though there seems no natural azimuthal angle to define this decay relative to; this arises because our definition of the transverse polarizations vectors defines an azimuthal angle, $\phi = 0$ is the direction of the real parts of $\epsilon_\pm$. Now define,

$$b_j = r_j e^{i\omega_j} \quad . \qquad \text{(H.11)}$$

We have,

$$2r_L^2 + r_+^2 + r_-^2 = 1 \quad . \qquad \text{(H.12)}$$

To extract the values of r_i, ω_i from an ensemble of V's we use the following method. Define λ_i such that,

$$\frac{1}{\sigma_0}\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} = \sum_i \lambda_i P_i(\Omega) \quad , \qquad \text{(H.13)}$$

where $P_i(\Omega)$ are orthogonal polynomials in Ω, that is

$$\int P_i(\Omega)P_j(\Omega)\mathrm{d}\Omega = \delta_{ij} l_i \quad , \qquad \text{(H.14)}$$

we find

$$\begin{aligned}
\lambda_i &= \frac{1}{l_i}\int \frac{1}{\sigma_0}\frac{\mathrm{d}\sigma}{\mathrm{d}\Omega} P_i(\Omega)\mathrm{d}\Omega \\
&= \frac{1}{l_i\sigma_0}\int P_i(\Omega)\mathrm{d}\sigma \quad .
\end{aligned} \qquad \text{(H.15)}$$

This last quantity we can approximate by Monte Carlo integration,

$$\lambda_i = \frac{1}{l_i}\langle P_i(\Omega_j)\rangle \quad , \tag{H.16}$$

where

$$\langle Y\rangle = \frac{1}{\text{no of events}} \sum_{\text{events}} Y \quad , \tag{H.17}$$

where Ω_j is the Ω of each V in our ensemble. The λ_i have a covariance matrix defined by

$$\text{Cov}(\lambda_i, \lambda_j) = \frac{1}{l_i l_j}(\langle P_i(\Omega_j)P_j(\Omega_j)\rangle - \langle P_i(\Omega_j)\rangle\langle P_j(\Omega_j)\rangle) \quad , \tag{H.18}$$

the 1 standard deviation errors on λ_i being given by

$$\text{error}(\lambda_i) = \sqrt{\text{Cov}(\lambda_i, \lambda_i)} \tag{H.19}$$

We choose as our orthogonal polynomials,

$$\begin{aligned}
P_1 &= \cos\theta \\
P_2 &= \cos^2\theta \\
P_3 &= \cos\phi\sin\theta \\
P_4 &= \sin\phi\sin\theta \\
P_5 &= \cos\phi\sin\theta\cos\theta \\
P_6 &= \sin\phi\sin\theta\cos\theta \\
P_7 &= \cos 2\phi(1-\cos^2\theta) \\
P_8 &= \sin 2\phi(1-\cos^2\theta) \quad .
\end{aligned} \tag{H.20}$$

Which exhaust all the possible information contained in $\frac{d\sigma}{d\Omega}$. With these choices we find.

$$\begin{aligned}
\lambda_1 &= \frac{4g_V g_A}{g_V^2+g_A^2}(r_+^2 - r_-^2) \\
\lambda_2 &= -2r_L^2 + r_+^2 + r_-^2 \\
\lambda_3 &= 4\sqrt{2}\frac{g_V g_A}{g_V^2+g_A^2}(r_L r_- \cos(\omega_L-\omega_-) - r_L r_+ \cos(\omega_L-\omega_+)) \\
\lambda_4 &= 4\sqrt{2}\frac{g_V g_A}{g_V^2+g_A^2}(r_L r_- \sin(\omega_L-\omega_-) - r_L r_+ \sin(\omega_L-\omega_+)) \\
\lambda_5 &= -2\sqrt{2}(r_L r_- \cos(\omega_L-\omega_-) + r_L r_+ \cos(\omega_L-\omega_+)) \\
\lambda_6 &= -2\sqrt{2}(r_L r_- \sin(\omega_L-\omega_-) + r_L r_+ \sin(\omega_L-\omega_+)) \\
\lambda_7 &= -2r_+ r_- \cos(\omega_+-\omega_-) \\
\lambda_8 &= -2r_+ r_- \sin(\omega_+-\omega_-) \quad .
\end{aligned} \tag{H.21}$$

Solving these equations we can find r_L, r_+, r_-, $\omega_L-\omega_+$, $\omega_L-\omega_-$, $\omega_+-\omega_-$.

