MATHEMATICS AND APPLICATIONS OF FEYNMAN DIAGRAMS

AN ABSTRACT

SUBMITTED ON THE FOURTH DAY OF MAY 2021

TO THE DEPARTMENT OF MATHEMATICS

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

OF THE SCHOOL OF SCIENCE AND ENGINEERING

OF TULANE UNIVERSITY

FOR THE DEGREE

OF

MASTER OF SCIENCE

BY

JUNHYUP SUNG

APPROVED: _____

VICTOR H. MOLL, PH.D.

CHAIR

_____

KYLE K. ZHAO, PH.D.

_____

SCOTT MCKINLEY, PH.D.

Abstract

The Feynman diagrams have become a highly valued tool for complex calculations and understanding the physics of elementary particles within the framework of quantum field theory. In this thesis, we present an overview on constructing and utilizing Feynman diagrams in quantum field theory along with an overview of quantum field theory itself. We begin with a review of prerequisite topics then progress to discussing symmetries using Lie groups, algebras, and representation theory. We then use the representations of the Lorentz group to derive the fields in a classical context then proceed with quantization to create the corresponding quantum fields while providing a thorough analysis of each quantum field. Path integrals are constructed for each quantum field by deriving their propagators then the formulas for scattering are derived in the context of quantum field theory. Quantum symmetries are briefly explored with the intention of quantizing classical results such as Noether's theorem. Then we construct interacting quantum field theories and introduce the Feynman diagrams and Feynman rules for different interaction theories and provide examples and applications of the Feynman diagrams. The physics behind the diagrams is carefully analyzed and interpreted. Finally, we conclude this thesis with a summary of what we have covered along with possible routes of study after mastering the contents of this thesis that will lead to current research topics.

MATHEMATICS AND APPLICATIONS OF FEYNMAN DIAGRAMS

A THESIS

SUBMITTED ON THE FOURTH DAY OF MAY 2021

TO THE DEPARTMENT OF MATHEMATICS

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS

OF THE SCHOOL OF SCIENCE AND ENGINEERING

OF TULANE UNIVERSITY

FOR THE DEGREE

OF

MASTER OF SCIENCE

BY

JUNHYUP SUNG

APPROVED:

VICTOR H. MOLL, PH.D.

CHAIR

KYLE K. ZHAO, PH.D.

SCOTT MCKINLEY, PH.D.

Acknowledgements

I would like to start this thesis by showing my upmost gratitude towards my thesis advisor Dr. Victor H. Moll for his guidance and willingness in allowing me to cover a topic in which we both initially had little experience in. It was both a very fun and exciting process in learning about a topic alongside a mentor and I will forever be grateful for the lessons Dr. Moll instilled in me especially the importance of being concrete and honest with one's own knowledge. Along with this, I want to thank Dr. Moll for taking me on as his student and providing me with the resources and encouragement to succeed in writing this thesis.

I would like to thank my thesis committee member Dr. Kyle Zhao for both being a part of my committee but also for being my analysis professor. He presented the subject of mathematical analysis, a topic that I considered to be one of my weak points, in such a clear and lucid format that his method of approaching analysis proved to be extremely useful in writing this thesis. I have definitely become better at mathematics as a result of his class and his lessons will be something I will continue to use throughout my career.

I would also like to thank my other committee member Dr. Scott McKinley for both being a part of my committee as well but also for encouraging me to apply for the $4+1$ program. The amount of positive life changes and the amount of knowledge I have learned during the $4+1$ program is something I could have never dreamed about and I want to thank Dr. McKinley for steering me in this positive direction.

There have been many professors in the mathematics department and in other

departments who have influenced my life in so many positive ways. However, I would like to deeply thank Dr. Norbert Riedel for not just the amazing chats we would have, but for being the one who gave me my passion for mathematics and physics and ultimately inspired me to become a mathematician. I will forever be grateful to Dr. Riedel for helping me to find my passion towards mathematics and changing my life for the better.

Lastly, I would like to thank my all of friends and classmates for always encouraging me to do my best and giving me support whenever I struggled. Many of my friends know of my initial struggles with mathematics and I want to thank all of them for the kind words and encouragement during my journey thus far. I could not have become the person I am today without them and I will forever be grateful for the support they gave me over the years.

List of Figures

Table of Contents

Chapter 1

Introduction

In the early 20^{th} century, intense research into the property of light and matter at the smallest scales revealed that nature did not obey the laws of classical physics that many scientists had come to accept as the paradigm for modeling physical phenomena. Various contradictions and paradoxes arose in which current scientific theories of the time could not explain nor resolve [1]. In a time period that spanned several decades with the contributions of thousands of scientists all over the world, a new physical theory, radically different from all known scientific theories of the time, was slowly developed that could quantitatively explain model nature at the smallest scales. This new physical theory is called quantum mechanics and despite its controversial philosophical implications, it quickly proved to be remarkably accurate in describing physical phenomena at atomic scales. It was able to quantitatively explain the hydrogen spectrum, explain the structure of atoms with great accuracy, and much more. Quantum mechanics very quickly established itself as one of the most successful scientific theories in history. In fact, millions of experiments have confirmed the validity of quantum mechanics with an astonishing degree of accuracy.

In a radically opposite situation during the same time, another major scientific development occurred by virtually one person in a paper published on September 26, 1905 [1]. In this landmark paper titled “On the Electrodynamics of Moving Bodies”,

a young theoretical physicist named Dr. Albert Einstein formally established the theory of special relativity which reconciled the mathematical asymmetries of classical electromagnetism and corrected Newtonian mechanics to account for bodies that move at speeds close to or at the speed of light. Immediately afterwards, it was soon realized by Einstein and other notable scientists and mathematicians, such as Hermann Minkowski and David Hilbert, that the theory of relativity was a geometrical theory that dealt with the intimate relationship of space and time. This has tremendous implications on all existing and future theories left to be discovered in the fact that the laws of nature must obey relativity [2]. Then Albert Einstein developed in a span of ten years a general theory of relativity that incorporated gravity where the theory of special relativity he had initially developed was the special case. Much like quantum mechanics, special and general relativity proved to be extremely accurate theories that could explain the physics of very large bodies and bodies moving close to or at the speed of light. Experiments have shown that the predications and results of special relativity and general relativity have been correct to an extremely accurate degree and relativity established itself as one of the most accurate scientific theories of all time as well.

With the creation of two of the most powerful theories that form the basis of modern physics, it was only natural that scientists would try to fuse quantum mechanics and special relativity into a new physical theory that would provide the tools and framework for studying an even larger range of physical phenomena with results just as, if not more, accurate from any of the two theories alone. This proved to be an extremely difficult task because many of the attempts at fusing both theories ran into contradictions or inconsistencies that violated one or both of the fundamental tenets of the two theories. An example of this was the first relativistic quantum wave equation called the *Klein-Gordon equation*

$$-\hbar^2 \frac{\partial^2}{\partial t^2}\psi(\mathbf{x},t) = \Big(-\hbar^2 c^2 \nabla^2 + m^2 c^4 \Big)\psi(\mathbf{x},t),$$

which was formulated by replacing the energy E and momentum $\mathbf{p}$ in the relativistic

energy equation

$$E^2 = \mathbf{p}^2c^2 + m^2c^4,$$

with their quantum operator equivalents. Interesting, this equation was the first equation that Erwin Schrödinger formulated but he discarded this equation in favor of the nonrelativistic *Schrödinger equation* because this equation did not account for the spin of the electron. The Klein-Gordon equation, although manifestly covariant, is not consistent with quantum mechanics because probability is not conserved and predicts negative energy states. Many scientists were not sure if it was possible to combine both theories until British theoretical physicist Paul Dirac successfully derived and presented the beautiful *Dirac equation*

$$(-i\gamma^\mu\partial_\mu + m)\psi = 0,$$

which was a covariant quantum mechanical equation created to model the electron that seemed to successfully combine special relativity and quantum mechanics [1]. The Dirac equation gave scientists the promised highly accurate results. However, there were several inconsistencies that arose with one of the most glaring inconsistency being the existence of negative energy states like the Klein-Gordon equation. In a rather *ad hoc* manner, Dirac used the Pauli exclusion principle along with his remarkable physical intuition to describe these negative energy states as a type of particle with the same properties as the electron except that it possessed a positive electric charge and would annihilate instantly when it interacted with a regular, negatively charged electron. This was the first time that the concept of*antimatter* was presented and was experimentally verified shortly afterwards which contributed to the Nobel committee awarding him, along with Erwin Schrödinger, the Nobel Prize in Physics [1].

Despite the various attempts at reconciling any problems with the fusion of special relativity and quantum mechanics, it was clear that another physical theory was needed to successfully combine and reconcile the relativity and quantum

mechanics. After World War II, physicists all over the world contributed to this development and eventually a new physical framework emerged that combined classical field theory, special relativity, and quantum mechanics into a new theory called *quantum field theory*. Quantum field theory proved to be extremely successful as well but once again, serious flaws emerged in the form of mathematical infinities which gave nonsensical results. Eventually, various mathematical techniques collectively called *renormalization* were created that removed the infinities that plagued quantum field theory [1]. Unsurprising, the mathematics and physical interpretations of quantum field theory proved to be conceptually difficult to understand and explain. However, a brilliant theoretical physicist named Richard Feynman developed a computational tool called the *Feynman diagrams* which allowed one to make calculations in quantum field theory by using a series of diagrams with a set of rules that made calculations far easier and streamlined. Very quickly however, the Feynman diagrams proved to provide a deeper insight into the underlying physics at hand in quantum field theory and combined with their computational power, the Feynman diagrams essentially became an indispensable part of the subject of quantum field theory.

The goal of this thesis to show how the Feynman diagrams are derived and utilized in particle physics and in quantum field theory in general. We will formulate the foundations of quantum field theory from the ground up by first presenting a brief overview on the key concepts and results from the theory of special relativity, classical field theory, and nonrelativistic quantum mechanics. Then we will cover the mathematical basis of quantum field theory by delving into the study of Lie groups, Lie algebras, and their representations which will naturally lead to the discussion of the representations of the Lorentz group. We will cover how symmetry transformations are expressed mathematically and discuss their relationship to quantum field theory. We will then provide a thorough derivation and analysis of the types of quantum fields that are studied in practice then construct each field's respective path integrals. After a brief overview of scattering theory in quantum

field theory along with a short digression on the symmetries of quantum fields, we will begin our study of interactions which will naturally lead to constructing our first Feynman diagrams and deriving the Feynman rules for the other fields so that we can construct their respective diagrams. We will include pedagogical examples and real life applications when we construct the Feynman diagrams. The idea of renormalization will be hinted at and briefly covered but will not be fully developed. This will not hinder our discussion of the Feynman diagrams and quantum field theory in general because as we will see in this thesis, the topic of quantum field theory is such a vast and rich subject. We will assume that the reader has a basic knowledge of nonrelativistic quantum mechanics and special relativity along with a familiarity in using Dirac notation.

Chapter 2

Background

2.1 Special Relativity

The Need for Relativity

The special theory of relativity, or special relativity, in its most general sense is the physical theory that explains the structure and relationship of space and time. Originally formulated by Albert Einstein in the groundbreaking paper "On the Electrodynamics of Moving Bodies", the original goal of special relativity was to address the apparent asymmetries of Maxwell's equations when applied to moving bodies by correcting Newtonian mechanics to be compatible with classical electrodynamics [1]. It soon became apparent however that special relativity fundamentally dealt with the relationship of space and time and in particular, its geometry. If we were to imagine all of our physical laws and theories as actors in a play, then special relativity would be considered the background on which the actors perform on. The structure and form of the background ultimately decides how the actors can interact with each other and their surroundings [2]. From this simply analogy, it is clear that relativity plays a central role in physics. Indeed, physicists realized that special relativity was to be applied not just to classical mechanics, but to all of physics including all currently known physical theories and any undiscovered physical theories [2]. In this section, we will present derivations

and explanations of the key concepts from special relativity that will be central in formulating and explaining quantum field theories.

The Postulates of Special Relativity and the Lorentz Transformations

We state the postulates of special relativity according to Griffiths [2]:

Postulates of the Special Theory of Relativity

1. **Principle of Relativity**: The laws of physics apply in all inertial frames.
2. **Invariance of the Speed of Light**: The speed of light in a vacuum is a constant c in all inertial frames, regardless of the motion of the source.

The postulates of special relativity can be used to derive the relativity of simultaneity, time dilation, and length contraction [3]. The relativity of simultaneity says that two events that occur simultaneously in one inertial frame can be seen to occur non-simultaneously in another inertial frame. Time dilation is the phenomenon where a moving clock measures time slower than a clock that is in its rest frame. This implies that the time interval between two events is dependent on the relative velocity between the two inertial frames they occur in. Length contraction is the phenomenon where a moving body experiences a reduction in the dimension parallel to its velocity. This implies that the spatial distance parallel to the velocity between two events in one moving inertial frame can be found to have a different spatial distance in another inertial frame. The type of transformations between two inertial frames S and S' that encodes these three concepts are the *Lorentz transformations*

$$\begin{cases} t' = \gamma(t - \frac{v}{c^2}x), \\ x' = \gamma(x - vt), \\ y' = y, \\ z' = z, \end{cases} \tag{2.1.1}$$

where

$$\gamma = \frac{1}{\sqrt{1 - \frac{v^2}{c^2}}}, \tag{2.1.2}$$

is the *Lorentz factor*. The matrix structure of the Lorentz transformation easily allows us to write the Lorentz transformation as

$$\begin{pmatrix} t' \\ x' \\ y' \\ z' \end{pmatrix} = \begin{pmatrix} \gamma & -\gamma\frac{v}{c^2} & 0 & 0 \\ -\gamma v & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} t \\ x \\ y \\ z \end{pmatrix}, \tag{2.1.3}$$

The above Lorentz transformation matrix Λ is called a *standard boost* and does not represent the most general type of Lorentz transformation since a general Lorentz transformation Λ involves some rotations as well. A *boost*, or *pure boost*, Λ_{B} is a type of Lorentz transformation that leaves corresponding axes parallel without any rotation [4]. A *rotation*, or *pure rotation*, Λ_{R} is a type of Lorentz transformation that leaves the temporal component invariant while the spatial components are mixed by rotations. Since there are three spatial axes, there are three rotations and three boosts. Any Lorentz transformation Λ can be written as the product of a boost followed by a suitable rotation $\Lambda = \Lambda_R\Lambda_B$ [4].

Now if light travels from one spacetime point (t_0, x_0, y_0, z_0) to another point (t_1, x_1, y_1, z_1) in an inertial frame S, the distance traveled would be

$$c(t_1 - t_0) = \sqrt{(x_1 - x_0)^2 + (y_1 - y_0)^2 + (z_1 - z_0)^2},$$

or equivalently,

$$-c^2(t_1 - t_0)^2 + (x_1 - x_0)^2 + (y_1 - y_0)^2 + (z_1 - z_0)^2 = 0.$$

In another inertial frame S', the distance light travels would be [3]

$$-c^2(t'_1 - t'_0)^2 + (x'_1 - x'_0)^2 + (y'_1 - y'_0)^2 + (z'_1 - z'_0)^2 = 0.$$

From the second postulate of relativity, the invariance of the speed of light in all inertial frames implies that

$$\begin{aligned} &-c^2(t_1 - t_0)^2 + (x_1 - x_0)^2 + (y_1 - y_0)^2 + (z_1 - z_0)^2 \\ &= -c^2(t'_1 - t'_0)^2 + (x'_1 - x'_0)^2 + (y'_1 - y'_0)^2 + (z'_1 - z'_0)^2. \end{aligned}$$

This naturally leads to the concept of the *spacetime interval* between two events,

$$(\Delta s)^2 = -c^2(\Delta t)^2 + (\Delta x)^2 + (\Delta y)^2 + (\Delta z)^2, \tag{2.1.4}$$

and if we set the spacetime coordinates in frame S' to coincide with the origin, we get

$$s^2 = -c^2t^2 + x^2 + y^2 + z^2. \tag{2.1.5}$$

Equation (2.1.5) is a quadratic form and given the matrix structure of the Lorentz transformations, we can construct a mathematical vector that combines the three spatial coordinates with time to create the *four-vector*

$$\bar{x} = (ct, x, y, z). \tag{2.1.6}$$

The bar notation on the four-vector is added to distinguish the spatial x-coordinate but we will abandon this bar notation shortly in favor of the more commonly used notation. Since the dimensions of ct is length, we introduce the standard notation of four-vectors by labeling $x^0 = ct$, $x^1 = x$, $x^2 = y$, and $x^3 = z$ then

$$x^\mu = (x^0, x^1, x^2, x^3), \tag{2.1.7}$$

where $\mu = 0, 1, 2, 3$. It is standard for Greek indices, such as μ and ν, to have values $\mu = 0, 1, 2, 3$ while Latin indices, such as i and j, to have values $i = 1, 2, 3$. In this thesis, we will adopt this convention as well. If we let $\beta = v/c$, the standard Lorentz transformation takes on the simpler form

$$\begin{pmatrix} x^{0\prime} \\ x^{1\prime} \\ x^{2\prime} \\ x^{3\prime} \end{pmatrix} = \begin{pmatrix} \gamma & -\gamma\beta & 0 & 0 \\ -\gamma\beta & \gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix}. \tag{2.1.8}$$

If we let $\beta = \cosh\eta$ and use the hyperbolic relation

$$\sinh^2\eta - \cosh^2\eta = 1, \tag{2.1.9}$$

the Lorentz factor becomes

$$\cosh^2\eta = \frac{1}{\sqrt{1 - \tanh^2}}, \tag{2.1.10}$$

and the standard Lorentz transformation becomes

$$\begin{pmatrix} x^{0\prime} \\ x^{1\prime} \\ x^{2\prime} \\ x^{3\prime} \end{pmatrix} = \begin{pmatrix} \cosh\eta & -\sinh\eta & 0 & 0 \\ -\sinh\eta & \cosh\eta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix}, \tag{2.1.11}$$

where the parameter η is called the *rapidity* [2]. The hyperbolic form of the Lorentz transformation (2.1.11) can be interpreted as a hyperbolic rotation in four dimensions which reflects the geometric nature of the theory of special relativity

[4]. The four-vector notation allows us to write the quadratic spacetime interval as

$$\begin{aligned}(x^\mu)^T \cdot x^\mu &= -(x^0)^2 + (x^1)^2 + (x^1)^2 + (x^3)^2 \\ &= \begin{pmatrix} x^0 & x^1 & x^2 & x^3 \end{pmatrix} \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix} \\ &= \begin{pmatrix} x^0 & x^1 & x^2 & x^3 \end{pmatrix} g_{\mu\nu} \begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix} \\ &= x^\mu g_{\mu\nu} x^\nu,\end{aligned}$$

where $g^{\mu\nu} = g_{\mu\nu}$ is the *Minkowski metric*. The *metric signature* of $g^{\mu\nu}$ is $(-,+,+,+)$ but in other literature it can also be $(+,-,-,-)$. We will adopt the signature $(-,+,+,+)$ for this thesis. We see that

$$\begin{aligned}g_{\mu\nu} x^\nu &= \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix} \\ &= (-x_0, x_1, x_2, x_3) \\ &= x_\mu, \end{aligned} \tag{2.1.12}$$

so we introduce lowering a Lorentz index as

$$x_\mu = \sum_{\nu=0}^{3} g_{\mu\nu} x^\mu = (-x_0, x_1, x_2, x_3), \tag{2.1.13}$$

and raising a Lorentz index as

$$x^{\mu} = \sum_{\nu=0}^{3} g^{\mu\nu} x_{\mu} = (x^0, x^1, x^2, x^3). \tag{2.1.14}$$

We call the four-vectors x_{μ} and x^{μ} *covariant* and *contravariant* four-vectors, respectively, thus raising a Lorentz index is equivalent to making a sign change in the temporal component while leaving the spatial components invariant [2]

$$(x_0, x_1, x_2, x_3) = (-x^0, x^1, x^2, x^3). \tag{2.1.15}$$

The Lorentz transformation matrix can be denoted as $\Lambda^{\mu}{}_{\nu}$ where μ and ν are the row and column numbers, respectively, then a Lorentz transformation can be compactly written as

$$x^{\mu\prime} = \sum_{\nu=0}^{3} \Lambda^{\mu}{}_{\nu} x^{\nu}. \tag{2.1.16}$$

To make equation (2.1.16) even more compact and elegant, we adopt the Einstein summation convention where any pair of repeated indices imply summation over the repeated indices which allows us to remove the summation symbol

$$x^{\mu\prime} = \Lambda^{\mu}{}_{\nu} x^{\nu} \equiv \sum_{\nu=0}^{3} \Lambda^{\mu}{}_{\nu} x^{\nu}. \tag{2.1.17}$$

Any pair of repeated indices must have one superscript and one subscript and are said to *contracted*. With this convention, we see that [5]

$$g^{\mu\nu} g_{\nu\rho} = \delta^{\mu}{}_{\rho}, \tag{2.1.18}$$

where $\delta^{\mu}{}_{\rho}$ is the Kronecker delta and the spacetime interval can now be written as

$$s^2 \equiv x^2 = x^{\mu} x_{\mu} = x_{\mu} x^{\mu} = -(x^0)^2 + \mathbf{x}^2. \tag{2.1.19}$$

A Lorentz transformation leaves the spacetime interval invariant.

Proof. Consider the spacetime interval

$$s^2 = -(x^0)^2 + (x^1)^2 + (x^2)^2 + (x^3)^2.$$

The last three terms on the right hand side is the Euclidean norm in $\mathbb{R}^3$ which is invariant under three dimensional rotations. Without loss of generality, under a standard boost we have

$$\begin{aligned} s^{2\prime} &= -(x^{0\prime})^2 + (x^{1\prime})^2 + (x^{2\prime})^2 + (x^{3\prime})^2 \\ &= -\gamma^2(1-\beta^2)(x^0)^2 + x^1 + x^2 + \gamma^2(1-\beta^2)(x^1)^2 \\ &= -(x^0)^2 + (x^1)^2 + (x^2)^2 + (x^3)^2 \\ &= s^2, \end{aligned}$$

so the interval is invariant under standard boosts. Since any Lorentz transformation consists of a product of boosts and rotations, the spacetime interval s is invariant under Lorentz transformations [4]. □

Now we have a formal definition for a Lorentz transformation [5].

