A SEARCH FOR THE STANDARD MODEL HIGGS DECAYING TO TWO MUONS AT THE CMS EXPERIMENT

By

ANDREW CARNES

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

2017

ACKNOWLEDGMENTS

I would like to thank my Mom and Dad for just being all around great parents. Without all of their support, the completion of this Ph.D would never have been possible. I would also like to thank Professors Paul Avery and Darin Acosta for their help. In addition, I would like to thank Andrew Brinkerhoff, Pierluigi Bortignon, and Andrea Marini for all of their support and hard work on the project. I'd also like to thank my friends Stany, Leo, Sean, Sophie, and David for making Geneva a great place to live.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Abstract of Dissertation Presented to the Graduate School
of the University of Florida in Partial Fulfillment of the
Requirements for the Degree of Doctor of Philosophy

A SEARCH FOR THE STANDARD MODEL HIGGS DECAYING TO TWO MUONS AT THE CMS EXPERIMENT

By

Andrew Carnes

December 2017

Chair: Paul Avery
Major: Physics

In 2012 two collaborations at the Large Hadron Collider announced the discovery of a new particle with properties similar to the Standard Model Higgs Boson. In order to determine whether the boson discovered with a mass of 125 GeV is actually the Standard Model Higgs, all of the different ways the particle can decay need to be investigated. If the probabilities for the different decays do not match the predictions of the Standard Model then this would imply new physics.

This dissertation presents the search for the Standard Model Higgs Boson decaying to $\mu^+\mu^-$. The search uses the 35.9 ± 0.9 fb^{-1} of $\sqrt{\mathrm{s}} = 13$ TeV proton-proton collision data recorded by the CMS detector in 2016. The observed and expected upper limits on the rate at a 95 % confidence level are presented for Higgs masses in the range 120 to 130 GeV. The expected and observed upper limits at a mass of 125 GeV are x.xx and $1.98^{+0.81}_{-0.57} \times$ SM respectively. These results provide the best results to date on the Higgs coupling to second generation fermions. No deviations from the Standard Model are observed.

CHAPTER 1
INTRODUCTION

The Standard Model (SM) of particle physics is an extremely successful theory shown to correctly predict the behavior of the particles and forces which make up the most basic constituents of the universe. In fact, it correctly describes all of the forces known except for gravity. In particular, the SM predicts that the massive particles of the theory acquire their mass by interacting with a scalar particle called the Higgs boson. On July 4, 2012 two collaborations at the Large Hadron Collider (LHC), the A Toroidal LHC Apparatus (ATLAS) and Compact Muon Solenoid (CMS), announced the discovery of a new boson at 125 GeV with properties similar to the Standard Model Higgs **???**. This discovery was fueled by the investigation into the Higgs decays to the vector bosons ZZ and $\gamma\gamma$. Soon after, evidence for the Higgs coupling to matter was found through the $\tau^{+}\tau-$ and $\mathrm{b\bar{b}}$ decays **????**. Whether the newly discovered boson is indeed the expected Standard Model Higgs remains to be determined. Insofar, all of the different decay modes will be investigated to search for deviations from the Standard Model predictions.

This leads to the study of the Higgs decay to $\mu^{+}\mu^{-}$. Although this decay is the smallest branching fraction expected to be detected **??**, the dimuon decay offers high efficiency and excellent momentum resolution, which should lead to a narrow peak over the falling background, mostly Drell Yan events. The tiny branching fraction enables greater sensitivity to small deviations from the predicted decay rate and in this respect offers an advantage over other channels where a miniscule deviation could be drowned out. Furthermore, the Higgs coupling to second generation fermions remains to be determined.

This dissertation presents the search for the Standard Model Higgs Boson decaying to $\mu^{+}\mu^{-}$ using the proton-proton collision data recorded by the CMS experiment in 2016. In order to maximize the data available for the search, the first machine learning in the L1 Trigger system at the LHC was developed and deployed for 2016 data collection. To further maximize the sensitivity of the search, an additional machine learning technique was invented to categorize events based upon the detector resolution and the event kinematics. The search looks for a Higgs boson with a mass between 120 and 130 GeV and presents the expected and observed upper limits in this range as well as the best fit for the rate of production.

The dissertation first covers the LHC which is responsible for accelerating the colliding the protons. Then the dissertation presents the CMS detector responsible for measuring the paths, momentum, and energy of the emerging particles. Next, the dissertation explains the theory underlying the Standard Model and its predictions of the Higgs particle. After, the machine learning implementation in the L1 trigger that reduced the number of fakes in the data by a factor of three is detailed. And finally, the search for H to $\mu^+\mu^-$ is presented.

CHAPTER 2
THE LARGE HADRON COLLIDER AND THE CMS EXPERIMENT

2.1 Large Hadron Collider

The Large Hadron Collider is a particle collider near Geneva, Switzerland run by the European Organization for Nuclear Research (CERN). The LHC is the largest and most powerful particle collider ever built, designed to collide protons with a center of mass energy of 14 TeV and a luminosity of $10^{34}\mathrm{cm}^{-2}\mathrm{s}^{-1}$ **?**. The luminosity is given by

$$\mathrm{L} = \frac{\mathrm{n_b f N_p^2} \gamma}{4\pi\epsilon_\mathrm{n}\beta^*} \tag{2–1}$$

where $\mathrm{n_b}$ is the number of bunches in each ring, f is the frequency for a bunch to circle the ring, $\mathrm{N_p}$ gives the number of protons in a bunch, and γ is the Lorentz factor. ϵ_n is the normalized transverse emittance, a measure of the spread of the beam in momentum and position space. β^* measures the focus of the beam at the interaction point. $\epsilon_\mathrm{n}\beta^*$ represents the transverse area at the point of interaction. A large luminosity is characterized by a high frequency of bunch crossings with lots of protons in each bunch packed as densely as possible, and a large luminosity results in a high rate of collisions. With many collisions at high energy, the detectors can collect enough events from yet unexplored regimes of physics to discover new physics or to verify or discard the predictions different hypothesis, so these parameters are very important.

The collider itself is 26.7 km in circumference 45-170 m underground. 8.3 T supercooled superconducting magnets operating at 2 K steer the high energy proton beams. In order to save money the LHC not only reuses the tunnels of a previous collider, the Large Electron Positron Collider (LEP), but also reuses older accelerators which were state of the art at their time. These older accelerators ramp up the energy of the protons and inject them into the LHC. All of this together makes up the CERN accelerator complex.

First, the protons are created from a source of Hydrogen gas. The hydrogen atoms of the gas are placed into a large electric field that separates the atoms into unbound protons

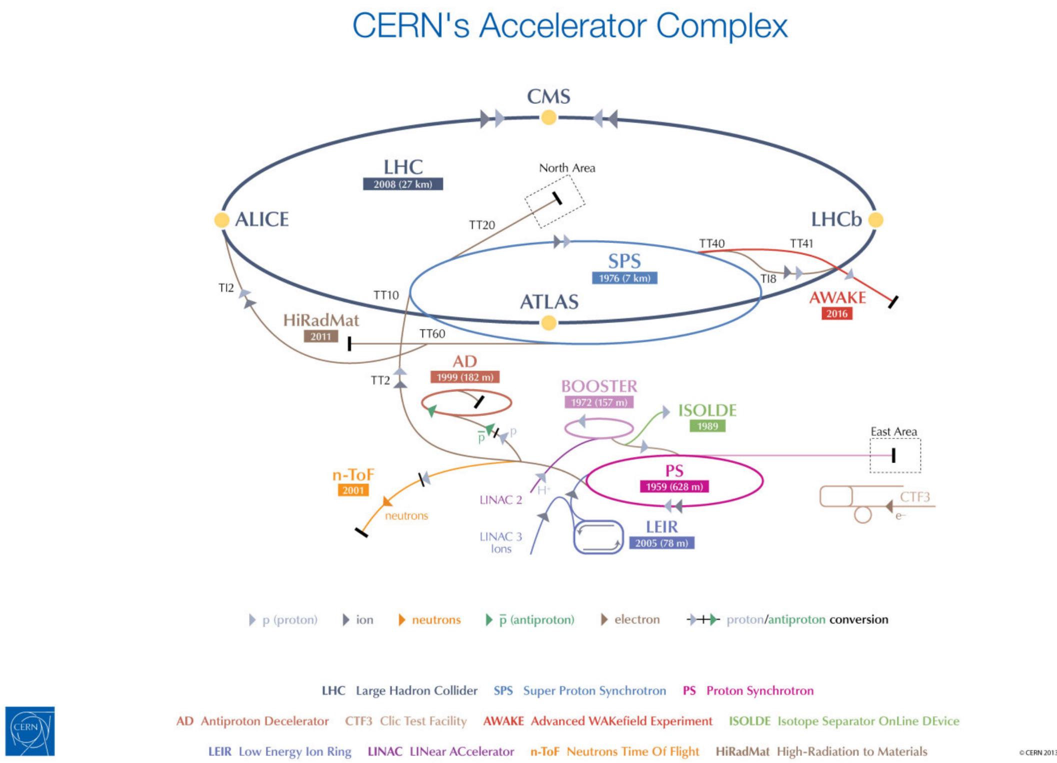


Figure 2-1. The CERN Accelerator Complex **?**

and electrons. The protons are then sent to a radio frequency quadrapole which focuses the protons and accelerates them. The radio frequency field is stronger for the protons in the back than in the front and consequently squeezes them into a tighter bunch. The protons then proceed to a linear accelerator, LINAC2, where they are accelerated to 50 MeV or 5% of the speed of light (c). The protons then enter a series of synchrotrons. A synchrotron is a device that accelerates particles by guiding them around a fixed circular path with a magnetic field and boosting their speed with an electric field as they pass a certain point. Since a faster particle bends less in the same magnetic field, the magnetic field strength is synchronized with the speed of the accelerating particles to keep them in the fixed circular path.

After LINAC2 the protons enter the first of the synchrotrons, the Proton Synchrotron Booster (PSB) accelerating the protons to 1.4 GeV (0.81c). From here the protons are injected into the Proton Synchrotron (PS) and accelerate to 25 GeV (0.999c). The PS then injects the protons into the Super Proton Synchrotron (SPS) further accelerating them to 450 GeV (0.99999c). Finally the protons are injected into the LHC where they accelerate up to 6.5 TeV (0.99999999c). Once accelerated to the appropriate collision energy, the proton beams are made to collide in the different detectors located around the ring. By colliding enough protons at large enough energies it is possible to probe corners of physics that have never been seen before. The two general purpose detectors at the LHC, ALTAS and CMS, are used to look for signs of new physics like the Higgs boson, dark matter, and extra dimensions by measuring the energy, the momentum, and the paths of the particles coming out of the collisions.

2.2 Compact Muon Solenoid Detector

The CMS detector, located in Cessy, France, is 21.6 m long, 15 m in diamater, and weighs more than the Eiffel Tower. Not only is the detector a massive and complex device it's also run by a huge collaboration involving approximately 3,800 people from 200 institutes spanning 43 different countries **?**. The greatest achievement of the collaboration to date is the discovery of a Higgs like particle in 2012, a feat shared with ATLAS.

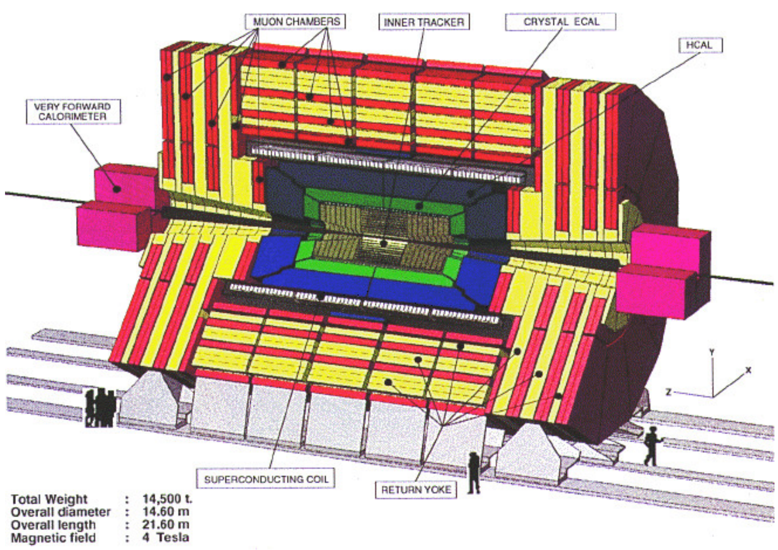


Figure 2-2. The CMS detector **?**

CMS was built primarily to look for the Standard Model Higgs and signs of Beyond Standard Model (BSM) physics like Supersymmetry, extra dimensions, or new heavy weak bosons **?**. Because BSM and Higgs decays to muons and electrons often have the highest signal to background ratio, CMS is designed to identify and measure these particles with a high accuracy. In layman's terms a high signal to background ratio just means that these events have fewer look-alikes. Jets [1] and photons are measured to a high degree of accuracy as well. In order to measure the energy, momentum, and location of the different types of particles CMS deploys a variety of subdetectors working in concert. The defining feature of the detector is an extremely powerful solenoid which enables the accurate measurement of momentum for charged particles. The tracker and calorimeters fit snugly within the 6 m diameter solenoid. The muon detectors reside outside the magnet but within the return yoke.

2.2.1 Silicon Tracker

The 3.8T magnetic field inside the solenoid enables the tracker to measure the transverse momentum of charged particles based upon the curvature of the track. Charged particles with lower transverse momentum ($\mathrm{p_t}$) bend more in a magnetic field than high $\mathrm{p_t}$ particles. As such, a measurement of the deviation of a curved track from a straight line, the sagitta, can be used to measure the curvature and determine the momentum **?**.

$$\mathrm{p_t} \cong \frac{\mathrm{L^2qB}}{\mathrm{8s}} \tag{2–2}$$

Here L is the length of the straight line between the first and last position measurments, q is the charge of the particle, B is the magnetic field, and s is the sagitta.

The equation for the error in the momentum measurement shows that a higher magnetic field enables better $\mathrm{p_t}$ resolution, illuminating the design choice for a powerful magnet.

$$\frac{\delta \mathrm{p_t}}{\mathrm{p_t}} \propto \frac{\mathrm{p_t}}{\mathrm{L^2B}} \tag{2–3}$$

[1] When a quark or gluon is created it can't exist alone, since it has color charge, and pulls other quarks from the vacuum creating a tight cone of composite colorless particles as well as their decays. This cone of particles is called a jet.

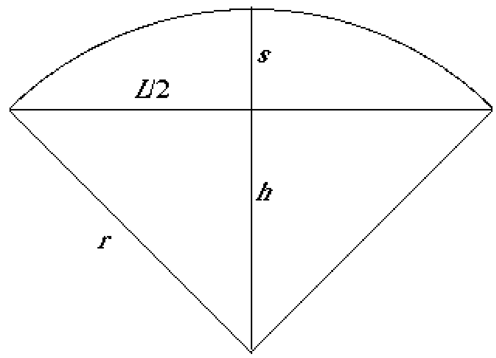


Figure 2-3. The sagitta measurement

The silicon tracker is made of tiny reverse biased bipolar diodes. When a charged particle travels through one of these diodes the ionization force of the particle releases electron hole pairs beyond the electrostatic equilibrium, inciting a current to flow. The tracker needs to be small enough such that the particles flowing through it don't deposit much energy. Energy deposition in the tracker would throw off energy measurements in the calorimeters. This means that the tracker needs to be smaller than a few radiation lengths [2] . The tracker at the thickest part is one radiation length. The tracker is placed nearest the collision point in order to identify primary and secondary vertices and to measure the momentum of particles before they are tainted by interactions with other detectors. [3] Being so near the collision point, the silicon tracker is bombarded by a constant flux of high intensity radiation. As such, the tracker is carefully designed to be robust to this radiation rich environment.

2.2.2 Calorimeters

The Electromagnetic Calorimeter (ECAL) is right outside the tracker and its main goal is to measure the energy of electrons and photons. It's designed to contain entire electromagnetic

[2] the length scale over which an electron deposits a substantial amount of energy into the material

[3] Vertex is shorthand for the location of the collision or decay that produced a set of particles.

showers for these particles and is consequently many radiation lengths thick. The ECAL is made of lead tungstenate scintillating crystals which release an amount of light proportional to the energy deposition. The light is collected and the total energy is calculated. The separation into individual crystals allows some spatial resolution as well. Particles with larger mass deposit less energy per unit distance into a solid. Many of the hadronic particles make it through the ECAL for this reason.

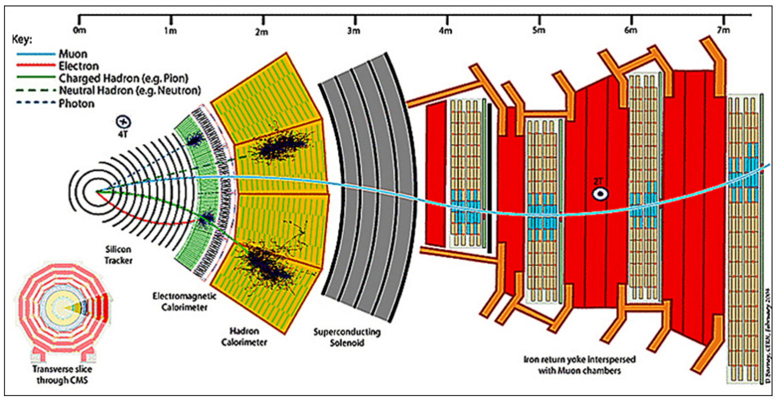


Figure 2-4. A slice of the CMS detector **?**

The HCAL is placed outside the ECAL to collect the energy of the particles that survived the other subsystems, mostly strongly interacting hadronic particles from jets. The HCAL works in a similar manner to the ECAL except that layers of plastic scintillating material are interspersed with layers of a dense passive absorber like brass or steel. The density of the passive absorber increases the chance of interaction and shower production thus reducing the total length of the hadronic shower and enabling the measurement of the total energy for most of the cascades. Again the scintillation light is collected to determine the energy. If the ECAL and HCAL were placed outside the magnet the particles would interact with the solenoid material before entering the calorimeters, throwing off their measurements. The showers in the calorimeters are a consequence of the electromagnetic and strong forces, which means that particles without these interactions pass through the materials undetected, e.g. neutrinos or BSM weakly interacting particles. Since the momentum in an interaction is conserved any imbalance means that some particles escaped the detector. If there is an excess of missing momentum beyond the amount expected due to neutrinos this may indicate the existence of dark matter or some other BSM particle. In order to measure the missing energy correctly it's important that the HCAL is built without any gaps and that it is dense enough to collect the energy of the strongly and electrically interacting particles.

While the momentum resolution in the tracker is proportional to the momentum, the energy resolution in the calorimeters decreases with increasing energy **?**.

$$\frac{\delta \mathrm{E}}{\mathrm{E}} = \sqrt{\left(\frac{\mathrm{S}}{\sqrt{\mathrm{E}}}\right)^2 + \left(\frac{\mathrm{N}}{\mathrm{E}}\right)^2 + \mathrm{C}} \tag{2–4}$$

The first term in the square root describes statistical fluctuations. The energy measured is proportional to the number of photons captured which has poisonnian fluctuations and the error ($\delta\mathrm{E}$) for this term is $\propto \sqrt{\mathrm{E}}$. The second term describes noise in the electronics whose error is energy independent, and the last term describes the errors in energy calibration which are proportional to energy.

2.2.3 Muon System

Neutrinos aren't the only Standard Model particles that make it through the tracker, ECAL, and HCAL. Muons have a relatively long lifetime $\sim 10^{-6}\mathrm{s}$ with $c\tau \sim 100\mathrm{m}$. The large gamma factor in combination with their long lifetime enables them to travel hundreds of kilometers on average, well through the entire CMS detector before decaying. Muons are charged so their tracks show up in the tracker and some energy is deposited in the calorimeters but, muons are so much more massive than electrons that the energy deposition in the ECAL is minimal. Making it through the ECAL the muons enter the HCAL. The HCAL is designed to stop strongly interacting hadronic particles and collect their energy. But muons don't interact with the strong force and make it through the HCAL as well. This enables the muon system to be placed outside the magnet.

The muon system consists of a few different types of detectors which all involve the same basic principle. The charged muon ionizes some gas and the ionized particles are attracted to charged surfaces initiating a current in the surfaces. With a large enough voltage differential between the charged surfaces the the ionized particles may gain enough kinetic energy to further ionize other atoms in the gas initiating an avalanche effect and reducing the need for signal amplification later. The muon system uses this strategy in the different detectors. The types of detectors in the muon system are the Cathode Strip Chambers (CSC), the Drift Tubes (DT), and the Resistive Plate Chambers (RPC) **?**.

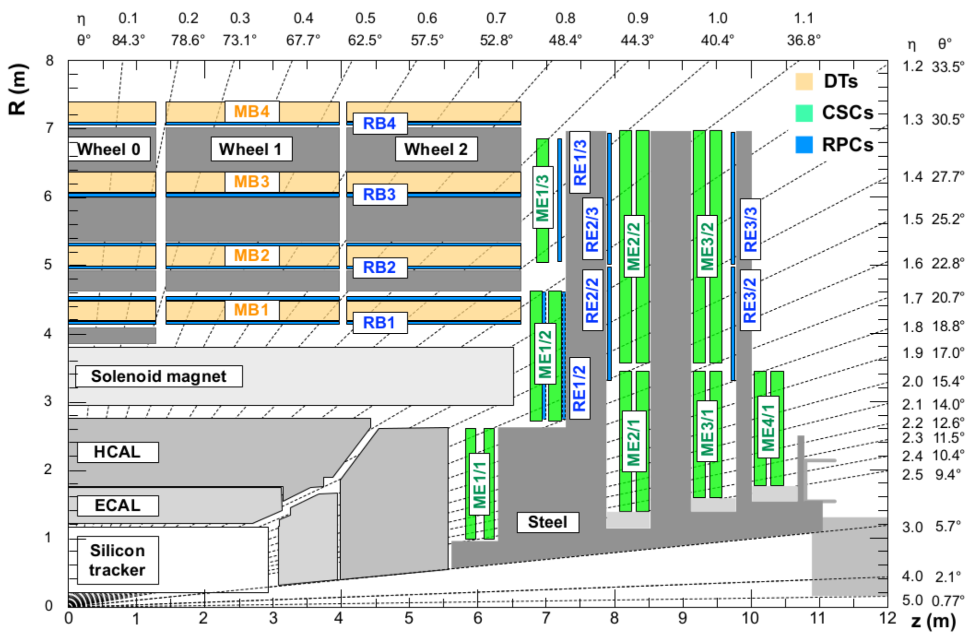


Figure 2-5. A Look at the Muon System **?**

2.2.3.1 Drift Tubes

The drift tubes are located in the barrel portion of CMS. Throughout the majority of the barrel the magnetic field is basically uniform. The drift tubes have aluminum plates on the top and bottom separated by aluminum I-beams shown in Figure 2-6. A wire acts as the anode and the I-beams are the cathodes. The tubes are designed to provide a constant drift velocity throughout each tube.

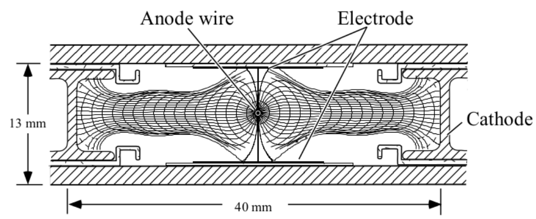


Figure 2-6. A Drift Tube **?**

When a charged particle flies through the tube it ionizes the gas inside. The electrons drift at constant velocity to the anode. The distance from the anode is deduced from the drift time, utilizing the fact that the ionized electrons drift with a constant velocity. This calculation does however require a reference time. In each chamber the drift tubes are placed in layers and the average crossing time in the chamber is used as the reference time.

2.2.3.2 Cathode Strip Chambers

The CSCs are located in the endcaps of the detector which range in $|\eta|$ from 0.8 to 2.4. One of the reasons the endcaps use CSCs instead of DTs is the nonuniform magnetic field which would adversely affect the drift times in the DT system. In this system there are oppositely charged strips and wires running roughly perpendicular to eachother.

When a muon flies through the CSC it induces charge on the wires and the strips and ionizes gas in the chamber. The ionized particles in the gas float to the charged strips and wires initiating a current in the nearby wires and strips. The induced charge from the muon itself also contributes to the currents. The most intense currents should be those associated with the location of the muon. The position resolution in the phi direction is roughly 100 μm.

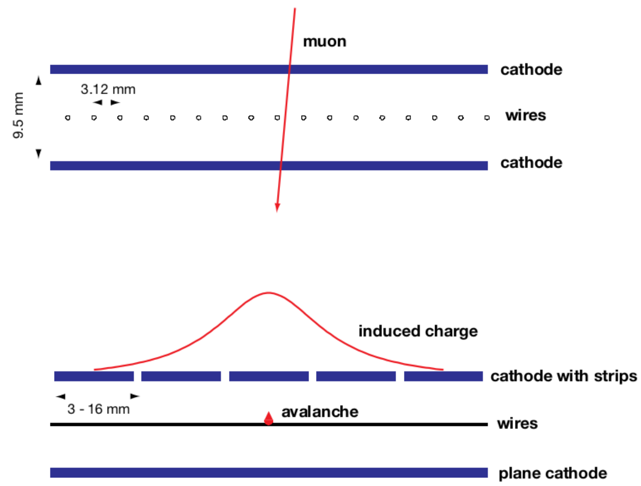


Figure 2-7. A Cathode Strip Chamber **?**

2.2.3.3 Resistive Plate Chambers

The RPCs are located both in the barrel and in the end caps. The RPCs have excellent timing resolution on the order of 1 ns. The RPCs use their excellent timing resolution to determine each particle's bunch crossing of origin. The accurate and rapid timing information helps with the online selection of muons, a huge priority for CMS considering that many interesting collisions produce muons. For this reason the RPCs focus on efficient online selection of muons instead of accurate offline reconstruction **?**.

In this way the RPCs complement the DTs and CSCs. The RPCs consist of two high resistance parallel plates surrounding a volume of gas. The outsides of the plates are painted with graphite paint forming the electrodes. A large voltage differential is kept between the electrodes. When a charged particle crosses the plates it induces an electrical discharge in the plates which remains localized in time and space due to the large resistivity.

2.2.4 Trigger System

Collision events come at a rate of 10 MHz with each event taking up roughly a MB of information. If the detector had to store all of the information from each event this would amount to pushing terabytes of information into a storage system every second, which is remarkably infeasible. To deal with this issue CMS utilizes a trigger system, which selects only interesting events cutting the rate down from 10 MHz to 1 KHz **?**. Since bunch crossings happen every 25 ns the trigger needs to operate at an incredibly high rate.

CMS tackled this issue by dividing the trigger into different tiers. The Level 1 Trigger is the first stage of the trigger system made from custom hardware which can operate at fantastic speed. The Level 1 Trigger reduces the rate from 10 MHz to 100 KHz and the events passing the L1 Trigger go onto the High Level Trigger (HLT) which further reduces the rate to 1 KHz. Due to the lower input rate the HLT can operate in software.

2.2.4.1 L1 Trigger

The L1 Trigger is made of up different subsystems that work together to decide whether to keep the data from a beam crossing for further processing. The University of Florida works with the Level 1 muon trigger system, the Endcap Muon Track Finder (EMTF) in particular. The muon system needs to determine the transverse momenta of muons and their location and choose the best candidates. Each of the different muon detectors have their own local triggers which send their best muon tracks to the Global Muon Trigger (GMT). The GMT chooses the best muon candidates from that set and passes these on to the Global Trigger (GT). The GT combines the information from the calorimeter triggering system. The GT uses this combined information to check whether the bunch crossing should be sent to the HLT or discarded. The L1 Trigger has many different trigger criteria defining separate triggers which are the trigger bits. If an event passes any of the triggers then it is forwarded for further processing.