However the above method assumes that we can assemble an ensemble of V's with identical polarization vectors – this typically means that the momentum and spins of all particles other than the V decay products must be identical for each V in the ensemble. This is clearly too restrictive – and has vanishing rate for all physical process. As a result we wish to extend equation (H.16) to ensembles of V's that no longer have a unique polarization common to all the V's. To evaluate what equation (H.16) tends to as the number of V's in the ensemble tends to infinity note that we can treat the λ_i and the $P_i(\Omega_j)$ as probability distributions of which we know the means (and variances); these tell us that the LHS of equation (H.21) tend to the RHS averaged over the ensemble of V's. Because of this averaging over the quantities on the RHS it is hard to state a relation between say $\overline{r_+ r_- \cos(\omega_+ - \omega_-)}$, $\overline{r_+ r_- \sin(\omega_+ - \omega_-)}$ and $\overline{r_+^2}$, $\overline{r_-^2}$: we suggest the following.

If we replace P_7 and P_8 by,

$$\begin{aligned} P_9 &= \cos(2\phi - \phi_0)(1 - \cos^2\theta) \\ P_{10} &= \sin(2\phi - \phi_0)(1 - \cos^2\theta) \end{aligned} \quad , \tag{H.22}$$

we have,

$$\begin{pmatrix} P_9 \\ P_{10} \end{pmatrix} = \begin{pmatrix} \cos\phi_0 & \sin\phi_0 \\ -\sin\phi_0 & \cos\phi_0 \end{pmatrix} \begin{pmatrix} P_7 \\ P_8 \end{pmatrix} \quad , \tag{H.23}$$

and

$$\begin{pmatrix} \lambda_9 \\ \lambda_{10} \end{pmatrix} = \begin{pmatrix} \cos\phi_0 & \sin\phi_0 \\ -\sin\phi_0 & \cos\phi_0 \end{pmatrix} \begin{pmatrix} \lambda_7 \\ \lambda_8 \end{pmatrix} \quad . \tag{H.24}$$

Now by setting $\tan\phi_0 = \lambda_8/\lambda_7$ we can arrange for $\lambda_{10} = 0$, ie,

$$\sum_j \sin(2\phi_j - \phi_0)(1 - \cos^2\phi_j) = 0 \quad , \tag{H.25}$$

then

$$\begin{aligned} \lambda_9 &= \cos\phi_0 \lambda_7 + \sin\phi_0 \lambda_8 = \sqrt{\lambda_7^2 + \lambda_8^2} \\ &= -2\overline{r_+ r_- \cos(\omega_+ - \omega_- - \phi_0)} \end{aligned} \quad . \tag{H.26}$$

Now $\lambda_9^2 \leq 4\overline{r_+ r_-}^2$ with equality iff $\phi_0 = \omega_+ - \omega_-$ for all the V's in the ensemble.

We also have,

$$\overline{r_+ r_-}^2 \leq \overline{r_+^2}\,\overline{r_-^2} \quad , \tag{H.27}$$

with equality iff $r_+/r_- =$ constant for all the V's in the ensemble.

Because of these inequalities we define the purity of our ensemble as,

$$\mathcal{P}_{+-} = \frac{\overline{r_+ r_- \cos(\omega_+ - \omega_-)}^2 + \overline{r_+ r_- \sin(\omega_+ - \omega_-)}^2}{\overline{r_+^2}\,\overline{r_-^2}} \quad . \tag{H.28}$$

So we have $0 \leq \mathcal{P}_{+-} \leq 1$ with $\mathcal{P}_{+-} = 1$ iff $\omega_+ - \omega_- = \phi_{+-} =$ constant and $r_+/r_- =$ constant for each V in the ensemble. We also define the average value of $\omega_+ - \omega_- = \phi_{+-}$ as,

$$\tan \phi_{+-} = \lambda_8/\lambda_7 \quad . \tag{H.29}$$

We also define $\mathcal{P}_{L+}$,ϕ_{L+} and $\mathcal{P}_{L-}$,ϕ_{L-} analogously.

References

1. F.E. Close, J. Korner, R.J.N. Phillips, and D.J. Summers, Presented at U.K. B Physics Workshop, Edinburgh, Scotland, December 1991 , J. Phys. G: Nucl. Part. Phys. 18 (1992) 1716.

2. D.J.Summers, *Phys.Lett.* B302 (1993) 326.

3. T.P.Cheng and L.F.Li *Gauge theory of elementary particle physics*, Oxford University Press, (1984).
   F.Mandl and G.Shaw, *Quantum Field Theory*, John Wiley and Sons, (1984).
   C.Itzykson and J.Zuber, *Introduction to quantum field theory*, McGraw–Hill, (1980).
   F.Halzen and A.D.Martin, *Quarks & Leptons*, John Wiley and Sons, (1984).