Definition 2.1.1. A **Lorentz transformation** is a linear, homogeneous change of coordinates from one inertial frame to another inertial frame

$$x^{\mu\prime} = \Lambda^{\mu}{}_{\nu} x^{\nu},$$

that leaves invariant the interval

$$x^{\mu} x_{\mu} = -(x^0)^2 + \boldsymbol{x}^2.$$

With our definition, a Lorentz transformation obeys the following

$$g_{\mu\nu}\Lambda^{\mu}{}_{\rho}\Lambda^{\nu}{}_{\sigma} = g_{\rho\sigma} \tag{2.1.20}$$

from which we can identify the inverse Lorentz transformation [5]

$$(\Lambda^{-1})^{\rho}{}_{\nu} = \Lambda_{\nu}{}^{\rho}, \tag{2.1.21}$$

which also implies that

$$g^{\mu\nu}\Lambda^{\rho}{}_{\mu}\Lambda^{\sigma}{}_{\nu} = g_{\rho\sigma}. \tag{2.1.22}$$

We will show later that the set of all Lorentz transformations forms a group.

The form of the spacetime interval very clearly resembles an inner product thus the spacetime interval is also known as the *Minkowski inner product*, or *invariant scalar product*. However, despite the fact that the invariant scalar product has many of the same properties of an inner product, it is not an inner product by definition. To see this, when we first derived the spacetime interval, the interval value was zero or, more formally,

$$x^{\mu}x_{\mu} = 0. \tag{2.1.23}$$

Two events whose spacetime interval is zero are called *lightlike* events and we call the four-vector x^{μ} lightlike as well. If we consider two events where their spatial interval is greater than their temporal interval, we will have $x^{\mu}x_{\mu} > 0$ and are called *spacelike* events and we call the four-vector x^{μ} spacelike as well. Analogously, if two events have a greater temporal interval than their spatial interval, we will have $x^{\mu}x_{\mu} < 0$ and are called *timelike* events and we call the four-vector x^{μ} timelike as well. Formally speaking, the invariant scalar product can have values

$$x^{\mu}x_{\mu} > 0, \tag{2.1.24}$$

$$x^{\mu}x_{\mu} = 0, \tag{2.1.25}$$

$$x^{\mu}x_{\mu} < 0, \tag{2.1.26}$$

thus the invariant scalar product is indefinite which violates the positive definiteness property of the inner product. Thus we have a formal definition.

Definition 2.1.2. The **Minkowski inner product**, **or invariant scalar product**, is a nondegenerate, symmetric bilinear form $[\cdot,\cdot]_{3,1} : \mathbb{R}^4 \times \mathbb{R}^4 \to \mathbb{R}$ defined as

$$[x, y]_{3,1} = x_1y_1 + x_2y_2 + x_3y_3 - x_4y_4,$$

which satisfies for all $x, y, z \in \mathbb{R}^4$ and for all scalars $a \in \mathbb{R}$:

1. Bilinearity

$$[ax + y, z]_{3,1} = a[x, z]_{3,1} + [y, z]_{3,1}$$

2. Symmetry

$$[x, y]_{3,1} = [y, x]_{3,1}$$

3. Non-degenerancy

$$[x, y]_{3,1} = 0;\ \forall y \in \mathbb{R}^4, \text{ then } x = 0$$

By now, it is clear that we are not working in the Euclidean space $\mathbb{R}^4$ because equipping the smooth manifold $\mathbb{R}^4$ with the Minkowski inner product creates a non-Euclidean manifold. The space that we are working in is now defined.

Definition 2.1.3. Minkowski space is the smooth four-dimensional manifold $\mathbb{R}^4$ equipped with the Minkowski inner product. Depending on the metric signature, it is denoted as $\mathbb{R}^{1,3}$ or $\mathbb{R}^{3,1}$ for signatures $(-,+,+,+)$ and $(+,-,-,-)$, respectively.

Our choice of the metric signature was $(-,+,+,+)$ thus we will be using the notation $\mathbb{R}^{1,3}$ for this thesis. All of the quantum fields that we will construct will have their domains in $\mathbb{R}^{1,3}$ so that they are relativistically invariant.

We say that an equation is relativistically invariant if it holds in all inertial frames, which is in accordance with the principle of relativity [4]. Before we end this section, we introduce the four-gradient which will aid us in subsequent chapters

[5]

$$\partial_\mu \equiv \frac{\partial}{\partial x^\mu} = \Big(\frac{\partial}{\partial t}, \nabla\Big), \tag{2.1.27}$$

$$\partial^\mu \equiv \frac{\partial}{\partial x_\mu} = \Big(-\frac{\partial}{\partial t}, \nabla\Big). \tag{2.1.28}$$

We see that

$$\partial^\mu x^\nu = g^{\mu\nu}, \tag{2.1.29}$$

and it can easily be proved that under Lorentz transformations, the four-gradients transform as [5]

$$\overline{\partial}^\mu = \Lambda^\mu{}_\nu \partial^\nu, \tag{2.1.30}$$

$$\overline{\partial}^2 = \partial^2. \tag{2.1.31}$$

2.2 Classical Field Theory

As the name of the topic suggests, fields are crucial in the formulation and understanding of quantum field theory. Field theory is used in the study of objects that form a continuous medium, such as water, air, and the electromagnetic field [6]. A field $\phi(\mathbf{x}, t)$ is a physical quantity that gives a value for any point in spacetime. The type of fields that we will deal with will be based on the representations of the Lorentz group which we will cover later on. For now, we will review the dynamics of fields and how the equations that govern their dynamics are derived.

Recall from Lagrangian mechanics the action functional S which is defined as

$$S = \int_{t_1}^{t_2} dt\, L(q_i, \dot{q}_i), \tag{2.2.1}$$

where $L(q_i, \dot{q}_i)$ is the *Lagrangian* function whose arguments q_i and $\dot{q}_i$ are called the *generalized coordinates* and *generalized velocities*, respectively [4]. The *Euler-*

Lagrange equations of motion are derived by using the *principle of stationary action* to minimize the action S to get

$$\frac{d}{dt}\Big(\frac{\partial L}{\partial \dot{q}_i}\Big) - \frac{\partial L}{\partial q_i} = 0, \tag{2.2.2}$$

where

$$\frac{\partial L}{\partial \dot{q}_i} = p_i, \tag{2.2.3}$$

$$\frac{\partial L}{\partial q_i} = F_i, \tag{2.2.4}$$

are the *generalized momenta* and *generalized forces*, respectively. The transition to Hamiltonian mechanics involves a Legendre transform of the Lagrangian $L(q_i, \dot{q}_i)$ to construct the *Hamiltonian* $H(q_i, \dot{p}_i)$ whose arguments $r_i = (q_i, p_i)$ are called the *canonical coordinates* [7]. The Hamiltonian is

$$H = \sum_i \dot{q}_i \frac{\partial L}{\partial \dot{q}_i} - L = \sum_i p_i \dot{q}_i - L, \tag{2.2.5}$$

where

$$p_i = \frac{\partial L}{\partial \dot{q}_i}, \tag{2.2.6}$$

are the generalized momenta from Lagrangian mechanics but are called the *canonical* or *conjugate momenta* in Hamiltonian mechanics [4]. Then *Hamilton's equations* are

$$\frac{\partial H}{\partial q_i} = -\dot{p}_i, \quad \frac{\partial H}{\partial p_i} = \dot{q}_i, \quad \frac{\partial H}{\partial t} = -\frac{\partial L}{\partial t}. \tag{2.2.7}$$

These formulations of classical mechanics easily generalize to field theory with some minor alterations. Lagrangian and Hamiltonian mechanics deals with systems of discrete particles but fields are continuous systems over $\mathbb{R}^n$. Thus for a set of fields

φ_a, we introduce the *Lagrangian density*

$$\mathcal{L}(\varphi_a, \nabla\varphi_a, \partial\varphi_a/\partial t, \mathbf{x}, t) = \mathcal{L}(\varphi_a, \partial_\mu\varphi_a), \tag{2.2.8}$$

and the *total Lagrangian* L is computed by integrating over all of $\mathbb{R}^3$

$$L = \int d^3x\, \mathcal{L}, \tag{2.2.9}$$

then the action is

$$\mathcal{S} = \int dt\, L = \int dt\, d^3x\, \mathcal{L} = \int d^4x\, \mathcal{L}, \tag{2.2.10}$$

where $d^4x = dt\, d^3x$ is the Lorentz invariant measure. To see that d^4x is Lorentz invariant, let $x^{\mu\prime} = \Lambda^\mu{}_\nu x^\nu$ then $d^4x' = |\det\Lambda|\, d^4x = d^4x$ since $|\det\Lambda| = 1$ [5]. Varying the action and applying the stationary action principle with vanishing boundary conditions at spatial and temporal infinity yields the field-theoretic analogue of the Euler-Lagrange equations

$$\partial_\mu\Big(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi_a)}\Big) - \frac{\partial\mathcal{L}}{\partial\varphi_a} = 0. \tag{2.2.11}$$

Analogously, we define the *Hamiltonian density*

$$\mathcal{H}(\varphi_a, \nabla\varphi_a, \partial\varphi_a/\partial t, \mathbf{x}, t) = \mathcal{H}(\varphi_a, \partial_\mu\varphi_a), \tag{2.2.12}$$

and the *total Hamiltonian* is computed in the same manner

$$H = \int d^3x\, \mathcal{H}. \tag{2.2.13}$$

Deriving the Hamiltonian density is very similar to the discrete case,

$$\mathcal{H} = \sum_i \frac{\partial\mathcal{L}}{\partial(\partial_0\varphi_a)}\varphi_a - \mathcal{L} = \sum_a \Pi_a\dot{\varphi}_a - \mathcal{L}, \tag{2.2.14}$$

where for each field φ_a

$$\Pi_a = \frac{\partial \mathcal{L}}{\partial(\partial_0 \varphi_a)} = \frac{\partial \mathcal{L}}{\partial \dot{\varphi}_a}, \tag{2.2.15}$$

are the *conjugate momenta densities* [5]. Then the field-theoretic analogue of Hamiltonian's equations are

$$\frac{\delta \mathcal{H}}{\delta \Pi_a} = \dot{\varphi}_a, \quad \frac{\delta \mathcal{H}}{\delta \varphi_a} = -\dot{\Pi}_a, \tag{2.2.16}$$

where the operator δ is the variational derivative. It is a common convention to refer to the Lagrangian and Hamiltonian density as the Lagrangian and Hamiltonian, respectively, and in general to drop the term density from all other names. We will adopt this convention in this thesis as well unless specified otherwise. In this thesis, we will place more emphasis on the Lagrangian formulation of field theory. The main advantages of this are that it is manifestly covariant, preserves the symmetries of the equation and naturally leads to the path integral formulation of quantum field theory [8].

2.3 Quantum Mechanics

We will present a brief overview of nonrelativistic quantum mechanics and present the main results that will be frequently used throughout this thesis. As mentioned in the introduction, quantum mechanics is the fundamental theory of physics that studies atomic and subatomic systems. Quantum mechanics differs from classical physics in many ways with the most notable difference being that the ordering of a series of measurements matters in quantum mechanics. A famous example of this is that it is not possible to know the position and momentum of a particle simultaneously which is in stark contrast with classical physics. The massive success of quantum mechanics has led scientists to believe that there is an underlying quantum theory for every classical theory and indeed this has been the case with a famous example being quantum electrodynamics. Like relativity,

it is believed that quantum mechanics is a fundamental aspect of nature thus all theories should obey quantum mechanics at the right size ranges. For this section and thesis, the reader is assumed to possess knowledge of and flexibility in using Dirac notation.

Postulates of Quantum Mechanics

We present the postulates of quantum mechanics according to Cohen-Tannoudji [9].

Postulates of Quantum Mechanics

1. At a fixed time t_0, the state of an isolated physical system is defined by specifying a ket $|\psi(t_0)\rangle$ belonging to a Hilbert space $\mathcal{H}$ called the state space.

2. Every measurable physical quantity A is described by a Hermitian operator $\hat{A}$ acting in $\mathcal{H}$; this operator is an observable.

3. The result of the measurement of a physical quantity A is one of the eigenvalues of the corresponding observable $\hat{A}$.

4. When the physical quantity A is measured on a system in the normalized state $|\psi\rangle$, the probability $P(a_n)$ of obtaining an eigenvalue (a_n for discrete and α for continuous) of $\hat{A}$ is:

$$P(a_n) = |\langle\phi_n|\psi\rangle|^2, \qquad \text{(Discrete, nondegenerate spectrum)}$$

$$P(a_n) = \sum_i^{g_n} \left|\langle\phi_n^i|\psi\rangle\right|^2, \qquad \text{(Discrete, degenerate spectrum)}$$

$$\mathrm{d}P(\alpha) = |\langle\phi_\alpha|\psi\rangle|^2 \, \mathrm{d}\alpha, \qquad \text{(Continuous, nondegenerate spectrum)}$$

5. If the measurement of the physical quantity on the system in the state $|\psi\rangle$ gives the result a_n, the state of the system immediately after the measurement

is the normalized projection,

$$\frac{P_n|\psi\rangle}{\sqrt{\langle\psi|P_n|\psi\rangle}},$$

of $|\psi\rangle$ onto the eigensubspace associated with a_n.

6. The time evolution of the state vector $|\psi(t)\rangle$ is governed by the Schrödinger equation:

$$i\hbar\frac{d}{dt}|\psi(t)\rangle = \hat{H}(t)|\psi(t)\rangle,$$

where $\hat{H}(t)$ is the observable associated with the total energy of the system (called the Hamiltonian).

In this thesis, we will adopt the convention where "hatted" quantities $\hat{A}$ denote Hermitian, quantum operators and "unhatted" quantities A are classical quantities. This will be made clearer as we progress further. The postulates of quantum mechanics also imply that two vectors that differ by a scalar multiplicative constant represent the same pure state thus the Hilbert space contains normalizable *rays* rather than vectors [7]. We now present a more mathematical description of the quantum mechanics [10].

Definition 2.3.1. For a vector $\psi \in \mathcal{H}$, the corresponding **ray** is the set

$$\underline{\psi} = \{e^{i\alpha}\psi : \alpha \in \mathbb{R}\}.$$

An element $\phi \in \underline{\psi}$ is called the **representative** of $\underline{\psi}$.

Definition 2.3.2. The **ray product** of two rays $\underline{\phi}$ and $\underline{\psi}$ is defined as

$$\underline{\phi} \cdot \underline{\psi} = |\langle\phi|\psi\rangle|,$$

and is independent of the choice of representatives from the sets $\underline{\phi}$ and $\underline{\psi}$.

Definition 2.3.3. The set of all rays equipped with the ray product is called the ray space $\underline{\mathcal{H}}$.

With these definitions we are ready to define what a symmetry transformation in quantum mechanics is [10].

Definition 2.3.4. A **symmetry transformation** is a bijective mapping $\underline{S} : \underline{\mathcal{H}} \to \underline{\mathcal{K}}$ that preserves the ray product

$$\underline{S}(\underline{\phi}) \cdot \underline{S}(\underline{\psi}) = \underline{\phi} \cdot \underline{\psi}, \quad \forall \underline{\phi}, \underline{\psi} \in \underline{\mathcal{H}}.$$

We now present one of the most fundamental theorems in quantum mechanics [10].

Theorem 2.3.1 (**Wigner's theorem**). *For every symmetry transformation $\underline{S} : \underline{\mathcal{H}} \to \underline{\mathcal{K}}$, there exists a transformation $S : \mathcal{H} \to \mathcal{K}$ such that*

1. $\underline{S}(\underline{\psi}) = \underline{S(\psi)}, \quad \forall \psi \in \mathcal{H}$ *and,*

2. *the transformation is either unitary or antiunitary.*

Wigner's theorem gives us the recipe for quantizing symmetries from classical physics to quantum physics and will be frequently invoked to create the necessary quantum symmetry transformations. The postulates can also be used to derive the canonical commutation relations

$$\begin{aligned} [\hat{X}_i, \hat{X}_j] &= 0, \\ [\hat{P}_i, \hat{P}_j] &= 0, \\ [\hat{X}_i, \hat{P}_j] &= i\hbar\delta_{ij}, \end{aligned} \tag{2.3.1}$$

and the scalar product of the position and momentum eigenkets

$$\langle x|p\rangle = \frac{1}{(2\pi\hbar)^{3/2}} e^{i\mathbf{x}\cdot\mathbf{p}/\hbar}. \tag{2.3.2}$$

The canonical commutation relations provide the method for quantizing a classical theory through *canonical quantization* which is to promote the classical dynamical

variables of position x (or its equivalent) and momentum p (or its equivalent) to Hermitian operators $\hat{X}$ and $\hat{P}$, respectively, and impose the canonical commutation relations (2.3.1) on the operators.

2.3.1 Formulations of Quantum Mechanics

Schrödinger Picture vs. Heisenberg Picture

In the *Schrödinger picture*, operators are time-independent while the state kets are time-dependant. The equation that governs the time evolution is the Schrödinger equation

$$i\hbar\frac{d}{dt}|\psi(t)\rangle = \hat{H}|\psi(t)\rangle. \tag{2.3.3}$$

In the *Heisenberg picture*, the Schrödinger operators undergo a unitary transformation which makes the operators time-dependent while the state kets become time-independent, formally,

$$\hat{A}^H(t) = \hat{U}^\dagger(t)\hat{A}^S\hat{U}(t), \tag{2.3.4}$$

where $\hat{U}(t) = e^{-i\hat{H}t/\hbar}$ is the unitary *time evolution operator* and the superscipts S and H denote the Schrödinger operator and Heisenberg operator, respectively. The equation that governs the time evolution of the Heisenberg operators is the *Heisenberg equation of motion*

$$\frac{d\hat{A}^H(t)}{dt} = \frac{1}{i\hbar}[\hat{A}^H(t), \hat{H}], \tag{2.3.5}$$

for any observable $\hat{A}$ in the Heisenberg picture.

Path Integral Formulation

We follow the methodology of Srednicki in the following [5]. Consider the general quantum Hamiltonian for one particle

$$\hat{H}(\hat{X},\hat{P}) = \frac{\hat{P}^2}{2m} + V(\hat{X}). \tag{2.3.6}$$

We will let $\hbar = 1$ for clarity for this subsection. Let $|x\rangle$ be the eigenket of $\hat{X}$ such that $\hat{X}|x\rangle = x|x\rangle$. In the Heisenberg picture, we make the necessary adjustments

$$\hat{X} \to \hat{X}(t) = e^{i\hat{H}t}\hat{X}e^{-i\hat{H}t}, \tag{2.3.7}$$

$$\hat{P} \to \hat{P}(t) = e^{i\hat{H}t}\hat{P}e^{-i\hat{H}t}, \tag{2.3.8}$$

$$|x,t\rangle \to |x\rangle. \tag{2.3.9}$$

Then in the Heisenberg picture $\hat{X}(t)|x,t\rangle = x|x,t\rangle$. To calculate the probability amplitude $\langle x'',t''|x',t'\rangle$, namely, the probability that a particle at position x' at time t' to move to position x'' at the later time t'', let us divide the time interval $t''-t'$ into $N+1$ equal subintervals of duration $\delta t = T/(N+1)$. Then let us insert N position intermediate eigenstates to get

$$\langle x'',t''|x',t'\rangle = \int \prod_{j=1}^{N} dx_j \, \langle x''|e^{-i\hat{H}\delta t}|x_N\rangle\langle x_N|e^{-i\hat{H}\delta t}|x_{N-1}\rangle \ldots \langle x_1|e^{-i\hat{H}\delta t}|x'\rangle. \tag{2.3.10}$$

Using the Baker-Campbell-Hausdorf formula

$$\exp(A+B) = \exp(A)\exp(B)\exp\Big(-\frac{1}{2}[A,B] + \ldots\Big), \tag{2.3.11}$$

we get up to second order in δt

$$\exp(-\hat{H}\delta t) = \exp\big[-i(\delta t/2m)\hat{P}^2\big]\exp\big[-i\delta t V(\hat{X})\big]\exp\big[O(\delta t^2)\big], \tag{2.3.12}$$

where in the limit of small δt, the last term goes to one [5]. Consider the probability amplitude $\langle x_2|e^{-i\hat{H}\delta t}|x_1\rangle$ and observe that

$$\begin{aligned}\langle x_2|e^{-i\hat{H}\delta t}|x_1\rangle &= \langle x_2|e^{-i(\delta t/2m)\hat{P}^2}e^{-i\delta t V(\hat{X})}|p_1\rangle \\ &= \langle x_2|e^{-i(\delta t/2m)\hat{P}^2}|p_1\rangle\langle p_1|e^{-i\delta t V(\hat{X})}|x_1\rangle \\ &= \int dp_1\, e^{-i(\delta t/2m)p_1^2}e^{-i\delta t V(x_1)}\langle x_2|p_1\rangle\langle p_1|x_1\rangle \\ &= \int \frac{dp_1}{2\pi}\, e^{-i(\delta t/2m)p_1^2}e^{-i\delta t V(x_1)}\langle x_2|p_1\rangle\langle p_1|x_1\rangle e^{ip_1(x_2-x_1)} \\ &= \int \frac{dp_1}{2\pi}\, e^{-iH(x_1,p_1)\delta t}e^{ip_1(x_2-x_1)}, \end{aligned} \tag{2.3.13}$$

where we used $\langle x|p\rangle = (2\pi)^{1/2}e^{i\mathbf{x}\cdot\mathbf{p}}$. It is easy to generalize to N cases but for more general Hamiltonians, we will adopt *Weyl ordering* and replace $H(x_1, p_1)$ with $H(\bar{x}_1, p_1)$ where $\bar{x}_1 = \frac{1}{2}(x_1 + x_2)$ [5]. Then we generalize equation (2.3.13) into (2.3.10) to get

$$\langle x'', t''|x', t'\rangle = \int \prod_{k=1}^{N} dx_k \prod_{j=1}^{N} \frac{dp_j}{2\pi}\, e^{ip_j(x_{j+1}-x_j)}e^{-iH(p_j,\bar{x}_j)\delta t}, \tag{2.3.14}$$

where $\bar{x}_j = \frac{1}{2}(x_j + x_{j+1})$, $x_0 = x'$, and $x_{N+1} = x''$. Let $\dot{x}_j = (x_{j+1} - x_j)/\delta t$ then in the limit $\delta t \to 0$ we have

$$\begin{aligned}\langle x'', t''|x', t'\rangle &= \int \mathcal{D}x\, \mathcal{D}p \exp\Big[i\int_{t'}^{t''} dt\, \big(p(t)\dot{x}(t) - H(x(t), p(t))\big)\Big] \\ &= \int \mathcal{D}x\, \mathcal{D}p \exp\Big[i\int_{t'}^{t''} dt\, L(\dot{x}(t), p(t))\Big] \end{aligned} \tag{2.3.15}$$