The Track Finders (TF) play an important role in the L1 Trigger system. The EMTF combines the location and direction information from the different CSC stations into muon tracks and calculates the transverse momenta for the different tracks. The EMTF chooses the best candidates (highest momentum and highest quality) to send to the GMT. The Drift Tube Track Finder (DTTF) performs a similar process for muons in the DT system. The RPC system calculates the location and direction and forms tracks in the same stage. In the process the RPC trigger system assigns transverse momenta and quality, and like the others chooses the best tracks to send to the GMT.

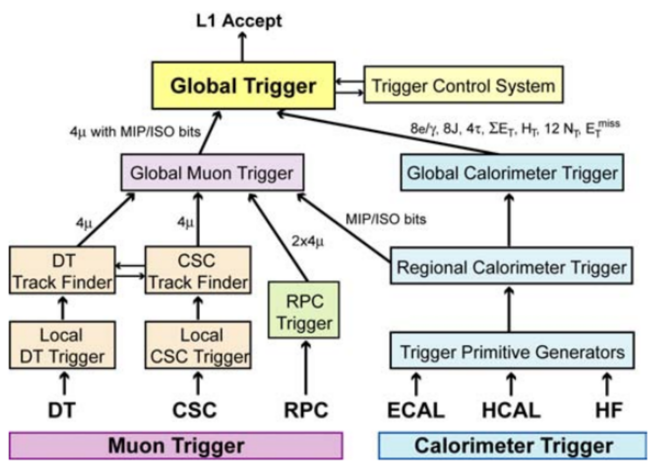


Figure 2-8. The L1 Trigger Architecture **?**

CHAPTER 3
THE STANDARD MODEL

The Standard Model (SM) of particle physics is an incredibly successful theory that correctly describes the physics of all known particles and forces that make up the universe, excluding gravity **?**. The particles of the SM come in two types: fermions and bosons. Fermions are the spin $\frac{1}{2}$ particles that make up the different types of matter, and bosons are the integer spin particles responsible for the different forces. Electrons are the most familiar type of fermion, but there are more exotic kinds like the up and down quarks that make up protons and neutrons. Electrons, protons, and neutrons make up atoms, accounting for nearly all of the matter in our day to day experience, but there are actually many other matter particles. In fact, there are three generations of quarks and leptons [1] with each generation heavier than the next. The up and down quarks are the first generation of quarks, charmed and strange are the next, and top and bottom are the third generation. For the leptons, the electron and electron neutrino are the first generation, the muon and muon neutrino are the second, and the tau and the tau neutrino the third. Each fermion also has a corresponding antiparticle. As an example, the positron is the antiparticle for the electron.

The force carrying particles that allow matter to interact and form more complex objects like atoms, molecules, and even people are the spin 1 bosons. These force carriers are the gluons, photons, and the W and Z particles. Gluons mediate the strong force, photons the electromagnetic force, and the W and Z bosons mediate the weak force. Every force has an associated charge: particles with electric charge can interact through the electromagnetic force, those with color charge may interact via the strong force, and those with isospin or weak hypercharge may interact through the weak force. The fundamental forces and particles interact to make the familiar composite objects that surround us in our daily lives. The strong

[1] Leptons are fermions like the electron that aren't quarks. The quarks interact with the strong force that binds nuclei together and the leptons do not.

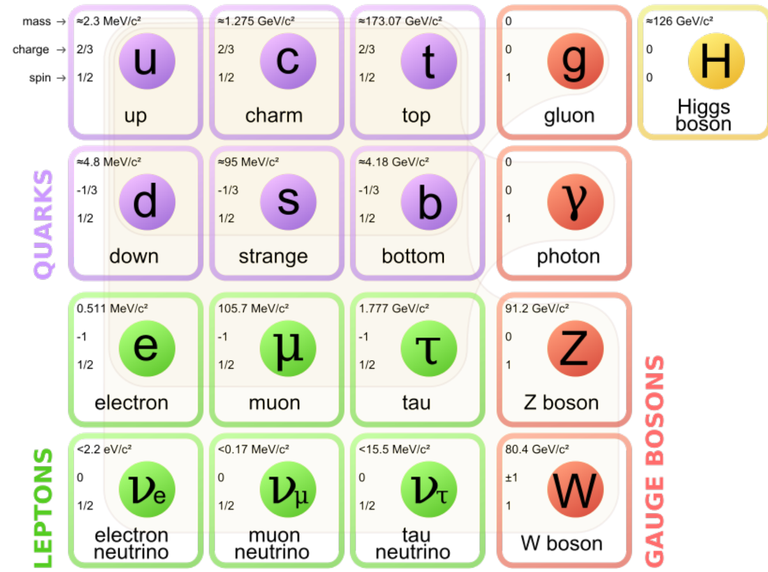


Figure 3-1. The Standard Model Particles

force binds quarks to form protons and neutrons, and even binds the protons and neutrons together to form nuclei, and the electromagnetic force binds electrons and nuclei to form atoms. The size of the composite objects gives an idea of the relative strength of the forces. A proton is 10^{-15} meters in size while an atom is 10^{-10} meters and a solar system is 10^{12} meters. The more tightly bound the stronger the force. But that isn't quite exact, in fact, the ratio of the strength of the forces is like so 1:10^{-3}:10^{-16}:10^{-41}, strong : electromagnetic : weak :gravitational [2] .

Of all the particles predicted by the SM, there is only one spin 0 particle, the Higgs boson, and it plays a special role in the theory. As the universe cooled from the Big Bang the Higgs field went through a phase transition and settled into a nonzero ground state forming a condensate. The electron, muon, tau and the W and Z particles of the SM interact with the Higgs condensate and acquire mass. With such a large role in the SM, finding this particle or a BSM Higgs has been a huge priority for the CMS collaboration **?**. In 2012 a Higgs particle

[2] Gravity is just included for perspective. The Standard Model does not describe this force and reconciling gravity with quantum mechanics is an open problem.

with a mass of 125 GeV was found and to date remains consistent with the Standard Model. However, the properties need to be investigated further before declaring the discovered Higgs the Higgs of the Standard Model.

In order to lay out the Standard Model in mathematical terms, a bunch of background information needs to be covered, and since the Standard Model is described by Quantum Field Theory (QFT), the mathematical formulation of QFT will be covered first. After developing the necessary mathematical formalism, the Standard Model and the Higgs mechanism will be derived and explained. Afterwards the experimental search for Higgs to dimuons will be described.

In this dissertation $\hbar$ and the speed of light, c, are set to 1. Moreover, 0,1,2,3 and t,x,y,z are used interchangeably to label the components of a four vector. When relevant 0 represents time and 1,2,3 represent x,y,z respectively. Einstein summation notation is used indicating that repeated indices are summed over, so $\mathrm{x_i x_i y_j y_j}$ is shorthand for $\sum_\mathrm{i} \sum_\mathrm{j} \mathrm{x_i x_i y_j y_j}$. Repeated Greek indices assume sums over all four space and time components, while repeated Roman indices assume sums over only the spatial components.

3.1 Quantum Field Theory

The mathematical framework used to describe the physics of the SM as well as other Beyond Standard Model (BSM) field theories is called Quantum Field Theory (QFT). QFT enables the predictions of different measurable probabilities. One of the most important is the probability for a set of particles to emerge from a collision of an initial set of particles. Another important one is the probability for a single particle to decay into another set of particles. These probabilities are encompassed in the cross sections and branching fractions. For example, the theory of the SM predicts the cross section for two protons to collide and create a Higgs. As another example, the SM also predicts the branching fraction for a Z boson to decay into to two muons. These probabilities can be measured simply by colliding particles and counting the outcomes which in turn means that the theory can be tested. In fact, any QFT model can be tested in this manner. Quantum Field Theory is written down in terms of a

Lagrangian, and the math for the various predictions follows from there, but first a brief aside about particles.

3.1.1 Particles, Symmetries, and Labels

Since QFT makes quantitative predictions in terms of particle collisions and particle decays it's interesting to contemplate what a particle really is. Consider an observer in a frame x with particle p and an observer in another frame x'. If the observer in x' can't identify particle p as well then it doesn't make sense to call p a particle. More concretely, consider a world where in frame x an observer sees a neatly stacked deck of cards, but in x' the observer sees the cards scattered all over the place. Calling the deck of cards a particle doesn't really make sense. On the other hand, both parties can still agree on the individual cards which kept the same suit and value. The suit and value are conserved quantities identifiable between frames. If the two observers get together later and compare notes they can see what happened to each card upon transforming from x to x' and work out a set of rules. The king of hearts may do one thing and the 10 of clubs another. They can then add the different forces into play repeat the process and compare again. Figuring this all out allows the two to determine the laws of physics for the fundamental pieces called particles.

This idea leads to Wigner's view: a particle is an object with conserved quantities that observers can agree on between frames. In our universe some of these labels are the mass, charge, spin, color, isospin, and weak hypercharge with each label attached to a specific symmetry or group of symmetries. And because different observers can agree on these quantities they can compare notes and work out the laws of physics for the different types of particles. The deep connection between conserved quantities, symmetries, and labels for particles will be seen later. For now the focus is to work out laws that people in different frames can confirm, and this is where the Lagrangian formalism comes into play.

3.1.2 The Lagrangian Formalism

The Lagrangian formalism is a mathematical device that allows physicists to describe the evolution of a physical system over time, and it's within this formalism that QFT can be built. But before building the full mathematics of QFT, a simple example describing the free Newtonian particle is covered. The Newtonian example serves to illuminate the fact that the laws of physics are indistinguishable in different inertial frames. Moving to the full relativistic theory reveals that the symmetry goes deeper. Investigating the symmetries of the relativistic theory end up leading the way to QFT, a relativistic description of quantum mechanics.

Getting into the framework, the goal of physics is to describe how a physical system evolves over time, and this evolution is usually given by some differential equation describing the state the system will take in the next interval of time given the current time. Moving from state to state from one interval of time to the next, the system traces out a path in some abstract space of possible states. As a concrete example, the differential equations describing the motion of a Newtonian particle determine the position and velocity of the particle at the next instant of time based upon the values in the previous instant of time. So how does one get the appropriate differential equation? At classical scales, nature tries to minimize the difference between the energy spent [3] and the energy available to spend and it minimizes the action S. At the more fundamental level of quantum mechanics, all possible paths contribute with different phases, but those close to the minimum add constructively, contributing the most. As the action gets larger, the paths away from the minimum count less and less and the quantum rule agrees with the classical rule. The classical case is analyzed first in order to begin investigating the symmetries that lead to QFT and an appropriate description of nature.

Equation 3–1 presents the action S, where L is the Lagrangian, T is the kinetic energy and U is the potential energy.

$$\mathrm{S} = \int \mathrm{Ldt} = \int \mathrm{T} - \mathrm{Udt} \tag{3–1}$$

At an extremum of S, $\delta\mathrm{S} = 0$ if S is smooth: S must decrease, go through a slope of zero, and then increase (or vice versa). For a true extremum, $\delta\mathrm{S}$ must be 0 in all directions. So to get

[3] Spent here just means used as kinetic energy.

the equations of motion vary the parameters of L and solve the system of equations for the values that yield zero change in the action.

$$
\begin{aligned}
\frac{\delta S}{\delta z_1} &= \int L(z_1 + dz_1, z_2, ...)dt - \int L(z_1, z_2, ...)dt = 0 \\
\frac{\delta S}{\delta z_2} &= \int L(z_1, z_2 + dz_2, ...)dt - \int L(z_1, z_2, ...)dt = 0 \\
&...
\end{aligned}
\tag{3–2}
$$

Following this process provides the Euler-Lagrange equations, describing how the parameters z evolve over time. The z's may be the position and velocity, or the quantum fields, or the temperature and volume or some other set of parameters that describe the system. The Lagrangian for a Newtonian free particle in one dimension is pretty simple and gets the point across.

$$
S = \frac{1}{2}\int m\dot{x}^2 dt \tag{3–3}
$$

If the action is at an extremum, perturbing the path x(t) by adding the infinitesimal ϵ(t) leaves the action unchanged.

$$
S' = \frac{1}{2}\int m(\dot{x} + \dot{\epsilon})^2 dt = \frac{1}{2}\int m(\dot{x}^2 + 2\dot{x}\dot{\epsilon} + \dot{\epsilon}^2)dt = \frac{1}{2}\int m(\dot{x}^2 + 2\dot{x}d\dot{x})dt \tag{3–4}
$$

$$
\delta S = S' - S = 0 = \frac{1}{2}\int m(\dot{x}^2 + 2\dot{x}\dot{\epsilon})dt - \frac{1}{2}\int m\dot{x}^2 dt = \int m\dot{x}\dot{\epsilon}dt \tag{3–5}
$$

With x(t) fixed at the boundaries of the integral, ϵ must be zero at t_o and t_f, so integrating by parts yields the following equation

$$
\delta S = 0 = \epsilon(t_f)\dot{x}(t_f) - \epsilon(t_o)\dot{x}(t_o) + \int m\ddot{x}\epsilon dt = 0\dot{x}(t_f) - 0\dot{x}(t_o) + \int m\ddot{x}\epsilon dt = \int m\ddot{x}\epsilon dt. \tag{3–6}
$$

And this equation must be zero for any infinitesimal deviation ϵ,

$$
\delta S = 0 \rightarrow m\ddot{x} = 0, \tag{3–7}
$$

implying that a free particle keeps the same velocity over time. Note that a Newtonian boost by constant velocity $\mathrm{v} \rightarrow \mathrm{v}' = \mathrm{v} + \mathrm{u}$ [4] leaves the equations of motion consistent. In the unprimed frame, the particle has velocity v with 0 acceleration. In the primed frame, the particle has velocity v + u with 0 acceleration. Both observers see the particle act as if there are zero forces in play,

$$\mathrm{S} = \frac{1}{2}\int \mathrm{m}(\mathrm{v}+\mathrm{u})^2 \mathrm{dt} = \frac{1}{2}\int \mathrm{m}(\mathrm{v}')^2 \mathrm{dt} \rightarrow \delta\mathrm{S} = 0 \rightarrow \mathrm{m}\frac{\mathrm{d}}{\mathrm{dt}}(\mathrm{v}+\mathrm{u}) = \mathrm{m}\frac{\mathrm{d}}{\mathrm{dt}}(\mathrm{v}') = 0. \tag{3–8}$$

If u is not constant but a function of time, u(t), then the equations of motion do not describe the same time evolution,

$$\mathrm{m}\frac{\mathrm{d}}{\mathrm{dt}}(\mathrm{v}+\mathrm{u}) = \mathrm{m}\dot{\mathrm{v}} + \mathrm{m}\dot{\mathrm{u}} = 0 \rightarrow \dot{\mathrm{v}} = -\dot{\mathrm{u}}. \tag{3–9}$$

In the case where u(t) depends upon time, the difference between the primed and unprimed frames' equations of motion is then $\delta\mathrm{F}$.

$$\delta\mathrm{F} = \mathrm{m}\frac{\mathrm{d}}{\mathrm{dt}}(\mathrm{v}+\mathrm{u}) - \mathrm{m}\frac{\mathrm{dv}}{\mathrm{dt}} = \mathrm{m}\dot{\mathrm{v}} + \mathrm{m}\dot{\mathrm{u}} - \mathrm{m}\dot{\mathrm{v}} = \mathrm{m}\dot{\mathrm{u}} \tag{3–10}$$

In the unprimed frame, the particle identified by the mass moves with constant velocity, $\dot{\mathrm{v}} = 0$. The observer in the primed frame looks at the particle with the same mass and sees it change velocity given by the equation $\dot{\mathrm{v}} = -\dot{\mathrm{u}}$. As an example, set v and u, and $\dot{\mathrm{u}}$ to zero for all times before t=0, and let u,and $\dot{\mathrm{u}}$ turn on after time 0. Both observers will agree that the particle is stationary up until time 0. After which, the observer in the primed frame will see the particle accelerate in strange ways. Meanwhile, the unprimed frame will continue to observe a stationary particle.

In general, every inertial frame finds $\delta\mathrm{F} = 0$ and every accelerating frame finds an extra force $\delta\mathrm{F}$ unique to its acceleration. In this way, no observer in an inertial frame can perform an experiment and determine which inertial frame he or she is in. On the other hand, each

[4] Renaming $\dot{\mathrm{x}}$ as v.

accelerating frame is identified by its $\delta\mathrm{F}$. Put another way, in every inertial frame, a ball released at rest remains at rest. In an accelerating frame, the ball will accelerate according to the motion of the frame $\delta\mathrm{F}$ and this change in the laws of physics identifies the frame in a unique way. Conversely, the laws of physics remain the same boosting between inertial frames, and this invariance is a symmetry of physics. Of course this example is Newtonian and the correct way to boost is given by the Lorentz transformation from Special Relativity, but this gets the point across.

Delving further along the path of symmetry, the fundamental forces depend only on the distance from the charge and not the direction implying that rotations are also a symmetry. This can be seen by looking at the Lagrangian,

$$\mathrm{L} = \frac{1}{2}\mathrm{m}\dot{\vec{\mathrm{x}}}^2 - \mathrm{U}((\vec{\mathrm{x}} - \vec{\mathrm{x}}')^2). \tag{3–11}$$

Rotations leave dot products and consequently the magnitude of vectors unchanged so the Lagrangian is invariant under this transformation. Naturally if the Lagrangian is invariant the equations of motion will be as well,

$$\mathrm{m}\frac{\mathrm{d}\vec{\mathrm{v}}}{\mathrm{dt}} = \vec{\nabla}\mathrm{U}. \tag{3–12}$$

In the equations of motion of 3–12, both sides are vectors and vectors transform the same way under rotations so the equations of motion are invariant. Note that in the case of rotations both the Lagrangian and the equations of motion are invariant, while for Newtonian boosts, only the equations of motion were invariant. This is due to the fact that Newtonian mechanics is the low velocity limit of relativistic mechanics. In the theory of Special Relativity, the action for a massive free particle is written like so,

$$\mathrm{S} = \int \frac{\mathrm{m}}{2}\mathrm{u}^{\mu}\mathrm{u}_{\mu}\mathrm{d}\tau. \tag{3–13}$$

Just as rotations preserve the dot product, Lorentz transformations (boosts and rotations) preserve the four vector product. Building a Lagrangian out of four vector products then gaurantees that both the Lagrangian and the equations of motion will remain invariant under

Lorentz transformations. However, four vectors aren't the most fundamental objects with invariant products. By studying the properties of the Lorentz group, it's possible to find even more fundamental building blocks called spinors. Spinors, vectors, and scalars are necessary to describe the different types of observed particles and create a Lagrangian that can describe the real world.

3.1.3 QFT From Symmetry

As just explored, the laws of physics are invariant under boosts and rotations, and the Lagrangian provides a mathematical framework for physical predictions. These facts together imply that there's a good shot at building a proper QFT by creating the appropriate invariant Lagrangian. Four vector products remain invariant under Lorentz transformations so they are a natural ingredient, but, as aforementioned, there are other mathematical objects that could be used as well. In this vein, the symmetries under rotations and boosts are investigated in order to look for other building blocks. The goal is to find three different representations of the Lorentz group and use one representation for scalar spin 0 particles, another for the spin $\frac{1}{2}$ fermions, and another for the spin 1 bosons.

3.1.3.1 Rotations

Rotations in three dimensions are described by the SO(3) group. Rotations preserve the lengths of vectors and the angles between them, which means that dot products between vectors remain invariant as well. In three dimensions, one can rotate about any of the three axes. The rotations about the x, y, and z axes may be characterized by the matrices below.

$$\mathrm{R_x} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\theta_\mathrm{x} & -\sin\theta_\mathrm{x} \\ 0 & \sin\theta_\mathrm{x} & \cos\theta_\mathrm{x} \end{pmatrix} \tag{3–14}$$

$$\mathrm{R_y} = \begin{pmatrix} \cos\theta_\mathrm{y} & 0 & \sin\theta_\mathrm{y} \\ 0 & 1 & 0 \\ -\sin\theta_\mathrm{y} & 0 & \cos\theta_\mathrm{y} \end{pmatrix} \tag{3–15}$$

$$R_z = \begin{pmatrix} \cos\theta_z & -\sin\theta_z & 0 \\ \sin\theta_z & \cos\theta_z & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{3–16}$$

These rotations may be built up from repeated rotations by an infinitesimally small angle $d\theta$. The matrices characterizing an infinitesimal rotation are produced by taking the limit as θ goes to zero.

$$dR_x = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & -d\theta_x \\ 0 & d\theta_x & 1 \end{pmatrix} = 1 - id\theta_x \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{pmatrix} = 1 - id\theta_x J_x \tag{3–17}$$

$$dR_y = \begin{pmatrix} 1 & 0 & d\theta_y \\ 0 & 1 & 0 \\ -d\theta_y & 0 & 1 \end{pmatrix} = 1 - id\theta_y \begin{pmatrix} 0 & 0 & i \\ 0 & 0 & 0 \\ -i & 0 & 0 \end{pmatrix} = 1 - id\theta_y J_y \tag{3–18}$$

$$dR_z = \begin{pmatrix} 1 & -d\theta_z & 0 \\ d\theta_z & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = 1 - id\theta_z \begin{pmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = 1 - id\theta_z J_z \tag{3–19}$$

Repeating an infinitesimal rotation many times builds the finite rotation, and in this way, the J matrices generate rotations along their respective axes. As such, they are aptly referred to as the generators of the group. Consider any of the J matrices,

$$R = (1 - i\frac{\theta}{N}J)^N = 1 + (-id\theta J) + \frac{1}{2!}(-id\theta J)^2 + \frac{1}{3!}(-id\theta J)^3 + ... = e^{-i\theta J}. \tag{3–20}$$

Notice that even powers of J yield J^2 and that odd powers of J return J,

$$= 1 - J^2 + J^2(1 + \frac{i^2}{2!}d\theta^2 + \frac{i^4}{4!}d\theta^4 + ...) - iJ(d\theta + \frac{i^2}{3!}d\theta^3 + \frac{i^4}{5!}d\theta^5 + ...) = (1 - J^2) + J^2\cos\theta - iJ\sin\theta. \tag{3–21}$$

Plugging in $\mathrm{J_z}$ reveals that this process does in fact rebuild the rotation matrix $\mathrm{R_z}$,

$$\begin{aligned}\mathrm{R_z} &= \left(\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix} - \begin{pmatrix}1&0&0\\0&1&0\\0&0&0\end{pmatrix}\right) + \begin{pmatrix}1&0&0\\0&1&0\\0&0&0\end{pmatrix}\cos\theta_\mathrm{z} + \begin{pmatrix}0&-1&0\\1&0&0\\0&0&0\end{pmatrix}\sin\theta_\mathrm{z}\\ &= \begin{pmatrix}\cos\theta_\mathrm{z}&-\sin\theta_\mathrm{z}&0\\ \sin\theta_\mathrm{z}&\cos\theta_\mathrm{z}&0\\0&0&1\end{pmatrix}.\end{aligned} \tag{3–22}$$

Similarly, the other generators rebuild their respective rotation matrices. The generators of the group are actually more fundamental than the rotation matrices. The multiplication table for the generators describes the algebra of the group, which describes the behavior of rotations at a local level. In fact the SO(3) rotation matrices are just one of the groups with this local algebra, and a specific group obeying the local algebra is analagous to a specific solution of a differential equation: each solution has a different global behavior yet each obeys the same physics at the differential scale. Moreover, the multiplication table can be specified without declaring any particular representation for the generators.

$$\begin{aligned}\mathrm{J_x} * \mathrm{J_y} &= \mathrm{iJ_z} + \mathrm{J_y} * \mathrm{J_x}\\ \mathrm{J_y} * \mathrm{J_z} &= \mathrm{iJ_x} + \mathrm{J_z} * \mathrm{J_y}\\ \mathrm{J_z} * \mathrm{J_x} &= \mathrm{iJ_y} + \mathrm{J_x} * \mathrm{J_z}\end{aligned} \tag{3–23}$$

The multiplication table can be specified in a more compact notation using the commutator[5] , $[\mathrm{a},\mathrm{b}] = \mathrm{ab} - \mathrm{ba}$, which closes the group, and the antisymmetric tensor ϵ.

$$[\mathrm{J_k}, \mathrm{J_l}] = \mathrm{i}\epsilon_{\mathrm{klm}}\mathrm{J_m} \tag{3–24}$$

[5] The Lie Bracket defines the multiplication for a Lie Algebra and this reduces to the commutator for Lie groups of matrices like SO(3). Using the commutator below returns another member of the group and thus the group is closed under commutation.

Finding a 3x3 representation of the generators – that obey the algebra – and then repeatedly applying the infinitesimal transormations builds the SO(3) rotation group. The group acts on real 3x1 objects called vectors, and these 3x1 vectors are a suitable candidate for a Newtonian Lagrangian. Finding another representation obeying this algebra will provide a more fundamental ingredient for the Lagrangian and allow the construction of a proper QFT. Similar to the way real numbers are built from the squares of imaginary numbers, vectors are built from spinors.

Looking for the lowest order nontrivial nxn matrices satisfying the algebra gives the 2x2 Pauli matrices. The 1x1 matrices are the trival solution: 1x1 matrices are simply scalar complex numbers, which commute and therefore fail to satisfy the algebra unless all of the J matrices are 0. The objects these 1x1 operators act on are 1x1 numbers called scalars which remain invariant under rotations and correspond to spin 0. The solution for the 2x2 case, the Pauli matrices are given by

$$\sigma_{\mathrm{x}} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \sigma_{\mathrm{y}} = \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix}, \sigma_{\mathrm{y}} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \tag{3–25}$$

However, plugging these into the commutator reveals a factor of two difference,

$$[\sigma_{\mathrm{k}}, \sigma_{\mathrm{l}}] = 2\mathrm{i}\epsilon_{\mathrm{klm}}\sigma_{\mathrm{m}}. \tag{3–26}$$

Defining $\mathrm{J}_{\mathrm{k}} = \frac{1}{2}\sigma_{\mathrm{k}}$ fixes this. These matrices act on an array of 2x1 complex numbers called spinors, and this new rotation group is called SU(2). By starting with vectors and analyzing the SO(3) rotation group along with its underlying algebra, new mathematical objects have been discovered. The 1x1 matrices satisfying the rotation algebra make up the spin 0 representation of SU(2), the complex 2x2 matrices acting on complex 2x1 objects satisfying the algebra make up the spin $\frac{1}{2}$ representation, and the complex 3x3 matrices acting on complex 3x1 objects satisfying the algebra make up the spin 1 representation. The pattern continues on. There are in fact many representations of SU(2). It's now possible to use these representations to build rotationally invariant Lagrangians, using the 2x1 spinors of SU(2) to model fermions and the 3x1 representation to model bosons. While rotationally invariant Lagrangians are important for nonrelativistic theories, the real goal is to break down four vectors in the same way to find the most fundamental ingredients for relativistic Lagrangians.