4. M.Froissart, *Phys. Rev.* 123, (1961) 1053.

5. S.Weinberg, *Phys. Rev. Lett.* 19. (1967) 1264.
   A.Salam *Elementary particle physics (Nobel symp No. 8)* (ed. N.Svartholm), Alemqvist and Wilsell, (1968).

6. A.A. Carter, Institute of Physics Conference on Nuclear and Particle Physics, Glasgow, March 1993.

7. D.J.Summers *Phys. Lett.* B274 (1992) 209.

8. J.D.Bjorjen, *Proc. of Summer Institute on Particle Physics,* SLAC - 198 ,p.1 ed M.Zipf

9. J.F.Gunion, G.L.Kane, H.E.Haber, and S.Dawson, *The Higgs Hunter's Guide*, Addison Wesley, (1990).

10. S.L.Wu *et al.* ECFA workshop on LEP200, eds A.Böhm and W.Hoogland, Vol II (1987) 312.

11. Z.Kunszt and W.J.Stirling, Aachen ECFA Workshop (1990) 428.

12. N.Brown *Z.Phys.* C49 (1991) 657.

13. P.Janot, preprint LAL 91-61, Talk given at ICFA Workshop, Saariselka, Finland, Sep 9-14, 1991
P.Janot, preprint LAL 92-27, Presented at 27th Rencontres de Moriond: Electroweak Interactions and Unified Theories, Les Arcs, France, Mar 15-22, 1992.

14. D.J.Summers, *Phys.Lett.* B277 (1992) 366.

15. D.J.Summers, *Phys.Lett.* 306B (1993) 129.

16. W.J.Stirling and D.J.Summers *Phys. Lett.* B283 (1992) 411.

17. See for example : W.J.Stirling *Topics in Modern Phenomenology*, Lecture Notes - HEP Summer School, RAL (1987)

18. A. Caner for the CDF Collaboration, XXVIIIth Rencontres de Moriond, Mar. 1993, to be published.

19. D.Froidevaux, Proceedings of the Large Hadron Collider Workshop, Aachen, CERN 90-10 Vol. II (1990) 444.

20. I would like to thank Zoltan Kunszt for forwarding the program that calculates the branching ratios of a Standard Model Higgs.

21. M. Della Negra *et al.* Proceedings of the Large Hadron Collider Workshop, Aachen, CERN 90-10 Vol. II (1990) 509.

22. C.Seez *et al.* Proceedings of the Large Hadron Collider Workshop, Aachen, CERN 90-10 Vol. II (1990) 474.

23. R.Kleiss, Z.Kunszt, W.J.Stirling, *Phys.Lett.* B253 (1991)269.

24. J.Ohnemus and W.J.Stirling, *Phys.Rev.* D47 (1992) 336.

25. A.Ballestrero and E.Maina, *Phys. Lett.* B268 (1991) 437.

26. J.Ohnemus and W.J.Stirling, FermiLab preprint, FERMILAB–PUB–92/278–T, DTP/92/74 (1992).
H.Baer, B.Bailey and J.F.Owens, Florida State Univsersity preprint, FSU–HEP–920915.

27. J.Smith, D.Thomas, and W.L.van Neervern, *Z.Phys.* C44 (1989) 267.
J.Ohnemus, *Phys.Rev.* D47 (1993) 940.

28. H.Y. Zhou and Y.P. Kuang, Tsinghua University preprint, TUIMP–TH-92/50, (1993).

29. M. Narian for the D0 Collaboration, XXVIIIth Rencontres de Moriond, Mar. 1993, to be published.

30. A.D.Linde, *JETP Lett.* 23 (1976) 64; *Phys. Lett.* 62B (1976) 435.
S.Weinberg, *Phys. Rev. Lett.* 36 (1976) 294.

31. M.J.Duncan, R. Philippe, and M.Sher, *Phys. Lett.* 153B (1985)165.

32. H.Goldstein *Classical mechanics*, Addison-Wesley, (1950).

33. R. Kleiss and W.J. Stirling, *Nucl. Phys.* B262 (1985) 235.

34. R. Kleiss and W.J. Stirling, *Phys.Lett.* 179B (1986) 159.

35. J.F.Gunion and Z.Kunszt, *Phys.Lett.* 161B (1985) 333.

36. S.D.Ellis, R.Kleiss and W.J.Stirling, Comp. Phys. Comm. 40 (1986) 359.

37. A.D. Martin, R.G. Roberts, and W.J. Stirling, Durham University preprint, DTP/92/80.

38. W.H.Press, B.P.Flannery, S.A.Teukolsky, and W.T.Vetterling, *Numerical Recipes*, Cambridge University Press, 1987.

39. G.Arfken, *Mathematical Methods for Physicists*, Academic Press, (1985).

40. M.J.Duncan, *Phys. Lett.* bf 179B (1986) 393.