Expression (2.3.15) is called *Feynman's path integral* and we see that we are integrating over all possible paths in phase space [11]. If the momenta in $H(x, p)$ is quadratic at most then the integration over p is gaussian and can be performed in an analytic form [5]. The constant that arises from this integration can then be absorbed into the functional measure $\mathcal{D}x$. Now at some time $t_1 \in (t', t'')$, consider

$$\langle x'', t''|\hat{X}(t_1)|x', t'\rangle = \langle x''|e^{-i\hat{H}(t''-t_1)}\hat{X}e^{-i\hat{H}(t_1-t')}|x'\rangle. \tag{2.3.16}$$

If we insert intermediate position and momentum eigenstates, we get

$$\begin{aligned}\langle x'', t''|\hat{X}(t_1)|x', t'\rangle &= \langle x'', t''|x(t_1)|x', t'\rangle \\ &= \int \prod_{j=1}^{N} dx_j \, x(t_1) \langle x''|e^{-i\hat{H}\delta t}|x_N\rangle\langle x_N|e^{-i\hat{H}\delta t}|x_{N-1}\rangle \\ &\ldots \langle x_1|e^{-i\hat{H}\delta t}|x'\rangle \\ &= \int \mathcal{D}x\, \mathcal{D}p\, x(t_1)\, e^{iS}. \end{aligned} \tag{2.3.17}$$

If we now consider $\hat{X}(t_1)$ and $\hat{X}(t_2)$, the ordering of the operator's times matter thus we have

$$\langle x'', t''|\mathrm{T}\hat{X}(t_1)\hat{X}(t_2)|x', t'\rangle = \int \mathcal{D}x\, \mathcal{D}p\, x(t_1)x(t_2)\, e^{iS}, \tag{2.3.18}$$

where T is the *time-ordering symbol*: a product of operators is implied to be written with all operators at later times to the left of the operators at earlier times. Later on, we will calculate many expressions such as (2.3.18) which are called *correlation functions*. To calculate them, suppose we made the replacement

$$H(x, p) \to H(x, p) - f(t)x(t) - h(t)p(t), \tag{2.3.19}$$

where $f(t)$ and $h(t)$ are specified functions [5]. Then the probability amplitude becomes

$$\langle x'', t''|x', t'\rangle_{f,h} = \int \mathcal{D}x\, \mathcal{D}p \exp\Big[i \int_{t'}^{t''} dt\big(p\dot{x} - H + fx + hp\big)\Big]. \tag{2.3.20}$$

Now we introduce a new derivative [5].

Definition 2.3.5. The **functional derivative** is a functional analogue of the differential operator from differential calculus that acts on functions such that

$$\frac{\delta}{\delta f(t_1)} f(t_2) = \delta(t_1 - t_2),$$

where $\delta(t)$ is the Dirac delta function. The normal rules of differential operators from differential calculus applies to the functional derivative.

Then observe that [5]

$$\begin{aligned}
\frac{1}{i}\frac{\delta}{\delta f(t_1)}\langle x'', t''|x', t'\rangle_{f,h} &= \int \mathcal{D}x\,\mathcal{D}p\, x(t_1)\, e^{i\int dt\,[p\dot{x}-H+fx+hp]}, \\
\frac{1}{i}\frac{\delta}{\delta f(t_1)}\frac{1}{i}\frac{\delta}{\delta f(t_2)}\langle x'', t''|x', t'\rangle_{f,h} &= \int \mathcal{D}x\,\mathcal{D}p\, x(t_1)x(t_2)\, e^{i\int dt\,[p\dot{x}-H+fx+hp]}, \\
\frac{1}{i}\frac{\delta}{\delta f(t_1)}\frac{1}{i}\frac{\delta}{\delta h(t_1)}\langle x'', t''|x', t'\rangle_{f,h} &= \int \mathcal{D}x\,\mathcal{D}p\, p(t_1)\, e^{i\int dt\,[p\dot{x}-H+fx+hp]}.
\end{aligned} \tag{2.3.21}$$

If we set $f(t) = h(t) = 0$, then we get the correlation functions thus

$$\langle x'', t''|\mathrm{T}\hat{X}(t_1)\ldots\hat{P}(t_n)\ldots|x', t'\rangle = \frac{1}{i}\frac{\delta}{\delta f(t_1)}\ldots\frac{1}{i}\frac{\delta}{\delta h(t_n)}\ldots\langle x'', t''|x', t'\rangle_{f,h}\Big|_{f=h=0}. \tag{2.3.22}$$

For later calculations, we will be interested in the ground state as the initial and final states which is called the *vacuum expectation value*. To compute this using the path integral approach, we let $t' \to -\infty$ and $t'' \to +\infty$ to get

$$\begin{aligned}
\langle 0|0\rangle_{f,h} &= \lim_{\substack{t'\to-\infty \\ t''\to+\infty}} \int dx''\,dx'\langle 0|x''\rangle\langle x'', t''|x', t'\rangle_{f,h}\langle x'|0\rangle \\
&= \lim_{\substack{t'\to-\infty \\ t''\to+\infty}} \int dx''\,dx'\psi_0^*(x'')\langle x'', t''|x', t'\rangle_{f,h}\psi_0(x').
\end{aligned} \tag{2.3.23}$$

To simplify the above expression we make the replacement

$$H \to (1 - i\epsilon)H, \tag{2.3.24}$$

where ϵ is a small positive infinitesimal. This allows the Hamiltonian to pick out only the ground state when operated on by $\langle x'', t''|$ and $|x', t'\rangle$ in the limits $t'' \to +\infty$ and $t' \to -\infty$ with ϵ fixed [5]. Thus we have

$$\langle 0|0\rangle_{f,h} = \int \mathcal{D}x\,\mathcal{D}p \exp\Big[i\int dt\,\big(p\dot{x} - (1 - i\epsilon)H + fx + hp\big)\Big]. \tag{2.3.25}$$

Given a perturbation $\hat{H}_1$, we have that the total Hamiltonian is $\hat{H} = \hat{H}_0 + \hat{H}_1$ thus

$$\begin{aligned}\langle 0|0\rangle_{f,h} &= \int \mathcal{D}x\,\mathcal{D}p \exp\Big[i\int dt\,\big(p\dot{x} - H_0 - H_1 + fx + hp\big)\Big] \\ &= \exp\Big[-i\int dt\, H_1\Big(\frac{1}{i}\frac{\delta}{\delta h(t)}, \frac{1}{i}\frac{\delta}{\delta f(t)}\Big)\Big] \\ &\times \int \mathcal{D}x\,\mathcal{D}p \exp\Big[i\int dt\,\big(p\dot{x} - H_0 + fx + hp\big)\Big], \end{aligned} \tag{2.3.26}$$

where we have suppressed $i\epsilon$. If the perturbation H_1 depends only on x, the $\hat{P}$ in the total Hamiltonian $\hat{H}$ is at most quadratic with no dependence on $\hat{X}$, and if we desire to compute only the vacuum expectation values involving time-ordered products of $\hat{X}$, we get the simplification

$$\langle 0|0\rangle_f = \exp\Big[i\int dt\, L_1\Big(\frac{1}{i}\frac{\delta}{\delta f(t)}\Big)\Big]\int \mathcal{D}x \exp\Big[i\int dt\,\big(L_0(\dot{x}, x) + fx\big)\Big], \tag{2.3.27}$$

where $L_1(x) = -H_1(x)$. The use of the functional integrals for the path integral to calculate or perhaps more appropriately, generate, the vacuum expectation values has led to the path integral to also be called the *generating functional* or the *partition function*.

Quantum Harmonic Oscillator

We will present the derivation of the solution to the quantum harmonic oscillator. One of the most commonly used methods to solve this problem is to bring the Hamiltonian into position-space and solve the differential equation using a power series expansion. Another very commonly used method is to use the ladder operator method which allows one to find the energy eigenvalues and eigenstates without ever solving the differential equation. Lastly, we can also solve this problem using path integrals. We will first demonstrate the ladder operator method and then the path integral method because the results and concepts that we get from both methods will generalize to all the quantum fields.

The Hamiltonian of the quantum harmonic oscillator is

$$\hat{H} = \frac{\hat{P}^2}{2m} + \frac{1}{2}m\omega^2\hat{X}^2. \tag{2.3.28}$$

We can non-dimensionalize the Hamiltonian by dividing by $\hbar\omega$ to get

$$\begin{aligned}\frac{\hat{H}}{\hbar\omega} &= \frac{\hat{P}^2}{2\hbar\omega m} + \frac{m\omega}{2\hbar}\hat{X}^2 \\ &= \frac{m\omega}{2\hbar}\Big(\hat{X}^2 + \frac{\hat{P}^2}{m^2\omega^2}\Big) \\ &= \frac{m\omega}{2\hbar}\Big(\hat{X} - \frac{i\hat{P}}{m\omega}\Big)\Big(\hat{X} + \frac{i\hat{P}}{m\omega}\Big) + + \frac{i[\hat{X},\hat{P}]}{4\hbar\omega m} \\ &= \sqrt{\frac{m\omega}{2\hbar}}\Big(\hat{X} - \frac{i\hat{P}}{m\omega}\Big)\sqrt{\frac{m\omega}{2\hbar}}\Big(\hat{X} - \frac{i\hat{P}}{m\omega}\Big) - \frac{1}{4m\omega}.\end{aligned} \tag{2.3.29}$$

We now define the following operators

$$\hat{a}^\dagger = \sqrt{\frac{m\omega}{2\hbar}}\Big(\hat{X} - \frac{i\hat{P}}{m\omega}\Big), \tag{2.3.30}$$

$$\hat{a} = \sqrt{\frac{m\omega}{2\hbar}}\Big(\hat{X} + \frac{i\hat{P}}{m\omega}\Big), \tag{2.3.31}$$

which are the *creation* and *annihilation* operators, respectively. Then observe that

$$\hat{a}^\dagger\hat{a} = \frac{\hat{H}}{\hbar\omega} - \frac{1}{2}$$

hence,

$$\hat{H} = \hbar\omega\Big(\hat{a}^\dagger\hat{a} + \frac{1}{2}\Big). \tag{2.3.32}$$

It is trivial to prove that $\hat{a}^\dagger\hat{a}$ is a Hermitian operator so we let $\hat{a}^\dagger\hat{a} = \hat{N}$, which is called the *number operator*, then we get

$$\hat{H} = \hbar\omega\Big(\hat{N} + \frac{1}{2}\Big), \tag{2.3.33}$$

thus the Hamiltonian and number operator have the same eigenvalues and eigenkets.

It is trivial to see that

$$\begin{aligned}[\hat{a}, \hat{a}^\dagger] &= 1, \\ [\hat{N}, \hat{a}^\dagger] &= \hat{a}^\dagger, \\ [\hat{N}, \hat{a}] &= -\hat{a}. \end{aligned} \tag{2.3.34}$$

Let $|n\rangle$ be an eigenket of $\hat{N}$ such that $\hat{N}|n\rangle = n|n\rangle$ then note that

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

hence the operator $\hat{a}^\dagger$ raises the eigenvalue, or quanta, by one. Similar calculations reveal that

$$\hat{N}\hat{a}|n\rangle = (n-1)|n\rangle, \tag{2.3.35}$$

hence the operator $\hat{a}$ lowers the eigenvalue by one. To find the eigenvalue for $\hat{a}^\dagger$, observe that

$$\begin{aligned}\langle n|\hat{a}\Big(\hat{a}^\dagger|n\rangle &= c_+|n+1\rangle\Big) \\ \langle n|(\hat{a}^\dagger\hat{a}+1)|n\rangle &= c_+\langle n|\hat{a}|n+1\rangle \\ (n+1) &= |c_+|^2 \\ c_+ &= \sqrt{n+1}.\end{aligned}$$

Similar calculations show that $c_- = \sqrt{n}$. Thus we have

$$\hat{a}^\dagger|n\rangle = \sqrt{n+1}|n+1\rangle, \tag{2.3.36}$$

$$\hat{a}|n\rangle = \sqrt{n}|n-1\rangle. \tag{2.3.37}$$

From the definition of the inner product, we have that

$$\begin{aligned} 0 &\leq \langle n|\hat{N}|n\rangle \\ &= n\langle n|n\rangle \\ &= n, \end{aligned}$$

so $n \geq 0$. Then we have that

$$\hat{H}|n\rangle = \hbar\omega\Big(\hat{N} + \frac{1}{2}\Big)|n\rangle = \hbar\omega\Big(n + \frac{1}{2}\Big)|n\rangle, \tag{2.3.38}$$

thus the energy eigenvalues are

$$E_n = \hbar\omega\Big(n + \frac{1}{2}\Big), \quad n = 0, 1, 2, ... \tag{2.3.39}$$

To find the eigenstates, note that

$$\begin{aligned} |1\rangle &= \hat{a}^\dagger|0\rangle, \\ |2\rangle &= \Big[\frac{(\hat{a}^\dagger)^2}{\sqrt{2}}\Big]|0\rangle, \\ |3\rangle &= \Big[\frac{(\hat{a}^\dagger)^3}{\sqrt{3!}}\Big]|0\rangle, \\ &\vdots \\ |n\rangle &= \Big[\frac{(\hat{a}^\dagger)^n}{\sqrt{n!}}\Big]|0\rangle. \end{aligned} \tag{2.3.40}$$

Observe that

$$\begin{aligned} \langle x|\Big(\hat{a}|0\rangle = 0\Big) \\ \sqrt{\frac{m\omega}{2\hbar}}\Big[x + \frac{i}{m\omega}\Big(-i\hbar\frac{d}{dx}\Big)\Big]\langle x|0\rangle = 0 \\ \frac{d}{dx}\langle x|0\rangle = -\frac{m\omega}{\hbar}x\langle x|0\rangle, \end{aligned} \tag{2.3.41}$$

which yields (after normalization) [11]

$$\langle x|0\rangle = \frac{1}{\pi^{1/4}\sqrt{x_0}}\exp\Big[-\frac{1}{2}\Big(\frac{x}{x_0}\Big)^2\Big], \tag{2.3.42}$$

where

$$x_0 = \sqrt{\frac{\hbar}{m\omega}}. \tag{2.3.43}$$

Then the eigenstates are [11]

$$\langle x|n\rangle = \Big(\frac{1}{\pi^{1/4}\sqrt{2^n n!}}\Big)\left(\frac{1}{x_0^{n+1/2}}\right)\Big(x - x_0^2\frac{d}{dx}\Big)^n \exp\Big[-\frac{1}{2}\Big(\frac{x}{x_0}\Big)^2\Big]. \tag{2.3.44}$$

The key idea of this section is that the creation operators create one unit of energy while the annihilation operator destroys one unit of energy. This concept will be extremely important once we consider particle creation and annihilation in quantum field theory.

Now let us rewrite the quantum harmonic oscillator Hamiltonian as

$$\hat{H}(\hat{X},\hat{P}) = \frac{\hat{P}^2}{2m} + \frac{1}{2}m\omega^2\hat{X}^2. \tag{2.3.45}$$

Then the ground state to ground state transition amplitude is

$$\langle 0|0\rangle_f = \int \mathcal{D}x\,\mathcal{D}p \exp\Big[i\int dt\,\big[(p\dot{x} - (1-i\epsilon)H + fx\big)\Big], \tag{2.3.46}$$

where multiplying $(1-i\epsilon)$ to the Hamiltonian results in $m\omega^2 \to (1-i\epsilon)m\omega^2$. We will set $m=1$ for brevity then we have in the Lagrangian formulation [5]

$$\langle 0|0\rangle_f = \int \mathcal{D}x \exp\Big[i\int dt\,\Big(\frac{1}{2}\dot{x}^2 - \frac{1}{2}(1-i\epsilon)\omega^2x^2 + fx\Big)\Big]. \tag{2.3.47}$$

Focusing on the action term, we can write the action term with the Fourier

transformed variables

$$\tilde{x}(E) = \int dt\, e^{iEt}\, x(t),$$
$$x(t) = \int \frac{dE}{2\pi}\, e^{-iEt}\, \tilde{x}(E), \tag{2.3.48}$$

and integrate over E' to get

$$\begin{aligned} S &= \frac{1}{2}\int \frac{dE}{2\pi}\Bigg[\Big((1+i\epsilon)E^2 - (1-i\epsilon)\omega^2\Big)\tilde{x}(E)\tilde{x}(-E) + \tilde{f}(E)\tilde{x}(-E) \\ &\quad + \tilde{f}(-E)\tilde{x}(E)\Big] \\ &= \frac{1}{2}\int \frac{dE}{2\pi}\Big[(E^2 - \omega^2 + i\epsilon)\tilde{x}(E)\tilde{x}(-E) + \tilde{f}(E)\tilde{x}(-E) + \tilde{f}(-E)\tilde{x}(E)\Big], \end{aligned} \tag{2.3.49}$$

where we absorbed the positive coefficient of $E^2 - \omega^2 + i(E^2 + \omega^2)\epsilon$ into ϵ to get $E^2 - \omega^2 + i\epsilon$. Next we shift integration variables by

$$\tilde{q}(E) = \tilde{x}(E) + \frac{\tilde{f}(E)}{E^2 - \omega^2 + i\epsilon}, \tag{2.3.50}$$

and the integration measure becomes $\mathcal{D}x = \mathcal{D}q$ then the action becomes [5]

$$S = \frac{1}{2}\int \frac{dE}{2\pi}\Big[\tilde{q}(E)(E^2 - \omega^2 + i\epsilon)\tilde{q}(-E) - \frac{\tilde{f}(E)\tilde{f}(-E)}{E^2 - \omega^2 + i\epsilon}\Big], \tag{2.3.51}$$

hence

$$\begin{aligned} \langle 0|0\rangle_f &= \exp\Big[\frac{i}{2}\int \frac{dE}{2\pi}\,\frac{\tilde{f}(E)\tilde{f}(-E)}{E^2 - \omega^2 + i\epsilon}\Big] \\ &\quad \times \int \mathcal{D}q \exp\Big[\frac{i}{2}\int \frac{dE}{2\pi}\,\tilde{q}(E)(E^2 - \omega^2 + i\epsilon)\tilde{q}(-E)\Big]. \end{aligned} \tag{2.3.52}$$

Now if we set $f = 0$, the first term equals 1 and since the ground state is normalized as $\langle 0|0\rangle_{f=0} = 1$, the bracketed expression in the second term must be zero hence

the second term is equal to 1. Then the transition amplitude becomes

$$\langle 0|0\rangle_f = \exp\Big[\frac{i}{2}\int dt\, dt'\, f(t)G(t-t')f(t')\Big], \tag{2.3.53}$$

where

$$G(t-t') = \int \frac{dE}{2\pi}\,\frac{e^{-iE(t-t')}}{-E^2+\omega^2-i\epsilon}, \tag{2.3.54}$$

is called the *propagator* of the harmonic oscillator [11]. The propagator is essentially the Green's function of the harmonic oscillator equation which can be seen by

$$\Big(\frac{\partial^t}{\partial t^2}+\omega^2\Big)G(t-t') = \delta(t-t'). \tag{2.3.55}$$

Contour integrating in the complex E plane and using the residue theorem yields [5]

$$G(t-t') = \frac{i}{2\omega}\exp\Big(-i\omega\,|t-t'|\,\Big). \tag{2.3.56}$$

Now we can calculate the vacuum expectation values using functional derivatives in the following manner

$$\begin{aligned}\langle 0|\mathrm{T}\hat{X}(t_1)\hat{X}(t_2)|0\rangle &= \frac{1}{i}\frac{\delta}{\delta f(t_1)}\frac{1}{i}\frac{\delta}{\delta f(t_2)}\langle 0|0\rangle_f\Big|_{f=0}\\ &= \frac{1}{i}G(t_2-t_1),\end{aligned} \tag{2.3.57}$$

$$\begin{aligned}\langle 0|\mathrm{T}\hat{X}(t_1)\hat{X}(t_2)\hat{X}(t_3)\hat{X}(t_4)|0\rangle &= \frac{1}{i}\frac{\delta}{\delta f(t_1)}\frac{1}{i}\frac{\delta}{\delta f(t_2)}\frac{1}{i}\frac{\delta}{\delta f(t_3)}\frac{1}{i}\frac{\delta}{\delta f(t_4)}\langle 0|0\rangle_f\Big|_{f=0}\\ &= \frac{1}{i^2}\Big[G(t_1-t_2)G(t_3-t_4)\\ &+ G(t_1-t_3)G(t_2-t_4)\\ &+ G(t_1-t_4)G(t_2-t_3)\Big],\end{aligned} \tag{2.3.58}$$

or more generally [5]

$$\langle 0|\mathrm{T}\hat{X}(t_1)\dots\hat{X}(t_{2n})|0\rangle = \frac{1}{i^n}\sum_{\text{pairings}} G(t_{i_1}-t_{i_2})\dots G(t_{i_{2n-1}}-t_{i_{2n}}). \tag{2.3.59}$$

Equation (2.3.59) is known as *Wick's theorem* and it will be extremely useful for computing vacuum expectation values. We will give the proof later on once we have introduced quantum fields.