3.1.3.2 The Lorentz Group

Four vector products are invariant with regards to rotations and boosts. This statement is defined by the mathematical equation below, where the Λ matrices represent the rotation/boost matrices of the Lorentz Group [6] and the η matrix is the Minkowski metric $\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$.

$$\mathrm{x}'_\mu \mathrm{x}'^\mu = \mathrm{x}_\mu \mathrm{x}^\mu \rightarrow \eta_{\sigma\rho}\Lambda^\sigma_\mu \Lambda^\rho_\nu \mathrm{x}^\mu \mathrm{x}^\nu = \eta_{\mu\nu}\mathrm{x}^\mu \mathrm{x}^\nu \rightarrow \eta_{\sigma\rho}\Lambda^\sigma_\mu \Lambda^\rho_\nu = \eta_{\mu\nu} \tag{3–27}$$

In this 3 + 1 dimensional space the rotations and their corresponding generators are now given by the following R and J matrices,

$$\mathrm{R_x} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \cos\theta_\mathrm{x} & -\sin\theta_\mathrm{x} \\ 0 & 0 & \sin\theta_\mathrm{x} & \cos\theta_\mathrm{x} \end{pmatrix}, \mathrm{J_x} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -\mathrm{i} \\ 0 & 0 & \mathrm{i} & 0 \end{pmatrix} \tag{3–28}$$

$$\mathrm{R_y} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta_\mathrm{y} & 0 & \sin\theta_\mathrm{y} \\ 0 & 0 & 1 & 0 \\ 0 & -\sin\theta_\mathrm{y} & 0 & \cos\theta_\mathrm{y} \end{pmatrix}, \mathrm{J_y} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \mathrm{i} \\ 0 & 0 & 0 & 0 \\ 0 & -\mathrm{i} & 0 & 0 \end{pmatrix} \tag{3–29}$$

[6] The Lorentz group dealt with here is the proper orthochronous Lorentz Group SO(1,3).

$$\mathrm{R_z} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\theta_\mathrm{z} & -\sin\theta_\mathrm{z} & 0 \\ 0 & \sin\theta_\mathrm{z} & \cos\theta_\mathrm{z} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \mathrm{J_z} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -\mathrm{i} & 0 \\ 0 & \mathrm{i} & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \tag{3–30}$$

where the R matrices satisfy equation $3--27$. The boosts are given by the B matrices,

$$\mathrm{B_x} = \begin{pmatrix} \cosh\omega_\mathrm{x} & \sinh\omega_\mathrm{x} & 0 & 0 \\ \sinh\omega_\mathrm{x} & \cosh\omega_\mathrm{x} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{3–31}$$

$$\mathrm{B_y} = \begin{pmatrix} \cosh\omega_\mathrm{y} & 0 & \sinh\omega_\mathrm{y} & 0 \\ 0 & 1 & 0 & 0 \\ \sinh\omega_\mathrm{y} & 0 & \cosh\omega_\mathrm{y} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{3–32}$$

$$\mathrm{B_z} = \begin{pmatrix} \cosh\omega_\mathrm{z} & 0 & 0 & \sinh\omega_\mathrm{z} \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \sinh\omega_\mathrm{z} & 0 & 0 & \cosh\omega_\mathrm{z} \end{pmatrix}. \tag{3–33}$$

These also leave the four vector product invariant. Looking at the differential boosts yields the generators K.

$$\mathrm{dB_x} = \begin{pmatrix} 1 & \mathrm{d}\omega_\mathrm{x} & 0 & 0 \\ \mathrm{d}\omega_\mathrm{x} & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = 1 + \mathrm{d}\omega_\mathrm{x} \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = 1 + \mathrm{d}\omega_\mathrm{x}\mathrm{K_x} \tag{3–34}$$

$$dB_y = \begin{pmatrix} 1 & 0 & d\omega_y & 0 \\ 0 & 1 & 0 & 0 \\ d\omega_y & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} = 1 + d\omega_y \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} = 1 + d\omega_y K_y \tag{3–35}$$

$$dB_z = \begin{pmatrix} 1 & 0 & 0 & d\omega_z \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ d\omega_z & 0 & 0 & 1 \end{pmatrix} = 1 + d\omega_z \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} = 1 + d\omega_z K_z \tag{3–36}$$

As before with the algebra for rotations in three dimensions the Lorentz algebra is defined by its multiplication table, but now there are rotations and boosts. The multiplication table is given by the commutation relations below, which can be confirmed by brute force computation.

$$[J_i, J_j] = i\epsilon_{ijk} J_k \tag{3–37}$$

$$[J_i, K_j] = i\epsilon_{ijk} K_k \tag{3–38}$$

$$[K_i, K_j] = -i\epsilon_{ijk} J_k \tag{3–39}$$

Notice that the commutator between J matrices returns another J matrix, but that the commutator between K matrices returns a J matrix. This means that the J operators form their own subgroup, but the K operators don't. On the other hand, mixing the Js and Ks up by defining the $Y^{\pm}$ operators allows the Lorentz algebra to be represented by two independent subgroups.

$$Y^{\pm} = \frac{1}{2}(J_i \pm iK_i) \tag{3–40}$$

$$[Y_i^{\pm}, Y_j^{\pm}] = i\epsilon_{ijk} Y_k^{\pm} \tag{3–41}$$

$$[Y_i^{\pm}, Y_j^{\mp}] = 0 \tag{3–42}$$

Both of the Y groups have the same commutation relations as SU(2). In this way, the Lorentz algebra can be viewed as if two orthogonal SU(2) rotation algebras have been glued together.

This is similar to the way in which orthogonal basis vectors are stuck together to create a larger dimensional space. Now, in order to figure out how to use this space to build the appropriate Lagrangians, the individual subspaces must be investigated. So looking at Y^+ alone is akin to looking along the Y^+ axis by setting Y^- to zero. Using $(\mathrm{y}_+, \mathrm{y}_-)$ to label the representation, the simplest nontrivial case along the Y^+ axis is spin $\frac{1}{2}$ x spin 0, given by $(\frac{1}{2}, 0)$.

Using $\mathrm{Y}_\mathrm{i}^- = \frac{1}{2}(\mathrm{J}_\mathrm{i} - \mathrm{i}\mathrm{K}_\mathrm{i}) = 0$ implies that $\mathrm{J}_\mathrm{i} = \mathrm{i}\mathrm{K}_\mathrm{i}$, and since Y_i^+ is the 2x2 representation obeying the SU(2) algebra, $\mathrm{Y}_\mathrm{i}^+ = \frac{\sigma_\mathrm{i}}{2}$. Putting this together,

$$\begin{aligned}
&\mathrm{J}_\mathrm{i} = \mathrm{i}\mathrm{K}_\mathrm{i} \\
&\quad \rightarrow \mathrm{Y}_\mathrm{i}^+ = \frac{1}{2}(\mathrm{J}_\mathrm{i} + \mathrm{i}\mathrm{K}_\mathrm{i}) = \frac{\sigma_\mathrm{i}}{2} = \frac{1}{2}(\mathrm{J}_\mathrm{i} + \mathrm{J}_\mathrm{i}) = \mathrm{J}_\mathrm{i} \\
&\quad \rightarrow \mathrm{J}_\mathrm{i} = \frac{1}{2}\sigma_\mathrm{i} \\
&\quad \rightarrow \mathrm{K}_\mathrm{i} = \frac{-\mathrm{i}\sigma_\mathrm{i}}{2}.
\end{aligned} \tag{3–43}$$

Finally, the finite Lorentz transformations for the $(\frac{1}{2}, 0)$ representation are given by

$$\mathrm{R}^{(\mathrm{L})} = \mathrm{e}^{\mathrm{i}\theta_\mathrm{i}\mathrm{J}_\mathrm{i}} = \mathrm{e}^{\mathrm{i}\theta_\mathrm{i}\frac{\sigma_\mathrm{i}}{2}} \tag{3–44}$$

for rotations, and

$$\mathrm{B}^{(\mathrm{L})} = \mathrm{e}^{\mathrm{i}\phi_\mathrm{i}\mathrm{K}_\mathrm{i}} = \mathrm{e}^{\phi_\mathrm{i}\frac{\sigma_\mathrm{i}}{2}} \tag{3–45}$$

for boosts. These act on 2x1 objects called left-chiral spinors, $\mathcal{L}$. A general Lorentz transformation on a left-chiral spinor can be written

$$\Lambda^{(\mathrm{L})} = \mathrm{e}^{\frac{\mathrm{i}}{2}\theta_\mathrm{i}\sigma_\mathrm{i} + \frac{1}{2}\phi_\mathrm{i}\sigma_\mathrm{i}} \tag{3–46}$$

. Replacing K_i with $-\mathrm{K}_\mathrm{i}$ takes Y_i^+ to Y_i^-, and gives the finite Lorentz transformations for the $(0, \frac{1}{2})$ representation

$$\mathrm{R}^{(\mathrm{R})} = \mathrm{e}^{\mathrm{i}\theta_\mathrm{i}\mathrm{J}_\mathrm{i}} = \mathrm{e}^{\frac{\mathrm{i}}{2}\theta_\mathrm{i}\sigma_\mathrm{i}} \tag{3–47}$$

for rotations, and

$$\mathrm{B}^{(\mathrm{R})} = \mathrm{e}^{\mathrm{i}\phi_\mathrm{i}\mathrm{K}_\mathrm{i}} = \mathrm{e}^{-\phi_\mathrm{i}\frac{\sigma_\mathrm{i}}{2}} \tag{3–48}$$

for boosts. These act on 2x1 objects called right-chiral spinors, $\mathcal{R}$. The general Lorentz transformation on a right-chiral spinor can be written

$$\Lambda^{(\mathrm{R})} = \mathrm{e}^{\frac{\mathrm{i}}{2}\theta_\mathrm{i}\sigma_\mathrm{i} - \frac{1}{2}\phi_\mathrm{i}\sigma_\mathrm{i}} \tag{3–49}$$

Last but not least, rank 2 spinors are given by the $(\frac{1}{2}, \frac{1}{2})$ representation. This representation is a tensor combining two spinors via outer product, and as will be shown later, these objects are actually four vectors.

$$\alpha = \mathcal{L}\mathcal{R}^\mathrm{T} \tag{3–50}$$

In order to transform α, both $\mathcal{L}$ and $\mathcal{R}$ must be transformed.

$$\alpha^{'} = \Lambda^{(\mathrm{L})}\mathcal{L}\mathcal{R}^\mathrm{T}\Lambda^{(\mathrm{R})\mathrm{T}} = \mathrm{e}^{\frac{\mathrm{i}}{2}\theta_\mathrm{i}\sigma_\mathrm{i} + \frac{1}{2}\phi_\mathrm{i}\sigma_\mathrm{i}}\mathcal{L}\mathcal{R}^\mathrm{T}\mathrm{e}^{\frac{\mathrm{i}}{2}\theta_\mathrm{i}\sigma_\mathrm{i}^\mathrm{T} - \frac{1}{2}\phi_\mathrm{i}\sigma_\mathrm{i}^\mathrm{T}} \tag{3–51}$$

However it would be nice if the transformation term on the right side was the Hermitian conjugate of the transformation on the left side. This would be the case if $\sigma_\mathrm{i}^\mathrm{T}$ was $-\sigma_\mathrm{i}^\dagger = -(\sigma_\mathrm{i}^*)^\mathrm{T}$, and this requires a transformation that turns σ into $-\sigma^*$. So $\mathcal{R}$ is rearranged such that $\mathcal{R} \to \tilde{\mathcal{R}} = \mathrm{t}\mathcal{R}$ where $\mathrm{t}\sigma_\mathrm{i}\mathrm{t}^{-1} = -\sigma_\mathrm{i}^*$. The matrix $\mathrm{t} = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$ satisfies the requirements. This change of basis for $\mathcal{R}$ redefines α,

$$\alpha = \mathcal{L}\tilde{\mathcal{R}}^\mathrm{T}. \tag{3–52}$$

Defined in this manner, the $(\frac{1}{2}, \frac{1}{2})$ representation now transforms like so,

$$\alpha^{'} = \mathrm{e}^{\frac{\mathrm{i}}{2}\theta_\mathrm{i}\sigma_\mathrm{i} + \frac{1}{2}\phi_\mathrm{i}\sigma_\mathrm{i}}\mathcal{L}\tilde{\mathcal{R}}^\mathrm{T}\mathrm{e}^{-\frac{\mathrm{i}}{2}\theta_\mathrm{i}\sigma_\mathrm{i}^\dagger + \frac{1}{2}\phi_\mathrm{i}\sigma_\mathrm{i}^\dagger}. \tag{3–53}$$

Considering the fact that $\sigma_\mathrm{i}^\dagger = \sigma_\mathrm{i}$, the transformation reduces further,

$$\alpha^{'} = \mathrm{e}^{\frac{\mathrm{i}}{2}\theta_\mathrm{i}\sigma_\mathrm{i} + \frac{1}{2}\phi_\mathrm{i}\sigma_\mathrm{i}}\alpha\mathrm{e}^{-\frac{\mathrm{i}}{2}\theta_\mathrm{i}\sigma_\mathrm{i} + \frac{1}{2}\phi_\mathrm{i}\sigma_\mathrm{i}}. \tag{3–54}$$

A transformation of the form $\mathrm{M}' = \mathrm{HMH}^\dagger$ where H is a Hermitian matrix, preserves the Hermitivity of the matrix M. Namely, M' will be Hermitian if M is Hermitian,

$$\mathrm{M}'^\dagger = (\mathrm{HMH}^\dagger)^\dagger = \mathrm{HM}^\dagger\mathrm{H}^\dagger = \mathrm{HMH}^\dagger = \mathrm{M}'. \tag{3–55}$$

On the other hand, M' will be anti-Hermitian if M is anti-Hermitian,

$$\mathrm{M}'^\dagger = (\mathrm{HMH}^\dagger)^\dagger = \mathrm{HM}^\dagger\mathrm{H}^\dagger = \mathrm{H}(-\mathrm{M})\mathrm{H}^\dagger = -\mathrm{HMH}^\dagger = -\mathrm{M}'. \tag{3–56}$$

The transformation of the $(\frac{1}{2}, \frac{1}{2})$ object α is of this type. With this in mind, note that any complex matrix can be broken up into a Hermitian piece and an anti-Hermitian piece, and that these pieces remain independent under the $(\frac{1}{2}, \frac{1}{2})$ Lorentz transformations constructed here. Then note that a general complex 2x2 matrix has 8 free parameters, 4 from the Hermitian part and 4 from the anti-Hermitian part. Meanwhile, α has only 4, two from each spinor [7] . This means that α may be represented by the Hermitian space alone. An appropriate basis for this space is the collection of Pauli matrices plus the identity,

$$\sigma_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \sigma_2 = \begin{pmatrix} 0 & -\mathrm{i} \\ \mathrm{i} & 0 \end{pmatrix}, \text{ and } \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \tag{3–57}$$

Thus any $(\frac{1}{2}, \frac{1}{2})$ object α may be written in terms of its four independent parameters like so,

$$\begin{aligned} \alpha &= \alpha_0\sigma_0 + \alpha_1\sigma_1 + \alpha_2\sigma_2 + \alpha_3\sigma_3 \\ \alpha &= \begin{pmatrix} \alpha_0 + \alpha_3 & \alpha_1 - \mathrm{i}\alpha_2 \\ \alpha_1 + \mathrm{i}\alpha_2 & \alpha_0 - \alpha_3 \end{pmatrix}. \end{aligned} \tag{3–58}$$

[7] A spinor has 2 complex components yielding 4 parameters. Two constraints reduce the number of free parameters to 2. One constraint requires a magnitude of 1, and the other requires that the overall phase doesn't matter.

Boosting α along the z direction hints that the four components transform like a four vector.

$$\alpha^{'} = \mathrm{e}^{\frac{1}{2}\phi_3\sigma_3}\begin{pmatrix}\alpha_0+\alpha_3 & \alpha_1-\mathrm{i}\alpha_2\\ \alpha_1+\mathrm{i}\alpha_2 & \alpha_0-\alpha_3\end{pmatrix}\mathrm{e}^{\frac{1}{2}\phi_3\sigma_3} \tag{3–59}$$

Exponentiating the σ_3 matrix and multiplying everything yields,

$$\begin{aligned}\begin{pmatrix}\alpha_0'+\alpha_3' & \alpha_1'-\mathrm{i}\alpha_2'\\ \alpha_1'+\mathrm{i}\alpha_2' & \alpha_0'-\alpha_3'\end{pmatrix} &= \begin{pmatrix}\mathrm{e}^{\frac{1}{2}\phi_3} & 0\\ 0 & \mathrm{e}^{-\frac{1}{2}\phi_3}\end{pmatrix}\begin{pmatrix}\alpha_0+\alpha_3 & \alpha_1-\mathrm{i}\alpha_2\\ \alpha_1+\mathrm{i}\alpha_2 & \alpha_0-\alpha_3\end{pmatrix}\begin{pmatrix}\mathrm{e}^{\frac{1}{2}\phi_3} & 0\\ 0 & \mathrm{e}^{-\frac{1}{2}\phi_3}\end{pmatrix}\\ &= \begin{pmatrix}\mathrm{e}^{\phi_3}(\alpha_0+\alpha_3) & (\alpha_1-\mathrm{i}\alpha_2)\\ (\alpha_1+\mathrm{i}\alpha_2) & \mathrm{e}^{-\phi_3}(\alpha_0-\alpha_3)\end{pmatrix}.\end{aligned} \tag{3–60}$$

Then solving the systems of equations makes the transformation clearer,

$$\begin{aligned}\alpha_1' &= \alpha_1\\ \alpha_2' &= \alpha_2\\ \alpha_0' &= (\cosh\phi_3)\alpha_0 + (\sinh\phi_3)\alpha_3\\ \alpha_3' &= (\sinh\phi_3)\alpha_0 + (\cosh\phi_3)\alpha_3.\end{aligned} \tag{3–61}$$

Finally the transformation can be written as a 4x4 matrix acting on the 4x1 four vector,

$$\alpha' = \begin{pmatrix}\cosh\phi_3 & 0 & 0 & \sinh\phi_3\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\\ \sinh\phi_3 & 0 & 0 & \cosh\phi_3\end{pmatrix}\begin{pmatrix}\alpha_0\\ \alpha_1\\ \alpha_2\\ \alpha_3\end{pmatrix} \tag{3–62}$$

The other transformations reproduce the usual 4x4 Lorentz transformations as well. In this way, four vectors are just rearranged versions of the $(\frac{1}{2}, \frac{1}{2})$ rank 2 spinors. In the way complex numbers are the square root of real numbers, spinors are the square root of four vectors.
In practice it's easier to work with 4 vectors as 4x1 column vectors, though one could use the $(\frac{1}{2}, \frac{1}{2})$ rank 2 spinor representation. With four vectors, spinors, and scalars in hand, the

invariant Lagrangians describing the different types of particles can now be built. Before doing this, however, there is a brief aside about group representations, particles, and labels.

3.1.3.3 The Poincare Group and Particle Labels

The Poincare group is the Lorentz Group plus translations. Physics experiments should be the same with a rotated apparatus, an apparatus moving at a different constant velocity, and also at another location. Adding translations to the group amounts to adding another the set of translation generators $\mathrm{P_i}$. In the "Particles, Symmetries, and Labels" section the argument was made that particles are things that observers agree on between frames. Distinguishable particles have labels that remain invariant. A spin $\frac{1}{2}$ particle with mass m remains a spin $\frac{1}{2}$ particle with mass m for all observers. As it turns out these labels relate to different subspaces in the representation of the group.

Consider a group with certain generators. When the generators of the group are represented a certain way, say as some NxN matrices, there is a likelihood that there are some invariant subspaces. As an example, a representation of a group may act on a 5 dimensional space spanned by vectors $\mathrm{e_1, e_2, e_3, e_4, e_5}$. The group may always transform vectors in the $\mathrm{e_1, e_2, e_3}$ subspace into one another, and those in $\mathrm{e_4, e_5}$ into one another, but never mix up $\mathrm{e_1, e_2, e_3}$ with $\mathrm{e_4, e_5}$. In this example the 5x5 operators could be decomposed into a 3x3 operator acting only on the $\mathrm{e_1, e_2, e_3}$ subspace and a 2x2 operator acting only on the $\mathrm{e_4, e_5}$ subspace. Each 5x5 member of the group could be written as a block diagonal matrix with a 3x3 block and a 2x2 block. The same goes for the generators. Since these subspaces retain their identity under the transformations of the group[8] , they may be considered different particles. The question is whether there are labels for the subspaces.

Labels in physics need to be measurable, and by the current understanding of quantum mechanics these must be eigenvalues of Hermitian operators. The operators that label these

[8] This assumes that the 3x3 and 2x2 pieces have no invariant subspaces besides themselves and 0.

subspaces are called the Casimir operators and must be built from the generators of the group. The operators should give the same values for $\mathrm{e_1, e_2, and e_3}$ since they transform into one another and represent the same particle, which implies that the Casimir labeling operator must be proportional to the identity in that subspace. The same goes for the operator in the $\mathrm{e_4, e_5}$ subspace. The Casimir operator for SU(2) is the J^2 operator which labels the spin of the particle. The Poincare group has two Casimir operators coresponding to two labels the mass, m, and the spin, j. Looking at the irreducible representations[9] of the Poincare group, the mass and spin arise naturally as labels for particles.

3.1.3.4 Building the Lagrangian for a Free Scalar Particle

Analyzing the symmetries of the SO(3) rotation group revealed the SU(2) group, the 2x2 complex Pauli matrices, and the complex 2x1 spinors. Analyzing the symmetries of the Lorentz group revealed the $(\frac{1}{2}, 0)$ left-chiral spinors, the $(0, \frac{1}{2})$ right chiral spinors, and the four vectors encoded in the $(\frac{1}{2}, \frac{1}{2})$ representation. Now, these pieces are put together to form relativistically invariant Lagrangians for free particles.

The $(0, 0)$ representation of the Lorentz group is a scalar, which means that it remains the same under Lorentz transformations. This representation is used for spin 0 particles. The equations of motion must relate the change in the scalar field Φ at one moment in space and time to the value of Φ at the next moment in space and time, so the Lagrangian must include both the field itself and the four vector derivative, ∂_μ. Note that both Newton's equations of motion, $\mathrm{F = m\ddot{x}}$, and the Schrodinger equation for the free particle, $\mathrm{i}\partial_{\mathrm{t}}\psi = \frac{1}{2\mathrm{m}}(-\mathrm{i}\partial_{\mathrm{x}})(-\mathrm{i}\partial_{\mathrm{x}})\psi$[10] have at most second order derivatives. The same holds for Maxwell's equations of electromagnetism. Hence, as an assumption, the derivative term in the Lagrangian

[9] This is the group theory term for the block diagonal pieces and the corresponding invariant subspaces. The irreducible representations, irreps, are those that can't be broken down in terms of smaller invariant subspaces and smaller block diagonal matrices. An arbitrary representation of the group is built from these irreps.

[10] Here the one dimensional case is presented for simplicity

will be the lowest order possible. As a Lorentz invariant scalar, any power of Φ can be included. On the other hand, ∂_μ is a four vector, and must be paired with a ∂^μ to form the invariant four vector product. Cross terms between these invariant pieces like $\Phi\partial_\mu\partial^\mu\Phi$ or $\partial_\mu\Phi\partial^\mu\Phi$ are also invariant, but lead to feedback between the derivatives and the value of the function. This means that the derivatives $\mathrm{i}\partial_\mathrm{t} = \mathrm{E}$ and $-\mathrm{i}\partial_\mathrm{i} = \mathrm{P_i}$ will change over time, but E and $\vec{\mathrm{P}}$ should remain constant for a free particle. The cross terms are thrown out to prevent this. The Φ terms with an order different than Φ^2 cause the same problem, and these are thrown out too. All of these choices lead to the following action,

$$\mathrm{S} = \int \mathrm{d}^4\mathrm{x} \left(\mathrm{c}_0 + \mathrm{c}_1\Phi^2 + \mathrm{c}_2\partial_\mu\Phi\partial^\mu\Phi + \mathrm{c}_3\partial_\mu\partial^\mu\Phi\right). \tag{3–63}$$

Note that the c_3 term is a total derivative and by the divergence theorem depends only on the values at the boundary, which are fixed. This implies that the contribution to the action from the c_3 term is the same regardless of how Φ changes in the volume. Because the Euler-Lagrange equations depend only on the variation in the volume, this term cannot contribute to $\delta\mathrm{S}$ and c_3 may be set to zero. The c_0 term is a more obvious constant and does not affect $\delta\mathrm{S}$ either, so it may be set to zero as well, leaving,

$$\mathrm{S} = \int \mathrm{d}^4\mathrm{x} \left(\mathrm{c}_1\Phi^2 + \mathrm{c}_2\partial_\mu\Phi\partial^\mu\Phi\right). \tag{3–64}$$

Finding Φ such that $\delta\mathrm{S} = 0$ amounts to applying the Euler-Lagrange equations $\partial_\mu\frac{\partial\mathrm{L}}{\partial(\partial_\mu\Phi)} = \frac{\partial\mathrm{L}}{\partial\Phi}$. These yield

$$\begin{aligned} &2\mathrm{c}_2\partial_\mu\partial^\mu\Phi = -2\mathrm{c}_1\Phi \\ &\rightarrow (-\mathrm{E}^2 + \vec{\mathrm{P}}^2)\Phi = \frac{\mathrm{c}_1}{\mathrm{c}_2}\Phi = -\mathrm{m}^2\Phi \\ &\rightarrow \frac{\mathrm{c}_1}{\mathrm{c}_2} = -\mathrm{m}^2 \end{aligned} \tag{3–65}$$

In order to get the correct dispersion relation for a relativistic particle, c_1 is set to $\frac{-1}{2}\mathrm{m}^2$ and c_2 is set to $\frac{1}{2}$. Notice that including any Φ^n with n$\neq$2 in the Lagrangian would have contributed to the differential equation via $\frac{\partial\mathrm{L}}{\partial\Phi}$ and that E, $\vec{\mathrm{P}}$ wouldn't be constant. Thus the equation wouldn't work for a free particle. The resulting equation of motion for the scalar particle, $\partial_\mu \partial^\mu \Phi = \mathrm{m}^2\Phi$, is called the Klein-Gordon equation, and provides the correct description for spin 0 particles. This was all derived using symmetry, a reasonable assumption about the order of the derviatives, and the fact that the Energy shouldn't change over time for a free particle. The final Lagrangian for the scalar particle is

$$\mathrm{S} = \int \mathrm{d}^4\mathrm{x}\frac{1}{2}\left(\partial_\mu\Phi\partial^\mu\Phi - \mathrm{m}^2\Phi^2\right). \tag{3–66}$$

3.1.3.5 Building the Lagrangian for a Free Spin $\frac{1}{2}$ Particle

With the action for the free scalar particle in hand, the free spin $\frac{1}{2}$ particle is up next. The spin $\frac{1}{2}$ action must combine the $\mathcal{L}$ and $\mathcal{R}$ spinors of the $(\frac{1}{2}, 0)$ and $(0, \frac{1}{2})$ representations and the four vector, ∂_μ, in a Lorentz invariant way. Moving forward in this regard, the $\mathcal{L}$, $\mathcal{R}$, and ∂_μ transformations are now analyzed to find the lowest order invariant combinations. The transformation for the left-chiral spinor is

$$\Lambda^{(\mathrm{L})} = \mathrm{e}^{\frac{\mathrm{i}}{2}\theta_\mathrm{i}\sigma_\mathrm{i} + \frac{1}{2}\phi_\mathrm{i}\sigma_\mathrm{i}}, \tag{3–67}$$

and the transformation for the right-chiral spinor is,

$$\Lambda^{(\mathrm{R})} = \mathrm{e}^{\frac{\mathrm{i}}{2}\theta_\mathrm{i}\sigma_\mathrm{i} - \frac{1}{2}\phi_\mathrm{i}\sigma_\mathrm{i}}. \tag{3–68}$$

Taking the Hermitian conjugate of the left-chiral transformation gives,

$$(\Lambda^{(\mathrm{L})})^\dagger = \mathrm{e}^{-\frac{\mathrm{i}}{2}\theta_\mathrm{i}\sigma_\mathrm{i}^\dagger + \frac{1}{2}\phi_\mathrm{i}\sigma_\mathrm{i}^\dagger} = \mathrm{e}^{-\frac{\mathrm{i}}{2}\theta_\mathrm{i}\sigma_\mathrm{i} + \frac{1}{2}\phi_\mathrm{i}\sigma_\mathrm{i}}. \tag{3–69}$$