2.4 Natural Units

Natural units are defined as setting $\hbar = c = 1$. Then we can express time T and length L in units of mass M, namely, $T = c^{-2}\hbar M^{-1}$ and $L = \hbar c^{-1}M^{-1}$, respectively. Thus any physical quantity X will have a unit of mass M raised to some power [5]. From the mass-energy equivalence relation, energy in natural units is $E = m$ so any physical quantity can also be expressed in units of energy as well, which will usually be the case in most equations. Let $[\cdot]$ be the mapping that gives the number to which the mass unit is raised to for a physical quantity X. We can easily see that [5]

$$\begin{aligned}
[m] &= +1, \\
[x^\mu] &= -1, \\
[\partial^\mu] &= +1, \\
[d^d x] &= -d.
\end{aligned} \tag{2.4.1}$$

We can easily deduce the dimensions of the path integral using equation (2.4.1). Instead of the Lagrangian L for discrete systems, let us instead consider the continuous Lagrangian density $\mathcal{L}$ which we will be using for quantum fields. Since the action $\mathcal{S}$ is in the exponential, we see that

$$[\mathcal{S}] = 0, \tag{2.4.2}$$

and since

$$S = \int d^d x \, \mathcal{L}, \tag{2.4.3}$$

we see that

$$[\mathcal{L}] = d, \tag{2.4.4}$$

for d spacetime dimensions [5].

The benefit of natural units is that manipulating equations become significantly easier since most equations will either be dimensionless or will be in dimensions of mass/energy. Along with this, the ubiquity of $\hbar$ and c in the equations of quantum field theory can make many equations and expressions cluttered thus setting them to 1 significantly reduces the equations to simpler, more elegant forms. We will adopt natural units throughout this thesis unless specified otherwise. We present the exact and approximate values of some physical values and constants in both SI (except for the elementary charge e which is in Lorentz-Heaviside units) and natural units that will aid us in explicit calculations:

$$\begin{aligned}
&1\,\mathrm{eV} = 1.602176634 \times 10^{-19}\,\mathrm{J} \approx 1.602 \times 10^{-19}\,\mathrm{J},\\
&1\,\mathrm{eV}/c^2 = 1.78266192 \times 10^{-36}\,\mathrm{kg} \approx 1.783 \times 10^{-36}\,\mathrm{kg},\\
&\hbar = \frac{h}{2\pi} = 1.054571817 \times 10^{-34}\,\mathrm{J}\cdot\mathrm{s} = 6.582119569 \times 10^{-16}\,\mathrm{eV}\cdot\mathrm{s}\\
&\approx 6.582 \times 10^{-16}\,\mathrm{eV}\cdot\mathrm{s},\\
&c = 299,792,458\,\mathrm{m/s} \approx 3 \times 10^8\,\mathrm{m/s},\\
&\hbar c \approx 200\,\mathrm{MeV}\cdot\mathrm{fm},\\
&\alpha = \frac{e^2}{4\pi} \approx \frac{1}{137},\\
&e = \sqrt{4\pi\alpha} \approx 0.30282212088,\\
&m_e = 0.51099895\,\mathrm{MeV}/c^2 \approx 0.511\,\mathrm{MeV}/c^2,\\
&m_p = 938.27208816\mathrm{MeV}/c^2 \approx 938.3\,\mathrm{MeV}/c^2,
\end{aligned}$$

where m_e and m_p denotes the mass of the electron and proton, respectively [1, 7].

To recover the original units, a simple dimensional analysis can be used to find where to properly insert $\hbar$ and c. For example, converting energy E back to mass requires that we divide by c^2 since $E/c^2 = m$. Note that the electric charge e in natural units becomes dimensionless since it can now be expressed in terms of the dimensionless *fine-structure constant* α. We will see later on the importance having a dimensionless electric charge e for quantum electrodynamics and how natural units plays a role in constructing viable, or *renormalizable*, quantum field theories.

Chapter 3

Symmetries

3.1 Lie Groups, Lie Algebras, and Representations

In this section we will introduce the important concepts from the mathematical theory of Lie groups, Lie algebras, and representation theory. The concepts from these mathematical theories have an intimate relationship with the study of quantum field theory and developing the necessary background will be crucial in fully understanding quantum field theory. We use the definitions and theorems from [12, 13, 14]:

Definition 3.1.1. A **group** $\langle G, * \rangle$ is a set G that is closed under the binary operation $*$ such that the following axioms are satisfied:

1. Associativity of $*$: For all $a, b, c \in G$,

$$(a * b) * c = a * (b * c),$$

2. Identity element e for $*$: There exists an identity element $e \in G$ such that for all $x \in G$,

$$x * e = e * x = x,$$

3. Inverse under $*$: For every $a \in G$, there exists an element $a' \in G$ such that

$$a * a' = a' * a = e.$$

Definition 3.1.2. [13] A subset $M \subset \mathbb{R}^m$ is a **smooth manifold** of dimension $n \geq 0$ if, for each $x \in M$ there exists a smooth function

$$h : U \to \mathbb{R}^m,$$

defined on an open set $U \subset \mathbb{R}^n$ such that

1. h maps U homeomorphically onto an open neighborhood V of $x \in M$ and,

2. for each $u \in U$ the matrix $[\partial h_\alpha(u)/\partial u_j]$ rank has n.

Definition 3.1.3. Let $M_n(\mathbb{R})$ and $M_n(\mathbb{C})$ be the space of all $n \times n$ matrices with real and complex entries, respectively.

We see that $M_n(\mathbb{R})$ and $M_n(\mathbb{C})$ can be identified as $\mathbb{R}^{n^2}$ and $\mathbb{C}^{n^2}$, respectively.

Definition 3.1.4. [12] A **Lie group** G is a smooth manifold and a group such that the group product

$$G \times G \to G$$

and the inverse map $G \to G$ are smooth.

Definition 3.1.5. The **general linear group** $\mathrm{GL}(n, \mathbb{F})$ of degree n over a field $\mathbb{F}$ is the set of $n \times n$ invertible matrices with matrix multiplication as its binary operation. For any A in $\mathrm{GL}(n, \mathbb{F})$, the entries of A belong to $\mathbb{F}$.

Theorem 3.1.1. $\mathrm{GL}(n, \mathbb{R})$ *is a Lie group under matrix multiplication.*

Proof. Choose any number of arbitrary $n \times n$ matrices in $\mathrm{GL}(n, \mathbb{R})$ then note that the product of any number of matrices in $\mathrm{GL}(n, \mathbb{R})$ yields a $n \times n$ matrix with a

nonzero determinant because of the property of determinants for square matrices A_i

$$\det\Big(\prod_i A_i\Big) = \prod_i \det(A_i).$$

Thus $\mathrm{GL}(n,\mathbb{R})$ is closed. It is trivial to see that the matrix product is associative and since all elements in $\mathrm{GL}(n,\mathbb{R})$ have a nonzero determinant, there exists a unique inverse for each element in $\mathrm{GL}(n,\mathbb{R})$. We trivially see that the unique $n \times n$ identity matrix is in $\mathrm{GL}(n,\mathbb{R})$ as well. Thus $\mathrm{GL}(n,\mathbb{R})$ is a group under matrix multiplication.

The set of all $n \times n$ real matrices $M_n(\mathbb{R})$ forms a real vector space of dimension n^2. Consider the determinant mapping $\det(\cdot) : \mathbb{R}^{n^2} \to \mathbb{R}$. The determinant is a polynomial map thus the determinant is a C^∞, or smooth, map. Then

$$\det^{-1}(\mathbb{R}/\{0\}) = GL(n,\mathbb{R}),$$

is an open subset of $\mathbb{R}^{n^2}$ which is trivially a smooth manifold. Thus $\mathrm{GL}(n,\mathbb{R})$ is smooth manifold of dimension n^2 therefore $\mathrm{GL}(n,\mathbb{R})$ is a Lie group. □

Theorem 3.1.2. $\mathrm{GL}(n,\mathbb{C})$ *is a Lie group under matrix multiplication.*

Proof. For all $Z \in \mathrm{GL}(n,\mathbb{C})$, Z can be decomposed as

$$Z = X + iY,$$

where $X, Y \in M_n(\mathbb{R})$. Now define a map f that constructs a $2n \times 2n$ real matrix such that the entries in Z of the form $a + ib$ are replaced with the block matrix

$$\begin{pmatrix} a & -b \\ b & a \end{pmatrix},$$

where $a, b \in \mathbb{R}$. It is trivial to see that this map is bijective and a group homomorphism thus the map is an isomorphism and therefore $\mathrm{GL}(n,\mathbb{C}) \cong \mathrm{GL}(2n,\mathbb{R})$. Thus we see that $\mathrm{GL}(n,\mathbb{C})$ is subgroup of $\mathrm{GL}(2n,\mathbb{R})$ and is a subset of $\mathbb{R}^{(2n)^2}$ therefore

$\mathrm{GL}(n, \mathbb{C})$ is a Lie group. □

Definition 3.1.6. [12] A **matrix Lie group** G is a subgroup of $\mathrm{GL}(n, \mathbb{C})$ where if any sequence A_m in G converges to some matrix A, then either A is in G or is not invertible.

Throughout this thesis, virtually all of the Lie groups we will encounter will be matrix Lie groups which are Lie groups as well. Thus we will focus mostly on concepts that relate to matrix Lie groups.

We present the next theorem without proof as an aid for identifying and proving more Lie groups [14].

Theorem 3.1.3 (**Closed Subgroup Theorem**). *Every closed subgroup H of a Lie group G is a Lie subgroup.*

By closed, we mean that the subgroup H is topologically a closed subset of G. This theorem and the definition of the matrix Lie group are equivalent [12]. In general, it is much more difficult to prove the smooth manifold condition of a proposed Lie group and so the Closed Subgroup theorem and the definition of a matrix Lie group significantly reduces our task to just proving only if a subgroup H is a closed subset of a Lie group G.

Definition 3.1.7. The **unitary group** $\mathrm{U}(n)$ is the group of all $n \times n$ complex matrices that satisfies $U^\dagger = U^{-1}$. The subset of unitary matrices with unit determinant is called the **special unitary group** $\mathrm{SU}(n)$.

Definition 3.1.8. The **orthogonal group** $\mathrm{O}(n)$ is the group of all $n \times n$ real matrices that satisfies $A^{\mathrm{T}} = A^{-1}$. The subset of orthogonal matrices with unit determinant is called the **special orthogonal group** $\mathrm{SO}(n)$.

Theorem 3.1.4. $\mathrm{U}(n)$ *and* $\mathrm{O}(n)$ *are Lie groups. Similarly,* $\mathrm{SU}(n)$ *and* $\mathrm{SO}(n)$ *are Lie groups.*

Proof. If $A_m \in \mathrm{U}(n)$ is a sequence of matrices with determinant ± 1 and converges to A, then A also has determinant ± 1 since the determinant is a continuous

function thus $\mathrm{U}(n)$ and $\mathrm{SU}(n)$ are matrix Lie groups. The same argument applies for $\mathrm{O}(n)$ and $\mathrm{SO}(n)$ therefore all four groups are Lie groups. □

Theorem 3.1.5. *The Lie group* $\mathrm{U}(1)$ *is isomorphic to the Lie group* $\mathrm{SO}(2)$.

Proof. Consider the map $f : \mathbb{C} \to \mathbb{R}^4$ defined as

$$f(e^{i\theta}) = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}.$$

If we consider $z_1, z_2 \in \mathbb{C}$ such that $f(z_1) = f(z_2)$, we get that

$$f(e^{i\theta_1}) = f(e^{i\theta_2})$$
$$\begin{pmatrix} \cos\theta_1 & -\sin\theta_1 \\ \sin\theta_1 & \cos\theta_1 \end{pmatrix} = \begin{pmatrix} \cos\theta_2 & -\sin\theta_2 \\ \sin\theta_2 & \cos\theta_2 \end{pmatrix},$$

hence

$$\cos\theta_1 = \cos\theta_2,$$
$$\sin\theta_1 = \sin\theta_2,$$

which implies that $\theta_1 = \theta_2$ thus f is injective. Now the values of y come from the interval $[-1, 1]$ and the domain of the phase angle θ is $[0, 2\pi]$ thus for all values $y \in \mathbb{R}^4$, there exists a $z \in \mathbb{C}$ such that $f(z) = y$. Thus the map f is surjective

therefore the map is a bijection. Now observe that

$$\begin{aligned} f(e^{i\theta_1}e^{i\theta_2}) &= \begin{pmatrix} \cos(\theta_1+\theta_2) & -\sin(\theta_1+\theta_2) \\ \sin(\theta_1+\theta_2) & \cos(\theta_1+\theta_2) \end{pmatrix} \\ &= \begin{pmatrix} \cos\theta_1\cos\theta_2 - \sin\theta_1\sin\theta_2 & -\sin\theta_1\cos\theta_2 - \cos\theta_1\sin\theta_2 \\ \sin\theta_1\cos\theta_2 + \cos\theta_1\sin\theta_2 & \cos\theta_1\cos\theta_2 - \sin\theta_1\sin\theta_2 \end{pmatrix} \\ &= \begin{pmatrix} \cos\theta_1 & -\sin\theta_1 \\ \sin\theta_1 & \cos\theta_1 \end{pmatrix} \times \begin{pmatrix} \cos\theta_2 & -\sin\theta_2 \\ \sin\theta_2 & \cos\theta_2 \end{pmatrix} \\ &= f(e^{i\theta_1}) \times f(e^{i\theta_2}), \end{aligned}$$

thus the map f is group homomorphism therefore f is an isomorphism and the two groups U(1) and SO(2) are isomorphic. □

Definition 3.1.9. [12] The **generalized orthogonal group** $\mathrm{O}(n,k)$ is the set of all real $(n+k)\times(n+k)$ matrices that preserve the symmetric bilinear form

$$[x,y]_{n,k} = x_1y_1 + \cdots + x_ny_n - x_{n+1}y_{n+1} - \cdots - x_{n+k}y_{n+k}, \tag{3.1.1}$$

where $n,k \in \mathbb{Z}^+$ and $x,y \in \mathbb{R}^{n+k}$. The subset of $\mathrm{O}(n,k)$ with unit determinant is defined as the **special generalized orthogonal group** denoted as $\mathrm{SO}(n,k)$.

The group $\mathrm{O}(n,k)$ will be incredibly important once we cover the mathematical theory of Lorentz transformations because we will identify the group of Lorentz transformations as a subgroup of $\mathrm{O}(n,k)$.

Theorem 3.1.6. *$\mathrm{O}(n,k)$ and $\mathrm{SO}(n,k)$ are Lie groups.*

Proof. It trivial to prove that $\mathrm{O}(n,k)$ and $\mathrm{SO}(n,k)$ are subgroups $\mathrm{GL}(n+k,\mathbb{R}^{n+k})$. If there exists a sequence of matrices $A_m \in \mathrm{O}(n,k)$ that preserves the symmetric form $[\cdot,\cdot]$ and converges to some matrix A, then A also preserves the symmetric bilinear form $[\cdot,\cdot]$ because $[\cdot,\cdot]$ is a function of polynomials thus is a continuous function [12]. A similar argument applies for $\mathrm{SO}(n,k)$. □

Definition 3.1.10. [12] Let G and H be Lie groups. A map $\Phi : G \to H$ is a **Lie group homomorphism** if Φ is a continuous group homomorphism. If Φ is bijective with a continuous inverse map, then Φ is a **Lie group isomorphism** and we say that Lie groups G and H are **isomorphic**.

Definition 3.1.11. [12] A **real or complex Lie algebra** is a real or complex vector space $\mathfrak{g}$ equipped with a map called a **bracket** $[\cdot,\cdot] : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ that obey the following properties:

1. $[\cdot,\cdot]$ is bilinear.

2. $[\cdot,\cdot]$ is skew symmetric: $[X,Y] = -[Y,X]$ for all $X, Y \in \mathfrak{g}$.

3. The **Jacobi identity** holds:

$$[X,[Y,Z]] + [Y,[Z,X]] + [Z,[X,Y]] = 0,$$

for all $X, Y, Z \in \mathfrak{g}$.

Two elements X and Y of a Lie algebra are said to **commute** if $[X,Y] = 0$ and a Lie algebra is **abelian** if $[X,Y] = 0$ for all $X, Y \in \mathfrak{g}$. Throughout this text, the bracket we will use will be the **matrix commutator** defined as

$$[X,Y] = XY - YX, \quad \forall A, B \in M_n(\mathbb{F}). \tag{3.1.2}$$

Definition 3.1.12. If $\mathfrak{g}$ and $\mathfrak{h}$ are Lie algebras, a linear map $\phi : \mathfrak{g} \to \mathfrak{h}$ is a **Lie algebra homomorphism** if $\phi([X,Y]) = [\phi(X), \phi(Y)]$ for all $X, Y \in \mathfrak{g}$. If ϕ is bijective then ϕ is a **Lie algebra isomorphism**.

Definition 3.1.13. Let G be a matrix Lie group. The **Lie algebra** of G, denoted as $\mathfrak{g}$, is the set of all matrices X such that $e^{itX} \in G$ for all $t \in \mathbb{R}$.

The definition we have chosen to use is the physicist's convention for the Lie algebra of a Lie group G. Mathematicians use the map $t \to e^{tX}$ [12].

Definition 3.1.14. [12] If V is a real vector space, then the **complexification** of V, denoted as $V_{\mathbb{C}}$, is the space of formal linear combinations

$$v_1 + iv_2,$$

with $v_1, v_2 \in V$.

An important concept that we will use is the Casimir operator of a Lie algebra.

Definition 3.1.15. A **Casimir operator** $\mathcal{C}$ is an element of a Lie algebra such that

$$[\mathcal{C}, X] = 0, \quad \forall X \in \mathfrak{g},$$

namely the Casimir operator commutes with all elements of a Lie algebra.

The Casimir operator's eigenvalues will be used to label representations of Lie groups and the state vectors belonging to a Hilbert space.

Definition 3.1.16. [12] Let G be a matrix Lie group. A **finite-dimensional complex representation** of G is a Lie group homomorphism

$$\Pi : G \to \mathrm{GL}(V),$$

where V is a finite-dimensional complex vector space. A **finite-dimensional real representation** of G is defined similarly except with V as a finite-dimensional real vector space.

Definition 3.1.17. [12] A **finite-dimensional complex representation** of a real or complex Lie algebra $\mathfrak{g}$ is a Lie algebra homomorphism

$$\pi : G \to \mathrm{gl}(V),$$

where V is a finite-dimensional complex vector space. If $\mathfrak{g}$ is a real Lie algebra

then a **finite-dimensional real representation** is defined similarly except with V as a finite-dimensional real vector space.

Definition 3.1.18. [12] Let Π be a finite dimensional real or complex representation of a matrix Lie group G acting on a space V. A subspace W of V is called **invariant** if $\Pi(A)w \in W$ for all $w \in W$ and for all $A \in G$. An invariant subspace W is called **nontrivial** if $W \neq \{0\}$ and $W \neq V$. A representation with no nontrivial invariant subspaces is called **irreducible**.

Definition 3.1.19. [12] If V is a finite-dimensional inner product space and G is a matrix Lie group, a representation $\Pi : G \to \mathrm{GL}(V)$ is **unitary** if $\Pi(A)$ is a unitary operator on V for every $A \in G$.

3.2 Representations of the Lorentz Group

Let us apply our knowledge of Lie groups, Lie algebras, and representation theory to the Lorentz transformations. As was hinted in the background chapter, the set of Lorentz transformations forms a group under matrix multiplication called the *Lorentz group* which leaves the spacetime interval $s = -(x^0)^2 + (x^1)^2 + (x^2)^2 + (x^3)^2$ invariant.

Proof. It is trivial to show the associative property for any three arbitrary Lorentz transformations. The identity element is $\Lambda^{\mu}{}_{\rho} = \delta^{\mu}{}_{\nu}$. Recall that any Lorentz transformation Λ can be constructed from pure boosts K and pure rotations J. Now let Λ_A and Λ_B be arbitrary Lorentz transformations then observe that

$$\begin{aligned}\Lambda_A\Lambda_B &= (K_1K_2\dots)(J_1J_2\dots)\\ &= \Lambda_C.\end{aligned}$$

Recall that for any matrix $\Lambda^{\mu}{}_{\rho}$,

$$g_{\mu\nu}\Lambda^{\mu}{}_{\rho}\Lambda^{\nu}{}_{\sigma} = g_{\rho\sigma},$$

which implies that

$$\det\Big[g_{\mu\nu}\Lambda^{\mu}{}_{\rho}\Lambda^{\nu}{}_{\sigma} = g_{\rho\sigma}\Big]$$
$$-\det[\Lambda^2] = -1$$
$$\det[\Lambda] = \pm 1,$$

therefore all matrices $\Lambda^{\mu}{}_{\rho}$ are nonsingular thus an inverse element exists for all Lorentz transformations which is $(\Lambda^{-1})^{\rho}{}_{\nu} = \Lambda_{\nu}{}^{\rho}$. □

From the previous section, we immediately identify the Lorentz group as $\mathrm{O}(3,1)$, namely, the group of 4×4 real matrices Λ that preserve the symmetric bilinear form

$$[x, y]_{3,1} = x_1 y_1 + x_2 y_2 + x_3 y_3 - x_4 y_4, \tag{3.2.1}$$

which is just the spacetime interval for Lorentz transformations [15]. Thus we now define the Lorentz group.

Definition 3.2.1. The **Lorentz group** is the group of linear, homogeneous coordinate transformations

$$x^{\mu} \to \bar{x}^{\mu} = \Lambda^{\mu}{}_{\nu} x^{\nu},$$

that leave invariant the interval

$$x^2 = x^{\mu} x_{\mu} = -t^2 + \boldsymbol{x}^2.$$

The determinant of the Lorentz transformation implies that we have two types of the Lorentz transformations, namely, the *proper* Lorentz transformations and the *improper* Lorentz transformations which have $\det[\Lambda] = +1$ and $\det[\Lambda] = -1$, respectively [5]. It is trivial to show that the proper Lorentz transformations form

a subgroup $\mathrm{SO}(3,1)$ of the Lorentz group $\mathrm{O}(3,1)$. Now from the relation

$$g_{\mu\nu}\Lambda^{\mu}{}_{\rho}\Lambda^{\nu}{}_{\sigma} = g_{\rho\sigma}, \tag{3.2.2}$$

we get for $\mu = \nu = 0$

$$(\Lambda^{0}{}_{0})^2 - \Lambda^{i}{}_{0}\Lambda^{i}{}_{0} = 1,$$

which implies that either $\Lambda^{0}{}_{0} \geq +1$ or $\Lambda^{0}{}_{0} \leq -1$. Thus the proper Lorentz subgroup contains two disconnected components [15]. Lorentz transformations with $\Lambda^{0}{}_{0} \geq +1$ are called *orthochronous* Lorentz transformations and Lorentz transformations with $\Lambda^{0}{}_{0} \leq -1$ are called *nonorthochronous* Lorentz transformations [5]. One should note that any nonorthochronous transformation can be written as the product of an orthochronous transformation and a discrete temporal inversion $(x^0, x^1, x^2, x^3) \to (-x^0, x^1, x^2, x^3)$ or a discrete spatial inversion $(x^0, x^1, x^2, x^3) \to (x^0, -x^1, -x^2, -x^3)$ known as *time reversal* and *parity* transformations, respectively [15]. The orthochronous Lorentz transformations form a subgroup as well and when a theory is said to be Lorentz invariant, it is invariant under the proper orthochronous Lorentz subgroup [5]. We will define the proper orthochronous Lorentz subgroup to be the group $\mathrm{SO}(3,1)$ with $\Lambda^{0}{}_{0} \geq +1$ and will be denoted as $\mathrm{SO}(3,1)^{+}$. We will consider parity and time reversal transformations separately since it is well known now that there are physical processes that violate parity and/or time reversal.

Since the Lorentz group $\mathrm{O}(3,1)$ is a Lie group, we can write the infinitesimal Lorentz transformation up to first order as [5]

$$\Lambda^{\mu}{}_{\nu} = \delta^{\mu}{}_{\nu} + \delta\omega^{\mu}{}_{\nu}. \tag{3.2.3}$$

Then from equation (3.2.2)

$$g_{\mu\nu}\Lambda^{\mu}{}_{\rho}\Lambda^{\nu}{}_{\sigma} = g_{\rho\sigma}$$
$$g_{\mu\nu}(\delta^{\mu}{}_{\rho} + \delta\omega^{\mu}{}_{\rho})(\delta^{\nu}{}_{\sigma} + \delta\omega^{\nu}{}_{\sigma}) = g_{\rho\sigma}$$
$$\delta\omega_{\rho\sigma} = -\delta\omega_{\sigma\rho}, \tag{3.2.4}$$

where second order terms are neglected hence $\delta\omega$ is an antisymmetric matrix. From elementary linear algebra, the vector space of $n \times n$ antisymmetric matrices have dimension $\frac{1}{2}n(n-1)$ so in four spacetime dimensions, the Lorentz transformations have six independent parameters. These parameters are very obvious. We stated before that a pure rotation, which leaves the temporal component invariant, is identified as the group SO(3) thus three of the parameters are rotations in the (x, y), (x, z), and (y, z) planes. A pure boost, which leaves $-(x^0)^2 + \mathbf{x}^2$ invariant, acts on the (t, x), (t, y), and (t, z) planes so the three boosts are the other three parameters [15]. For rotations by an angle $\delta\theta$ about the unit vector $\hat{\mathbf{n}}$, we have $\delta\omega_{ij} = -\varepsilon_{ijk}\hat{n}_k\delta\theta$ and for a boost in the direction $\hat{\mathbf{n}}$ by rapidity $\delta\eta$, we have $\delta\omega_{i0} = \hat{n}_i\delta\eta$ [5]. One of the parameters of the Lorentz group is the modulus of the boost velocity $|\mathbf{v}| \in [0, 1)$ and we see that it ranges over a non-compact interval thus the Lorentz group is a non-compact Lie group [16, 12].

Now let us see how quantum operators should transform under Lorentz transformations. From Wigner's theorem, symmetries are represented as unitary or antiunitary operators in quantum physics so for each $\Lambda \in \mathrm{SO}(3,1)^+$ we define a unitary operator $\hat{U}(\Lambda)$ that obeys the composition property. An infinitesimal transformation can be written as

$$\hat{U}(1 + \delta\omega) = I + \frac{i}{2}\delta\omega_{\mu\nu}\hat{M}^{\mu\nu}, \tag{3.2.5}$$

where the second rank antisymmetric, Hermitian four-tensor operator, or tensor operator for short, $\hat{M}^{\mu\nu}$ is the generator of the Lorentz group [15]. The components of $\hat{M}^{\mu\nu}$ are an antisymmetric set of Hermitian operators that are the

generators of the Lorentz group as well. For some $\Lambda' = I + \delta\omega'$, let us expand $\hat{U}(\Lambda)^{-1}\hat{U}(\Lambda')\hat{U}(\Lambda) = \hat{U}(\Lambda^{-1}\Lambda'\Lambda)$,

$$\begin{aligned}
\hat{U}(\Lambda)^{-1}\hat{U}(\Lambda')\hat{U}(\Lambda) &= \hat{U}(\Lambda^{-1}\Lambda'\Lambda) \\
\hat{I} + \frac{i}{2}\delta\omega_{\mu\nu}\hat{U}(\Lambda)^{-1}\hat{M}^{\mu\nu}\hat{U}(\Lambda) &= \hat{U}(I + \Lambda^{-1}\delta\omega\Lambda) \\
\hat{I} + \frac{i}{2}\delta\omega_{\mu\nu}\hat{U}(\Lambda)^{-1}\hat{M}^{\mu\nu}(\Lambda) &= \hat{I} + \frac{i}{2}\delta\omega_{\mu\nu}\Lambda^{\mu}{}_{\rho}\Lambda^{\nu}{}_{\sigma}\hat{M}^{\rho\sigma} \\
\delta\omega_{\mu\nu}\hat{U}(\Lambda)^{-1}\hat{M}^{\mu\nu}\hat{U}(\Lambda) &= \delta\omega_{\mu\nu}\Lambda^{\mu}{}_{\rho}\Lambda^{\nu}{}_{\sigma}\hat{M}^{\rho\sigma}.
\end{aligned}$$

Since $\delta\omega_{\mu\nu}$ is arbitrary, we get the Lorentz transformation rule for antisymmetric tensors

$$\hat{U}(\Lambda)^{-1}\hat{M}^{\mu\nu}\hat{U}(\Lambda) = \Lambda^{\mu}{}_{\rho}\Lambda^{\nu}{}_{\sigma}\hat{M}^{\rho\sigma}, \tag{3.2.6}$$

where we see that after a Lorentz transformation, the second rank tensor operator $\hat{M}^{\rho\sigma}$ remains antisymmetric. It is trivial to see that this result applies to second rank symmetric four-tensor operators as well, namely, a second rank symmetric tensor operator remains a second rank symmetric tensor operator after a Lorentz transformation [15]. To find the Lie algebra of the Lorentz group, let $\Lambda = I + \delta\omega$ in equation (3.2.19) then expand to first order to get

$$[\hat{M}^{\mu\nu}, \hat{M}^{\rho\sigma}] = i(g^{\mu\rho}\hat{M}^{\nu\sigma} - g^{\nu\rho}\hat{M}^{\mu\sigma} - g^{\mu\sigma}\hat{M}^{\nu\rho} + g^{\nu\sigma}\hat{M}^{\mu\rho}),$$

which is the Lie algebra of the Lorentz group [5]. It will prove fruitful in the analysis of Lorentz symmetry to let $\hat{J}_i = \frac{1}{2}\varepsilon_{ijk}\hat{M}^{jk}$ and $\hat{K}_i = \hat{M}^{i0}$. Then the Lorentz Lie algebra in terms of $\hat{J}_i$ and $\hat{K}_i$ is

$$\begin{aligned}
[\hat{J}_i, \hat{J}_j] &= i\varepsilon_{ijk}\hat{J}_k, \\
[\hat{K}_i, \hat{K}_j] &= -i\varepsilon_{ijk}\hat{J}_k, \\
[\hat{J}_i, \hat{K}_j] &= i\varepsilon_{ijk}\hat{K}_k.
\end{aligned} \tag{3.2.7}$$

We recognize the first algebra to be the Lie algebra of Lie group SO(3) which is

the group of rotations in three-dimensions that was identified as one of the sets of generators of the Lorentz group. Thus the components $\hat{J}_i$ are to be identified as the components of the angular momentum operator $\hat{\mathbf{J}}$. The algebras that involve the components $\hat{K}_i$ are to be identified as the components of the boost operator $\hat{\mathbf{K}}$ but they do not form a closed algebra since boosts do not form a group. The second line says that a series of boosts is equivalent to a rotation while the third line says that the boost operator $\hat{\mathbf{K}}$ transforms as a three-vector under rotations [5].

From special relativity and nonrelativistic quantum mechanics, we can easily deduce that first rank four-tensor operators, or four-vector operators, transform under Lorentz transformations in the following manner

$$\hat{U}(\Lambda)^{-1}\hat{P}^{\mu}U(\Lambda) = \Lambda^{\mu}{}_{\nu}\hat{P}^{\nu} \tag{3.2.8}$$

where $\hat{P}^{\mu}$ is the Hermitian four-momentum operator [5]. Based on our knowledge of the components of the four-momentum from relativity, we identify $\hat{P}^0 = \hat{H}$ as the Hamiltonian and $\hat{P}^i$ is the momentum operator. Recall that the generators of spatial and temporal translational symmetries are the linear momentum and Hamiltonian, respectively. The four-momentum operator suggests that we can combine these two symmetries into the general spacetime translational symmetry and thus we construct the unitary *spacetime translation operator* $\hat{T}(a)$. An infinitesimal spacetime translation can be written as

$$\hat{T}(a) = \hat{I} - i\delta a_{\mu}\hat{P}^{\mu}. \tag{3.2.9}$$

and compounding multiple infinitesimal spacetime translations will result in a finite transformation which yields the following form for the spacetime translation operator

$$\hat{T}(a) = \exp\big(-i\hat{P}^{\mu}a_{\mu}\big) \equiv \exp\big(-i\hat{P}a\big), \tag{3.2.10}$$

where it is implied that $\hat{P}^\mu a_\mu = \hat{P}a$ [5]. We can immediately identify the commutation relations among the components of $\hat{P}^\mu$ from nonrelatvisitic quantum mechanics

$$\begin{aligned}[\hat{P}_i, \hat{P}_j] &= 0, \\ [\hat{P}_i, \hat{H}] &= 0. \end{aligned} \tag{3.2.11}$$

By applying an infinitesimal Lorentz transformation on equation (3.2.8), expanding and equating the coefficients on both sides yields

$$[\hat{P}^\mu, \hat{M}^{\rho\sigma}] = i(g^{\mu\sigma}\hat{P}^\rho - g^{\mu\rho}\hat{P}^\sigma). \tag{3.2.12}$$

Then we can use the algebra of angular momentum and boosts to write (3.2.12) in terms of $\hat{\mathbf{J}}$ and $\hat{\mathbf{K}}$

$$\begin{aligned}[\hat{J}_i, \hat{H}] &= 0, \\ [\hat{J}_i, \hat{P}_j] &= i\varepsilon_{ijk}\hat{P}_k, \\ [\hat{K}_i, \hat{H}] &= i\hat{P}_i, \\ [\hat{K}_i, \hat{P}_j] &= i\delta_{ij}\hat{H}. \end{aligned} \tag{3.2.13}$$

Compiling the algebra of spacetime translations and Lorentz transformations yields the closed algebra

$$\begin{gathered}[\hat{J}_i, \hat{H}] = 0, \quad [\hat{J}_i, \hat{P}_j] = i\varepsilon_{ijk}\hat{P}_k, \quad [\hat{K}_i, \hat{H}] = i\hat{P}_i, \quad [\hat{K}_i, \hat{P}_j] = i\delta_{ij}\hat{H}, \\ [\hat{J}_i, \hat{J}_j] = i\varepsilon_{ijk}\hat{J}_k, \quad [\hat{K}_i, \hat{K}_j] = -i\varepsilon_{ijk}\hat{J}_k, \quad [\hat{J}_i, \hat{K}_j] = i\varepsilon_{ijk}\hat{K}_k, \\ [\hat{P}_i, \hat{P}_j] = 0, \quad [\hat{P}_i, \hat{H}] = 0. \end{gathered} \tag{3.2.14}$$

The above commutation relations forms the Lie algebra of the *Poincaré group* [5]. We see that the generators of the Poincaré group consist of the generators of the Lorentz group and the translational group. Like the Lorentz group, the Poincaré

group is a non-compact Lie group [15].

We now seek to find how quantum field operators transform under Lorentz transformations. It is natural to work in the Heisenberg picture to show that time and space are of equal importance but instead of the usual definition of promoting Schrödinger operators to Heisenberg operators, we use the relativistic generalization

$$\hat{T}(a)\hat{\varphi}(0)\hat{T}^{-1}(a) = \hat{\varphi}(x), \tag{3.2.15}$$

where $\hat{T}(a)$ is the spacetime translation operator to get the relativisitc Heisenberg field operator. Again, it is crucial to recognize that the argument of the field is a four-vector. Consider a field operator with no Lorentz indices $\hat{\varphi}(x)$, namely, a quantum scalar field. From our derivation on how Hermitian operators transform under Lorentz transformations, we can easily deduce that scalar fields transform as

$$\hat{U}(\Lambda)^{-1}\hat{\varphi}(x)\hat{U}(\Lambda) = \hat{\varphi}(\Lambda^{-1}x). \tag{3.2.16}$$

With our knowledge on how four-gradients transform under Lorentz transformations, we can also see that

$$\begin{aligned}\hat{U}(\Lambda)^{-1}\partial^{\mu}\hat{\varphi}(x)\hat{U}(\Lambda) &= \Lambda^{\mu}{}_{\nu}\overline{\partial}^{\rho}\hat{\varphi}(\Lambda^{-1}x),\\ \hat{U}(\Lambda)^{-1}\partial^{2}\hat{\varphi}(x)\hat{U}(\Lambda) &= \Lambda^{\mu}{}_{\nu}\overline{\partial}^{2}\hat{\varphi}(\Lambda^{-1}x),\end{aligned} \tag{3.2.17}$$

where $\overline{\partial}$ indicates that we differentiate with respect to $\overline{x} = \Lambda^{-1}x$ [5]. A four-gradient acting on a scalar field creates a vector field so from the first line of eequation (3.2.17), we immediately identify the transformation property of a vector field $\hat{A}^{\mu}(x)$ as

$$\hat{U}(\Lambda)^{-1}\hat{A}^{\mu}(x)\hat{U}(\Lambda) = \Lambda^{\mu}{}_{\rho}\hat{A}^{\rho}(\Lambda^{-1}x), \tag{3.2.18}$$

and for tensor fields,

$$\hat{U}(\Lambda)^{-1}\hat{B}^{\mu\nu}(x)\hat{U}(\Lambda) = \Lambda^{\mu}{}_{\rho}\Lambda^{\nu}{}_{\sigma}\hat{B}^{\rho\sigma}(\Lambda^{-1}x). \tag{3.2.19}$$

In classical physics, there are many kinds of fields other than scalar and vector fields that are used to model physical phenomena. In quantum field theory, we find this to be true as well but quantum fields must transform in a Lorentz invariant manner. Thus we wish to find how fields with n Lorentz indices transform under Lorentz transformations, namely,

$$\hat{U}(\Lambda)^{-1}\hat{\varphi}_A(x)\hat{U}(\Lambda) = D_A{}^B(\Lambda)\hat{\varphi}_B(\Lambda^{-1}x), \tag{3.2.20}$$

where $\hat{\varphi}_A(x)$ and $\hat{\varphi}_B(x)$ are fields with an arbitrary number of Lorentz indices and $D_A{}^B(\Lambda)$ is a matrix that depends on Λ. Mathematically, we seek to find representations of the Lorentz group which will leave the field Lorentz invariant [16]. We will drop the operator hats in the when we state the definitions of the field representations since the field representations apply to the classical field from which we can easily quantize later on. To find the generators of a representations, if we consider an infinitesimal transformation

$$D_A{}^B(1+\delta\omega) = \delta_A{}^B + \frac{i}{2}\delta\omega_{\mu\nu}(S^{\mu\nu})_A{}^B \tag{3.2.21}$$

then equation (3.2.20) becomes

$$[\hat{\varphi}_A(x), \hat{M}^{\mu\nu}] = L^{\mu\nu}\hat{\varphi}_A(x) + (S^{\mu\nu})_A{}^B\hat{\varphi}_B(x), \tag{3.2.22}$$

where $L^{\mu\nu} = \frac{1}{i}(x^{\mu}\partial^{\nu} - x^{\nu}\partial^{\mu})$ and $S^{\mu\nu}$ each separately obey the commutations relations of the generators. Now there are many representations that can be found for the Lorentz group. One example is the tensor representation defined in equation (3.2.19) [15]. A generic four-tensor forms a basis for a 16-dimensional representation but this representation is reducible. Recall that symmetric and

antisymmetric tensors remain symmetric and antisymmetric, respectively, under Lorentz transformations thus the symmetric and antisymmetric components of $B^{\mu\nu}$ do not mix. Thus the representation reduces to a 10-dimensional symmetric representation $S^{\mu\nu}$ and an antisymmetric representation $A^{\mu\nu}$ [15]. Furthermore, if we take the trace of a symmetric tensor $T(x) \equiv g_{\mu\nu}S^{\mu\nu}$, we get that under Lorentz transformations,

$$\begin{aligned}\hat{U}(\Lambda)^{-1}\hat{T}(x)\hat{U}(\Lambda) &= g_{\mu\nu}\Lambda^{\mu}{}_{\rho}\Lambda^{\nu}{}_{\sigma}S^{\rho\sigma}(\Lambda^{-1}x) \\ &= g_{\rho\sigma}S^{\rho\sigma}(\Lambda^{-1}x) \\ &= \hat{T}(\Lambda^{-1}x), \end{aligned} \tag{3.2.23}$$

thus the trace of a symmetric tensor transforms as a one-dimensional Lorentz scalar. Thus the symmetric representation reduces into the irreducible, nine-dimensional traceless representation $S^{\mu\nu} - (1/4)T$ and the one dimensional scalar representation T [15]. By dimensions, the tensor representation is written as

$$\mathbf{4} \otimes \mathbf{4} = \mathbf{1} \oplus \mathbf{6} \oplus \mathbf{9}. \tag{3.2.24}$$

Thus the tensor product of two four-vector representations is the direct sum of the **1**, **6**, and **9** representations [15]. We now see that some representations are reducible and will have irreducible subcomponents. It is clear that finding and analyzing irreducible representations would prove more fundamental than finding reducible representations since the properties of reducible representations can be understood from its irreducible parts. In summary, we seek irreducible representations of the Lorentz group [5]. From the Lie algebra of the Lorentz group (3.2.7), we see that the SO(3) algebra is the same as the SU(2) Lie algebra. Indeed recall from nonrelativistic quantum mechanics that the representations of SU(2) are three Hermitian matrices that obey the same commutation relation of SO(3) and has a quadratic Casimir operator $\hat{\mathbf{J}}^2$ whose eigenvalues j can be used to label physical states. The index j can have values $j = 0, \frac{1}{2}, 1, \frac{3}{2}$.... However for $j = \frac{1}{2}$, a 2π

rotation results in a minus sign for the quantum state but a 2π rotation should leave the state unchanged for a representation of SO(3) [15]. Thus the half-integer representations serve as inequivalent, irreducible representations for SU(2). We want to include such representations in a relativistic quantum theory since it is known that there exist particles with spin-1/2 [15]. To do this, we complexify the Lorentz Lie algebra in the following manner [5]

$$\hat{\mathbf{N}} = \frac{\hat{\mathbf{J}} - i\hat{\mathbf{K}}}{2}, \tag{3.2.25}$$

$$\hat{\mathbf{N}}^{\dagger} = \frac{\hat{\mathbf{J}} + i\hat{\mathbf{K}}}{2}, \tag{3.2.26}$$

then the Lorentz Lie algebra becomes

$$[\hat{N}_i, \hat{N}_j] = i\varepsilon_{ijk}\hat{N}_k, \tag{3.2.27}$$

$$[\hat{N}_i^{\dagger}, \hat{N}_j^{\dagger}] = i\varepsilon_{ijk}\hat{N}_k^{\dagger}, \tag{3.2.28}$$

$$[\hat{N}_i, \hat{N}_j^{\dagger}] = 0. \tag{3.2.29}$$

We see that the Lorentz algebra can be written as two different SU(2) algebras that are related by Hermitian conjugation so the Lorentz algebra can be written as the algebra of $\mathrm{SU}(2) \times \mathrm{SU}(2)$. The quadratic Casimir operator is then $(\hat{\mathbf{N}}^2, \hat{\mathbf{N}}^{\dagger\,2})$ so we can label these representations as (n, n') and the dimension of the representation is $(2n+1)(2n'+1)$ [5]. Since $\hat{J}_i = \hat{N}_i + \hat{N}_i^{\dagger}$, we can label different components of a representation by their angular momentum representations. This is Clebsch-Gordon theory or addition of angular momentum from nonrelativistic quantum mechanics and the result is that the allowed values of j are $|n-n'|, |n-n'|+1, ..., |n+n'|$ [11]. The most common representations that are used and will be covered in this

paper are [5]

$$\begin{aligned}(0,0) &= \text{scalar},\\ \left(\frac{1}{2},0\right) &= \text{left-handed spinor},\\ \left(0,\frac{1}{2}\right) &= \text{right-handed spinor},\\ \left(\frac{1}{2},\frac{1}{2}\right) &= \text{vector}.\end{aligned} \tag{3.2.30}$$

Now we have explicit fields that can use to formulate quantum field theory. A scalar representation is easily defined.