This reveals that $(\Lambda^{(\mathrm{L})})^\dagger$ is the inverse of $\Lambda^{(\mathrm{R})}$. Similarly, $(\Lambda^{(\mathrm{R})})^\dagger$ is the inverse of $\Lambda^{(\mathrm{L})}$. Thus, $\mathcal{L}^\dagger\mathcal{R}$ and $\mathcal{R}^\dagger\mathcal{L}$ are Lorentz invariants, which can be seen below,

$$(\mathcal{L}^\dagger\mathcal{R})' = \mathcal{L}^\dagger(\Lambda^{(\mathrm{L})})^\dagger\Lambda^{(\mathrm{R})}\mathcal{R} = \mathcal{L}^\dagger(\Lambda^{(\mathrm{R})})^{-1}\Lambda^{(\mathrm{R})}\mathcal{R} = \mathcal{L}^\dagger\mathcal{R} \tag{3–70}$$

$$(\mathcal{R}^\dagger\mathcal{L})' = \mathcal{R}^\dagger(\Lambda^{(\mathrm{R})})^\dagger\Lambda^{(\mathrm{L})}\mathcal{L} = \mathcal{R}^\dagger(\Lambda^{(\mathrm{L})})^{-1}\Lambda^{(\mathrm{L})}\mathcal{L} = \mathcal{R}^\dagger\mathcal{L}. \tag{3–71}$$

These are the lowest order invariant pieces involving the field alone. In order to couple the derivative to the field, the $\mathcal{L}$ and $\mathcal{R}$ spinors must attach to ∂_μ in an invariant way. Recall that a four vector may be expressed as an outer product of left and right spinors,

$$\alpha_{\mathrm{L\tilde{R}}} = \mathcal{L}_\alpha \tilde{\mathcal{R}}_\alpha^{\mathrm{T}}. \tag{3–72}$$

The following also works,

$$\alpha_{\mathrm{R\tilde{L}}} = \mathcal{R}_\alpha \tilde{\mathcal{L}}_\alpha^{\mathrm{T}}. \tag{3–73}$$

Recall that the $\sim$ transformation sent σ_{i} to $-\sigma_{\mathrm{i}}*$. So the transformations for the $\sim$ transposed spinors are,

$$(\Lambda^{(\tilde{\mathrm{L}})})^{\mathrm{T}} = \mathrm{e}^{\frac{\mathrm{i}}{2}\theta_{\mathrm{i}}(-\sigma_{\mathrm{i}}*)^{\mathrm{T}} + \frac{1}{2}\phi_{\mathrm{i}}(-\sigma_{\mathrm{i}}*)^{\mathrm{T}}} = \mathrm{e}^{-\frac{\mathrm{i}}{2}\theta_{\mathrm{i}}\sigma_{\mathrm{i}}^\dagger - \frac{1}{2}\phi_{\mathrm{i}}\sigma_{\mathrm{i}}^\dagger} = \mathrm{e}^{-\frac{\mathrm{i}}{2}\theta_{\mathrm{i}}\sigma_{\mathrm{i}} - \frac{1}{2}\phi_{\mathrm{i}}\sigma_{\mathrm{i}}} \tag{3–74}$$

and

$$(\Lambda^{(\tilde{\mathrm{R}})})^{\mathrm{T}} = \mathrm{e}^{\frac{\mathrm{i}}{2}\theta_{\mathrm{i}}(-\sigma_{\mathrm{i}}*)^{\mathrm{T}} - \frac{1}{2}\phi_{\mathrm{i}}(-\sigma_{\mathrm{i}}*)^{\mathrm{T}}} = \mathrm{e}^{-\frac{\mathrm{i}}{2}\theta_{\mathrm{i}}\sigma_{\mathrm{i}}^\dagger + \frac{1}{2}\phi_{\mathrm{i}}\sigma_{\mathrm{i}}^\dagger} = \mathrm{e}^{-\frac{\mathrm{i}}{2}\theta_{\mathrm{i}}\sigma_{\mathrm{i}} + \frac{1}{2}\phi_{\mathrm{i}}\sigma_{\mathrm{i}}}. \tag{3–75}$$

The equations above show that the $(\Lambda^{(\tilde{\mathrm{L}})})^{\mathrm{T}}$ transformation is the inverse of $\Lambda^{(\mathrm{L})}$, and that the $(\Lambda^{(\tilde{\mathrm{R}})})^{\mathrm{T}}$ transformation is the inverse of $\Lambda^{(\mathrm{R})}$. So the invariant pieces coupling the spinor to the four vector are,

$$\mathcal{R}^\dagger \alpha_{\mathrm{L\tilde{R}}} \mathcal{R} = \mathcal{R}^\dagger \mathcal{L}_\alpha \tilde{\mathcal{R}}_\alpha^{\mathrm{T}} \mathcal{R} \tag{3–76}$$

and

$$\mathcal{L}^\dagger \alpha_{\mathrm{R\tilde{L}}} \mathcal{L} = \mathcal{L}^\dagger \mathcal{R}_\alpha \tilde{\mathcal{L}}_\alpha^{\mathrm{T}} \mathcal{L}. \tag{3–77}$$

The two types of four vectors have slightly different transformations

$$\alpha_{\mathrm{L\tilde{R}}}^{'} = \mathrm{e}^{\frac{\mathrm{i}}{2}\theta_{\mathrm{i}}\sigma_{\mathrm{i}} + \frac{1}{2}\phi_{\mathrm{i}}\sigma_{\mathrm{i}}} \alpha_{\mathrm{L\tilde{R}}} \mathrm{e}^{-\frac{\mathrm{i}}{2}\theta_{\mathrm{i}}\sigma_{\mathrm{i}} + \frac{1}{2}\phi_{\mathrm{i}}\sigma_{\mathrm{i}}} \tag{3–78}$$

$$\alpha_{\mathrm{R\tilde{L}}}^{'} = \mathrm{e}^{\frac{\mathrm{i}}{2}\theta_{\mathrm{i}}\sigma_{\mathrm{i}} - \frac{1}{2}\phi_{\mathrm{i}}\sigma_{\mathrm{i}}} \alpha_{\mathrm{R\tilde{L}}} \mathrm{e}^{-\frac{\mathrm{i}}{2}\theta_{\mathrm{i}}\sigma_{\mathrm{i}} - \frac{1}{2}\phi_{\mathrm{i}}\sigma_{\mathrm{i}}}. \tag{3–79}$$

Both transform with positive θ under rotations, but the boosts are a different story. The $\mathrm{L\tilde{R}}$ four vector transforms with positive ϕ while the $\mathrm{R\tilde{L}}$ four vector transforms with negative

ϕ. Behaving the same way under rotations implies that the x,y,z components mix up the same way in both types of four vector. In this respect, rotating x,y,z or -x,-y,-z are both valid. Exemplifying this, an infinitesimal rotation around z by $\mathrm{d}\theta_\mathrm{z}$ gives

$$\begin{aligned} \mathrm{x}' &= \mathrm{x} - \mathrm{y}\mathrm{d}\theta_\mathrm{z} \\ \mathrm{y}' &= \mathrm{y} + \mathrm{x}\mathrm{d}\theta_\mathrm{z} \end{aligned} \tag{3–80}$$

for positive x,y,z and

$$\begin{aligned} -\mathrm{x}' &= -\mathrm{x} - (-\mathrm{y})\mathrm{d}\theta_\mathrm{z} \\ -\mathrm{y}' &= -\mathrm{y}' + (-\mathrm{x})\mathrm{d}\theta_\mathrm{z} \end{aligned} \tag{3–81}$$

for -x,-y,-z. These transformations are equivalent. Moving onto boosts, the opposite ϕ sign naively implies that the x,y,z terms mix up with t in the opposite way for the two types of rank 2 spinors. However, if both $\mathrm{L}\tilde{\mathrm{R}}$ and $\mathrm{R}\tilde{\mathrm{L}}$ represent a four vector, then the boost must yield the same transformation on the components. Therefore, to get from $\mathrm{L}\tilde{\mathrm{R}}$ to $\mathrm{R}\tilde{\mathrm{L}}$ take t,x,y,z to t,-x,-y,-z. Illuminating this, an infinitesimal boost along x by $\mathrm{d}\phi_\mathrm{x}$ gives

$$\begin{aligned} \mathrm{t}' &= \mathrm{t} + \mathrm{x}\mathrm{d}\phi_\mathrm{x} \\ \mathrm{x}' &= \mathrm{x} + \mathrm{t}\mathrm{d}\phi_\mathrm{x} \end{aligned} \tag{3–82}$$

for the $\mathrm{L}\tilde{\mathrm{R}}$ representation and

$$\begin{aligned} \mathrm{t}' &= \mathrm{t} + (-\mathrm{x})(-\mathrm{d}\phi_\mathrm{x}) \\ -\mathrm{x}' &= -\mathrm{x} + \mathrm{t}(-\mathrm{d}\phi_\mathrm{x}) \end{aligned} \tag{3–83}$$

for $\mathrm{R}\tilde{\mathrm{L}}$. The negative ϕ and the negative spatial components counteract to transform the four vector correctly. The $\mathrm{R}\tilde{\mathrm{L}}$ rank 2 spinor encodes a four vector in the same way as a $\mathrm{L}\tilde{\mathrm{R}}$ rank 2 spinor, except that the spatial components are coded into the rank 2 spinor with the opposite sign. Writing the representations out in terms of the four vector components yields

$$\begin{aligned}
\alpha_{\mathrm{L\tilde{R}}} &= \alpha_0\sigma_0 + \alpha_1\sigma_1 + \alpha_2\sigma_2 + \alpha_3\sigma_3 \\
\alpha_{\mathrm{L\tilde{R}}} &= \alpha_\mu\sigma^\mu \\
\alpha_{\mathrm{L\tilde{R}}} &= \begin{pmatrix} \alpha_0+\alpha_3 & \alpha_1 - \mathrm{i}\alpha_2 \\ \alpha_1+\mathrm{i}\alpha_2 & \alpha_0-\alpha_3 \end{pmatrix}
\end{aligned} \tag{3–84}$$

and

$$\begin{aligned}
\alpha_{\mathrm{R\tilde{L}}} &= \alpha_0\sigma_0 - \alpha_1\sigma_1 - \alpha_2\sigma_2 - \alpha_3\sigma_3 \\
\alpha_{\mathrm{R\tilde{L}}} &= \alpha_\mu\bar{\sigma}^\mu \\
\alpha_{\mathrm{R\tilde{L}}} &= \begin{pmatrix} \alpha_0-\alpha_3 & \alpha_1 + \mathrm{i}\alpha_2 \\ \alpha_1-\mathrm{i}\alpha_2 & \alpha_0+\alpha_3 \end{pmatrix}.
\end{aligned} \tag{3–85}$$

The action for the spin $\frac{1}{2}$ particle is built from the lowest order invariant combinations of the spinors themselves and the lowest order invariant term coupling ∂_μ to the spinors. These ingredients are $\mathcal{L}^\dagger\mathcal{R}$, $\mathcal{R}^\dagger\mathcal{L}$, $\mathcal{R}^\dagger\sigma^\mu\partial_\mu\mathcal{R}$, and $\mathcal{L}^\dagger\bar{\sigma}^\mu\partial_\mu\mathcal{L}$. In the scalar particle case m coupled the field to itself as potential energy [11] with the same power as ∂_μ [12]. Following that example and throwing an i into the derivative terms[13] to make them Hermitian gives the following Lagrangian,

$$\mathrm{S} = \int \mathrm{d}^4\mathrm{x}\left(\mathcal{R}^\dagger\sigma^\mu\mathrm{i}\partial_\mu\mathcal{R} + \mathrm{i}\mathcal{L}^\dagger\bar{\sigma}^\mu\mathrm{i}\partial_\mu\mathcal{L} - \mathrm{m}\mathcal{L}^\dagger\mathcal{R} - \mathrm{m}\mathcal{R}^\dagger\mathcal{L}\right). \tag{3–86}$$

[11] The mass term had a negative sign in the Lagrangian.

[12] E, P, and m must have the same power to get $\mathrm{E}^2 = \mathrm{m}^2 + \vec{\mathrm{P}}^2$ right, and $\mathrm{E} \sim \partial_0$, $\mathrm{P_i} \sim \partial_\mathrm{i}$ so m must have the same power as ∂_μ.

[13] This gives the P_μ operator which equals $\mathrm{i}\partial_\mu$.

The Euler-Lagrange equations reveal how the fields change over time,

$$
\begin{aligned}
&\sigma^\mu \mathrm{i}\partial_\mu \mathcal{R} = \mathrm{m}\mathcal{L} \\
&\bar{\sigma}^\mu \mathrm{i}\partial_\mu \mathcal{L} = \mathrm{m}\mathcal{R}. \\
&\sigma^\mu \mathrm{i}\partial_\mu \mathcal{R}^\dagger = -\mathrm{m}\mathcal{L}^\dagger \\
&\bar{\sigma}^\mu \mathrm{i}\partial_\mu \mathcal{L}^\dagger = -\mathrm{m}\mathcal{R}^\dagger.
\end{aligned}
\tag{3–87}
$$

Notice that the mass term couples the left and right chiral spinors. At rest, P=0 and the equations reduce to $\mathrm{i}\partial_0\mathcal{R} = \mathrm{m}\mathcal{L}$ and $\mathrm{i}\partial_0\mathcal{L} = \mathrm{m}\mathcal{R}$ showing that $\partial_0^2\mathcal{R} = -\mathrm{m}^2\mathcal{R}$ and $\partial_0^2\mathcal{L} = -\mathrm{m}^2\mathcal{L}$. The mass is actually a frequency determining how quickly a particle oscillates between its right-chiral and left-chiral states. Another interesting point is that swapping the sign of m here swaps the roles of the fields and the conjugate fields.

Plugging $\mathcal{R} = \frac{1}{\mathrm{m}}\bar{\sigma}^\nu \mathrm{i}\partial_\nu\mathcal{L}$ into $\sigma^\mu \mathrm{i}\partial_\mu\mathcal{R} = \mathrm{m}\mathcal{L}$ shows that the constants chosen for the terms in the Lagrangian provide the correct dispersion relation,

$$
\begin{aligned}
&\sigma^\mu \mathrm{i}\partial_\mu \frac{1}{\mathrm{m}}\bar{\sigma}^\nu \mathrm{i}\partial_\nu \mathcal{L} = \mathrm{m}\mathcal{L} \\
&\to \sigma^\mu\bar{\sigma}^\nu \mathrm{P}_\mu \mathrm{P}_\nu = \mathrm{m}^2 \\
&\to \eta^\mu_\nu \mathrm{P}_\mu \mathrm{P}_\nu = \mathrm{m}^2 \\
&\to \mathrm{P}^\mu \mathrm{P}_\mu = \mathrm{m}^2 \\
&\to \mathrm{E}^2 - \vec{\mathrm{P}}^2 = \mathrm{m}^2
\end{aligned}
\tag{3–88}
$$

The action for the spin $\frac{1}{2}$ particle can be rewritten into its more compact, canonical form,

$$
\mathrm{S} = \int \mathrm{d}^4\mathrm{x}\; \bar{\psi}\left(\mathrm{i}\gamma^\mu\partial_\mu - \mathrm{m}\right)\psi, \tag{3–89}
$$

after defining

$$
\gamma^\mu = \begin{pmatrix} 0 & \sigma^\mu \\ \bar{\sigma}^\mu & 0 \end{pmatrix}, \; \psi = \begin{pmatrix} \mathcal{L} \\ \mathcal{R} \end{pmatrix}, \text{ and } \bar{\psi} = \psi^\dagger\gamma_0. \tag{3–90}
$$

The spin $\frac{1}{2}$ Lagrangian is called the Dirac Lagrangian and the resulting equations of motion are the Dirac equation.

3.1.3.6 Building the Lagrangian for a Free Spin 1 Particle

Last but not least is the Lagrangian for a spin 1 force carrying particle. The main ingredients are the A_μ and ∂_μ four vectors. The lowest order possible invariants are $\partial^\mu \mathrm{A}^\nu \partial_\mu \mathrm{A}_\nu$, $\partial^\mu \mathrm{A}^\nu \partial_\nu \mathrm{A}_\mu$, $\mathrm{A}^\mu \mathrm{A}_\mu$, and $\partial^\mu \mathrm{A}_\mu$. However the last term is a total derivative and won't be included. The resulting action and equations of motion are

$$\mathrm{S} = \int \mathrm{d}^4\mathrm{x} \left(\mathrm{c}_0 \mathrm{i}\partial^\mu \mathrm{A}^\nu \mathrm{i}\partial_\mu \mathrm{A}_\nu + \mathrm{c}_1 \mathrm{i}\partial^\mu \mathrm{A}^\nu \mathrm{i}\partial_\nu \mathrm{A}_\mu + \mathrm{c}_2 \mathrm{A}^\mu \mathrm{A}_\mu\right). \tag{3–91}$$

$$\mathrm{c}_2 \mathrm{A}^\nu = -\partial_\mu \left(\mathrm{c}_0 \partial^\mu \mathrm{A}^\nu + \mathrm{c}_1 \partial^\nu \mathrm{A}^\mu\right). \tag{3–92}$$

This is the same form as the equation for the four vector potential in electromagnetism, which has $\mathrm{c}_0 = \frac{1}{2}$ and $\mathrm{c}_1 = -\frac{1}{2}$. The remaining term, c_2 looks like a mass term, so c_2 is set to $\frac{1}{2}\mathrm{m}^2$ [14] . The action and equations of motion with the appropriate constants become,

$$\mathrm{S} = \int \mathrm{d}^4\mathrm{x} \left(\frac{-1}{2}\partial^\mu \mathrm{A}^\nu \partial_\mu \mathrm{A}_\nu + \frac{1}{2}\partial^\mu \mathrm{A}^\nu \partial_\nu \mathrm{A}_\mu + \frac{\mathrm{m}^2}{2}\mathrm{A}^\mu \mathrm{A}_\mu\right) \tag{3–93}$$

$$-\mathrm{m}^2 \mathrm{A}^\nu = \partial_\mu \left(\partial^\mu \mathrm{A}^\nu - \partial^\nu \mathrm{A}^\mu\right). \tag{3–94}$$

Taking another derivative reveals that the equations reduce further and take on the correct dispersion relation, confirming the choice of constants,

$$\begin{aligned}
&-\mathrm{m}^2 \partial_\nu \mathrm{A}^\nu = \partial_\nu \partial_\mu \partial^\mu \mathrm{A}^\nu - \partial_\nu \partial_\mu \partial^\nu \mathrm{A}^\mu = 0 \\
&\rightarrow \partial_\nu \mathrm{A}^\nu = 0 \\
&\rightarrow -\mathrm{m}^2 \mathrm{A}^\nu = \partial_\mu \partial^\mu \mathrm{A}^\nu \\
&\rightarrow \mathrm{m}^2 \mathrm{A}^\nu = \mathrm{i}\partial_\mu \mathrm{i}\partial^\mu \mathrm{A}^\nu \\
&\rightarrow \mathrm{m}^2 = \mathrm{E}^2 - \vec{\mathrm{P}}^2.
\end{aligned} \tag{3–95}$$

[14] The order of the mass should match the order of the derivatives.

The action for the spin 1 particle is called the Proca action. It's normally written in an equivalent form in terms of the tensor $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$

$$S = \int d^4x \left(\frac{-1}{4} F_{\mu\nu} F^{\mu\nu} + \frac{m^2}{2} A^\mu A_\mu \right). \tag{3–96}$$

3.1.4 Lagrangians in Quantum Mechanics and QFT

The Lagrangians for the different types of free particles have been derived. Now it's important to know how to use them to predict results that can be verified by experiment. Quantum behavior is different than the classical behavior covered earlier. Classically, each particle takes a single well defined path, but in the quantum regime particles behave as if they take all possible paths from some initial point to some final point. Analogously, the fields from which particles arise behave as if they take all possible configurations between states. Each possibility contributes to the net probability as a unit complex number with a phase given by the action. These contributions are called probability amplitudes,

$$\mathcal{A}_j = e^{iS_j}. \tag{3–97}$$

Adding the amplitudes for every possibility gives the net amplitude, and squaring the net amplitude gives the probability,

$$P = |\sum_j \mathcal{A}_j|^2. \tag{3–98}$$

Consider a free non-relativistic quantum particle in one dimension with the Lagrangian $\mathcal{L} = \frac{1}{2} m \dot{x}^2$. The amplitude for the particle to start at x_0, t_0 and end at x_f, t_f can be calculated by discretizing time, $t \in t_0, t_1, ..., t_f$. A path is then determined by specifying x at each point in time, $x(t) \in x(t_0), x(t_1), ..., x(t_f)$. In this way, the action for a path x, $S(x)$, can be written

$$S(x) = \int dt L = \frac{m}{2} \int_{t_0}^{t_f} dt \dot{x}^2 = \frac{m}{2} \sum_i \frac{(x(t_i) - x(t_{i-1}))^2}{(t_i - t_{i-1})^2} (t_i - t_{i-1}) = \frac{m}{2} \sum_i \frac{(x(t_i) - x(t_{i-1}))^2}{\Delta t}. \tag{3–99}$$

The amplitude for the path is then given by,

$$\mathcal{A}(x) = e^{i\frac{m}{2}\sum_i \frac{(x(t_i)-x(t_{i-1}))^2}{(t_i-t_{i-1})^2}} = e^{i\frac{m}{2}\frac{(x(t_f)-x(t_{f-1}))^2}{\Delta t}} e^{i\frac{m}{2}\frac{(x(t_{f-1})-x(t_{f-2}))^2}{\Delta t}} ... e^{i\frac{m}{2}\frac{(x(t_1)-x(t_0))^2}{\Delta t}}. \tag{3–100}$$

Summing over all paths, $x(t)$, with fixed endpoints, $x(t_0) = x_0$ and $x(t_f) = x_f$, provides the net amplitude

$$\begin{aligned}\mathcal{A} &= \sum_x \mathcal{A}(x) = \sum_{x(t_f-1)} \sum_{x(t_f-2)} ... \sum_{x(t_1)} e^{i\frac{m}{2}\frac{(x(t_f)-x(t_{f-1}))^2}{\Delta t}} e^{i\frac{m}{2}\frac{(x(t_{f-1})-x(t_{f-2}))^2}{\Delta t}} ... e^{i\frac{m}{2}\frac{(x(t_1)-x(t_0))^2}{\Delta t}} \\ &= \mathcal{C}\int dx_{t_f-1}dx_{t_f-2}...dx_{t_1} e^{i\frac{m}{2}\frac{(x(t_f)-x(t_{f-1}))^2}{\Delta t}} e^{i\frac{m}{2}\frac{(x(t_{f-1})-x(t_{f-2}))^2}{\Delta t}} ... e^{i\frac{m}{2}\frac{(x(t_1)-x(t_0))^2}{\Delta t}} \\ &= \int \mathcal{D}[x(t)] e^{i\frac{m}{2}\int_{t_0}^{t_f} dt \partial_t x \partial_t x}.\end{aligned} \tag{3–101}$$

Integrating by parts and using the average velocity $v = \frac{x_f - x_0}{t_f - t_0}$ for the boundary velocities provides the net amplitude for the particle to go from one place to another,

$$\begin{aligned}\mathcal{A} &= \int \mathcal{D}[x(t)] e^{i\frac{m}{2} x\partial_t x|_{t_0}^{t_f} + i\frac{m}{2}\int_{t_0}^{t_f} dt(-x\partial_t^2 x)} \\ &= e^{i\frac{m}{2}(x(t_f)-x(t_0))v} \int \mathcal{D}[x(t)] e^{i\frac{m}{2}\int_{t_0}^{t_f} dt(-x\partial_t^2 x)} \\ &= \mathcal{N} e^{\frac{-im(x_f-x_0)^2}{2(t_f-t_0)}} \\ &= \sqrt{\frac{m}{2i\pi(t_f-t_0)}} e^{\frac{-im(x_f-x_0)^2}{2(t_f-t_0)}}.\end{aligned} \tag{3–102}$$

The normalization $\mathcal{N}$ is fixed by requiring that the probability for the particle to be somewhere at t_f is one. In other words, integrating over all x_f for constant x_0, t_0, t_f should give one. A more rigorous derivation of this propagator can be found in Feynman and Hibbs for example.

In general the amplitude to go from $\phi(t_0) = \phi_0$ to $\phi(t_f) = \phi_f$ is determined by the path integral,

$$\mathcal{A}(\phi_f, t_f; \phi_0, t_0) = \langle \phi_f, t_f | \phi_0, t_0 \rangle = \int \mathcal{D}[\cdot] e^{i\int_{t_0}^{t_f} dt L}. \tag{3–103}$$

Furthermore, the expectation value for an operation performed at time t_1 when the system starts in the state $\phi(t_0) = \phi_0$ and ends in $\phi(t_f) = \phi_f$ is

$$|\langle \phi_f, t_f | \mathcal{O}(t_1) | \phi_0, t_0 \rangle|^2 = \left| \int \mathcal{D}[\cdot] \mathcal{O}(t_1) e^{i\int_{t_0}^{t_f} dt L} \right|^2. \tag{3–104}$$

Of particular importance for QFT are scattering events. When two particles are measured at $(t_1, \vec{x}_1)$ and $(t_2, \vec{x}_2)$ some time in the past, what is the amplitude to measure two particles at $(t_3, \vec{x}_3)$ and $(t_4, \vec{x}_4)$ some time in the future? Considering that particles are in fact disturbances in the field, an equivalent question may be asked. How is the field amplitude correlated between spacetime points $(t_4, \vec{x}_4), (t_3, \vec{x}_3), (t_2, \vec{x}_2)$, and $(t_1, \vec{x}_1)$. So for two to two scattering the amplitude is given by

$$\langle 0| \psi_i(t_4, \vec{x}_4)\psi_j(t_3, \vec{x}_3)\psi_k(t_2, \vec{x}_2)\psi_l(t_1, \vec{x}_1) |0\rangle = \frac{\int \mathcal{D}[\cdot]\psi_i(t_4, \vec{x}_4)\psi_j(t_3, \vec{x}_3)\psi_k(t_2, \vec{x}_2)\psi_l(t_1, \vec{x}_1)e^{i\int_{-\infty}^{\infty} dt L}}{\int \mathcal{D}[\cdot]e^{i\int_{-\infty}^{\infty} dt L}}, \tag{3–105}$$

where ψ is a field, 0 denotes the vacuum state, and i, j, k, l specify the types of particles observed. Calculating this path integral for the full Lagrangian determines the amplitude for the scattering event. Usually the correlation functions can only be calculated for the free fields where the Lagrangian density is quadratic in the fields. Interactions generally spoil this simplicity, but even if the Lagrangian is more complicated, the path integral for the full theory can usually be expanded in terms of the correlation functions from the free field. However, this only works in practice if the additional terms in the Lagrangian are small compared to the free part of the Lagrangian. This method of expansion is called perturbation theory.

3.1.5 Perturbation Theory and Feynman Rules

Unfortunately, the correlation functions determining the scattering and decay amplitudes for most of the interesting theories are not directly solvable. On the other hand, they can often be expanded in terms of the free solutions. When the magnitude of the additional complexity is small compared to the free part of the Lagrangian, only the first few terms of the series expansion are needed and the technique is called perturbation theory.