Definition 3.2.2. The $(0,0)$ field representation of the Lorentz algebra is called a **scalar field** $\varphi(x)$ and transforms as

$$U(\Lambda)^{-1}\varphi(x)U(\Lambda) = \varphi(\Lambda^{-1}x),$$

under Lorentz transformations.

To find the $(\frac{1}{2},0)$ representation matrix for the left-handed spinor field $\hat{\psi}_a(x)$ where a can take on two values, consider

$$\hat{U}(\Lambda)^{-1}\hat{\psi}_a(x)\hat{U}(\Lambda) = (D_L)_a{}^b(\Lambda)\hat{\psi}_b(\Lambda^{-1}x) \tag{3.2.31}$$

and the infinitesimal transformation

$$(D_L)_a{}^b(I+\delta\omega) = \delta_a{}^b + \frac{i}{2}\delta\omega_{\mu\nu}(S_L^{\mu\nu})_a{}^b, \tag{3.2.32}$$

where $(S_L^{\mu\nu})_a{}^b = -(S_L^{\nu\mu})_a{}^b$. From nonrelativisitic quantum mechanics, the $j=\frac{1}{2}$ matrices are the *Pauli matrices* σ^i where

$$\sigma^1 = \begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}, \quad \sigma^2 = \begin{pmatrix}0 & -i\\ i & 0\end{pmatrix}, \quad \sigma^3 = \begin{pmatrix}1 & 0\\ 0 & -1\end{pmatrix}, \tag{3.2.33}$$

thus

$$(S_L^{ij})_a{}^b = \frac{1}{2}\varepsilon_{ijk}\sigma_k. \tag{3.2.34}$$

The SU(2) algebras are exchanged under Hermitian conjugation thus the Hermitian conjugate of a left-handed spinor $\hat{\psi}_a$ is the $(0, \frac{1}{2})$ right handed spinor $\hat{\psi}_{\dot{a}}$ where

$$(\hat{\psi}_a)^\dagger = \hat{\psi}^\dagger_{\dot{a}}. \tag{3.2.35}$$

We will denote a right-handed spinor with a dot on top of its Lorentz index. It is trivial to see then that

$$[(S_L^{\mu\nu})_a{}^b]^* = -(S_R^{\mu\nu})_{\dot{a}}{}^{\dot{b}}. \tag{3.2.36}$$

Compounding multiple infinitesimal transformations then yields the $(\frac{1}{2}, 0)$ and $(0, \frac{1}{2})$ finite transformation

$$D_L(\Lambda) = e^{-\frac{i}{2}\omega_{\mu\nu}S_L^{\mu\nu}}, \tag{3.2.37}$$

$$D_R(\Lambda) = e^{\frac{i}{2}\omega_{\mu\nu}S_R^{\mu\nu}}. \tag{3.2.38}$$

Now we can properly define left and right-handed spinor fields.

Definition 3.2.3. The $(\frac{1}{2}, 0)$ field representation of the Lorentz algebra is called the **left-handed spinor field**, or **left-handed Weyl field**, $\psi_a(x)$ where the index a has two values. The $(0, \frac{1}{2})$ field representation is called the **right-handed spinor field**, or **right-handed Weyl field**, $\psi_{\dot{a}}(x)$ where the index $\dot{a}$ has two values. The fields transform as

$$U(\Lambda)^{-1}\psi_a(x)U(\Lambda) = (D_L)_a{}^b(\Lambda)\psi_b(\Lambda^{-1}x)$$
$$U(\Lambda)^{-1}\psi_{\dot{a}}(x)U(\Lambda) = (D_R)_{\dot{a}}{}^{\dot{b}}(\Lambda)\psi_{\dot{b}}(\Lambda^{-1}x),$$

where

$$
\begin{aligned}
D_L(\Lambda) &= e^{-\frac{i}{2}\omega_{\mu\nu}S_L^{\mu\nu}}, \\
D_R(\Lambda) &= e^{\frac{i}{2}\omega_{\mu\nu}S_R^{\mu\nu}}, \\
(S_L^{\mu\nu})_a{}^b &= +\frac{i}{4}(\sigma^\mu\bar{\sigma}^\nu - \sigma^\nu\bar{\sigma}^\mu)_a{}^b, \\
(S_R^{\mu\nu})^{\dot{a}}{}_{\dot{b}} &= -\frac{i}{4}(\bar{\sigma}^\mu\sigma^\nu - \bar{\sigma}^\nu\sigma^\mu)^{\dot{a}}{}_{\dot{b}}
\end{aligned}
$$

To manipulate the spinor indices, we introduce the antisymmetric tensor ε_{ab} which has the property

$$(D_L)_a{}^c(\Lambda)(D_L)_b{}^d(\Lambda)\varepsilon_{cd} = \varepsilon_{ab}, \tag{3.2.39}$$

thus ε_{ab} is an invariant symbol of the Lorentz group like the metric tensor $g^{\mu\nu}$ [5]. We define ε_{ab} the following way

$$\varepsilon^{12} = \varepsilon_{21} = 1, \quad \varepsilon^{21} = \varepsilon_{12} = -1, \tag{3.2.40}$$

then we see that [5]

$$\varepsilon_{ab}\varepsilon^{bc} = \delta_a{}^c, \quad \varepsilon^{ab}\varepsilon_{bc} = \delta^a{}_c. \tag{3.2.41}$$

With this, we can lower and raise indices in the following manner

$$\psi^a(x) = \varepsilon^{ab}\psi_b(x). \tag{3.2.42}$$

For vector fields that possess one dotted and one undotted index $A_{a\dot{a}}(x)$, we can use the invariant symbol $\sigma^\mu_{a\dot{a}}$ such that [5]

$$A_{a\dot{a}}(x) = \sigma^\mu_{a\dot{a}}A_\mu(x). \tag{3.2.43}$$

By convention we choose $\sigma^{\mu}_{a\dot{a}}$ to have the value

$$\sigma^{\mu}_{a\dot{a}} = (I, \boldsymbol{\sigma}), \tag{3.2.44}$$

where $\boldsymbol{\sigma}$ is the Pauli matrix vector. Along with this, we introduce the following symbol

$$\overline{\sigma}^{\mu\dot{a}a} = \varepsilon^{ab}\varepsilon^{\dot{a}\dot{b}}\sigma^{\mu}_{b\dot{b}} = (I, -\boldsymbol{\sigma}). \tag{3.2.45}$$

It is rather cumbersome to consistently use these indices for the Weyl fields thus an index free notation method that we will use is

$$\chi\psi = \chi^a\psi_a, \quad \chi^\dagger\psi^\dagger = \chi^\dagger_{\dot{a}}\psi^{\dagger\dot{a}}, \tag{3.2.46}$$

namely, missing contracted undotted and dotted indices are implicitly understood to be written as ${}^{c}{}_{c}$ and ${}_{\dot{c}}{}^{\dot{c}}$ [5]. With this, we have the simplification

$$\begin{aligned}(\chi\psi)^\dagger &= (\chi^a\psi_a)^\dagger \\ &= (\psi_a)^\dagger(\chi^a)^\dagger \\ &= \psi^\dagger_{\dot{a}}\chi^{\dagger\dot{a}} \\ &= \psi^\dagger\chi^\dagger.\end{aligned} \tag{3.2.47}$$

By now, the astute reader may have noticed that for two left-handed (right-handed) Weyl fields we have

$$\chi_a(x)\psi_a(y) = -\psi_b(y)\chi_a(x), \tag{3.2.48}$$

namely, the Weyl fields anticommute and since they in the $(\frac{1}{2}, 0)$ $((\frac{1}{2}, 0))$ representation, the Weyl fields will describe fermions. We will delve further into this later on once we derive the fermion fields.

The vector field is now defined.

Definition 3.2.4. The $(\frac{1}{2}, \frac{1}{2})$ field representation of the Lorentz algebra is called

a **vector field** $A^\mu(x)$ and under Lorentz transformations transforms as

$$U(\Lambda)^{-1}A^\mu(x)U(\Lambda)^{-1} = D_V(\Lambda)A^\rho(\Lambda^{-1}x),$$

where

$$D_V(\Lambda) = e^{-i\theta S_V^{\mu\nu}},$$
$$(S_V^{\mu\nu})^\rho{}_\tau = \frac{1}{i}(g^{\mu\rho}\delta^\nu{}_\tau - g^{\nu\rho}\delta^\mu{}_\tau).$$

We will encounter two more field representations that will describe other spin-1/2 fields.

3.3 Symmetries of Nature

Noether's Theorem and Continuous Symmetries

In the last section we covered many of the symmetries that are part of the Lorentz group. In this section, we will explore and give physical meaning to the mathematical abstraction of the previous section and explore the fundamental importance of symmetries in quantum field theory. We begin with the definition of a symmetry.

Definition 3.3.1. Let $\varphi_a(x)$ be a set of fields for $a = 1, \ldots, n$ and consider the infinitesimal transformation for each of the fields

$$x_\mu \to x'_\mu \qquad = x_\mu + \delta\omega_{\mu\nu}x^\nu,$$
$$\varphi_a(x) \to \varphi'_a(x') = \varphi_a(x) + \delta\varphi_a(x),$$

where $\delta\omega_{\mu\nu} = -\delta\omega_{\nu\mu}$. If the transformation leaves the action $\mathcal{S}$ invariant, the transformation is called a **symmetry transformation**, or a **symmetry**. Equivalently,

a symmetry transformation leaves the Lagrangian invariant up to a four-divergence

$$\mathcal{L}(x) \to \delta\mathcal{L}(x) = \mathcal{L}(x) + \partial_\mu K^\mu(x),$$

for some $K^\mu(x)$.

In physics, the symmetries of a system give rise to corresponding conservation laws. This fact is encoded in Noether's theorem.

Theorem 3.3.1 (**Noether's Theorem**). *Every differentiable symmetry of the action of a physical system has a corresponding conserved current.*

Proof. Let $\varphi_a(x)$ be a set of fields for $a = 1, \ldots, n$ with Lagrangian $\mathcal{L}(x)$. By varying the Lagrangian with respect to the fields we get

$$\begin{aligned}
\delta\mathcal{L}(x) &= \frac{\partial\mathcal{L}}{\partial\varphi_a}\delta\varphi_a + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi_a)}\delta(\partial_\mu\varphi_a) + \frac{\partial\mathcal{L}}{\partial x_\mu}\delta x_\mu \\
&= \partial_\mu\Big(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi_a)}\Big)\delta\varphi_a + \frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi_a)}\delta(\partial_\mu\varphi_a) + \partial^\nu\mathcal{L}(x) \\
&= \partial_\mu\Big(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi_a)}\delta\varphi_a\Big) + g^{\mu\nu}\partial_\mu\mathcal{L}(x) \\
&= \partial_\mu\Big(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi_a)}\delta\varphi_a + K^\mu(x)\Big),
\end{aligned}$$

where we used the Euler-Lagrange equations to get the first term of the second line. Then we have

$$\partial_\mu j_a^\mu(x) = \delta\mathcal{L}(x), \tag{3.3.1}$$

where

$$j_a^\mu(x) = \frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi_a)}\delta\varphi_a + K^\mu(x), \tag{3.3.2}$$

are the *Noether currents*. If the transformation leaves the Lagrangian invariant, namely, $\delta\mathcal{L} = 0$ and makes the four-divergence vanish $K^\mu = 0$ then we have

$$\partial_\mu\Big(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi_a)}\delta\varphi_a\Big) = 0 \quad (K^\mu = 0). \tag{3.3.3}$$

If the transformation leaves the Lagrangian invariant but the four-divergence does

not vanish, then we have

$$\partial_\mu\Big(\frac{\partial\mathcal{L}}{\partial(\partial_\mu\varphi_a)}\delta\varphi_a + K^\mu(x)\Big) = 0 \quad (K^\mu \neq 0). \tag{3.3.4}$$

In either cases, the Noether currents are conserved and each satisfies the continuity equation

$$\partial_\mu j^\mu(x) = \frac{\partial}{\partial t}j^0(x) + \nabla\cdot\mathbf{j}(x) = 0 \tag{3.3.5}$$

for $a = 1, \ldots, n$. The corresponding *Noether charges* are

$$Q_a = \int d^3x\, j_a^0(x). \tag{3.3.6}$$

Since this holds for n arbitrary fields, there are n conserved Noether currents. □

Note that since each field transforms differently depending on their representation, we can identify the following infinitesimal transformations

$$\begin{aligned}
\delta\varphi_a(x) &= -\frac{\partial\varphi_a}{\partial x_\mu}\delta x_\mu, && \text{(scalar fields)}\\
\delta\varphi_a(x) &= -\frac{i}{2}\delta\omega_{\mu\nu}S^{\mu\nu}\varphi_a(x) - \frac{\partial\varphi_a}{\partial x_\mu}\delta x_\mu, && \text{(spinor fields)}
\end{aligned} \tag{3.3.7}$$

where $\delta\omega_{\mu\nu} = -\delta\omega_{\nu\mu}$ and $S^{\mu\nu}$ is the generator of the Lorentz transformations for a general spinor field.

We call $j^0(x)$ and $\mathbf{j}(x)$ the *charge density* and *current density*, respectively, thus equation (3.3.5) is the *local* conservation law of this charge [5]. We should note that the names "charge" and "current" do not always refer to electric charge and electric current, respectively. The names refer to a general class of the generators of a symmetry transformation. One such example is the *color charge* of quarks which is a completely different property than the standard electric charge [1]. We see that the number of generators of a symmetry group, or its dimension, corresponds to the number of conservation laws a physical system is expected to have.

We now see the beautiful connection between symmetries and the conservation

laws that arise from them. Let us explore symmetries that will have relevance in later chapters. Consider the symmetry of spacetime translations. Given a set of scalar fields $\varphi_a(x)$, we have that

$$\delta\varphi_a = -\frac{\partial\varphi_a}{\partial x_\mu}\delta x_\mu = -\partial_\mu\varphi_a(x), \tag{3.3.8}$$

we expect that under spacetime translations

$$\begin{aligned}
\varphi_a(x) &\to \varphi_a(x-a) \\
&= \varphi_a(x) - a^\nu\partial_\nu\varphi_a(x), \\
\mathcal{L}(x) &\to \mathcal{L}(x-a) \\
&= \mathcal{L}(x) - a^\nu\partial_\nu\mathcal{L}(x),
\end{aligned} \tag{3.3.9}$$

thus $\delta\varphi_a(x) = -a^\nu\partial_\nu\varphi_a(x)$ and $\delta\mathcal{L}(x) = -a^\nu\partial_\nu\mathcal{L}(x) = -\partial_\nu(a^\nu\mathcal{L}(x))$ [5]. Then the total divergence is $K^\nu(x) = -a^\nu\mathcal{L}(x)$ thus the conserved current is

$$\begin{aligned}
j^\mu(x) &= \frac{\partial\mathcal{L}(x)}{\partial(\partial_\mu\varphi_a(x))}(-a^\nu\partial_\nu\varphi_a(x)) + a^\nu\mathcal{L}(x) \\
&= a_\nu T^{\mu\nu}(x),
\end{aligned} \tag{3.3.10}$$

where

$$T^{\mu\nu}(x) = -\frac{\partial\mathcal{L}(x)}{\partial(\partial_\mu\varphi_a(x))}\partial^\nu\varphi_a(x) + g^{\mu\nu}\mathcal{L}(x), \tag{3.3.11}$$

is the *energy-momentum tensor* [8]. To identify the conserved quantities that arise from spacetime translations, let us examine the components of the tensor. For

$\mu = \nu = 0$ we have

$$\begin{aligned} T^{00} &= -\frac{\partial \mathcal{L}}{\partial(\partial_0 \varphi_a)} \partial^0 \varphi_a + g^{00} \mathcal{L} \\ &= \frac{\partial \mathcal{L}}{\partial(\dot{\varphi}_a)} \dot{\varphi}_a - \mathcal{L} \\ &= \Pi_a \dot{\varphi}_a - \mathcal{L} \\ &= \mathcal{H}, \end{aligned} \tag{3.3.12}$$

thus we see that the Hamiltonian density is the conserved quantity under time translation, namely, the energy is the generator of time translations. Then by the definition of four-momentum, we immediately identify the other components $T^{0\mu}$ to be the *energy-momentum four-vector*

$$P^\mu = \int d^3x \, T^{0\mu}(x), \tag{3.3.13}$$

hence the *spatial momentum vector* is

$$\begin{aligned} P^i &= \int d^3x \, T^{0i} \\ &= -\int d^3x \, \Pi_i(x) \nabla^i \varphi(x), \end{aligned} \tag{3.3.14}$$

thus the linear momentum is the generator of spatial translations [5]. Then we see that the energy-momentum four-vector

$$P^\mu = \int d^3x \, T^{0\mu}(x),$$

is the conserved charge of spacetime translations. Thus spacetime translational symmetry gives rise to the conservation of energy and linear momentum. Another symmetry that will be important to consider is of course, Lorentz symmetry. We find that the Noether current is

$$\mathcal{M}^{\mu\nu\rho}(x) = x^\nu T^{\mu\rho}(x) - x^\rho T^{\mu\nu}(x), \tag{3.3.15}$$

with the conserved charges

$$M^{\nu\rho} = \int d^3x\, \mathcal{M}^{0\nu\rho}(x), \tag{3.3.16}$$

which are the generators of the Lorentz group. Thus Lorentz symmetry gives rise to six conservation laws which are angular momentum (in three dimensions) and the velocity of the center of mass (in three dimensions) which come from rotations and boosts, respectively [5]. If we consider a scalar field and let $\nu = i$ and $\rho = j$ in equation (3.3.15) then insert it into equation (3.3.16), we have

$$M^{ij} = \int d^3x\, x^i T^{0j} - x^j T^{0i}, \tag{3.3.17}$$

where T^{0i} are the spatial momentum densities. We recognize this expression to be field-theoretic analogue of orbital angular momentum thus we can rewrite (3.3.17) as

$$L^{ij} = \int d^3x\, x^i T^{0j} - x^j T^{0i}, \tag{3.3.18}$$

hence the field-theoretic total angular momentum for scalar fields is

$$J^{ij} = L^{ij}. \tag{3.3.19}$$

If we considered the Noether currents for Lorentz symmetry of a set of spinor fields $\varphi_a(x)$ such as the Weyl fields for $a = 1, \ldots, n$, the Noether currents would be [5]

$$\mathcal{M}^{\mu\nu\rho} = x^\nu T^{\mu\rho} - x^\rho T^{\mu\nu} - i\frac{\partial \mathcal{L}}{\partial(\partial_\mu \varphi_a)} S^{\nu\rho}\varphi_a, \tag{3.3.20}$$

with the Noether charges

$$\begin{aligned} M^{\nu\rho} &= \int d^3x\, \mathcal{M}^{0\nu\rho} \\ &= \int d^3x\, x^\nu T^{0\rho} - x^\rho T^{0\nu} - i\frac{\partial \mathcal{L}}{\partial(\partial_0 \varphi_a)} S^{\nu\rho}\varphi_a \\ &= \int d^3x\, x^\nu T^{0\rho} - x^\rho T^{0\nu} - i\Pi_a\, S^{\nu\rho}\varphi_a, \end{aligned} \tag{3.3.21}$$

where $S^{\mu\nu}$ depends on the representation. If we let $\nu = i$ and $\rho = j$ in equation (3.3.21) then we have

$$M^{ij} = \int d^3x\, x^i T^{0j} - x^j T^{0i} - i\Pi_a\, S^{ij}\varphi_a. \tag{3.3.22}$$

We recognize the first two terms in equation (3.3.22) as the orbital momentum L^{ij} and we identify the third term as the field-theoretic spin. Thus we have

$$S^{ij} = -\int d^3x\, i\Pi_a\, \Sigma^{ij}\varphi_a, \tag{3.3.23}$$

where we have merely switched the notations $S^{ij} \leftrightarrow \Sigma^{ij}$ for clarity [15]. The field-theoretic total angular momentum for spinor fields is

$$J^{ij} = L^{ij} + S^{ij}. \tag{3.3.24}$$

We now see that the concept of spin naturally arises from the nontrivial representations of the Lorentz group. For scalar fields, the total angular momentum lacks a spin component which is in agreement with the definition of the trivial representation of the Lorentz group. Along with this, we know from nonrelativistic quantum mechanics that scalar particles have spin-0 and fermions have spin-1/2 so we see that once we quantize the total angular momentum into a self-adjoint operator, we will naturally have spin (or no spin for quantized scalar fields) built into quantum field theory. It is now clear that the labeling of the representations of the Lorentz group serve as the labels for the spin of the fields.

If our Lagrangian consists of complex fields, we can explore a rather obscure but fundamental symmetry which is the invariance under a phase, or U(1), transformation

$$\varphi(x) \to e^{i\alpha}\varphi(x), \tag{3.3.25}$$

where $\alpha \in \mathbb{R}$. Repeating the same method as above, the U(1) Noether current is

$$j^{\mu} = \mathrm{Im}(\varphi^{\dagger}\overleftrightarrow{\partial^{\mu}}\varphi). \tag{3.3.26}$$

Recall that the Lie group U(1) is isomorphic to the Lie group SO(2) and a complex field $\varphi(x)$ can be written as the sum of two real fields

$$\varphi(x) = \frac{\varphi_1(x) + i\varphi_2(x)}{\sqrt{2}}, \tag{3.3.27}$$

thus we can consider SO(2) symmetry for a two-component field vector (φ_1, φ_2)

$$\begin{pmatrix} \varphi_1(x) \\ \varphi_2(x) \end{pmatrix} \to \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix} \begin{pmatrix} \varphi_1(x) \\ \varphi_2(x) \end{pmatrix}, \tag{3.3.28}$$

where α is the angle of rotation [5]. The SO(2) Noether current is then

$$j^{\mu} = (\varphi_1\overleftrightarrow{\partial^{\mu}}\varphi_2), \tag{3.3.29}$$

which is equivalent to the U(1) Noether current. The U(1) Noether charge is then

$$\begin{aligned} Q &= \int d^3x\, j^0(x) \\ &= \int d^3x\ = \mathrm{Im}(\varphi^{\dagger}\overleftrightarrow{\partial^0}\varphi), \end{aligned} \tag{3.3.30}$$

which is constant in time. When we make the transition to quantum field theory, the U(1) Noether charge will become an operator and the conservation law that is associated with U(1) symmetry will be clear.

Discrete Symmetries

So far we have only discussed continuous symmetries so let us now consider discrete symmetries. From our earlier analysis on Lorentz symmetry, we encountered two discrete symmetries, namely, parity and time reversal symmetry

$$\mathcal{P}^{\mu}{}_{\nu} = (\mathcal{P}^{-1})^{\mu}{}_{\nu} = \begin{pmatrix} +1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}, \tag{3.3.31}$$

(3.3.32)

$$\mathcal{T}^{\mu}{}_{\nu} = (\mathcal{T}^{-1})^{\mu}{}_{\nu} = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & +1 & 0 & 0 \\ 0 & 0 & +1 & 0 \\ 0 & 0 & 0 & +1 \end{pmatrix}, \tag{3.3.33}$$

respectively [5]. The operators for parity and time reversal are

$$\hat{P} \equiv \hat{U}(\mathcal{P}), \tag{3.3.34}$$

$$\hat{T} \equiv \hat{U}(\mathcal{T}). \tag{3.3.35}$$

From nonrelativistic quantum mechanics, recall that the operator for time reversal is an antiunitary operator while the operator for parity is unitary [11]. Since the parity and time reversal matrices are idempotent, applying a parity and time reversal transformation twice should leave all observables invariant so we see that under a parity and time reversal transformation, the field operators can transform as

$$\hat{P}^{-1}\hat{\varphi}(x)\hat{P} = \pm\hat{\varphi}(\mathcal{P}x), \tag{3.3.36}$$

$$\hat{T}^{-1}\hat{\varphi}(x)\hat{T} = \pm\hat{\varphi}(\mathcal{T}x). \tag{3.3.37}$$

If a field has a minus sign under these transformations, we say that the field transforms as a *pseudoscalar* and is odd under these transformations. If a field has a plus sign under these transformations, we say that the field transforms as a *scalar* and is even under these transformations. The guiding principle is that we must make our Lagrangian even under parity and time reversal

$$\hat{P}^{-1}\mathcal{L}(x)\hat{P} = +\mathcal{L}(\mathcal{P}x), \tag{3.3.38}$$

$$\hat{T}^{-1}\mathcal{L}(x)\hat{T} = +\mathcal{L}(\mathcal{T}x), \tag{3.3.39}$$

so we are at liberty to choose the sign of the transformation of the field operators as long as the Lagrangian is even under these transformations [5]. Then the action $\mathcal{S}$ is invariant under parity and time reversal thus both these symmetries are conserved.

Another class of discrete symmetries are symmetries that change the sign of fields without changing their spacetime argument is called Z_2 symmetry where Z_2 is the additive group of integers modulo 2 [5]. A unitary operator $\hat{Z}$ called the Z_2 *operator* has the transformation

$$\hat{Z}^{-1}\hat{\varphi}(x)\hat{Z} = \eta_a\hat{\varphi}_a(x), \tag{3.3.40}$$

where η_a is either $+1$ or -1 for each field $\hat{\varphi}_a(x)$. To give an explicit example, a Lagrangian that possesses U(1) symmetry will automatically have the discrete symmetry

$$\hat{\varphi}(x) \to \hat{\varphi}^\dagger(x), \tag{3.3.41}$$

which is equivalent to

$$\begin{pmatrix} \varphi_1(x) \\ \varphi_2(x) \end{pmatrix} \to \begin{pmatrix} +1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} \varphi_1(x) \\ \varphi_2(x) \end{pmatrix}, \tag{3.3.42}$$

for a Lagrangian with two real scalar fields. This discrete symmetry is called *charge*

conjugation and we denote the *charge conjugation operator* as $\hat{C}$ which transforms a complex field as

$$\hat{C}^{-1}\hat{\varphi}(x)\hat{C} = \hat{\varphi}^{\dagger}(x), \tag{3.3.43}$$

or

$$\hat{C}^{-1}\hat{\varphi}_1(x)\hat{C} = +\hat{\varphi}_1(x), \tag{3.3.44}$$

$$\hat{C}^{-1}\hat{\varphi}_2(x)\hat{C} = -\hat{\varphi}_2(x), \tag{3.3.45}$$

for SO(2) symmetries. For our discussion on fermion fields, it will prove fruitful to introduce the *charge conjugation matrix* [5]

$$\mathcal{C} = \begin{pmatrix} \varepsilon_{ab} & 0 \\ 0 & \varepsilon^{\dot{a}\dot{c}} \end{pmatrix}. \tag{3.3.46}$$

Now we will see later on that a Lagrangian that possesses U(1) will give rise to two types of particles, say, a-type particles and b-type particles. Charge conjugation means that the scattering amplitudes are invariant if we exchange a-type particles with b-type particles and vice versa which implies that both types of particles possess the same mass [5]. This means that the b-type particle is the antiparticle of the a-type particle. We will explore this further in the later chapters.

Experiments confirm that all Lagrangians obey another fundamental symmetry called *charge, parity, and time reversal symmetry*, or CPT symmetry for short. The combination of the operators CPT is the only combination where all Lagrangians are invariant. We will accept this fundamental symmetry without proof for all our Lagrangians.

3.4 Construction of an Ideal Quantum Field

Every Lagrangian for a quantum field must possess Lorentz symmetry. This means that every field must transform according to the representation of the Lorentz group

and every other term in the Lagrangian must transform in a Lorentz invariant way. By Noether's theorem, this means that all Lagrangians will have at least ten conservation laws along with other conservation laws from any additional symmetries. Along with this, all Lagrangians must be invariant under CPT symmetry. With these two rules, we can begin to construct our first quantum fields based on the field representations of the Lorentz algebra.

Chapter 4

Quantum Fields

4.1 Scalar Fields

Classical Real Scalar Field

We begin with the simplest model of a quantum field theory by first constructing a classical field theory then quantizing it using the canonical quantization rules. Many of the insights and results will be generalized to other quantum field theories thus the derivation and analysis of the real scalar field will prove to be quite insightful.

The simplest Lorentz invariant field is the trivial representation of the Lorentz algebra, namely, the $(0,0)$ representation which is the *classical* scalar field that we defined in the symmetry chapter. More specifically, we will focus on the *real* scalar field and since we are working in Minkowski space, we define the real scalar field as

$$\varphi : \mathbb{R}^{3,1} \to \mathbb{R}. \tag{4.1.1}$$

By explicit construction, given four vectors x^μ and $x^{\mu\prime} = \Lambda^\mu{}_\nu x^\nu + a^\nu$, the field transforms as a *Lorentz scalar*, namely, $\varphi'(x') = \varphi(x)$ which implies that the equations of motion must be the same for both $\varphi(x)$ and $\varphi'(x')$. A suitable equation

of motion that fulfills this condition is the previous Klein-Gordon equation

$$(-\partial^2 + m^2)\varphi(x) = 0. \tag{4.1.2}$$

For this classical analysis, we will treat the mass term m as an arbitrary parameter and will reinterpret it as mass once we construct our quantum the theory.

We seek a Lagrangian $\mathcal{L}$ for a real scalar field that transforms as a Lorentz scalar so that the action $\mathcal{S}$ is Lorentz invariant. We also require our Lagrangian to reproduce the (classical) Klein-Gordon equation as the equation of motion. Consider the ansatz,

$$\mathcal{L} = -\frac{1}{2}\partial^\mu\varphi\partial_\mu\varphi - \frac{1}{2}m^2\varphi^2. \tag{4.1.3}$$

We prove that our ansatz is Lorentz invariant.

Proof. Let $\mathcal{L}$ be defined as in equation (4.1.3). Under a Lorentz transformation, the following transforms as follows,

$$\begin{aligned} x^\mu \to x' &= \Lambda^\mu{}_\nu x^\nu \\ \varphi(x) \to \varphi'(x') &= \varphi(\Lambda^{-1}x) \\ \partial_\mu\varphi(x) \to \partial_\mu\varphi(\Lambda^{-1}x) &= (\Lambda^{-1})^\nu{}_\mu(\partial_\nu\varphi)(\Lambda^{-1}x). \end{aligned}$$

Then under a Lorentz transformation, the Klein-Gordon Lagrangian transforms as

$$\begin{aligned} \mathcal{L} \to & -\frac{1}{2}\partial^\mu\varphi'(x')\partial_\mu\varphi'(x') - \frac{1}{2}m^2(\varphi'(x'))^2 \\ &= -g^{\mu\nu}(\Lambda^{-1})^\rho{}_\mu(\Lambda^{-1})^\sigma{}_\mu\frac{1}{2}\partial_\rho\varphi(\Lambda^{-1}x)\partial_\sigma\varphi(\Lambda^{-1}x) - \frac{1}{2}m^2\varphi^2(\Lambda^{-1}x) \\ &= -\frac{1}{2}g^{\rho\sigma}\partial_\rho\varphi(\Lambda^{-1}x)\partial_\sigma\varphi(\Lambda^{-1}x) - \frac{1}{2}m^2\varphi^2(\Lambda^{-1}x) \\ &= -\frac{1}{2}\partial^\sigma\varphi(\Lambda^{-1}x)\partial_\sigma\varphi(\Lambda^{-1}x) - \frac{1}{2}m^2\varphi^2(\Lambda^{-1}x) \\ &= \mathcal{L}(\Lambda^{-1}x). \end{aligned}$$

Thus the Klein-Gordon Lagrangian transforms as a Lorentz scalar. □

We apply an infinitesimal variation $\delta\varphi$ and use the principle of stationary action on the Lagrangian to obtain

$$\begin{gathered}\delta\mathcal{S} = 0 \\ \delta \int d^4x\, \mathcal{L}(\varphi, \partial_\mu\varphi) = 0 \\ \int d^4x \Big[-\frac{1}{2}\partial^\mu\delta\varphi\partial_\mu\varphi - -\frac{1}{2}\partial^\mu\varphi\partial_\mu\delta\phi - m^2\varphi\delta\varphi\Big] = 0 \\ -\frac{1}{2}\partial_\mu\varphi\delta\varphi\Big| + \int d^4x \Big[\partial^\mu\partial_\mu\varphi - m^2\varphi\Big]\delta\varphi = 0\end{gathered}$$

where integration by parts was used on the first term of the third equation to derive the last line [5]. We impose the boundary condition that $\delta\varphi$ vanishes at the infinities in both directions of spacetime thus the first term of the last line vanishes and we have

$$\int d^4x \Big[\partial^\mu\partial_\mu\varphi - m^2\varphi\Big]\delta\varphi = 0. \tag{4.1.4}$$

Since $\delta\varphi$ only depends on some x, the above equation is true if and only if the bracketed expression equals zero thus we have our equation of motion

$$(-\partial^2 + m^2)\varphi = 0, \tag{4.1.5}$$

which is the desired Klein Gordon equation.

Similar to the free particle in nonrelativistic quantum mechanics, the Klein-Gordon equation in classical field theory admits plane wave solutions

$$\varphi(\mathbf{x}, t) = Ae^{i(\mathbf{p}\cdot\mathbf{x}\pm\omega t)},$$

where

$$\omega = (|\mathbf{p}|^2 + m^2)^{1/2}, \tag{4.1.6}$$

which is the relativistic energy in natural units ($E = \hbar\omega \equiv \omega$). The general solution

of the Klein-Gordon equation is the sum of plane waves,

$$\varphi(\mathbf{x},t) = \int \frac{d^3p}{f(p)}\Big[a(\mathbf{p})e^{i(\mathbf{p}\cdot\mathbf{x}-\omega t)} + b(\mathbf{p})e^{i(\mathbf{p}\cdot\mathbf{x}+\omega t)}\Big] \tag{4.1.7}$$

where $a(\mathbf{p})$ and $b(\mathbf{p})$ are functions of the momentum $\mathbf{p}$ and $f(p)$ is a function of the magnitude of $|\mathbf{p}|$ [5]. We seek a value of $f(p)$ so that the integration measure $\frac{d^3p}{f(p)}$ is Lorentz invariant. Consider the following identities

$$\delta(f(x) - f(x_0)) = \frac{1}{|f'(x)|}\delta(x - x_0), \tag{4.1.8}$$

$$\int_{-\infty}^{+\infty} \frac{1}{|f'(x_0)|}\,\delta(x - x_0) = \frac{1}{|f'(x_0)|}. \tag{4.1.9}$$

The measure $d^4p\,\delta(p^2 + m^2)\theta(p^0)$, where $\theta(p^0)$ is the unit step function, is Lorentz invariant so we integrate with respect to p^0 to get [5]

$$\int_{-\infty}^{+\infty} dp^0\,\delta(p^2 + m^2) = \frac{1}{2\omega}. \tag{4.1.10}$$

For purposes of normalization when we quantize the Klein-Gordon field, we will choose $f(p) = (2\pi)^3 2\omega$ and define our new Lorentz invariant measure as [5]

$$\frac{d^3p}{(2\pi)^3 2\omega} = d^3p_L \tag{4.1.11}$$

so that equation (4.1.7) now becomes

$$\varphi(\mathbf{x},t) = \int d^3p_L\,\Big[a(\mathbf{p})e^{i(\mathbf{p}\cdot\mathbf{x}-\omega t)} + b(\mathbf{p})e^{i(\mathbf{p}\cdot\mathbf{x}+\omega t)}\Big]. \tag{4.1.12}$$

Since we are considering real scalar fields, we impose the condition that $\varphi^* = \varphi$ from which we see that

$$\begin{aligned}\varphi^*(\mathbf{x},t) &= \int d^3p_L\,\Big[a^*(\mathbf{p})e^{-i(\mathbf{p}\cdot\mathbf{x}-\omega t)} + b^*(\mathbf{p})e^{-i(\mathbf{p}\cdot\mathbf{x}+\omega t)}\Big] \\ &= \int d^3p_L\,\Big[a^*(\mathbf{p})e^{-i\mathbf{p}\cdot\mathbf{x}+i\omega t} + b^*(-\mathbf{p})e^{i\mathbf{p}\cdot\mathbf{x}-i\omega t}\Big].\end{aligned} \tag{4.1.13}$$

Comparing equations (4.1.7) and (4.1.13) reveals that $a(\mathbf{p}) = b^*(-\mathbf{p})$ so the Klein-Gordan field can be rewritten as

$$\begin{aligned}\varphi(\mathbf{x},t) &= \int d^3p_L \left[a(\mathbf{p})e^{i\mathbf{p}\cdot\mathbf{x}-i\omega t} + a^*(-\mathbf{p})e^{i\mathbf{p}\cdot\mathbf{x}+i\omega t}\right] \\ &= \int d^3p_L \left[a(\mathbf{p})e^{i\mathbf{p}\cdot\mathbf{x}-i\omega t} + a^*(\mathbf{p})e^{-i\mathbf{p}\cdot\mathbf{x}+i\omega t}\right]\end{aligned}$$

hence,

$$\varphi(x) = \int d^3p_L \left[a(\mathbf{p})e^{ipx} + a^*(\mathbf{p})e^{-ipx}\right] \tag{4.1.14}$$

where $px = p^\mu x_\mu = \mathbf{p}\cdot\mathbf{x} - \omega t$ is the Lorentz invariant inner product of the four-momentum and four-position vectors (it is understood that $\varphi(x)$ involves a temporal component) [5]. The conjugate momentum of the Klein-Gordon Lagrangian ((4.1.3)) is

$$\begin{aligned}\Pi(\mathbf{x},t) &= \frac{\partial\mathcal{L}}{\partial_\mu(\partial_0\varphi)} \\ &= \frac{\partial}{\partial_\mu(\partial_0\varphi)}\left[\frac{1}{2}(\partial_0\varphi)^2 - \frac{1}{2}(\nabla\varphi)^2 - \frac{1}{2}m^2\varphi^2\right] \\ &= \partial_0\varphi,\end{aligned}$$

hence

$$\Pi(x) = \partial_0\varphi = \dot{\varphi} \tag{4.1.15}$$

We are now in an ideal position to construct the Hamiltonian of the Klein-Gordon field. Following the formalism introduced in the background chapter, we

get [5]

$$\begin{aligned}\mathcal{H} &= \Pi\dot{\varphi} - \mathcal{L} \\ &= \Pi^2 - \left[-\frac{1}{2}\partial^\mu\varphi\partial_\mu\varphi - \frac{1}{2}m^2\varphi^2 \right] \\ &= \Pi^2 - \left[\frac{1}{2}(\partial_0\varphi)^2 - \frac{1}{2}(\nabla\varphi)^2 - \frac{1}{2}m^2\varphi^2 \right] \\ &= \Pi^2 - \left[\frac{1}{2}\Pi^2 - \frac{1}{2}(\nabla\varphi)^2 - \frac{1}{2}m^2\varphi^2 \right] \\ &= \frac{1}{2}\Pi^2 + \frac{1}{2}(\nabla\varphi)^2 + \frac{1}{2}m^2\varphi^2\end{aligned}$$

hence,

$$\mathcal{H} = \frac{1}{2}\Pi^2 + \frac{1}{2}(\nabla\varphi)^2 + \frac{1}{2}m^2\varphi^2. \tag{4.1.16}$$

With our field $\varphi(x)$, its conjugate momentum $\Pi(x)$, and Hamiltonian $\mathcal{H}$, we are ready to quantize our field. However, rewriting our Hamiltonian in terms of the Fourier coefficients $a(\mathbf{p})$ and $a^*(\mathbf{p})$ then proceeding with quantization will prove to be advantageous for later calculations and provide a deeper understanding as to the nature of the quantum field. We apply the variable change $p \to p'$ to equation (4.1.14) to get

$$\varphi(x) = \int d^3p'_L \left[a(\mathbf{p}')e^{ip'x} + a^*(\mathbf{p}')e^{-ip'x} \right]$$

then

$$\begin{aligned}
\varphi(x)\, e^{-ipx} &= \int d^3p'_L \left[a(\mathbf{p}')e^{ip'x} + a^*(\mathbf{p}')e^{-ip'x}\right] e^{-ipx} \\
\int d^3x\, \varphi(x)\, e^{-ipx} &= \int d^3x \int d^3p'_L \left[a(\mathbf{p}')e^{ip'x} + a^*(\mathbf{p}')e^{-ip'x}\right] e^{-ipx} \\
\int d^3x\, \varphi(x)\, e^{-ipx} &= \int d^3p' \frac{1}{2\omega}\frac{1}{(2\pi)^3}\int d^3x \left[a(\mathbf{p}')e^{i(\mathbf{p}-\mathbf{p}')x - i(\omega-\omega')t}\right. \\
&\quad \left. + a^*(\mathbf{p}')e^{i(-\mathbf{p}-\mathbf{p}')x + i(\omega+\omega')t}\right] \\
\int d^3x\, \varphi(x)\, e^{-ipx} &= \int d^3p' \frac{1}{2\omega}\Big[a(\mathbf{p}')\delta^3(\mathbf{p}-\mathbf{p}')e^{-i(\omega-\omega')t} \\
&\quad + a^*(\mathbf{p}')\delta^3(\mathbf{p}+\mathbf{p}')e^{i(\omega-\omega')t}\Big] \\
\int d^3x\, \varphi(x)\, e^{-ipx} &= \frac{1}{2\omega}a(\mathbf{p}) + \frac{1}{2\omega}a^*(-\mathbf{p})e^{2i\omega t}
\end{aligned}$$

where we have used the Fourier definition of the Dirac delta function

$$\delta^n(\mathbf{x}-\mathbf{x}_0) = \frac{1}{(2\pi)^n}\int_{-\infty}^{+\infty} d^n p\, e^{i(\mathbf{x}-\mathbf{x}_0)\mathbf{p}} \tag{4.1.17}$$

and the symmetrical property of the delta function. Then

$$\begin{aligned}
&\partial_0\Big[\int d^3x\, \varphi(x)\, e^{-ipx} = \frac{1}{2\omega}a(\mathbf{p}) + \frac{1}{2\omega}a^*(-\mathbf{p})e^{2i\omega t}\Big] \\
&\int d^3x\, \partial_0\varphi(x)\, e^{-ipx} + i\omega\int d^3x\, \varphi(x)\, e^{-ipx} = ia^*(-\mathbf{p})e^{2i\omega t} \\
&\int d^3x\, \partial_0\varphi(x)\, e^{-ipx} = -i\omega\Big[\frac{1}{2\omega}a(\mathbf{p}) + \frac{1}{2\omega}a^*(-\mathbf{p})e^{2i\omega t}\Big] + ia^*(-\mathbf{p})e^{2i\omega t} \\
&\int d^3x\, \partial_0\varphi(x)\, e^{-ipx} = -\frac{i}{2}a(\mathbf{p}) + \frac{i}{2}a^*(-\mathbf{p})e^{2i\omega t}
\end{aligned}$$

hence we have

$$\begin{aligned}
\int d^3x\, \varphi(x)\, e^{-ipx} &= \frac{1}{2\omega}a(\mathbf{p}) + \frac{1}{2\omega}a^*(-\mathbf{p})e^{2i\omega t} \\
\int d^3x\, \partial_0\varphi(x)\, e^{-ipx} &= -\frac{i}{2}a(\mathbf{p}) + \frac{i}{2}a^*(-\mathbf{p})e^{2i\omega t}
\end{aligned}$$

which we then combine to get

$$\begin{aligned} a(\mathbf{p}) &= \int d^3x\, e^{-ipx}[i\partial_0\varphi(x) + \omega\varphi(x)] \\ &= i\int d^3x\, e^{-ipx} \overleftrightarrow{\partial_0}\, \varphi(x), \end{aligned} \tag{4.1.18}$$

where the operator $\overleftrightarrow{\partial_0}$ is defined as $f\overleftrightarrow{\partial_0}g = f(\partial_0 g) - (\partial_0 f)g$ [5]. Complex conjugating yields the other Fourier coefficient thus we finally get

$$a(\mathbf{p}) = i\int d^3x\, e^{-ipx} \overleftrightarrow{\partial_0}\, \varphi(x), \tag{4.1.19}$$

$$a^*(\mathbf{p}) = -i\int d^3x\, e^{ipx} \overleftrightarrow{\partial_0}\, \varphi(x). \tag{4.1.20}$$

Now we rewrite the Hamiltonian in terms of $a(\mathbf{p})$ and $a^*(\mathbf{p})$ to get [5]

$$\begin{aligned} H &= \int d^3x\, \mathcal{H} \\ &= \frac{1}{2}\int d^3p_L\, \omega\, [a^*(\mathbf{p})a(\mathbf{p}) + a(\mathbf{p})a^*(\mathbf{p})], \end{aligned} \tag{4.1.21}$$

where we have carefully evaluated the Hamiltonian in a noncommutative manner in anticipation for quantization [5].