The best way to showcase perturbation theory is to run over the process for the scalar particle. The first step is to figure out the free solutions. The two point correlation function for the free case is given by

$$\langle 0| \Phi(t_2, \vec{x}_2)\Phi(t_1, \vec{x}_1) |0\rangle = \frac{\int \mathcal{D}[\Phi]\Phi(t_2, \vec{x}_2)\Phi(t_1, \vec{x}_1)e^{i\int_{-\infty}^{\infty} d^4x \frac{1}{2}\left(\partial_\mu\Phi\partial^\mu\Phi - m^2\Phi^2\right)}}{\int \mathcal{D}[\Phi]e^{i\int_{-\infty}^{\infty} d^4x \frac{1}{2}\left(\partial_\mu\Phi\partial^\mu\Phi - m^2\Phi^2\right)}}, \tag{3–106}$$

and it describes a particle going from $(t_1, \vec{x}_1)$ to $(t_2, \vec{x}_2)$. Integrating by parts and throwing away the boundary term turns the path integrals into Gaussians

$$\langle 0| \Phi(t_2, \vec{x}_2)\Phi(t_1, \vec{x}_1) |0\rangle = \frac{\int \mathcal{D}[\Phi]\Phi(t_2, \vec{x}_2)\Phi(t_1, \vec{x}_1)e^{i\int_{-\infty}^{\infty} d^4x \frac{1}{2}\Phi(-\partial_\mu\partial^\mu - m^2)\Phi}}{\int \mathcal{D}[\Phi]e^{i\int_{-\infty}^{\infty} d^4x \frac{1}{2}\Phi(-\partial_\mu\partial^\mu - m^2)\Phi}}. \tag{3–107}$$

Focusing on the integral in the denominator and looking at the integral as a matrix multiplication simplifies things

$$\int \mathcal{D}[\Phi]e^{i\int_{-\infty}^{\infty} d^4x \frac{1}{2}\Phi(-\partial_\mu\partial^\mu - m^2)\Phi} = \int d\Phi_1 d\Phi_2 ... d\Phi_N e^{i\frac{1}{2}\Phi_i K_{ij}\Phi_j}. \tag{3–108}$$

Transforming to the eigenbasis yields a multidimensional Gaussian integral

$$\begin{aligned}\int d\Phi_1 d\Phi_2 ... d\Phi_N e^{i\frac{1}{2}\Phi_i K_{ij}\Phi_j} &= \int d\tilde{\Phi}_1 d\tilde{\Phi}_2 ... d\tilde{\Phi}_N e^{i\frac{1}{2}\tilde{\Phi}_i \tilde{K}_{ij}\tilde{\Phi}_j} = \int d\tilde{\Phi}_1 d\tilde{\Phi}_2 ... d\tilde{\Phi}_N e^{\frac{-1}{2i}\tilde{k}_i\tilde{\Phi}_i^2} \\ &= \prod_{i=1}^{N}\left(\frac{2i\pi}{\tilde{k}_i}\right)^{\frac{1}{2}} = \frac{(2i\pi)^{N/2}}{(\det K)^{\frac{1}{2}}} \equiv Z_0.\end{aligned} \tag{3–109}$$

The solution uses the fact that the determinant is the product of eigenvalues. The integral from the numerator can be dealt with by taking derivatives of the moment generating function

$$\int d\Phi_1 d\Phi_2 ... d\Phi_N \Phi_k \Phi_l e^{i\frac{1}{2}\Phi_i K_{ij}\Phi_j} = \frac{1}{i}\frac{\delta}{\delta J_k}\frac{1}{i}\frac{\delta}{\delta J_l}\Big|_{J=0}\int d\Phi_1 d\Phi_2 ... d\Phi_N e^{i\frac{1}{2}\Phi_i K_{ij}\Phi_j + iJ_i\Phi_i}. \tag{3–110}$$

And the integral on the right can be solved using the substitution $\Phi' = \Phi + K^{-1}J$ to complete the square,

$$\begin{aligned}\int d\Phi_1 d\Phi_2 ... d\Phi_N e^{i\frac{1}{2}\Phi_i K_{ij}\Phi_j + iJ_i\Phi_i} &= \int d\Phi'_1 d\Phi'_2 ... d\Phi'_N e^{i\frac{1}{2}\Phi'_i K_{ij}\Phi'_j - \frac{i}{2}J_i K^{-1}_{ij}J_j} \\ &= \frac{(2i\pi)^{N/2}}{(\det K)^{\frac{1}{2}}} e^{\frac{-i}{2}J_i K^{-1}_{ij}J_j} \\ &= Z_0 e^{\frac{-i}{2}\int d^4x d^4y J(y)K^{-1}(y,x)J(x)},\end{aligned} \tag{3–111}$$

where the last line reverts back to continuous functions and operators. Applying the derivatives provides the numerator of the two point correlation function

$$
\begin{aligned}
\int \mathrm{d}\Phi_1 \mathrm{d}\Phi_2 ... \mathrm{d}\Phi_N \Phi_k \Phi_l e^{i\frac{1}{2}\Phi_i K_{ij}\Phi_j} &= \frac{1}{i}\frac{\delta}{\delta J_k}\frac{1}{i}\frac{\delta}{\delta J_l}\Big|_{J=0}\int \mathrm{d}\Phi_1 \mathrm{d}\Phi_2 ... \mathrm{d}\Phi_N e^{i\frac{1}{2}\Phi_i K_{ij}\Phi_j + iJ_i\Phi_i} \\
&= \frac{1}{i}\frac{\delta}{\delta J_k}\frac{1}{i}\frac{\delta}{\delta J_l}\Big|_{J=0} Z_0 e^{\frac{-i}{2}J_i K^{-1}_{ij} J_j} \qquad (3\text{–}112) \\
&= Z_0 i K^{-1}_{ij}.
\end{aligned}
$$

Then plugging in the solutions for the numerator and denominator shows that the two point correlation function is given by the Feynman propagator, $\mathcal{D}_F$,

$$
\begin{aligned}
\langle 0|\,\Phi(x_2)\Phi(x_1)\,|0\rangle &= \frac{\int \mathcal{D}[\Phi]\Phi(x_2)\Phi(x_1) e^{i\int_{-\infty}^{\infty} d^4x \frac{1}{2}\Phi(-\partial_\mu\partial^\mu - m^2)\Phi}}{\int \mathcal{D}[\Phi] e^{i\int_{-\infty}^{\infty} d^4x \frac{1}{2}\Phi(-\partial_\mu\partial^\mu - m^2)\Phi}} \\
&= \frac{Z_0 i K^{-1}(x_2, x_1)}{Z_0} \qquad (3\text{–}113) \\
&= iK^{-1}(x_2, x_1) = i(-\partial_\mu\partial^\mu - m^2)^{-1} = \mathcal{D}_F(x_2, x_1).
\end{aligned}
$$

In momentum space K is diagonal and $\mathcal{D}_F = \frac{i}{p^2 - m^2 + i\epsilon}$. Converting to the spacetime basis yields, $\mathcal{D}_F(x_2, x_1) = \int \frac{d^4k}{\sqrt{2\pi}^4}\frac{ie^{-ik\cdot(x_2 - x_1)}}{p^2 - m^2 + i\epsilon}$, which represents a free particle propagating frome one place to another.

The Feynman propagators are the building blocks for the higher order correlation functions. Take the four point function as an example,

$$
\begin{aligned}
\langle 0|\,\Phi(x_4)\Phi(x_3)\Phi(x_2)\Phi(x_1)\,|0\rangle &= \frac{\int \mathcal{D}[\Phi]\Phi(x_4)\Phi(x_3)\Phi(x_2)\Phi(x_1) e^{i\int_{-\infty}^{\infty} d^4x \frac{1}{2}\Phi(-\partial_\mu\partial^\mu - m^2)\Phi}}{\int \mathcal{D}[\Phi] e^{i\int_{-\infty}^{\infty} d^4x \frac{1}{2}\Phi(-\partial_\mu\partial^\mu - m^2)\Phi}} \\
&= \frac{\frac{1}{i}\frac{\delta}{\delta J(x_4)}\frac{1}{i}\frac{\delta}{\delta J(x_3)}\frac{1}{i}\frac{\delta}{\delta J(x_2)}\frac{1}{i}\frac{\delta}{\delta J(x_1)}\Big|_{J=0}\int \mathcal{D}[\Phi] e^{i\int_{-\infty}^{\infty} d^4x \frac{1}{2}\Phi(-\partial_\mu\partial^\mu - m^2)\Phi + iJ\Phi}}{\int \mathcal{D}[\Phi] e^{i\int_{-\infty}^{\infty} d^4x \frac{1}{2}\Phi(-\partial_\mu\partial^\mu - m^2)\Phi}} \\
&= \frac{\delta}{\delta J(x_4)}\frac{\delta}{\delta J(x_3)}\frac{\delta}{\delta J(x_2)}\frac{\delta}{\delta J(x_1)}\Big|_{J=0} e^{-\frac{1}{2}\int d^4x d^4y J(y)\mathcal{D}_F(y,x)J(x)} \\
&= \frac{\delta}{\delta J(x_4)}\frac{\delta}{\delta J(x_3)}\frac{\delta}{\delta J(x_2)}\frac{\delta}{\delta J(x_1)}\Big|_{J=0} \sum_{n=0}^{\infty}\frac{1}{n!}\left(-\frac{1}{2}\int d^4x d^4y J(y)\mathcal{D}_F(y,x)J(x)\right)^n.
\end{aligned}
$$

(3–114)

When J=0 only the term with four Js and four derivatives survives. The four derivatives hit the Js and leave terms like $\mathcal{D}_F(d, c)\mathcal{D}_F(b, a)$. In this case, there are 4!=24 ways to rearrange a,b,c, and d. However, the order of the spacetime points within the propagator doesn't matter nor does the order of the propagators, so this reduces the total to $\frac{4!}{2!2^2} = 3$ ways. The $\frac{1}{n!}$ and the $\frac{1}{2^n}$ from the expansion naturally take care of the degeneracy,

$$\begin{aligned}
&\langle 0|\,\Phi(x_4)\Phi(x_3)\Phi(x_2)\Phi(x_1)\,|0\rangle \\
&= \frac{\delta}{\delta J(x_4)}\frac{\delta}{\delta J(x_3)}\frac{\delta}{\delta J(x_2)}\frac{\delta}{\delta J(x_1)}\Big|_{J=0}\frac{1}{8}\int d^4y_j d^4x_j d^4y_i d^4x_i J(y_j)J(x_j)J(y_i)J(x_i)\mathcal{D}_F(y_j,x_j)\mathcal{D}_F(y_i,x_i) \\
&= \mathcal{D}_F(x_4,x_3)\mathcal{D}_F(x_2,x_1) + \mathcal{D}_F(x_4,x_2)\mathcal{D}_F(x_3,x_1) + \mathcal{D}_F(x_3,x_2)\mathcal{D}_F(x_4,x_1).
\end{aligned} \tag{3–115}$$

Any higher order correlation functions are built from the propagators in the same way: add up all possible combinations of the propagators for the appropriate order and account for the degeneracies.

This solves the free theory, which is nice, but it isn't that useful. Particles interact and the interacting solutions are the ones needed to describe scattering and decays. Fortunately, the interacting theories can be expanded in terms of the free correlations. For example, the two point correlation function for an interacting theory is given by,

$$\begin{aligned}
\langle 0|\,\Phi(x_2)\Phi(x_1)\,|0\rangle &= \frac{\int \mathcal{D}[\Phi]\Phi(x_2)\Phi(x_1)e^{i\int_{-\infty}^{\infty} d^4x\frac{1}{2}\Phi(-\partial_\mu\partial^\mu - m^2)\Phi + \mathcal{L}_{int}}}{\int \mathcal{D}[\Phi]e^{i\int_{-\infty}^{\infty} d^4x\frac{1}{2}\Phi(-\partial_\mu\partial^\mu - m^2)\Phi + \mathcal{L}_{int}}} \\
&= \frac{\int \mathcal{D}[\Phi]\Phi(x_2)\Phi(x_1)e^{i\int_{-\infty}^{\infty} d^4x\frac{1}{2}\Phi(-\partial_\mu\partial^\mu - m^2)\Phi}e^{i\int_{-\infty}^{\infty} d^4x\mathcal{L}_{int}}}{\int \mathcal{D}[\Phi]e^{i\int_{-\infty}^{\infty} d^4x\frac{1}{2}\Phi(-\partial_\mu\partial^\mu - m^2)\Phi}e^{i\int_{-\infty}^{\infty} d^4x\mathcal{L}_{int}}} \\
&= \frac{\int \mathcal{D}[\Phi]\Phi(x_2)\Phi(x_1)e^{i\int_{-\infty}^{\infty} d^4x\frac{1}{2}\Phi(-\partial_\mu\partial^\mu - m^2)\Phi}\sum_{n=0}^{\infty}\frac{1}{n!}\left(i\int_{-\infty}^{\infty} d^4x\mathcal{L}_{int}\right)^n}{\int \mathcal{D}[\Phi]e^{i\int_{-\infty}^{\infty} d^4x\frac{1}{2}\Phi(-\partial_\mu\partial^\mu - m^2)\Phi}\sum_{n=0}^{\infty}\frac{1}{n!}\left(i\int_{-\infty}^{\infty} d^4x\mathcal{L}_{int}\right)^n},
\end{aligned} \tag{3–116}$$

where $\mathcal{L}_{int}$ may be written as a powerseries in Φ. Expanding this way, the numerator becomes a series of integrals of the form

$$\int d^4x_{i_1}d^4x_{i_2}...d^4x_{i_k}\int \mathcal{D}[\Phi]\Phi(x_{i_1})...\Phi(x_{i_k})\Phi(x_2)\Phi(x_1)e^{i\int_{-\infty}^{\infty} d^4x\frac{1}{2}\Phi(-\partial_\mu\partial^\mu - m^2)\Phi}, \tag{3–117}$$

which are just correlation functions of the free field. Likewise, the denominator is also an expansion of the free correlations. In this way, any solution of the full theory may be written in terms of the free field solutions. Furthermore, when the interaction is small compared to the free part of the Lagrangian, the series expansions converge, and only a few terms are needed.

The simplest interacting theory is $\mathcal{L}_{\mathrm{int}} = -\frac{\lambda}{4!}\Phi^4(\mathrm{x})$, and the two point correlation function describing the propagation of a particle is given by

$$\begin{aligned} &\langle 0|\,\Phi(\mathrm{x}_2)\Phi(\mathrm{x}_1)\,|0\rangle = \\ &\frac{\int \mathcal{D}[\Phi]\Phi(\mathrm{x}_2)\Phi(\mathrm{x}_1)\mathrm{e}^{\mathrm{i}\int_{-\infty}^{\infty}\mathrm{d}^4\mathrm{x}\frac{1}{2}\Phi(-\partial_\mu\partial^\mu-\mathrm{m}^2)\Phi}\sum_{\mathrm{n}=0}^{\infty}\frac{1}{\mathrm{n}!}\left(\int_{-\infty}^{\infty}\mathrm{d}^4\mathrm{z}\frac{-\mathrm{i}\lambda}{4!}\Phi(\mathrm{z})^4\right)^{\mathrm{n}}}{\int \mathcal{D}[\Phi]\mathrm{e}^{\mathrm{i}\int_{-\infty}^{\infty}\mathrm{d}^4\mathrm{x}\frac{1}{2}\Phi(-\partial_\mu\partial^\mu-\mathrm{m}^2)\Phi}\sum_{\mathrm{n}=0}^{\infty}\frac{1}{\mathrm{n}!}\left(\int_{-\infty}^{\infty}\mathrm{d}^4\mathrm{z}\frac{-\mathrm{i}\lambda}{4!}\Phi(\mathrm{z})^4\right)^{\mathrm{n}}}. \end{aligned} \tag{3–118}$$

Consider the numerator up to first order in λ

$$\begin{aligned} \mathrm{N}(\lambda^1) &= \int \mathcal{D}[\Phi]\Phi(\mathrm{x}_2)\Phi(\mathrm{x}_1)\mathrm{e}^{\mathrm{i}\int_{-\infty}^{\infty}\mathrm{d}^4\mathrm{x}\frac{1}{2}\Phi(-\partial_\mu\partial^\mu-\mathrm{m}^2)\Phi}\left(1+\int_{-\infty}^{\infty}\mathrm{d}^4\mathrm{z}\frac{-\mathrm{i}\lambda}{4!}\Phi(\mathrm{z})^4\right) \\ &= \mathrm{Z}_0\mathcal{D}_{\mathrm{F}}(\mathrm{x}_2,\mathrm{x}_1) + \int_{-\infty}^{\infty}\mathrm{d}^4\mathrm{z}\frac{-\mathrm{i}\lambda}{4!}\int \mathcal{D}[\Phi]\Phi(\mathrm{x}_2)\Phi(\mathrm{x}_1)\Phi(\mathrm{z})^4\mathrm{e}^{\mathrm{i}\int_{-\infty}^{\infty}\mathrm{d}^4\mathrm{x}\frac{1}{2}\Phi(-\partial_\mu\partial^\mu-\mathrm{m}^2)\Phi} \end{aligned}. \tag{3–119}$$

The first term is just the Feynman propagator, but the order λ term contains a correlation function with six Φs. Solving the six point integral requires adding up all combinations of terms of the form $\mathcal{D}_{\mathrm{F}}(\mathrm{a},\mathrm{b})\mathcal{D}_{\mathrm{F}}(\mathrm{c},\mathrm{d})\mathcal{D}_{\mathrm{F}}(\mathrm{e},\mathrm{f})$ while accounting for the appropriate degeneracies. There are $\frac{6!}{3!2^3} = 15$ possible propagator triples: 12 ways to put x and y in separate propagators and 3 ways to put them in the same one. With this information the perturbation series for the numerator is finally in hand,

$$\begin{aligned} \mathrm{N}(\lambda^1) = \mathrm{Z}_0\,&\underbrace{\mathcal{D}_{\mathrm{F}}(\mathrm{x}_2,\mathrm{x}_1)}_{\mathsf{A}} + \mathrm{Z}_0\frac{12}{4!}\underbrace{(-\mathrm{i}\lambda)\int_{-\infty}^{\infty}\mathrm{d}^4\mathrm{z}\mathcal{D}_{\mathrm{F}}(\mathrm{x}_2,\mathrm{z})\mathcal{D}_{\mathrm{F}}(\mathrm{x}_1,\mathrm{z})\mathcal{D}_{\mathrm{F}}(\mathrm{z},\mathrm{z})}_{\mathsf{B}} \\ &+ \mathrm{Z}_0\frac{3}{4!}\underbrace{(-\mathrm{i}\lambda)\mathcal{D}_{\mathrm{F}}(\mathrm{x}_2,\mathrm{x}_1)\int_{-\infty}^{\infty}\mathrm{d}^4\mathrm{z}\mathcal{D}_{\mathrm{F}}(\mathrm{z},\mathrm{z})\mathcal{D}_{\mathrm{F}}(\mathrm{z},\mathrm{z})}_{\mathsf{C}} \end{aligned}. \tag{3–120}$$

The perturbation series for any n-point correlation function can be represented in terms of Feynman diagrams, which assign symbols to the different pieces of math. In this simple scalar interacting theory, a line represents a factor of $\mathcal{D}_\mathrm{F}(\mathrm{a},\mathrm{b})$ ($\mathcal{D}_\mathrm{F}(\mathrm{z},\mathrm{z})$ if it's a loop) and an internal vertex representing an interaction, provides a factor of $-\mathrm{i}\lambda\int\mathrm{d}^4\mathrm{z}$. The Feynman diagrams for N(λ^1) are shown in Figure 3-2. In those diagrams, A and B are fully connected while C is disconnected. The separation reveals the fact that the pieces may be calculated independently and then multiplied. The fully connected pieces are the fundamental expressions from which any correlation function may be derived. There are two types of connected diagrams, those with external points and those with no external points. Those with zero external points represent vacuum fluctuations. The denominator is a normalization that serves

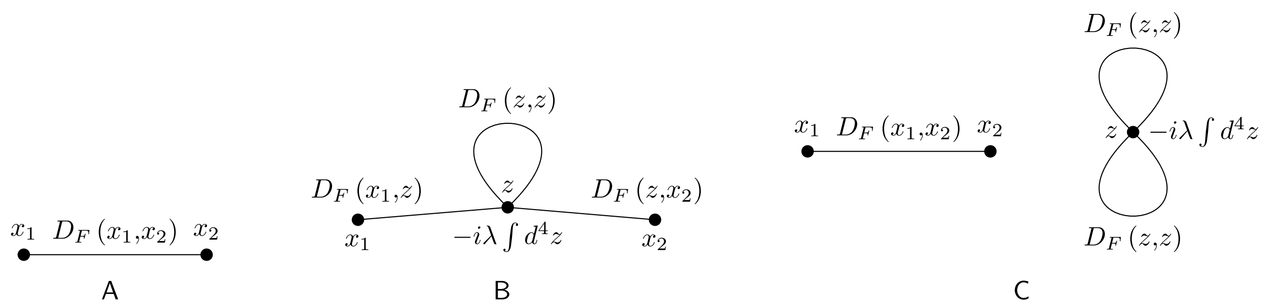


Figure 3-2. Feynman diagrams for N(λ^1) $= \mathrm{Z}_0$A $+ \mathrm{Z}_0\frac{12}{4!}$B $+ \mathrm{Z}_0\frac{3}{4!}$C.

to eliminate Z_0 and the disconnected diagrams with vacuum fluctuations like C. It removes the contributions where the vacuum fluctuates independently from the process of interest. Finally, the propagator for the fully interacting theory up to first order in λ is

$$\langle 0|\,\Phi(\mathrm{x}_2)\Phi(\mathrm{x}_1)\,|0\rangle_{\lambda^1} = \mathrm{A} + \frac{12}{4!}\mathrm{B} = \mathcal{D}_\mathrm{F}(\mathrm{x}_2,\mathrm{x}_1) + \frac{12}{4!}(-\mathrm{i}\lambda)\int_{-\infty}^{\infty}\mathrm{d}^4\mathrm{z}\mathcal{D}_\mathrm{F}(\mathrm{x}_2,\mathrm{z})\mathcal{D}_\mathrm{F}(\mathrm{x}_1,\mathrm{z})\mathcal{D}_\mathrm{F}(\mathrm{z},\mathrm{z}). \tag{3–121}$$

The factors out front are called symmetry factors and provide a weight for each diagram. In this theory the symmetry factor for a diagram is the number of nondegenerate ways to place pairs of points into the propagators with a factor of by $\frac{1}{4!}$ for every interaction vertex.

The Feynman rules for a theory can be used to build the perturbation series for any process, and this is why they are so ubiquitous. Any n-point correlation function is the weighted[15] sum of all possible diagrams with n external points – excluding those with disconnected vacuum fluctuations. For $2 \rightarrow 2$ scattering, the correlation function to first order requires the six diagrams shown in Figure 3-3. The first three diagrams are of the form

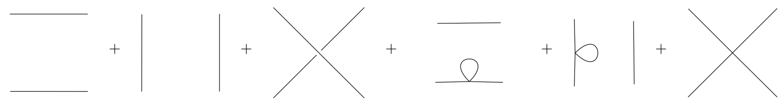


Figure 3-3. The Feynman diagrams for $\langle 0| \Phi(x_4)\Phi(x_3)\Phi(x_2)\Phi(x_1) |0\rangle_{\lambda^1}$ representing the matrix element for $2 \rightarrow 2$ scattering up to first order.

A*A, and these represent the three ways to put four points into $\mathcal{D}_F\mathcal{D}_F$. The next two terms are of the form A*B, and the last one is a new diagram where the particles actually scatter instead of propagating separately. And the last term is built from the following factors: the interaction vertex provides a factor of $-i\lambda \int d^4z$ and the four propagators provide a factor of $\mathcal{D}_F(x_1, z)\mathcal{D}_F(x_2, z)\mathcal{D}_F(x_3, z)\mathcal{D}_F(x_4, z)$. The symmetry factor is $1 = \frac{4!}{4!}$ with the $\frac{1}{4!}$ from the internal vertex and the 4! for the ways to place the zs among x_1, x_2, x_3, x_4. All of this results in a mathematical expression for the last diagram,

$$D = -i\lambda \int d^4z \mathcal{D}_F(x_1, z)\mathcal{D}_F(x_2, z)\mathcal{D}_F(x_3, z)\mathcal{D}_F(x_4, z). \tag{3–122}$$

Summing the six diagrams yields the full four point correlation function to first order,

$$\begin{aligned}\langle 0| \Phi(x_4)\Phi(x_3)\Phi(x_2)\Phi(x_1) |0\rangle_{\lambda^1} = {} & \mathcal{D}_F(x_1, x_3)\mathcal{D}_F(x_2, x_4) + \mathcal{D}_F(x_1, x_2)\mathcal{D}_F(x_3, x_4) + \mathcal{D}_F(x_1, x_4)\mathcal{D}_F(x_2, x_3) \\ & + \mathcal{D}_F(x_1, x_3)\frac{12}{4!}(-i\lambda)\int_{-\infty}^{\infty} d^4z \mathcal{D}_F(x_2, z)\mathcal{D}_F(x_4, z)\mathcal{D}_F(z, z) \\ & + \mathcal{D}_F(x_2, x_4)\frac{12}{4!}(-i\lambda)\int_{-\infty}^{\infty} d^4z \mathcal{D}_F(x_1, z)\mathcal{D}_F(x_3, z)\mathcal{D}_F(z, z) \\ & + -i\lambda \int d^4z \mathcal{D}_F(x_1, z)\mathcal{D}_F(x_2, z)\mathcal{D}_F(x_3, z)\mathcal{D}_F(x_4, z).\end{aligned} \tag{3–123}$$

[15] The weights are the symmetry factors.

The Feynman rules for the actual interactions of the Standard Model are derived from the interacting Lagrangians in a similar way.

3.1.6 Interactions

The spin 1 and spin $\frac{1}{2}$ Lagrangians were obtained after a lengthy adventure through the Lorentz group symmetries. Now the goal is to couple them to produce theories where fermions and bosons interact as observed in nature. The simplest Lorentz invariant term that couples a vector to the Dirac spinors is $\bar{\psi}\gamma^\mu \mathrm{A}_\mu \psi$, providing an interacting theory that looks like

$$\begin{aligned}\mathrm{S} &= \int \mathrm{d^4x} \frac{-1}{4} \mathrm{F}_{\mu\nu}\mathrm{F}^{\mu\nu} + \frac{\mathrm{m_A^2}}{2}\mathrm{A}^\mu \mathrm{A}_\mu + \bar{\psi}\left(\mathrm{i}\gamma^\mu \partial_\mu - \mathrm{m}\right)\psi + \mathrm{q}\bar{\psi}\gamma^\mu \mathrm{A}_\mu \psi \\ &= \int \mathrm{d^4x} \frac{-1}{4} \mathrm{F}_{\mu\nu}\mathrm{F}^{\mu\nu} + \frac{\mathrm{m_A^2}}{2}\mathrm{A}^\mu \mathrm{A}_\mu + \bar{\psi}\left[\mathrm{i}\gamma^\mu \left(\partial_\mu - \mathrm{iqA}_\mu\right) - \mathrm{m}\right]\psi\end{aligned} \tag{3–124}$$

To describe electromagnetism, the photon field should be massless and the interacting theory should be gauge invariant. Setting $\mathrm{m_A}$ to zero and checking whether $\mathrm{A}_\mu \to \mathrm{A}_\mu(\mathrm{x}) + \partial_\mu \alpha(\mathrm{x})$ is a symmetry, provides

$$\mathrm{S} = \int \mathrm{d^4x} \frac{-1}{4} \mathrm{F}_{\mu\nu}\mathrm{F}^{\mu\nu} + \bar{\psi}\left[\mathrm{i}\gamma^\mu \left(\partial_\mu - \mathrm{iqA}_\mu - \mathrm{iq}\partial_\mu \alpha\right) - \mathrm{m}\right]\psi, \tag{3–125}$$

which isn't gauge invariant. In effect, the gauge transformation on A_μ has shifted the derivative, $\partial_\mu \to \partial_\mu - \mathrm{iq}\partial_\mu \alpha$. In order to retain gauge invariance, a simultaneous transformation that sends $\partial_\mu \to \partial_\mu + \mathrm{iq}\partial_\mu \alpha$ is needed to cancel out the extra term. The correct transformation is $\bar{\psi}\partial_\mu \psi \to \bar{\psi}\mathrm{e}^{-\mathrm{iq}\alpha}\partial_\mu \mathrm{e}^{\mathrm{iq}\alpha}\psi$. So, if

$$\mathrm{A}_\mu(\mathrm{x}) \to \mathrm{A}_\mu(\mathrm{x}) + \partial_\mu \alpha(\mathrm{x}) \tag{3–126}$$

and

$$\psi(\mathrm{x}) \to \mathrm{e}^{\mathrm{iq}\alpha(\mathrm{x})}\psi(\mathrm{x}) \tag{3–127}$$

the Lagrangian is gauge invariant. The second equation reveals that the field is invariant under local transformations of the phase.

Requiring gauge invariance requires the fermion field to maintain U(1) invariance. By running the logic in reverse and demanding more complicated unitary invariances on the fermion field(s), Lagrangians may be produced that describe new forces. In the case of electromagnetism above, demanding that the fermion field is invariant under a local U(1) transformation, sends $\bar{\psi}\partial_\mu\psi \;\rightarrow\; \bar{\psi}\mathrm{e}^{-\mathrm{iq}\alpha}\partial_\mu\mathrm{e}^{\mathrm{iq}\alpha}\psi$ effectively shifting the derivative, $\partial_\mu \rightarrow \partial_\mu + \mathrm{iq}\partial_\mu\alpha$. Now A_μ needs to cancel the shift requiring

$$\partial_\mu + \mathrm{iq}\partial_\mu\alpha + \mathrm{c}(\mathrm{A}_\mu + \mathrm{a}_\mu) = \partial_\mu + \mathrm{cA}_\mu. \tag{3–128}$$

Therefore c=-iq and $\mathrm{a}_\mu = \partial_\mu\alpha$. Furthermore, $\mathrm{F}^{\mu\nu}\mathrm{F}_{\mu\nu}$ must remain invariant under $\mathrm{A}_\mu(\mathrm{x}) \rightarrow \mathrm{A}_\mu(\mathrm{x}) + \partial_\mu\alpha(\mathrm{x})$. These conditions plus Lorentz invariance force the Lagrangian to be of the form,

$$\mathrm{S} = \int \mathrm{d}^4\mathrm{x}\frac{-1}{4}(\partial_\mu\mathrm{A}_\nu - \partial_\nu\mathrm{A}_\mu)(\partial^\mu\mathrm{A}^\nu - \partial^\nu\mathrm{A}^\mu) + \bar{\psi}\left[\mathrm{i}\gamma^\mu\left(\partial_\mu - \mathrm{iqA}_\mu\right) - \mathrm{m}\right]\psi, \tag{3–129}$$

effectively deriving the Lagrangian for Quantum Electrodynamics (QED).

The QED U(1) symmetry has a single generator and, as a consequence, bypasses any commutation issues. The SU(2) transformation $\psi \rightarrow \mathrm{U}\psi$ with $\mathrm{U} = \mathrm{e}^{\frac{\mathrm{i}}{2}\mathrm{q}\alpha_\mathrm{i}(\mathrm{x})\sigma_\mathrm{i}}$ is a bit more complicated. If

$$\bar{\psi}\mathrm{U}^\dagger\mathrm{i}\gamma^\mu\left(\partial_\mu + \mathrm{cA}'_\mu\right)\mathrm{U}\psi = \bar{\psi}\mathrm{i}\gamma^\mu\left(\partial_\mu + \mathrm{cA}_\mu\right)\psi \tag{3–130}$$

is true when $\mathrm{A}_\mu \rightarrow \mathrm{A}'_\mu$ then the derivative term will be gauge invariant under SU(2). This is true when $\left(\partial_\mu + \mathrm{cA}'_\mu\right)\mathrm{U}\psi = \mathrm{U}\left(\partial_\mu + \mathrm{cA}_\mu\right)\psi$. The transformation $\mathrm{A}_\mu \rightarrow \mathrm{UA}_\mu\mathrm{U}^\dagger + \Delta\mathrm{A}_\mu$ helps simplify the conditions for gauge invariance,

$$\begin{aligned}
\left(\partial_\mu + \mathrm{cA}'_\mu\right)\mathrm{U}\psi &= \partial_\mu(\mathrm{U}\psi) + \mathrm{cA}'_\mu\mathrm{U}\psi \\
&= (\partial_\mu\mathrm{U})\psi + \mathrm{U}(\partial_\mu\psi) + \mathrm{cA}'_\mu\mathrm{U}\psi \\
&= (\partial_\mu\mathrm{U})\psi + \mathrm{U}(\partial_\mu\psi) + \mathrm{cUA}_\mu\mathrm{U}^\dagger\mathrm{U}\psi + \mathrm{c}\Delta\mathrm{A}_\mu \\
&= \mathrm{U}\left(\partial_\mu + \mathrm{cA}_\mu\right)\psi + (\partial_\mu\mathrm{U})\psi + \mathrm{c}\Delta\mathrm{A}_\mu \\
&\rightarrow \mathrm{c}\Delta\mathrm{A}_\mu = -(\partial_\mu\mathrm{U})\psi = -(\partial_\mu\mathrm{U})\mathrm{U}^\dagger\mathrm{U}\psi = \mathrm{U}\partial_\mu\mathrm{U}^\dagger\psi
\end{aligned} \tag{3–131}$$

The last line uses the fact that for a unitary matrix U, $\mathrm{UU^\dagger} = 1$ and $\partial_\mu(\mathrm{UU^\dagger}) = 0$. To preserve gauge invariance under a unitary transformation $\psi \to \mathrm{U}\psi$, cA_μ must transform as

$$\mathrm{cA}_\mu(\mathrm{x}) \to \mathrm{cUA}_\mu(\mathrm{x})\mathrm{U}^\dagger + \mathrm{U}\partial_\mu\mathrm{U}^\dagger. \tag{3–132}$$

For invariance under SU(2) the derivative can be calculated

$$\begin{aligned}\mathrm{cA}_\mu(\mathrm{x}) &\to \mathrm{cUA}_\mu(\mathrm{x})\mathrm{U}^\dagger + \mathrm{c}\Delta\mathrm{A}_\mu \\ &= \mathrm{cUA}_\mu(\mathrm{x})\mathrm{U}^\dagger - \mathrm{iq}\frac{\sigma_\mathrm{i}}{2}\partial_\mu\alpha_\mathrm{i}(\mathrm{x}).\end{aligned} \tag{3–133}$$

Choosing c=-iq and $\Delta\mathrm{A}_\mu = \frac{\sigma_\mathrm{i}}{2}\partial_\mu\alpha_\mathrm{i}(\mathrm{x})$ determines the coupling constant and the transformation for the SU(2) invariant Lagrangian.

A_μ is a matrix which may be expanded in terms of the SU(2) generators, $\mathrm{A}_\mu = \mathrm{A}^1_\mu\frac{\sigma_1}{2} + \mathrm{A}^2_\mu\frac{\sigma_2}{2} + \mathrm{A}^3_\mu\frac{\sigma_3}{2}$. Note that U(1) only requires a single particle while SU(2) requires three. For a general unitary transformation, there will be a force carrying particle for each generator, and an expansion $\mathrm{A}_\mu = \mathrm{T^c A^c_\mu}$. By representing A_μ in terms of the generators, the transformation for each component, $\mathrm{A}^\mathrm{c}_\mu$, may be calculated for a general unitary transformation,

$$\mathrm{A}^\mathrm{c}_\mu \to \mathrm{A}^\mathrm{c}_\mu - \mathrm{f^{abc}}\alpha^\mathrm{a}\mathrm{A}^\mathrm{b}_\mu + \partial_\mu\alpha^\mathrm{c}, \tag{3–134}$$

where $[\mathrm{T^a}, \mathrm{T^b}] = \mathrm{if^{abc}T^c}$ defines the Lie alegbra of the group. For SU(2) in particular, $\mathrm{A}_\mu = \frac{\sigma_\mathrm{c}}{2}\mathrm{A^c}$ and $\mathrm{f^{abc}} = \epsilon^\mathrm{abc}$. This defines the D_μ part of the SU(2) Lagrangian, but $\mathrm{F}_{\mu\nu}$ still needs to be defined so that it remains invariant under equation 3–134. Defining

$$\mathrm{F}_{\mu\nu} = \partial_\mu\mathrm{A}_\nu - \partial_\nu\mathrm{A}_\mu - \mathrm{iq}[\mathrm{A}_\mu, \mathrm{A}_\nu] \tag{3–135}$$

makes $\mathrm{F}_{\mu\nu}$ covariant, $\mathrm{F}_{\mu\nu} \to \mathrm{UF}_{\mu\nu}\mathrm{U}^\dagger$, and to make it invariant, the trace is taken. So finally, the full gauge invariant Lagrangian with massless fermions is given by

$$\mathrm{S} = \int \mathrm{d^4x}\mathcal{N}\mathrm{tr}(\mathrm{F}_{\mu\nu}\mathrm{F}^{\mu\nu}) + \bar{\psi}\mathrm{i}\gamma^\mu\mathrm{D}_\mu\psi, \tag{3–136}$$

where $\mathrm{D}_\mu = \partial_\mu - \mathrm{iqA}_\mu$ is the covariant derivative. The Lagrangian holds for any unitary gauge transformation, U(1) provides electromagnetism, adding on SU(2) provides the electroweak interactions, and adding SU(3) provides the strong force. All three together, describe the Standard Model. For SU(N) invariant theories the Lagrangian is given by,

$$\mathrm{S} = \int \mathrm{d^4x}\frac{-1}{2}\mathrm{tr}(\mathrm{F}_{\mu\nu}\mathrm{F}^{\mu\nu}) + \bar{\psi}\mathrm{i}\gamma^\mu \mathrm{D}_\mu\psi = \int \mathrm{d^4x}\frac{-1}{4}\mathrm{F}^{\mathrm{a}}_{\mu\nu}\mathrm{F}^{\mu\nu}_{\mathrm{a}} + \bar{\psi}\left[\mathrm{i}\gamma^\mu\left(\partial_\mu - \mathrm{iqT^iA^i_\mu}\right)\right]\psi. \quad (3–137)$$

where

$$\mathrm{F}^{\mathrm{a}}_{\mu\nu} = \partial_\mu \mathrm{A}^{\mathrm{a}}_\nu - \partial_\nu \mathrm{A}^{\mathrm{a}}_\mu + \mathrm{qf^{abc}A^b_\mu A^c_\nu}. \quad (3–138)$$

$\mathcal{N}$ is set to $-\frac{1}{2}$ in order to get the appropriate $-\frac{1}{4}$ in front of the $\mathrm{F}^{\mathrm{a}}_{\mu\nu}$ tensors[16] .

The SU(2) invariant Lagrangian of equation 3–137 provides the weak interaction. Unfortunately, the theory has massless fermions and massless weak force particles, while in real life these particles have mass. Adding mass terms directly ruins the SU(2) invariance, so another mechanism is needed, and this is where the Higgs comes into play.

3.1.7 The Higgs Mechanism

The W and Z bosons observed in nature are massive, but directly adding a mass term for a force carrying particle ruins the gauge symmetry,

$$\frac{1}{2}\mathrm{m^2_A A_\mu A^\mu} \rightarrow \frac{1}{2}\mathrm{m^2_A}(\mathrm{UA_\mu U^\dagger} + \Delta\mathrm{A}_\mu)(\mathrm{UA^\mu U^\dagger} + \Delta\mathrm{A}^\mu). \quad (3–139)$$

Meanwhile for the matter particles, the gauge transformation acts on a column of fermions which leaves the $\bar{\psi}\mathrm{m}\psi$ term invariant,

$$\begin{pmatrix}\bar{\psi}_1\\ \ldots\\ \bar{\psi}_{\mathrm{n}}\end{pmatrix}\mathrm{m}\begin{pmatrix}\psi_1 & \ldots & \psi_{\mathrm{n}}\end{pmatrix} \rightarrow \begin{pmatrix}\bar{\psi}_1\\ \ldots\\ \bar{\psi}_{\mathrm{n}}\end{pmatrix}\mathrm{U^\dagger mU}\begin{pmatrix}\psi_1 & \ldots & \psi_{\mathrm{n}}\end{pmatrix}, \quad (3–140)$$

[16] The calculation uses the fact that the SU(N) the generator matrices form an orthogonal basis, $\mathrm{tr}(\mathrm{T^aT^b}) = \frac{1}{2}\delta^{\mathrm{ab}}$.

but this term restricts the fermions to the same mass. In order to describe the massive fermions and massive W and Z particles as seen in nature the Higgs mechanism is needed. Relativistically, mass is energy, so the idea is to produce $\mathrm{m_A^2 A_\mu A^\mu}$ and $\bar{\psi}\mathrm{m}\psi$ terms via some interaction energy involving a new field. Because mass is a scalar, the interactions require a scalar field, and because the mass is derived from a nonzero interaction energy, the groundstate of the scalar field should be nonzero. For all of these reasons, the Higgs mechanism adds a scalar field with a ϕ^4 potential term to the Lagrangian,

$$\mathcal{L}_\phi^{\mathrm{toy}} = (\partial^\mu\phi - \mathrm{iqA}^\mu\phi)^\dagger(\partial_\mu\phi - \mathrm{iqA}_\mu\phi) + \frac{\mathrm{m_h^2}}{2}\phi^\dagger\phi - \frac{\lambda}{4}(\phi^\dagger\phi)^2. \tag{3–141}$$

As ϕ goes into the groundstate, $\phi \to \phi_0$, the force carrying field(s) acquire mass through the $\mathrm{A}^2\phi_0^2$ terms. Similarly, interaction terms coupling the left and right spinors provide the fermions with mass as $\phi \to \phi_0$

$$\mathcal{L}_\mathrm{I}^{\mathrm{toy}} = -\beta(\mathrm{L}^\dagger\phi\mathrm{R} + \mathrm{R}^\dagger\phi\mathrm{L}). \tag{3–142}$$

The toy examples of equations 3–141 and 3–142 cover the basic ideas behind the Higgs mechanism, but correctly describing the electroweak interaction requires a more complex and intricate theory, the U(1)xSU(2) Weinberg-Salam Lagrangian. The Weinberg-Salam Lagrangian includes a U(1) gauge field and an SU(2) gauge field.In the theory, the U(1) field and the third component of the SU(2) field mix up, with one orthogonal piece providing the massive Z boson and the other the massless photon. The remaining first and second components of the SU(2) field mix up to provide the massive W^+ and W^- bosons.

The Weinberg-Salam Lagrangian for the electron (e) and electron neutrino (ν_e) of equation 3–143 is written in terms of left and right handed Dirac spinors. Left and right handed Dirac spinors are defined in terms of the left and right handed spinors as follows, $\psi_\mathrm{L} = \begin{pmatrix}\mathrm{L}\\ 0\end{pmatrix}$ and $\psi_\mathrm{R} = \begin{pmatrix}0\\ \mathrm{R}\end{pmatrix}$. In nature, the weak force treats left and right handed particles differently. The W^+ and W^- particles interact only with left handed particles and interact with those of the same generation symmetrically. In this regard, the Lagrangian is written to respect the interchange of the left handed particles with the SU(2) transformation acting on $\mathrm{L} = \begin{pmatrix} \nu_\mathrm{e} \\ \mathrm{e_L} \end{pmatrix}$. The remaining right handed particles are denoted by $\mathrm{R} = \mathrm{e_R}$. In addition, only left handed neutrinos have been observed in nature, so ν_e is left handed. The electroweak Lagrangian is then,

$$\begin{aligned}\mathcal{L} &= (\partial^\mu\phi - \mathrm{i}\mathrm{q}_{\mathrm{w}\phi}\frac{\sigma_\mathrm{i}}{2}\mathrm{W}_\mathrm{i}^\mu\phi - \mathrm{i}\mathrm{q}_{\mathrm{b}\phi}\mathrm{B}^\mu\phi)^\dagger(\partial_\mu\phi - \mathrm{i}\mathrm{q}_{\mathrm{w}\phi}\frac{\sigma_\mathrm{i}}{2}\mathrm{W}_\mu^\mathrm{i}\phi - \mathrm{i}\mathrm{q}_{\mathrm{b}\phi}\mathrm{B}^\mu\phi) \\ &+ \bar{\mathrm{L}}\mathrm{i}\gamma^\mu(\partial_\mu - \mathrm{i}\mathrm{q}_\mathrm{wl}\frac{\sigma_\mathrm{i}}{2}\mathrm{W}_\mathrm{i}^\mu - \mathrm{i}\mathrm{q}_\mathrm{bl}\mathrm{B}_\mu)\mathrm{L} + \bar{\mathrm{R}}\mathrm{i}\gamma^\mu(\partial_\mu - \mathrm{i}\mathrm{q}_\mathrm{br}\mathrm{B}_\mu)\mathrm{R} \\ &- \beta(\bar{\nu}_\mathrm{e}\phi_+\mathrm{e_R} + \bar{\mathrm{e}}_\mathrm{R}\phi_+^*\nu_\mathrm{e} + \bar{\mathrm{e}}_\mathrm{L}\phi_-\mathrm{e_R} + \bar{\mathrm{e}}_\mathrm{R}\phi_-^*\mathrm{e_L}) - \frac{1}{4}\mathrm{G}_\mathrm{i}^{\mu\nu}\mathrm{G}_{\mu\nu}^\mathrm{i} - \frac{1}{4}\mathrm{F}^{\mu\nu}\mathrm{F}_{\mu\nu} \\ &+ \frac{\mathrm{m_h^2}}{2}\phi^\dagger\phi - \frac{\lambda}{4}(\phi^\dagger\phi)^2.\end{aligned} \tag{3–143}$$

The SU(2) field, W_μ, is a complex 2x2 matrix expanded in terms of the Pauli matrices which operates on 2x1 complex column vectors. Therefore ϕ is written

$$\phi = \begin{pmatrix} \phi_+ \\ \phi_- \end{pmatrix}. \tag{3–144}$$

To isolate the mass terms, the scalar field is expanded about its minimum, $|\phi| = \phi_0 = \sqrt{\frac{\mathrm{m_h^2}}{\lambda}}$. The minimum is degenerate, only requiring that $|\phi_+|^2 + |\phi_-|^2 = \phi_0^2$. The particular solution $\phi = \begin{pmatrix} 0 \\ \phi_0 + \frac{1}{\sqrt{2}}\mathrm{h(x)} \end{pmatrix}$ fixes the fermion mass terms such that the electron acquires mass in the correct way and that the strange mass-like coupling between the electron and neutrino disappears,

$$\beta(\bar{\mathrm{L}}\phi\mathrm{R} + \bar{\mathrm{R}}\phi^\dagger\mathrm{L}) = \beta(\bar{\nu}_\mathrm{e}\phi_+\mathrm{e_R} + \bar{\mathrm{e}}_\mathrm{R}\phi_+^*\nu_\mathrm{e} + \bar{\mathrm{e}}_\mathrm{L}\phi_-\mathrm{e_R} + \bar{\mathrm{e}}_\mathrm{R}\phi_-^*\mathrm{e_L}) \to \beta(\bar{\mathrm{e}}_\mathrm{L}\phi_0\mathrm{e_R} + \bar{\mathrm{e}}_\mathrm{R}\phi_0\mathrm{e_L}). \tag{3–145}$$

Fixing $\phi^+ = 0$ and $\mathrm{Im}\phi^- = 0$ comes at a cost, breaking the SU(2) symmetry of the Lagrangian. The coupling is no longer in the invariant form $\bar{\mathrm{L}}\phi\mathrm{R} + \bar{\mathrm{R}}\phi^\dagger\mathrm{L} \to \bar{\mathrm{L}}\mathrm{U}^\dagger\mathrm{U}\phi\mathrm{R} + \bar{\mathrm{R}}\phi^\dagger\mathrm{U}^\dagger\mathrm{U}\mathrm{L}$.

As in the toy example, the $(D^\mu\phi)^\dagger(D_\mu\phi)$ term bestows mass onto the force carriers,

$$\begin{aligned}
D_\mu\phi &= (\partial_\mu - iq_{w\phi}\frac{\sigma_i}{2}W^i_\mu - iq_{b\phi}B_\mu)\phi \\
&= \left[\begin{pmatrix} \partial_\mu - iq_{b\phi}B_\mu & 0 \\ 0 & \partial_\mu - iq_{b\phi}B_\mu \end{pmatrix} - \frac{i}{2}q_{w\phi}\begin{pmatrix} W^3_\mu & W^1_\mu - iW^2_\mu \\ W^1_\mu + iW^2_\mu & -W^3_\mu \end{pmatrix}\right]\begin{pmatrix} 0 \\ \phi_0 + \frac{1}{\sqrt{2}}h \end{pmatrix} \\
&= -\frac{i}{2}\begin{pmatrix} q_{w\phi}\phi_0(W^1_\mu - iW^2_\mu) + \frac{1}{\sqrt{2}}q_{w\phi}h(W^1_\mu - iW^2_\mu) \\ i\sqrt{2}\partial_\mu h + \phi_0(2q_{b\phi}B_\mu - q_{w\phi}W^3_\mu) + \frac{1}{\sqrt{2}}h(2q_{b\phi}B_\mu - q_{w\phi}W^3_\mu) \end{pmatrix}.
\end{aligned} \tag{3–146}$$

The ϕ_0^2 terms determine the masses,

$$(D^\mu\phi)^\dagger(D_\mu\phi) = q^2_{w\phi}\frac{\phi_0^2}{4}(W^1_\mu)^2 + q^2_{w\phi}\frac{\phi_0^2}{4}(W^2_\mu)^2 + \frac{\phi_0^2}{4}(q_{w\phi}W^3_\mu - 2q_{b\phi}B_\mu)^2 + \text{other terms.} \tag{3–147}$$

The orthogonal term $q_{w\phi}W^3_\mu + 2q_{b\phi}B_\mu$ is missing from the covariant derivative and remains massless, providing the photon field. This leaves

$$m_w = \frac{q_{w\phi}\phi_0}{\sqrt{2}},\ m_z = \frac{m_w}{q_{w\phi}},\ \text{and } m_\gamma = 0. \tag{3–148}$$

The photon, A_μ, is a linear combination of W^3_μ and B_μ. The U(1) symmetry corresponding to electromagnetism is then

$$\begin{aligned}
&U_A = e^{i(Q_{wi}T^3 + Q_{bi})\alpha(x)} \\
&W^3_\mu \to W^3_\mu + \frac{1}{g_w}\partial_\mu\alpha \\
&B_\mu \to B_\mu + \frac{1}{g_b}\partial_\mu\alpha
\end{aligned} \tag{3–149}$$

where T^3 is the third SU(2) generator for the given representation, and the Qs are the normalized charges defined by,

$$q_{wi} = Q_{wi}g_w \text{ and } q_{bi} = Q_{bi}g_b. \tag{3–150}$$

The U(1) gauge symmetry leads to conservation of electromagnetic charge, $Q = Q_wI_3 + Q_b$, implying that the electromagnetic charge for a given particle is $Q_i = Q_{wi}I_{3i} + Q_{bi}$. I_3 represents the eigenvalue of the T^3 generator denoting the particle eigenstate. For example, the operator $\mathrm{T}^3 = \frac{\sigma^3}{2}$ acting on L has two eigenstates, ν_e and $\mathrm{e_L}$, corresponding to eigenvalues $\frac{1}{2}$ and $\frac{-1}{2}$ respectively. Similarly, the eigenstates $\phi^+ = 0$ and $\phi^- = \phi_0 + \mathrm{h}$ correspond to eigenvalues $\frac{1}{2}$ and $\frac{-1}{2}$.

In order to agree with experiment, the left handed particles and the scalar field are assigned $\mathrm{Q_{wi}} = 1$, and the right handed particles are assigned $\mathrm{Q_{wi}} = 0$. The electromagnetic charge $\mathrm{Q_i}$ and the isospin $\mathrm{Q_{wi}}$ values fix the remaining electroweak $\mathrm{Q_{bi}}$ values,

$$\begin{aligned}
\mathrm{Q_{e_L}} &= \mathrm{Q_{we_L} I_3} + \mathrm{Q_{be_L}} = \frac{-1}{2} + \mathrm{Q_{be_L}} \\
\mathrm{Q_{e_R}} &= \mathrm{Q_{we_R} I_3} + \mathrm{Q_{be_R}} = 0 + \mathrm{Q_{be_R}} \\
\mathrm{Q_{\nu_e}} &= \mathrm{Q_{w\nu_e} I_3} + \mathrm{Q_{b\nu_e}} = \frac{1}{2} + \mathrm{Q_{b\nu_e}} \\
\mathrm{Q_\phi} &= \mathrm{Q_{w\phi} I_3} + \mathrm{Q_{b\phi}} = \frac{-1}{2} + \mathrm{Q_{b\phi}},
\end{aligned} \tag{3–151}$$

providing,

$$\mathrm{Q_{be_L}} = \frac{-1}{2},\ \mathrm{Q_{be_R}} = -1,\ \mathrm{Q_{b\nu_e}} = \frac{-1}{2},\ \text{and } \mathrm{Q_{b\phi}} = \frac{+1}{2}. \tag{3–152}$$

The electroweak Lagrangian reduces to,

$$\begin{aligned}
\mathcal{L} = {} & (\partial^\mu \phi - \mathrm{i g_w} \frac{\sigma_\mathrm{i}}{2} \mathrm{W_i^\mu} \phi - \frac{\mathrm{i}}{2} \mathrm{g_b B^\mu} \phi)^\dagger (\partial_\mu \phi - \mathrm{i g_w} \frac{\sigma_\mathrm{i}}{2} \mathrm{W^i_\mu} \phi - \frac{\mathrm{i}}{2} \mathrm{g_b B^\mu} \phi) \\
& + \bar{\mathrm{L}} \mathrm{i} \gamma^\mu (\partial_\mu - \mathrm{i g_w} \frac{\sigma_\mathrm{i}}{2} \mathrm{W_i^\mu} + \frac{\mathrm{i}}{2} \mathrm{g_b B_\mu}) \mathrm{L} + \bar{\mathrm{R}} \mathrm{i} \gamma^\mu (\partial_\mu + \mathrm{i g_b B_\mu}) \mathrm{R} \\
& - \beta (\bar{\mathrm{L}} \phi \mathrm{R} + \bar{\mathrm{R}} \phi^\dagger \mathrm{L}) - \frac{1}{4} \mathrm{G_i^{\mu\nu} G^i_{\mu\nu}} - \frac{1}{4} \mathrm{F^{\mu\nu} F_{\mu\nu}} + \frac{\mathrm{m_h^2}}{2} \phi^\dagger \phi - \frac{\lambda}{4} (\phi^\dagger \phi)^2
\end{aligned} \tag{3–153}$$

with $\phi = \begin{pmatrix} 0 \\ \phi_0 + \frac{1}{\sqrt{2}} \mathrm{h(x)} \end{pmatrix}$ and $\mathrm{L} = \begin{pmatrix} \nu_\mathrm{e} \\ \mathrm{e_L} \end{pmatrix}$. In equation 3–153, the W^+ particle is $\mathrm{W}^1_\mu + \mathrm{i W}^2_\mu$, and the W^- particle is $\mathrm{W}^1_\mu - \mathrm{i W}^2_\mu$. As for the neutral bosons, the photon is $\mathrm{g_w W^3_\mu} + \mathrm{g_b B_\mu}$, the Z boson is $\mathrm{g_w W^3_\mu} - \mathrm{g_b B_\mu}$, and the Higgs boson is h(x). Adding the next two generations of leptons, the three generations of quarks, and the SU(3) interactions to the electroweak Lagrangian defines the entire Standard Model.

3.2 The Standard Model Higgs and the LHC

The SM Higgs interacts with the massive particles of the SM: the non-neutrino leptons, the quarks, and the W^{+}, W^{-}, and Z particles. As such, it can be produced by colliding certain combinations of these particles, and it can decay into them as well. The cross sections are proportional to the probability for a production process, and consequently, they describe how likely each collision is to produce a Higgs of a certain mass. Some of these cross sections are shown in Figure 3-4 for LHC proton-proton collisions with a center of mass energy of 14 TeV.

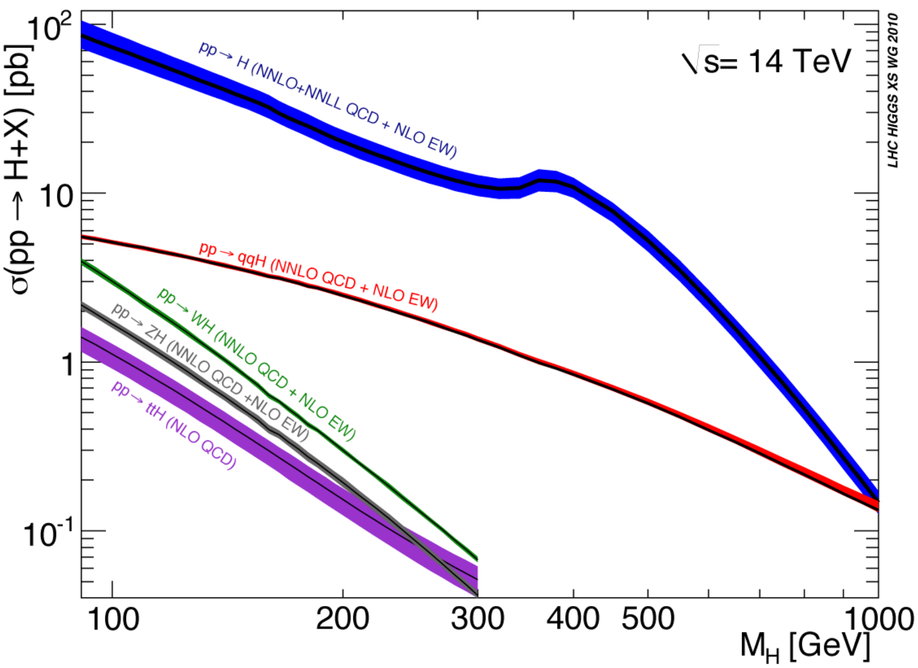


Figure 3-4. The highest production mode cross sections for the SM Higgs at 14 TeV **?**

The cross sections are functions of the mass of the Higgs as well as the energy of the collisons. For a given collision energy, as in Figure 3-4, the cross sections decrease as the Higgs mass increases. Naturally, there are fewer kinematic possibilities for a heavier particle when a larger portion of the collision energy was used to create the particle. On the other hand, for a specific Higgs mass, the cross section grows with collision energy at the LHC. This constrasts with cross sections involving collisions of fundamental particles, e.g. electron antielectron collisions, due to the fact that the LHC collides protons together.

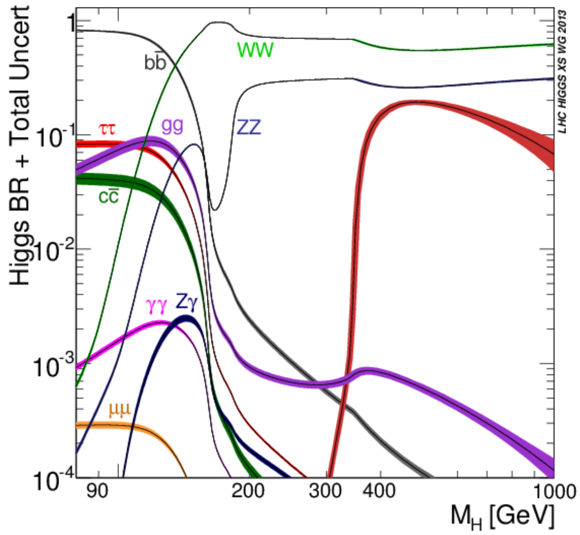
Protons behave like a quantum superposition of an infinite number of quark-antiquarks, an infinite number of gluons, and the usual uud. As a consequence, the total momentum of a proton in a collision is divided up amongst these constituents called partons. This experimentally verified phenomena is modelled by the parton distribution function, which describes the number of partons with a given fraction of the total momentum. In general, there are many partons with very little of the momentum. With a larger proton momentum, the minimum creation energy is a smaller fraction than before. Since there are more particles at a lower fraction, it's as if a larger number of partons with the necessary energy are colliding. This effective increase in the density of energetic partons results in a growth of the cross section with collision energy.

Beyond the production mechanisms, the SM Higgs is unstable and decays with a width of about 5 MeV at 125 GeV. The probability of each decay changes depending upon the mass of the decay products. In general, the Higgs couples more strongly to particles with higher mass, making the decays to heavier particles more likely.

The muon has the lowest mass – excluding the photon and gluon – of the particles in Figure 3-5 and consequently H→ $\mu^+\mu^-$ has the lowest branching fraction in the set. [17] The gluons and photons are massless and do not couple to the Higgs at leading order. These massless vector bosons interact with the Higgs through a loop of top quarks. The extremely heavy mass of the top quark, about 173 GeV, balances the fact that the loop production is a higher order mechanism.

The Higgs to massless vector boson coupling via the top loop is seen in the GF Feynman diagram in Figure 3-6. At $\mathrm{M_h} = 125$ GeV, $\sqrt{\mathrm{s}} =$ 13 TeV, the GF channel comprises 87% of the total Higgs production cross section, VBF 7%, VH 4%, and $\mathrm{t\tilde{t}}$H 1% **?**. Besides $\mathrm{t\tilde{t}}$H, the

[17] The Higgs also couples to the electron and the first generation quarks but the masses are so light that CMS does not expect to see the SM Higgs in those modes.



Higgs Branching Ratios	
$b\tilde{b}$	0.57
WW	0.22
gg	0.085
$\tau\tau$	0.065
ZZ	0.027
$c\tilde{c}$	0.027
$\gamma\gamma$	0.0023
$Z\gamma$	0.0016
$\mu^+\mu^-$	0.00022

Figure 3-5. The graphic on the top left presents the SM Higgs branching fractions as functions of mass while the table on the bottom right displays the branching fractions for a 125 GeV SM Higgs **?**.

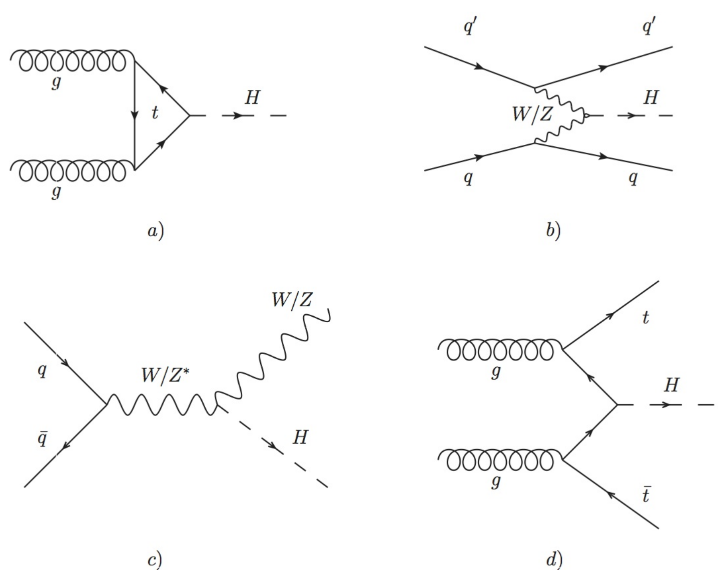


Figure 3-6. The SM production modes with the highest cross sections. a) Gluon Gluon Fusion (GF) b) Vector Boson Fusion (VBF) c) Associated Production with a Vector Boson (VH) d) $t\tilde{t}$H

process q + $\tilde{\mathrm{q}} \rightarrow$ H isn't considered due to its low cross section. The low masses of the other quarks suppress the process.

Quark gluon (qg) scattering is a major background for the Higgs to two jets decays since the process closely resembles GF in this mode.

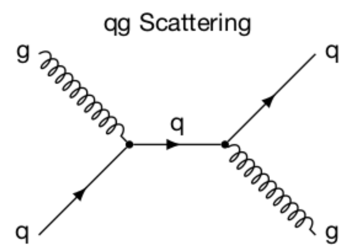


Figure 3-7. Quark gluon scattering creates many two jet events. This background looks very similar to GF when the Higgs decays to two jets. The colliding protons are made of quarks and gluons so this process is extremely common.

CHAPTER 4
RESULTS

4.1 Fusce Eget Tempus Lectus,

Algorithm 4.1. *Euclids algorithm*

```
procedure EUCLID(a, b)                ▷ The g.c.d. of a and b
    r ← a mod b
    while r ≠ 0 do                    ▷ We have the answer if r is 0
        a ← b
        b ← r
        r ← a mod b
    end while
    return b                          ▷ The gcd is b
end procedure
```

Proposition 4.1. *The Upsilon Function*

(1) If $\beta > 0$ and $\alpha \neq 0$, then for all $\mathrm{n} \geq -1$,

$$\mathrm{I_n(c;\alpha;\beta;\delta) = -\frac{e^{\alpha c}}{\alpha}\sum_{i=0}^{n}(\frac{\beta}{\alpha})^{n-i}Hh_i(\beta c-\delta)}$$

$$+(\frac{\beta}{\alpha})^{\mathrm{n}+1}\frac{\sqrt{2\pi}}{\beta}\mathrm{e}^{\frac{\alpha\delta}{\beta}+\frac{\alpha^2}{2\beta^2}}\phi(-\beta\mathrm{c}+\delta+\frac{\alpha}{\beta})$$

(2) If $\beta < 0$ and $\alpha < 0$, then for all $\mathrm{x} \geq -1$

$$\mathrm{I_n(c;\alpha;\beta;\delta) = -\frac{e^{\alpha c}}{\alpha}\sum_{i=0}^{n}(\frac{\beta}{\alpha})^{n-i}Hh_i(\beta c-\delta)}$$

$$-(\frac{\beta}{\alpha})^{\mathrm{n}+1}\frac{\sqrt{2\pi}}{\beta}\mathrm{e}^{\frac{\alpha\delta}{\beta}+\frac{\alpha^2}{2\beta^2}}\phi(\beta\mathrm{c}-\delta-\frac{\alpha}{\beta})$$

Proof. Case 1.

$\beta > 0$ and $\alpha \neq 0$. Since, for any constant α and $\mathrm{n} \geq 0$, $\mathrm{e^{\alpha x}Hh_n(\beta x-\delta)} \to 0$ as $\mathrm{x} \to \infty$ thanks to (B4), integration by parts leads to

$$\mathrm{I_n = -\frac{1}{\alpha}Hh(\beta c-\delta)e^{\alpha c} + \frac{\beta}{\alpha}\int_c^{\infty} e^{\alpha x}Hh_{n-1}(\beta c-\delta)dx}$$

In other words, we have a recursion, for $\mathrm{n} \geq 0$, $\mathrm{I_n} = -(\mathrm{e}^{\alpha \mathrm{c}}\alpha)\mathrm{Hh_n}(\beta \mathrm{c} - \delta) + (\frac{\beta}{\alpha})\mathrm{I_{n-1}}$ with

$$\mathrm{I_{-1}} = \sqrt{2\pi}\int_{\mathrm{c}} \infty \mathrm{e}^{\alpha \mathrm{x}}\varphi(-\beta \mathrm{x} + \delta)\mathrm{dx}$$

$$= \frac{\sqrt{2\pi}}{\beta}\mathrm{e}^{\frac{\alpha\delta}{\beta}+\frac{\alpha^2}{2\beta^2}}\phi(-\beta \mathrm{c} + \delta + \frac{\alpha}{\beta})$$

Solving it yields, for $\mathrm{n} \geq -1$,

$$\mathrm{I_n} = -\frac{\mathrm{e}^{\alpha \mathrm{c}}}{\alpha}\sum_{\mathrm{i}=0}^{\mathrm{n}}(\frac{\beta}{\alpha})^{\mathrm{i}}\mathrm{Hh_{n-i}}(\beta \mathrm{c} + \delta) + (\frac{\beta}{\alpha})^{\mathrm{n}+1}\mathrm{I_{-1}}$$

$$= -\frac{\mathrm{e}^{\alpha \mathrm{c}}}{\alpha}\sum_{\mathrm{i}=0}^{\mathrm{n}}(\frac{\beta}{\alpha})^{\mathrm{n-i}}\mathrm{Hh_i}(\beta \mathrm{c} + \delta)$$

$$+(\frac{\beta}{\alpha})^{\mathrm{n}+1}\frac{\sqrt{2\pi}}{\beta}\mathrm{e}^{\frac{\alpha\delta}{\beta}+\frac{\alpha^2}{2\beta^2}}\phi(-\beta \mathrm{c} + \delta + \frac{\alpha}{\beta})$$

where the sum over an empty set is defined to be zero. □

Proof. Case2. $\beta < 0$ and $\alpha < 0$. In this case, we must also have, for $\mathrm{n} \geq 0$ and any constant $\alpha < 0, \mathrm{e}^{\alpha \mathrm{x}}\mathrm{Hh_n}(\beta \mathrm{x} - \delta) \to 0$ as

$\mathrm{x} \to \infty$, thanks to (B5). Using integration by parts, we again have the same recursion, for $\mathrm{n} \geq 0, \mathrm{I_n} = -(\mathrm{e}^{\alpha \mathrm{c}}/\alpha)\mathrm{Hh_n}(\beta \mathrm{c} - \delta) + (\beta/\alpha)\mathrm{I_{n-1}}$, but with a different initial condition

$$\mathrm{I_{-1}} = \sqrt{2\pi}\int_{\mathrm{c}}^{\infty} \mathrm{e}^{\alpha \mathrm{x}}\varphi(-\beta \mathrm{x} + \delta)\mathrm{dx}$$

$$= -\frac{\sqrt{2\pi}}{\beta}\exp\{\frac{\alpha\delta}{\beta} + \frac{\alpha^2}{2\beta^2}\}\phi(\beta \mathrm{c} - \delta - \frac{\alpha}{\beta})$$

Solving it yields (B8), for $\mathrm{n} \geq -1$.

□

Finally, we sum the double exponential and the normal random variables

Proposition B.3.

Suppose $\{\xi_1, \xi_2, ...\}$ *is a sequence of i.i.d. exponential random variables with rate* $\eta > 0$*, and Z is a normal variable with distribution* $\mathrm{N}(0, \sigma^2)$*. Then for every* $\mathrm{n} \geq 1$*, we have: (1) The density functions are given by:*

$$f_{Z+\sum_{i=1}^{n}\xi_i}(t) = (\sigma\eta)^n \frac{e^{(\sigma\eta)^2/2}}{\sigma\sqrt{2\pi}} e^{-t\eta} \mathrm{Hh}_{n-1}(-\frac{t}{\sigma} + \sigma\eta)$$

$$f_{Z-\sum_{i=1}^{n}\xi_i}(t) = (\sigma\eta)^n \frac{e^{(\sigma\eta)^2/2}}{\sigma\sqrt{2\pi}} e^{-t\eta} \mathrm{Hh}_{n-1}(\frac{t}{\sigma} + \sigma\eta)$$

(2) The tail probabilities are given by

$$P(Z + \sum_{i=1}^{n} \xi_i \geq x) = (\sigma\eta)^n \frac{e^{(\sigma\eta)^2/2}}{\sigma\sqrt{2\pi}} e^{-t\eta} I_{n-1}(x; -\eta, -\frac{1}{\sigma}, -\sigma\eta)$$

$$P(Z - \sum_{i=1}^{n} \xi_i \geq x) = (\sigma\eta)^n \frac{e^{(\sigma\eta)^2/2}}{\sigma\sqrt{2\pi}} e^{-t\eta} I_{n-1}(x; \eta, \frac{1}{\sigma}, -\sigma\eta)$$

Proof. Case 1. The densities of $Z + \sum_{i=1}^{n} \xi_i$*, and* $Z - \sum_{i=1}^{n} \xi_i$*. We have*

$$f_{Z+\sum_{i=1}^{n}\xi_i}(t) = \int_{-\infty}^{\infty} f_{\sum_{i=1}^{n}\xi_i}(t-x) f_Z(x) dx$$

$$= e^{-t\eta}(\eta^n) \int_{-\infty} t \frac{e^{x\eta}(t-x)^{n-1}}{(n-1)!} \frac{1}{\sigma\sqrt{2\pi}} e^{-x^2/(2\sigma^2)} dx$$

$$= e^{-t\eta}(\eta^n) e^{(\sigma\eta)^2/(2)} \int_{-\infty} t \frac{(t-x)^{n-1}}{(n-1)!} \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\sigma^2\eta)^2/(2\sigma^2)} dx$$

Letting $y = (x - \sigma^2\eta)/\sigma$ *yields*

$$f_{Z+\sum_{i=1}^{n}\xi_i}(t) = e^{-t\eta}(\eta^n) e^{(\sigma\eta)^2/(2)} \sigma^{n-1}$$

$$\times \int_{-\infty}^{t/\sigma - \sigma\eta} \frac{(t/\sigma - y - \sigma\eta)^{n-1}}{(n-1)!} \frac{1}{\sqrt{2\pi}} e^{-y^2/2} dy$$

$$= \frac{e^{(\sigma\eta)^2/2}}{\sqrt{2\pi}}(\sigma^{n-1}\eta^n)e^{-t\eta}Hh_{n-1}(-t/\sigma+\sigma\eta)$$

because $(1/(n-1)!)\int_{-\infty} a(a-y)^{n-1}e^{-y^2/2}dy = Hh_{n-1}(a)$. *The derivation of* $f_{Z+\sum_{i=1}^n \xi_i}(t)$ *is similar.*

Case 2. $P(Z+\sum_{i=1}^n \xi_i \geq x)$ *and* $P(Z-\sum_{i=1}^n \xi_i \geq x)$. *From (B9), it is clear that*

$$P(Z+\sum_{i=1}^{n}\xi_i \geq x) = \frac{(\sigma\eta)^n e^{(\sigma\eta)^2/2}}{\sigma\sqrt{2\pi}}\int_x^\infty e^{(-i\eta)}Hh_{n-1}(-\frac{t}{\sigma}+\sigma\eta)dt$$

$$= \frac{(\sigma\eta)^n e^{(\sigma\eta)^2/2}}{\sigma\sqrt{2\pi}}I_{n-1}(x;-\eta,-\frac{1}{\sigma},-\sigma\eta)dt$$

by (B6). We can compute $P(Z-\sum_{i=1}^n \xi_i \geq x)$ *similarly.*

Theorem 4.1. *Theorem With* $\pi_n := P(N(t)=n) = e^{-\lambda T}(\lambda T)^n/n!$ *and* I_n *in Proposition* **??**. , *we have*

$$P(Z(T)\geq a) = \frac{e^{(\sigma\eta_1)^2 T/2}}{\sigma\sqrt{2\pi T}}\sum_{n=1}^{\infty}\pi_n\sum_{k=1}^{n}P_{n,k}(\sigma\sqrt{T}\eta_1)^k \times I_{k-1}(a-\mu T;-\eta_1,-\frac{1}{\sigma\sqrt{T}},-\sigma\eta_1\sqrt{T})$$

$$+\frac{e^{(\sigma\eta_2)^2 T/2}}{\sigma\sqrt{2\pi T}}\sum_{n=1}^{\infty}\pi_n\sum_{k=1}^{n}Q_{n,k}(\sigma\sqrt{T}\eta_2)^k$$

$$\times I_{k-1}(a-\mu T;\eta_2,\frac{1}{\sigma\sqrt{T}},-\sigma\eta_2\sqrt{T})$$

$$+\pi_0\phi(-\frac{a-\mu T}{\sigma\sqrt{T}})$$

Proof by the decomposition (B2)

$$P(Z(T)\geq a) = \sum_{n=0}^{\infty}\pi_n P(\mu T+\sigma\sqrt{T}Z+\sum_{j=1}^{n}Y_j \geq a)$$

$$= \pi_0 \mathrm{P}(\mu \mathrm{T} + \sigma\sqrt{\mathrm{T}}\mathrm{Z} \geq \mathrm{a})$$

$$+ \sum_{\mathrm{n}=1}^{\infty} \pi_{\mathrm{n}} \sum_{\mathrm{k}=1}^{\mathrm{n}} \mathrm{P}_{\mathrm{n,k}} \mathrm{P}(\mu \mathrm{T} + \sigma\sqrt{\mathrm{T}}\mathrm{Z} + \sum_{\mathrm{j}=1}^{\mathrm{n}} \xi_{\mathrm{j}}^{+} \geq \mathrm{a})$$

$$+ \sum_{\mathrm{n}=1}^{\infty} \pi_{\mathrm{n}} \sum_{\mathrm{k}=1}^{\mathrm{n}} \mathrm{Q}_{\mathrm{n,k}} \mathrm{P}(\mu \mathrm{T} + \sigma\sqrt{\mathrm{T}}\mathrm{Z} - \sum_{\mathrm{j}=1}^{\mathrm{n}} \xi_{\mathrm{j}}^{-} \geq \mathrm{a})$$

The result now follows via (B11) and (B12) for $\eta_1 > 1$ and $\eta_2 > 0$.

CHAPTER 5
SUMMARY AND CONCLUSIONS

5.1 Non Porttitor Tellus

Aliquam molestie sed urna quis convallis. Aenean nibh eros, aliquam non eros in, tempus lacinia justo. In magna sapien, blandit a faucibus ac, scelerisque nec purus. Praesent fermentum felis nec massa interdum, vel dapibus mi luctus. Cras id fringilla mauris. Ut molestie eros mi, ut hendrerit nulla tempor et. Pellentesque tortor quam, mattis a scelerisque nec, euismod et odio. Mauris rhoncus metus sit amet risus mattis, eu mattis sem interdum.

5.1.1 Nam Arcu Magna

Semper vel lorem eu, venenatis ultrices est. Nam aliquet ut erat ac scelerisque. Maecenas ut molestie mi. Phasellus ipsum magna, sollicitudin eu ipsum quis, imperdiet cursus turpis. Etiam pretium enim a fermentum accumsan. Morbi vel vehicula enim.

5.1.1.1 Ut pellentesque velit sede

Placerat cursus. Integer congue urna non massa dictum, a pellentesque arcu accumsan. Nulla posuere, elit accumsan eleifend elementum, ipsum massa tristique metus, in ornare neque nisl sed odio. Nullam eget elementum nisi. Duis a consectetur erat, sit amet malesuada sapien. Aliquam nec sapien et leo sagittis porttitor at ut lacus. Vivamus vulputate elit vitae libero condimentum dictum. Nulla facilisi. Quisque non nibh et massa ullamcorper iaculis.

APPENDIX A
THIS IS THE FIRST APPENDIX

Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Maecenas eget magna. Aenean et lorem. Ut dignissim neque at nisi. In hac habitasse platea dictumst. In porta ornare eros. Nunc eu ante. In non est vehicula tellus cursus suscipit. Proin sed libero. Sed risus enim, eleifend in, pellentesque ac, nonummy quis, nulla. Phasellus imperdiet libero nec massa. Ut sapien libero, adipiscing eu, volutpat porttitor, ultricies eget, nisi. Sed odio. Suspendisse potenti. Duis dolor augue, viverra id, porta in, dignissim id, nisl. Vivamus blandit cursus eros. Maecenas sit amet urna sit amet orci nonummy pharetra.

Praesent cursus nibh et mauris. In aliquam felis sit amet ligula. Nulla faucibus nisl eget nisl. Aliquam tincidunt. Mauris eget elit sed massa luctus posuere. Pellentesque suscipit. In odio urna, semper ut, convallis ut, porta et, nibh. Nulla sodales metus nec velit posuere gravida. Cras tristique. Etiam urna risus, accumsan ut, placerat sed, iaculis id, est.

Nullam mi. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Duis vitae metus in massa hendrerit rhoncus. Fusce tortor justo, laoreet eu, facilisis at, gravida et, felis. Donec imperdiet mollis erat. Integer tempus nulla ac lorem. Fusce porttitor. Aenean quis arcu. Morbi consectetuer, leo eu mollis elementum, urna massa malesuada risus, euismod tempor lorem elit ut mauris. Cras elit orci, facilisis ac, mattis iaculis, cursus ac, augue. Donec eget nisl. Pellentesque fermentum sodales nibh. Vivamus non risus. Donec est libero, tincidunt sit amet, pretium vitae, blandit sed, tellus. Nunc diam risus, interdum sed, laoreet quis, varius ac, turpis. In et purus eget nibh vehicula rhoncus. Aenean et neque. Praesent nisl nisi, tempus quis, nonummy ac, auctor a, neque. Suspendisse et metus. Suspendisse non metus eu mauris auctor sagittis.

APPENDIX B
AN EXAMPLE OF A HALF TITLE PAGE

LaTeX 2_{ε}

Figure B-1. LaTeX2ϵ. logo

This is how a section should look if the first page is a landscape page. Lorem ipsum dolor sit amet, consectetuer adipiscing elit. Ut sit amet nulla. Integer mauris turpis, dapibus ac, auctor non, vehicula sit amet, magna. Suspendisse eu tellus. Etiam porta. Donec magna. Donec ut dui. In hac habitasse platea dictumst. Nullam suscipit, mi at adipiscing commodo, lorem erat scelerisque erat, non pulvinar leo mi eu metus. Phasellus id felis. Sed quam purus, molestie quis, ultrices nec, dictum at, magna. Proin viverra viverra ante.

Maecenas sagittis magna quis ligula. Duis vestibulum mi a felis. Aenean accumsan mattis massa. Nullam lacus sem, consectetuer non, condimentum sit amet, pharetra ac, odio. Morbi nisi magna, tincidunt sed, placerat nec, tincidunt id, lectus. Donec ac dui non mauris vulputate aliquam. Nullam scelerisque congue pede. Integer ipsum. Vestibulum auctor. Suspendisse eget leo id libero cursus dictum. Sed malesuada. Aliquam imperdiet. Donec dui metus, porta eu, aliquet vel, vulputate vitae, lacus.

Nulla quis purus id turpis luctus feugiat. Fusce feugiat. Proin felis. Morbi elit est, fermentum in, tincidunt vitae, convallis vel, orci. Vestibulum justo. Suspendisse non nisl. Pellentesque pretium adipiscing elit. Phasellus fermentum consequat augue. Sed pede nisl, fermentum vel, vulputate id, sollicitudin sed, ligula. Cras suscipit, quam et euismod sagittis, nisl felis gravida felis, quis pulvinar purus est vel pede. Suspendisse mattis est ac nunc. Curabitur rutrum, turpis sit amet commodo tempus, metus lorem commodo lectus, eget fringilla justo nisi et purus. Ut quam sapien, vehicula quis, rhoncus non, sagittis nec, risus.

Donec eget augue ac lacus adipiscing porta. Maecenas pede. Vivamus molestie. Duis condimentum ligula auctor pede. Nullam ullamcorper rhoncus erat. Ut ornare interdum urna. Suspendisse potenti. Curabitur mattis mauris nec risus. Aenean iaculis turpis eu tortor. Donec nec ante non mauris pellentesque fringilla.

Phasellus vitae dui id orci sodales cursus. Curabitur sed nulla quis mauris tincidunt iaculis. Vivamus semper semper orci. Phasellus suscipit ante vitae leo. Sed arcu ipsum, condimentum id, luctus in, sodales eu, magna. In dictum, arcu quis pharetra vestibulum, ante enim placerat lacus, vitae placerat est leo vitae elit. Pellentesque bibendum enim vulputate eros. Nunc laoreet. Pellentesque habitant morbi tristique senectus et netus et malesuada fames ac turpis egestas. Praesent purus odio, euismod sit amet, aliquam a, volutpat in, augue. Phasellus id massa. Suspendisse suscipit ligula pharetra dolor. Pellentesque vel pede.

Aliquam pharetra est sit amet magna. Aliquam varius. Donec eu lectus et nisl iaculis porttitor. Morbi mattis, mauris sed luctus hendrerit, nulla velit molestie dolor, ac volutpat urna augue vel quam. Maecenas pellentesque libero et massa. Integer vestibulum, lacus at mattis euismod, nisl arcu commodo lectus, ut euismod dolor ligula sit amet libero. Nam in ligula sit amet ante eleifend aliquet. Phasellus feugiat erat at nulla. Proin in lectus. Proin laoreet leo laoreet leo congue lacinia. Quisque non diam sit amet enim ultrices commodo. Praesent fermentum lectus sed ligula. Integer pulvinar accumsan pede. Quisque molestie ligula eget odio. Vestibulum ante ipsum primis in faucibus orci luctus et ultrices posuere cubilia Curae;

APPENDIX C
DERIVATION OF THE Υ FUNCTION

Proposition C.1. *The Upsilon Function*

(1) If $\beta > 0$ and $\alpha \neq 0$, then for all $\mathrm{n} \geq -1$,

$$\mathrm{I_n}(\mathrm{c};\alpha;\beta;\delta) = -\frac{\mathrm{e}^{\alpha\mathrm{c}}}{\alpha}\sum_{\mathrm{i}=0}^{\mathrm{n}}(\frac{\beta}{\alpha})^{\mathrm{n-i}}\mathrm{Hh_i}(\beta\mathrm{c}-\delta)$$

$$+(\frac{\beta}{\alpha})^{\mathrm{n}+1}\frac{\sqrt{2\pi}}{\beta}\mathrm{e}^{\frac{\alpha\delta}{\beta}+\frac{\alpha^2}{2\beta^2}}\phi(-\beta\mathrm{c}+\delta+\frac{\alpha}{\beta})$$

(2) If $\beta < 0$ and $\alpha < 0$, then for all $\mathrm{x} \geq -1$

$$\mathrm{I_n}(\mathrm{c};\alpha;\beta;\delta) = -\frac{\mathrm{e}^{\alpha\mathrm{c}}}{\alpha}\sum_{\mathrm{i}=0}^{\mathrm{n}}(\frac{\beta}{\alpha})^{\mathrm{n-i}}\mathrm{Hh_i}(\beta\mathrm{c}-\delta)$$

$$-(\frac{\beta}{\alpha})^{\mathrm{n}+1}\frac{\sqrt{2\pi}}{\beta}\mathrm{e}^{\frac{\alpha\delta}{\beta}+\frac{\alpha^2}{2\beta^2}}\phi(\beta\mathrm{c}-\delta-\frac{\alpha}{\beta})$$

Proof. Case 1.

$\beta > 0$ and $\alpha \neq 0$. Since, for any constant α and $\mathrm{n} \geq 0$, $\mathrm{e}^{\alpha\mathrm{x}}\mathrm{Hh_n}(\beta\mathrm{x}-\delta) \to 0$ as $\mathrm{x} \to \infty$ thanks to (B4), integration by parts leads to

$$\mathrm{I_n} = -\frac{1}{\alpha}\mathrm{Hh}(\beta\mathrm{c}-\delta)\mathrm{e}^{\alpha\mathrm{c}} + \frac{\beta}{\alpha}\int_{\mathrm{c}}^{\infty}\mathrm{e}^{\alpha\mathrm{x}}\mathrm{Hh_{n-1}}(\beta\mathrm{c}-\delta)\mathrm{dx}$$

In other words, we have a recursion, for $\mathrm{n} \geq 0$, $\mathrm{I_n} = -(\mathrm{e}^{\alpha\mathrm{c}}\alpha)\mathrm{Hh_n}(\beta\mathrm{c}-\delta) + (\frac{\beta}{\alpha})\mathrm{I_{n-1}}$ with

$$\mathrm{I_{-1}} = \sqrt{2\pi}\int_{\mathrm{c}}\infty\mathrm{e}^{\alpha\mathrm{x}}\varphi(-\beta\mathrm{x}+\delta)\mathrm{dx}$$

$$= \frac{\sqrt{2\pi}}{\beta}\mathrm{e}^{\frac{\alpha\delta}{\beta}+\frac{\alpha^2}{2\beta^2}}\phi(-\beta\mathrm{c}+\delta+\frac{\alpha}{\beta})$$

Solving it yields, for $\mathrm{n} \geq -1$,

$$\mathrm{I_n} = -\frac{\mathrm{e}^{\alpha \mathrm{c}}}{\alpha}\sum_{\mathrm{i}=0}^{\mathrm{n}}(\frac{\beta}{\alpha})^{\mathrm{i}}\mathrm{Hh_{n-i}}(\beta \mathrm{c}+\delta) + (\frac{\beta}{\alpha})^{\mathrm{n}+1}\mathrm{I}_{-1}$$

$$= -\frac{\mathrm{e}^{\alpha \mathrm{c}}}{\alpha}\sum_{\mathrm{i}=0}^{\mathrm{n}}(\frac{\beta}{\alpha})^{\mathrm{n-i}}\mathrm{Hh_i}(\beta \mathrm{c}+\delta)$$

$$+(\frac{\beta}{\alpha})^{\mathrm{n}+1}\frac{\sqrt{2\pi}}{\beta}\mathrm{e}^{\frac{\alpha\delta}{\beta}+\frac{\alpha^2}{2\beta^2}}\phi(-\beta \mathrm{c}+\delta+\frac{\alpha}{\beta})$$

where the sum over an empty set is defined to be zero. □

Case2. $\beta < 0$ *and* $\alpha < 0$. *In this case, we must also have, for* $\mathrm{n} \geq 0$ *and any constant* $\alpha < 0, \mathrm{e}^{\alpha \mathrm{x}}\mathrm{Hh_n}(\beta \mathrm{x}-\delta) \to 0$ *as*

$\mathrm{x} \to \infty$, *thanks to (B5). Using integration by parts, we again have the same recursion, for* $\mathrm{n} \geq 0, \mathrm{I_n} = -(\mathrm{e}^{\alpha \mathrm{c}}/\alpha)\mathrm{Hh_n}(\beta \mathrm{c}-\delta) + (\beta/\alpha)\mathrm{I_{n-1}}$, *but with a different initial condition*

$$\mathrm{I}_{-1} = \sqrt{2\pi}\int_{\mathrm{c}}^{\infty}\mathrm{e}^{\alpha \mathrm{x}}\varphi(-\beta \mathrm{x}+\delta)\mathrm{dx}$$

$$= -\frac{\sqrt{2\pi}}{\beta}\exp\{\frac{\alpha\delta}{\beta}+\frac{\alpha^2}{2\beta^2}\}\phi(\beta \mathrm{c}-\delta-\frac{\alpha}{\beta})$$

Solving it yields (B8), for $\mathrm{n} \geq -1$.

Finally, we sum the double exponential and the normal random variables

Proposition B.3.

Suppose $\{\xi_1, \xi_2, ...\}$ *is a sequence of i.i.d. exponential random variables with rate* $\eta > 0$, *and Z is a normal variable with distribution* $\mathrm{N}(0, \sigma^2)$. *Then for every* $\mathrm{n} \geq 1$, *we have: (1) The density functions are given by:*

$$\mathrm{f}_{\mathrm{Z}+\sum_{\mathrm{i}=1}^{\mathrm{n}}\xi_\mathrm{i}}(\mathrm{t}) = (\sigma\eta)^{\mathrm{n}}\frac{\mathrm{e}^{(\sigma\eta)^2/2}}{\sigma\sqrt{2\pi}}\mathrm{e}^{-\mathrm{t}\eta}\mathrm{Hh_{n-1}}(-\frac{\mathrm{t}}{\sigma}+\sigma\eta)$$

$$\mathrm{f}_{\mathrm{Z}-\sum_{\mathrm{i}=1}^{\mathrm{n}}\xi_\mathrm{i}}(\mathrm{t}) = (\sigma\eta)^{\mathrm{n}}\frac{\mathrm{e}^{(\sigma\eta)^2/2}}{\sigma\sqrt{2\pi}}\mathrm{e}^{-\mathrm{t}\eta}\mathrm{Hh_{n-1}}(\frac{\mathrm{t}}{\sigma}+\sigma\eta)$$

(2) The tail probabilities are given by

$$\mathrm{P}(\mathrm{Z}+\sum_{\mathrm{i}=1}^{\mathrm{n}}\xi_{\mathrm{i}}\geq \mathrm{x})=(\sigma\eta)^{\mathrm{n}}\frac{\mathrm{e}^{(\sigma\eta)^2/2}}{\sigma\sqrt{2\pi}}\mathrm{e}^{-\mathrm{t}\eta}\mathrm{I}_{\mathrm{n}-1}(\mathrm{x};-\eta,-\frac{1}{\sigma},-\sigma\eta)$$

$$\mathrm{P}(\mathrm{Z}-\sum_{\mathrm{i}=1}^{\mathrm{n}}\xi_{\mathrm{i}}\geq \mathrm{x})=(\sigma\eta)^{\mathrm{n}}\frac{\mathrm{e}^{(\sigma\eta)^2/2}}{\sigma\sqrt{2\pi}}\mathrm{e}^{-\mathrm{t}\eta}\mathrm{I}_{\mathrm{n}-1}(\mathrm{x};\eta,\frac{1}{\sigma},-\sigma\eta)$$

Proof. Case 1. The densities of $\mathrm{Z}+\sum_{\mathrm{i}=1}^{\mathrm{n}}\xi_{\mathrm{i}}$, *and* $\mathrm{Z}-\sum_{\mathrm{i}=1}^{\mathrm{n}}\xi_{\mathrm{i}}$. *We have*

$$\mathrm{f}_{\mathrm{Z}+\sum_{\mathrm{i}=1}^{\mathrm{n}}\xi_{\mathrm{i}}}(\mathrm{t})=\int_{-\infty}^{\infty}\mathrm{f}_{\sum_{\mathrm{i}=1}^{\mathrm{n}}\xi_{\mathrm{i}}}(\mathrm{t}-\mathrm{x})\mathrm{f}_{\mathrm{Z}}(\mathrm{x})\mathrm{dx}$$

$$=\mathrm{e}^{-\mathrm{t}\eta}(\eta^{\mathrm{n}})\int_{-\infty}\mathrm{t}\frac{\mathrm{e}^{\mathrm{x}\eta}(\mathrm{t}-\mathrm{x})^{\mathrm{n}-1}}{(\mathrm{n}-1)!}\frac{1}{\sigma\sqrt{2\pi}}\mathrm{e}^{-\mathrm{x}^2/(2\sigma^2)}\mathrm{dx}$$

$$=\mathrm{e}^{-\mathrm{t}\eta}(\eta^{\mathrm{n}})\mathrm{e}^{(\sigma\eta)^2/(2)}\int_{-\infty}\mathrm{t}\frac{(\mathrm{t}-\mathrm{x})^{\mathrm{n}-1}}{(\mathrm{n}-1)!}\frac{1}{\sigma\sqrt{2\pi}}\mathrm{e}^{-(\mathrm{x}-\sigma^2\eta)^2/(2\sigma^2)}\mathrm{dx}$$

Letting $\mathrm{y}=(\mathrm{x}-\sigma^2\eta)/\sigma$ *yields*

$$\mathrm{f}_{\mathrm{Z}+\sum_{\mathrm{i}=1}^{\mathrm{n}}\xi_{\mathrm{i}}}(\mathrm{t})=\mathrm{e}^{-\mathrm{t}\eta}(\eta^{\mathrm{n}})\mathrm{e}^{(\sigma\eta)^2/(2)}\sigma^{\mathrm{n}-1}$$

$$\times\int_{-\infty}^{\mathrm{t}/\sigma-\sigma\eta}\frac{(\mathrm{t}/\sigma-\mathrm{y}-\sigma\eta)^{\mathrm{n}-1}}{(\mathrm{n}-1)!}\frac{1}{\sqrt{2\pi}}\mathrm{e}^{-\mathrm{y}^2/2}\mathrm{dy}$$

$$=\frac{\mathrm{e}^{(\sigma\eta)^2/2}}{\sqrt{2\pi}}(\sigma^{\mathrm{n}-1}\eta^{\mathrm{n}})\mathrm{e}^{-\mathrm{t}\eta}\mathrm{Hh}_{\mathrm{n}-1}(-\mathrm{t}/\sigma+\sigma\eta)$$

because $(1/(\mathrm{n}-1)!)\int_{-\infty}\mathrm{a}(\mathrm{a}-\mathrm{y})^{\mathrm{n}-1}\mathrm{e}^{-\mathrm{y}^2/2}\mathrm{dy}=\mathrm{Hh}_{\mathrm{n}-1}(\mathrm{a})$. *The derivation of* $\mathrm{f}_{\mathrm{Z}+\sum_{\mathrm{i}=1}^{\mathrm{n}}\xi_{\mathrm{i}}}(\mathrm{t})$ *is similar.*

Case 2. $\mathrm{P}(\mathrm{Z}+\sum_{\mathrm{i}=1}^{\mathrm{n}}\xi_{\mathrm{i}}\geq \mathrm{x})$ *and* $\mathrm{P}(\mathrm{Z}-\sum_{\mathrm{i}=1}^{\mathrm{n}}\xi_{\mathrm{i}}\geq \mathrm{x})$. *From (B9), it is clear that*

$$\mathrm{P}(\mathrm{Z}+\sum_{\mathrm{i}=1}^{\mathrm{n}}\xi_{\mathrm{i}}\geq \mathrm{x})=\frac{(\sigma\eta)^{\mathrm{n}}\mathrm{e}^{(\sigma\eta)^2/2}}{\sigma\sqrt{2\pi}}\int_{\mathrm{x}}^{\infty}\mathrm{e}^{(-\mathrm{i}\eta)}\mathrm{Hh}_{\mathrm{n}-1}(-\frac{\mathrm{t}}{\sigma}+\sigma\eta)\mathrm{dt}$$

$$= \frac{(\sigma\eta)^n e^{(\sigma\eta)^2/2}}{\sigma\sqrt{2\pi}} I_{n-1}(x; -\eta, -\frac{1}{\sigma}, -\sigma\eta)dt$$

by (B6). We can compute $P(Z - \sum_{i=1}^{n} \xi_i \geq x)$ *similarly.*

Theorem C.1. *Theorem With* $\pi_n := P(N(t) = n) = e^{-\lambda T}(\lambda T)^n/n!$ *and* I_n *in Proposition* **??**. *, we have*

$$P(Z(T) \geq a) = \frac{e^{(\sigma\eta_1)^2 T/2}}{\sigma\sqrt{2\pi T}} \sum_{n=1}^{\infty} \pi_n \sum_{k=1}^{n} P_{n,k}(\sigma\sqrt{T}\eta_1)^k \times I_{k-1}(a - \mu T; -\eta_1, -\frac{1}{\sigma\sqrt{T}}, -\sigma\eta_1\sqrt{T})$$

$$+\frac{e^{(\sigma\eta_2)^2 T/2}}{\sigma\sqrt{2\pi T}} \sum_{n=1}^{\infty} \pi_n \sum_{k=1}^{n} Q_{n,k}(\sigma\sqrt{T}\eta_2)^k$$

$$\times I_{k-1}(a - \mu T; \eta_2, \frac{1}{\sigma\sqrt{T}}, -\sigma\eta_2\sqrt{T})$$

$$+\pi_0\phi(-\frac{a - \mu T}{\sigma\sqrt{T}})$$

Proof by the decomposition (B2)

$$P(Z(T) \geq a) = \sum_{n=0}^{\infty} \pi_n P(\mu T + \sigma\sqrt{T}Z + \sum_{j=1}^{n} Y_j \geq a)$$

$$= \pi_0 P(\mu T + \sigma\sqrt{T}Z \geq a)$$

$$+\sum_{n=1}^{\infty} \pi_n \sum_{k=1}^{n} P_{n,k} P(\mu T + \sigma\sqrt{T}Z + \sum_{j=1}^{n} \xi_j^+ \geq a)$$

$$+\sum_{n=1}^{\infty} \pi_n \sum_{k=1}^{n} Q_{n,k} P(\mu T + \sigma\sqrt{T}Z - \sum_{j=1}^{n} \xi_j^- \geq a)$$

The result now follows via (B11) and (B12) for $\eta_1 > 1$ *and* $\eta_2 > 0$.

APPENDIX D
DERIVATION OF THE Υ FUNCTION

We first decompose the sum of the double exponential random variables.

The memoryless property of exponential random variables yields $(\xi^+ - \xi^- | \xi^+ > \xi^-) =^{\mathrm{d}} \xi^+$ and $(\xi^+ - \xi^- | \xi^+ < \xi^-) =^{\mathrm{d}} -\xi^-$, thus leading to the conclusion that

$$\xi^+ - \xi^- = \left\{ \begin{array}{cc} \xi^+ & \text{with probability } \eta_2/(\eta_1+\eta_2) \\ -\xi^- & \text{with probability } \eta_1/(\eta_1+\eta_2) \end{array} \right\}.$$

because the probabilities of the events $\xi^+ > \xi^-$ and $\xi^+ < \xi^-$ are $\eta_2/(\eta_1+\eta_2)$ and $\eta_1/(\eta_1+\eta_2)$, respectively. The following proposition extends (B.1.)

Proposition B.1. For every $\mathrm{n} \geq 1$, we have the following decomposition

$$\sum_{\mathrm{i}=1}^{\mathrm{n}} \mathrm{Y_i} =^{\mathrm{d}} \left\{ \begin{array}{cc} \sum_{\mathrm{i}=1}^{\mathrm{k}} \xi_{\mathrm{i}}^+ & \text{with probability } \mathrm{P_{n,k}}, \mathrm{k} = 1, 2, ..., \mathrm{n} \\ -\sum_{\mathrm{i}=1}^{\mathrm{k}} \xi_{\mathrm{i}}^- & \text{with probability } \mathrm{Q_{n,k}}, \mathrm{k} = 1, 2, ..., \mathrm{n} \end{array} \right\}.$$

where $\mathrm{P_{n,k}}$ and $\mathrm{Q_{n,k}}$ are given by

$$\mathrm{P_{n,k}} = \sum_{\mathrm{i}=\mathrm{k}}^{\mathrm{n}-1} \binom{\mathrm{n}-\mathrm{k}-1}{\mathrm{i}-\mathrm{k}} \binom{\mathrm{n}}{\mathrm{i}} (\frac{\eta_1}{\eta_1+\eta_2})^{\mathrm{i}-\mathrm{k}} (\frac{\eta_2}{\eta_1+\eta_2})^{\mathrm{n}-\mathrm{i}} \mathrm{p}^{\mathrm{i}} \mathrm{q}^{\mathrm{n}-\mathrm{i}}$$

$$1 \leq \mathrm{k} \leq \mathrm{n}-1$$

$$\mathrm{Q_{n,k}} = \sum_{\mathrm{i}=\mathrm{k}}^{\mathrm{n}-1} \binom{\mathrm{n}-\mathrm{k}-1}{\mathrm{i}-\mathrm{k}} \binom{\mathrm{n}}{\mathrm{i}} (\frac{\eta_1}{\eta_1+\eta_2})^{\mathrm{n}-\mathrm{i}} (\frac{\eta_2}{\eta_1+\eta_2})^{\mathrm{i}-\mathrm{k}} \mathrm{p}^{\mathrm{n}-\mathrm{i}} \mathrm{q}^{\mathrm{i}}$$

$$1 \leq \mathrm{k} \leq \mathrm{n}-1, \mathrm{P_{n,n}} = \mathrm{p^n}, \mathrm{Q_{n,n}} = \mathrm{q^n}$$

and $\binom{0}{0}$ is defined to be one. Hence ξ_{i}^+ and ξ_{i}^- are i.i.d. exponential random variables with rates η_1 and η_2, respectively.

As a key step in deriving closed-form solutions for call and put options, this proposition indicates that the sum of the i.i.d. double exponential random variable can be written, in distribution, as a randomly mixed gamma random variable. To prove Proposition B.1, the following lemma is needed.

Lemma B.1.

$$\sum_{i=1}^{n} \xi_i^{+} - \sum_{i=1}^{n} \xi_i^{-}$$

$$=^{d} \left\{ \begin{array}{cc} \sum_{i=1}^{k} \xi_i & \text{with probability } \binom{n-k+m-1}{m-1}(\frac{\eta_1}{\eta_1+\eta_2})^{n-k}(\frac{\eta_2}{\eta_1+\eta_2})^{m}, k = 1, ..., n \\ -\sum_{i=1}^{l} \xi_i & \text{with probability } \binom{n-l+m-1}{n-1}(\frac{\eta_1}{\eta_1+\eta_2})^{n}(\frac{\eta_2}{\eta_1+\eta_2})^{m-l}, l = 1, ..., m \end{array} \right\}.$$

We prove it by introducing the random variables $A(n, m) = \sum_{i=1}^{n} \xi_i - sum_{j=1}^{m} \tilde{\xi}_j$ Then

$$A(n, m) =^{d} \left\{ \begin{array}{cc} A(n-1, m-1) + \xi^{+} & \text{with probability } \eta_2/(\eta_1 + \eta_2) \\ A(n-1, m-1) - \xi^{-} & \text{with probability } \eta_1/(\eta_1 + \eta_2) \end{array} \right\}.$$

$$=^{d} \left\{ \begin{array}{cc} A(n, m-1) & \text{with probability } \eta_2/(\eta_1 + \eta_2) \\ A(n-1, m) & \text{with probability } \eta_1/(\eta_1 + \eta_2) \end{array} \right\}.$$

via B.1.. Now suppose horizontal axis that are representing the number of $\{\zeta_i^{+}\}$ and vertical axis representing the number of $\{\zeta_i^{-}\}$. Suppose we have a random walk on the integer lattice points. Starting from any point $(n, m), n, m \geq 1$, the random walk goes either one step to the left with probability $\eta_1/(\eta_1 + \eta_2)$ or one step down with probability $\eta_2/(\eta_1 + \eta_2)$, and the random walks stops once it reaches the horizontal or vertical axis. For any path from (n,m) to (k,0) , $1 \geq k \geq n$, it must reach (k,1) first before it makes a final move to (k,0). Furthermore, all the paths going from (n,m) to (k,1) must have exactly n-k lefts and m-1 downs, whence the total number of such paths is $\binom{n-k+m-1}{m-1}$. Similarly the total number of paths from (n,m) to (0,l) , $1 \geq l \geq m$, is $\binom{n-l+m-1}{n-1}$. Thus

$$A(n,m) =^d \left\{ \begin{array}{ll} \sum_{i=1}^{k} \xi_i & \text{with probability } \binom{n-k+m-1}{m-1}(\frac{\eta_1}{\eta_1+\eta_2})^{n-k}(\frac{\eta_2}{\eta_1+\eta_2})^m, k = 1, ..., n \\ -\sum_{i=1}^{l} \xi_i & \text{with probability } \binom{n-l+m-1}{n-1}(\frac{\eta_1}{\eta_1+\eta_2})^{n}(\frac{\eta_2}{\eta_1+\eta_2})^{m-l}, l = 1, ..., m \end{array} \right\}.$$

and the lemma is proven.

Now, let's prove the proposition B.1. By the same analogy used in Lemma B.1 to compute probability $P_{n,m}, 1 \geq k \geq n$, the probability weight assigned to $\sum_{i=1}^{k} \xi_i^+$ when we decompose $\sum_{i=1}^{k} Y_i$, it is equivalent to consider the probability of the random walk ever reach (k,0) starting from the point (i,n-i) being $\binom{n}{i} p^i q^{n-i}$. Note that the point (k,0) can only be reached from point (i,n-i) such that $k \geq i \geq n-1$, because the random walk can only go left or down, and stops once it reaches the horizontal axis. Therefore, for $1 \geq k \geq n-1$, (B3) leads to

$$P_{n,k} = \sum_{i=k} n - 1P(\text{goingfrom}(i, n-i)\text{to}(k, 0)).P(\text{startingfrom}(i, n-i))$$

$$= \sum_{i=k}^{n-1} \binom{i + (n-i) - k - 1}{(n-i) - 1} \binom{n}{i} (\frac{\eta_1}{\eta_1 + \eta_2})^{i-k} (\frac{\eta_2}{\eta_1 + \eta_2})^{n-i} p^i q^{n-i}$$

$$= \sum_{i=k}^{n-1} \binom{n-k-1}{n-i-1} \binom{n}{i} (\frac{\eta_1}{\eta_1 + \eta_2})^{i-k} (\frac{\eta_2}{\eta_1 + \eta_2})^{n-i} p^i q^{n-i}$$

$$= \sum_{i=k}^{n-1} \binom{n-k-1}{i-k} \binom{n}{i} (\frac{\eta_1}{\eta_1 + \eta_2})^{i-k} (\frac{\eta_2}{\eta_1 + \eta_2})^{n-i} p^i q^{n-i}$$

Of course $P_{n,n} = p^n$. Similarly, we can compute $Q_{n,k}$:

$$Q_{n,k} = \sum_{i=k} n - 1P(\text{goingfrom}(n-i, i)\text{to}(0, k)).P(\text{startingfrom}(n-i, i))$$

$$= \sum_{i=k}^{n-1} \binom{i + (n-i) - k - 1}{(n-i) - 1} \binom{n}{n-i} (\frac{\eta_1}{\eta_1 + \eta_2})^{n-i} (\frac{\eta_2}{\eta_1 + \eta_2})^{i-k} p^{n-i} q^i$$

$$= \sum_{i=k}^{n-1} \binom{n-k-1}{i-k} \binom{n}{i} (\frac{\eta_1}{\eta_1+\eta_2})^{n-i} (\frac{\eta_2}{\eta_1+\eta_2})^{i-k} p^{n-i} q^i$$

with $Q_{n,n} = q^n$. Incidentally, we have also got $\sum k = 1n(P_{n,k} + Q_{n,k}) = 1$

B.2. Let's develop now the results on Hh functions. First of all, note that $Hh_n(x) \to 0$, as $x \to \infty$, for $n \geq -1$; and $Hh_n(x) \to \infty$, as $x \to -\infty$, for $n \geq -1$; and $Hh_0(x) = \sqrt{2\pi}\phi(-x) \to \sqrt{2\pi}$, as $x \to -\infty$. Also, for every $n \geq -1$, as $x \to \infty$,

$$\lim Hh_n(x)/\{\frac{1}{x^{n+1}} e^{-\frac{x^2}{2}}\} = 1$$

and as $x \to \infty$

$$Hh_n(x) = O(|x|^n)$$

Here (B4) is clearly true for $n = -1$, while for $n \geq 0$ note that as $x \to_{\infty}$,

$$Hh_n(x) = \frac{1}{n!} \int_x \infty (t-x)^n e^{-\frac{t^2}{2}} dt$$

$$\leq \frac{2^n}{n!} \int_{-\infty}^{\infty} |t|^n e^{-t^2} 2dt + \frac{2^n}{n!} \int -\infty\infty |x|^n e^{-t^2} 2dt = O(|x|^n)$$

For option pricing it is important to evaluate the integral $I_n(c;\alpha;\beta;\delta)$,

$$I_n(c;\alpha;\beta;\delta) = \int_c \infty e^{\alpha x} Hh_n(\beta x - \delta) dx, n \geq 0$$

for arbitrary constants α, c and β.

REFERENCES

Garfinkle, David, Horowitz, Gary T, and Strominger, Andrew. "Charged black holes in string theory." *Physical Review D* 43 (1991).10: 3140.

Green, Karen L. "A wrinkle in time." *comiXology* (2008).

L'engle, Madeleine. *A Wrinkle in Time: 50th Anniversary Commemorative Edition*, vol. 1. Macmillan, 2012.

Strickler, Howard D, Rosenberg, Philip S, Devesa, Susan S, Hertel, Joan, Fraumeni Jr, Joseph F, and Goedert, James J. "Contamination of poliovirus vaccines with simian virus 40 (1955-1963) and subsequent cancer rates." *Jama* 279 (1998).4: 292–295.

BIOGRAPHICAL SKETCH

This section is where your biographical sketch is typed in the bio.tex file. It should be in third person, past tense. Do not put personal details such as your birthday in the file. Again, to make a full paragraph you must write at least three sentences.