Quantizing the Classical Real Scalar Field

Now that we have finished gathering everything we need from the classical physics of the Klein-Gordon field, we now finally promote the field $\varphi(\mathbf{x},t) \to \hat{\varphi}(\mathbf{x},t)$ and its conjugate momentum $\Pi(\mathbf{x},t) \to \hat{\Pi}(\mathbf{x},t)$ to operators in the Heisenberg picture and impose the canonical commutation relations at equal times,

$$\begin{aligned} [\hat{\varphi}(\mathbf{x},t), \hat{\varphi}(\mathbf{x}',t)] &= 0, \\ [\hat{\Pi}(\mathbf{x},t), \hat{\Pi}(\mathbf{x}',t)] &= 0, \\ [\hat{\varphi}(\mathbf{x},t), \hat{\Pi}(\mathbf{x}',t)] &= i\delta^3(\mathbf{x}-\mathbf{x}'), \end{aligned} \tag{4.1.22}$$

where the delta function replaces the identity operator since we are using continuous labels rather than discrete labels [5]. Since we have promoted the field and its conjugate momentum to operators, we see that the Fourier coefficients $a(\mathbf{p})$ and $a^*(\mathbf{p})$ now become operators $\hat{a}(\mathbf{p})$ and $\hat{a}^\dagger(\mathbf{p})$, respectively, where the dagger operation † denotes the Hermitian conjugation operation,

$$\hat{a}(\mathbf{p}) = \int d^3x\, e^{-ipx}[i\partial_0\hat{\varphi}(x) + \omega\hat{\varphi}(x)] = i\int d^3x\, e^{-ipx}\overleftrightarrow{\partial_0}\hat{\varphi}(x), \tag{4.1.23}$$

$$\hat{a}^\dagger(\mathbf{p}) = \int d^3x\, e^{ipx}[-i\partial_0\hat{\varphi}(x) + \omega\hat{\varphi}(x)] = -i\int d^3x\, e^{ipx}\overleftrightarrow{\partial_0}\hat{\varphi}(x). \tag{4.1.24}$$

We can compute the commutation relations of the coefficient operators by using the canonical commutation relations to get

$$
\begin{aligned}
[\hat{a}(\mathbf{p}), \hat{a}(\mathbf{p}')] &= \int d^3x\, d^3x'\, e^{-ipx} e^{-ip'x'} \Big([i\partial_0\hat{\varphi}(x) + \omega\hat{\varphi}(x), i\partial_0\hat{\varphi}(x') + \omega\hat{\varphi}(x')] \Big) \\
&= \int d^3x\, d^3x'\, e^{-ipx} e^{-ip'x'} \Big([i\hat{\Pi}(x) + \omega\hat{\varphi}(x), i\hat{\Pi}(x') + \omega\hat{\varphi}(x')] \Big) \\
&= \int d^3x\, d^3x'\, e^{-ipx} e^{-ip'x'} \Big((i)^2[\hat{\Pi}(x), \hat{\Pi}(x')] + i\omega[\hat{\Pi}(x), \hat{\varphi}(x')] \\
&+ i\omega[\hat{\varphi}(x), \hat{\Pi}(x')] + \omega^2[\hat{\varphi}(x), \hat{\varphi}(x')] \Big) \\
&= \int d^3x\, d^3x'\, e^{-ipx} e^{-ip'x'} \Big(\omega\delta^3(\mathbf{x} - \mathbf{x}') - \omega\delta^3(\mathbf{x} - \mathbf{x}') \Big) \\
&= 0, \\
[\hat{a}^\dagger(\mathbf{p}), \hat{a}^\dagger(\mathbf{p}')] &= \int d^3x\, d^3x'\, e^{ipx} e^{ip'x'} \Big([-i\partial_0\hat{\varphi}(x) + \omega\hat{\varphi}(x), -i\partial_0\hat{\varphi}(x') + \omega\hat{\varphi}(x')] \Big) \\
&= \int d^3x\, d^3x'\, e^{ipx} e^{ip'x'} \Big([-i\hat{\Pi}(x) + \omega\hat{\varphi}(x), -i\hat{\Pi}(x') + \omega\hat{\varphi}(x')] \Big) \\
&= \int d^3x\, d^3x'\, e^{ipx} e^{ip'x'} \Big((-i)^2[\hat{\Pi}(x), \hat{\Pi}(x')] - i\omega[\hat{\Pi}(x), \hat{\varphi}(x')] \\
&- i\omega[\hat{\varphi}(x), \hat{\Pi}(x')] + \omega^2[\hat{\varphi}(x), \hat{\varphi}(x')] \Big) \\
&= \int d^3x\, d^3x'\, e^{ipx} e^{ip'x'} \Big(-\omega\delta^3(\mathbf{x} - \mathbf{x}') + \omega\delta^3(\mathbf{x} - \mathbf{x}') \Big) \\
&= 0, \\
[\hat{a}(\mathbf{p}), \hat{a}^\dagger(\mathbf{p}')] &= \int d^3x\, d^3x'\, e^{-ip'x'} e^{ipx} \Big([i\partial_0\hat{\varphi}(x) + \omega\hat{\varphi}(x), -i\partial_0\hat{\varphi}(x') + \omega\hat{\varphi}(x')] \Big) \\
&= \int d^3x\, d^3x'\, e^{-ip'x'} e^{ipx} \Big([i\hat{\Pi}(x) + \omega\hat{\varphi}(x), -i\hat{\Pi}(x') + \omega\hat{\varphi}(x')] \Big) \\
&= \int d^3x\, d^3x'\, e^{-ip'x'} e^{ipx} \Big((i)(-i)[\hat{\Pi}(x), \hat{\Pi}(x')] + (i\omega)[\hat{\Pi}(x), \hat{\varphi}(x')] \\
&+ (-i\omega)[\hat{\varphi}(x), \hat{\Pi}(x')] + \omega^2[\hat{\varphi}(x), \hat{\varphi}(x')] \Big) \\
&= \int d^3x\, d^3x'\, e^{-ip'x'} e^{ipx} \Big(\omega\delta^3(\mathbf{x} - \mathbf{x}') + \omega\delta^3(\mathbf{x} - \mathbf{x}') \Big) \\
&= (2\omega) \int d^3x\, d^3x'\, e^{-ip'x'} e^{ipx} \delta^3(\mathbf{x} - \mathbf{x}') \\
&= (2\omega) \int d^3x\, e^{i(p-p')x} \\
&= (2\omega)(2\pi)^3\delta^3(\mathbf{p} - \mathbf{p}'),
\end{aligned}
$$

thus,

$$\begin{aligned}
[\hat{a}(\mathbf{p}), \hat{a}(\mathbf{p}')] &= 0, \\
[\hat{a}^\dagger(\mathbf{p}), \hat{a}^\dagger(\mathbf{p}')] &= 0, \\
[\hat{a}(\mathbf{p}), \hat{a}^\dagger(\mathbf{p}')] &= (2\omega)(2\pi)^3 \delta^3(\mathbf{p} - \mathbf{p}'). \qquad (4.1.25)
\end{aligned}$$

With the Fourier commutation relations, we can rewrite the quantum Hamiltonian as

$$\begin{aligned}
\hat{H} &= \frac{1}{2} \int d^3 p_L \, \omega \left(\hat{a}^\dagger(\mathbf{p}) \hat{a}(\mathbf{p}) + \hat{a}(\mathbf{p}) \hat{a}^\dagger(\mathbf{p}) \right) \\
&= \int d^3 p_L \, \omega \left(\hat{a}^\dagger(\mathbf{p}) \hat{a}(\mathbf{p}) + \frac{1}{2} [\hat{a}(\mathbf{p}), \hat{a}^\dagger(\mathbf{p})] \right) \\
&= \int d^3 p_L \, \omega \, \hat{a}^\dagger(\mathbf{p}) \hat{a}(\mathbf{p}) + E_0. \qquad (4.1.26)
\end{aligned}$$

We now come across a fascinating result: the algebra of the operators $\hat{a}(\mathbf{p})$ and $\hat{a}^\dagger(\mathbf{p})$ are equivalent to the algebra of the creation and annihilation operators of the uncoupled, simple harmonic oscillator from quantum mechanics. It can be further verified that the operators obey

$$[\hat{H}, \hat{a}^\dagger(\mathbf{p})] = \omega \hat{a}^\dagger(\mathbf{p}), \quad [H, \hat{a}(\mathbf{p})] = -\omega \hat{a}(\mathbf{p}).$$

Thus we associate $\hat{a}^\dagger(\mathbf{p})$ and $\hat{a}(\mathbf{p})$ as the creation and annihilation operators, respectively, for the Klein-Gordon field and the Hamiltonian physically represents the field as an infinite sum of uncoupled harmonic oscillators. Before we delve into further analysis of our finding, we must address the rather settling infinite number

$$E_0 = \frac{1}{2} \int d^3 p \, \omega \, \delta^3(0), \qquad (4.1.27)$$

in the Hamiltonian. This infinite number is quantum field-theoretic analogue of the zero-point energy for the simple harmonic oscillator from quantum mechanics. The stark difference is that since the Klein-Gordon field is an infinite ensemble of

these oscillators, E_0 is the sum of all the zero-point energies of each oscillator in this infinite ensemble. This number is actually to be expected in our calculations because we are integrating over an infinite volume. In actual experiments, this infinite energy term is not measured because experiments measure only energy differences from the ground state of the system in question [8]. Additionally, the appearance of E_0 gives indications that a quantum field theory becomes unsuitable at higher energies thus giving a clue that there may be an underlying theory that can account for such extremely high energy situations [5]. For our calculations, we can ignore these infinite terms by adding an arbitrary constant into the Lagrangian and selecting a value to cancel E_0 so that the ground state Hamilitonian gives a zero energy eigenvalue. Divergent integrals such as equation (4.1.27) are infamous in the study of quantum field theories.

Since we now are considering multiparticle states, the mathematical space that we work with is the *Fock space*. In analogy with the simple harmonic oscillator, we construct our Fock space by first defining a *vacuum state* $|0\rangle$ to be the state with no particles, or vacuum, normalized as $\langle 0|0\rangle = 1$, such that

$$\hat{a}(\mathbf{p})|0\rangle = 0, \tag{4.1.28}$$

and the one-particle state created by the creation operator acting on the vacuum state as

$$\hat{a}^\dagger(\mathbf{p})|0\rangle = (2\omega)^{1/2}\hat{a}^\dagger(\mathbf{p})|0\rangle = |p\rangle, \tag{4.1.29}$$

where $\omega = (\mathbf{p}^2 + m^2)^{1/2}$ [5]. The normalization of the one-particle state $\langle p|p'\rangle$ must be Lorentz invariant so we cannot use the normalization from nonrelativsitic quantum mechanics, namely, $\langle p|p'\rangle = \delta^3(\mathbf{p} - \mathbf{p}')$. To find the correct normalization, consider the following

$$\langle 0|[\hat{a}(\mathbf{p}), \hat{a}^\dagger(\mathbf{p}')]|0\rangle = \langle 0|\hat{a}(\mathbf{p})\hat{a}^\dagger(\mathbf{p}')|0\rangle - \langle 0|\hat{a}^\dagger(\mathbf{p}')\hat{a}(\mathbf{p})|0\rangle$$
$$(2\omega)(2\pi)^3\delta^3(\mathbf{p} - \mathbf{p}') = \langle p|p'\rangle,$$

hence the Lorentz invariant normalization for the one-particles state is

$$\langle p|p'\rangle = (2\omega)(2\pi)^3\delta^3(\mathbf{p}-\mathbf{p}'). \tag{4.1.30}$$

Then a multiparticle state is created by acting on the vacuum with multiple creation operators

$$\prod_{k=1}^{n}(2\omega_k)^{1/2}\,\hat{a}_1^\dagger(\mathbf{p}_1),\ldots\hat{a}_n^\dagger(\mathbf{p}_n)|0\rangle = |p_1\ldots p_n\rangle, \tag{4.1.31}$$

which is an eigenstate of the Hamiltonian with energy $(\omega_1+\omega_2+\ldots+\omega_n)$ [5]. Now we finally come into full circle and are ready to answer what the parameter m represents in the equation of motion for the Klein-Gordon field. Using Noether's theorem, the *total four-momentum operator* is constructed as

$$\begin{aligned}\hat{P}^\mu &= \int d^3x\,\hat{T}^{0\mu}\\ &= -\int d^3x\,\hat{\Pi}(x)\partial^\mu\hat{\varphi}(x)\\ &= \int \frac{d^3x}{(2\pi)^3}\,p^\mu\hat{a}^\dagger(\mathbf{p})\hat{a}(\mathbf{p}),\end{aligned} \tag{4.1.32}$$

from which the *total spatial momentum operator* is constructed as

$$\hat{P}^i = \int \frac{d^3x}{(2\pi)^3}\,p^i\,\hat{a}^\dagger(\mathbf{p})\hat{a}(\mathbf{p}). \tag{4.1.33}$$

Thus the creation operator $\hat{a}_k^\dagger(\mathbf{p})$ creates an excitation with momentum $\mathbf{p}_k$ and energy $\omega_k = (|\mathbf{p}_k|^2+m^2)^{1/2}$. The equation $\omega_k = (|\mathbf{p}_k|^2+m^2)^{1/2}$ is the relativistic energy equation so m is the mass of these individual excitations. These individual excitations are what we perceive as *particles* hence particles are interpretted to be the modes of excitations of the underlying quantum field. It is important to note that the creation operator $\hat{a}^\dagger(\mathbf{p})$ creates a particle in the vacuum with momentum $\mathbf{p}$ and not in some localized position [8]. Since we used commutators for quantizing

the scalar field, we have that for a two particle state

$$\begin{aligned} \hat{a}_1^\dagger \hat{a}_2^\dagger |0\rangle &= \hat{a}_2^\dagger \hat{a}_2^\dagger |0\rangle \\ |p_1, p_2\rangle &= |p_2, p_1\rangle, \end{aligned} \tag{4.1.34}$$

thus the interchange of two scalar particles results in the same state. This can be easily generalized to n particles, thus we find that the scalar field and the scalar particles obey *Bose-Einstein statistics*. We now promote the field-theoretic total angular momentum from the chapter on symmetries to a Hermitian operator which yields

$$\hat{\mathbf{J}} = \hat{\mathbf{L}}. \tag{4.1.35}$$

As mentioned previously, the total angular momentum operator lacks a spin component thus the scalar field and its excitations, the scalar particles, possess spin-0 which is expected of a Lorentz scalar field. In summary, the scalar field is a spin-0 bosonic field with its excitation, namely, the scalar particle being a spin-0 boson [5]. Now the resolution of the identity for one-particle states can now be defined as [8]

$$\hat{1} = \int d^3 p_L \, |\mathbf{p}\rangle\langle\mathbf{p}|, \tag{4.1.36}$$

then the field operator $\hat{\varphi}(\mathbf{x})$ in the Schrödinger picture acting on the vacuum creates the following state

$$\hat{\varphi}(\mathbf{x})|0\rangle = \int d^3 p_L \, e^{-i\mathbf{x}\cdot\mathbf{p}} |\mathbf{p}\rangle = |\mathbf{x}\rangle \tag{4.1.37}$$

and similarly we can get

$$\langle 0|\hat{\varphi}(\mathbf{x})|p\rangle = e^{i\mathbf{x}\cdot\mathbf{p}}. \tag{4.1.38}$$

Thus the field operator $\hat{\varphi}(\mathbf{x})$ acting on the vacuum creates a particle at position $\mathbf{x}$ [8].

We now see that the field is clearly the more fundamental and important object than the particle. A particle is nothing more than the excitation of the underlying field and given enough energy that satisfies the relativistic energy formula $\omega = (|\mathbf{p}|^2 + m^2)^{1/2}$, multiple particles can be created from the field. Thus we see that the field concept naturally gives rise to describing multiparticle states.

4.2 Fermion Fields

Classical Fermion Fields

From our analysis of the representations of the Lorentz group, we found that the simplest, nontrivial irreducible representations of the Lorentz Lie algebra were the $(\frac{1}{2}, 0)$ and $(0, \frac{1}{2})$ representations which were the two component left- $\psi_a(x)$ and right-handed $\psi^\dagger_{\dot{a}}(x)$ spinor fields, or Weyl fields, respectively. We now seek to construct a Lorentz invariant Lagrangian that is Hermitian and quadratic in ψ_a and $\psi^\dagger_{\dot{a}}$. A suitable Lagrangian is [5]

$$\mathcal{L} = i\psi^\dagger \bar{\sigma}^\mu \partial_\mu \psi - \frac{1}{2} m \psi\psi - \frac{1}{2} m \psi^\dagger \psi^\dagger, \tag{4.2.1}$$

where $\bar{\sigma}^\mu = (I, -\boldsymbol{\sigma})$ [5]. From the form of our Lagrangian, it is evident that the kinetic term $i\psi^\dagger \bar{\sigma}^\mu \partial_\mu \psi$ is not Hermitian. Indeed this can be seen in the following,

$$\begin{aligned}
(i\psi^\dagger \bar{\sigma}^\mu \partial_\mu \psi)^\dagger &= -i\partial_\mu \psi^\dagger_{\dot{c}} \bar{\sigma}^{\mu a \dot{c}} \psi_a \\
&= i\psi^\dagger_{\dot{c}} \bar{\sigma}^{\mu a \dot{c}} \partial_\mu \psi_a - i\partial_\mu (\psi^\dagger_{\dot{c}} \bar{\sigma}^{\mu a \dot{c}} \psi_a) \\
&= i\psi^\dagger \bar{\sigma}^\mu \partial_\mu \psi_a - i\partial_\mu (\psi^\dagger \bar{\sigma}^\mu \psi),
\end{aligned}$$

where the second line follows from the identity $\partial(AB) = A(\partial B) + (\partial A)B$ [5]. However, the second term in the third line is the total divergence and if we impose suitable boundary conditions, the term will vanish thus this kinetic term is suitable [5]. Now we follow our usual convention and apply an infinitesimal variation on

the action with respect to $\psi^\dagger$ along with the stationary action principle to get the equations of motion for ψ:

$$\begin{aligned} 0 = \frac{\delta S}{\delta \psi^\dagger} &= \frac{\delta}{\delta \psi^\dagger} \int d^4x \, \mathcal{L} \\ &= \int d^4x \left[i\delta\psi^\dagger \bar{\sigma}^\mu \partial_\mu \psi - m\psi^\dagger \delta\psi^\dagger \right] \\ &= \int d^4x \left[i\bar{\sigma}^\mu \partial_\mu \psi - m\psi^\dagger \right] \delta\psi^\dagger, \end{aligned}$$

which yields the equation of motions

$$\frac{\delta S}{\delta \psi^\dagger} = i\bar{\sigma}^\mu \partial_\mu \psi - m\psi^\dagger = 0.$$

Hermitian conjugating the above equation is equivalent to varying the Lagrangian with respect to ψ, namely, $\frac{\delta S}{\delta \psi}$ which yields the equations of motion for $\psi^\dagger$ thus we have the following equations of motion

$$\frac{\delta S}{\delta \psi^\dagger} = i\bar{\sigma}^{\mu\,\dot{a}c} \partial_\mu \psi_c - m\psi^{\dagger\,\dot{a}} = 0, \tag{4.2.2}$$

$$\frac{\delta S}{\delta \psi} = i\sigma^\mu_{a\dot{c}} \partial_\mu \psi^{\dagger\,\dot{c}} - m\psi_a = 0, \tag{4.2.3}$$

where we have restored spinor indices for clarity [5]. The above equations of motion are a bit troublesome in the fact that they are coupled equations. However, setting the mass term to zero decouples the equations and gives us the *Weyl equations*,

$$i\bar{\sigma}^\mu \partial_\mu \psi_L = 0, \tag{4.2.4}$$

$$i\sigma^\mu \partial_\mu \psi_R = 0, \tag{4.2.5}$$

where $\psi_L = \psi_c$ and $\psi_R = \psi^{\dagger\,\dot{c}}$ are the left- and right- handed Weyl fields, respectively (we have presented the Weyl equations in this manner in accordance with common literature) [5]. Note that setting the mass term to zero in Lagrangian (4.2.1) gives

us the Lagrangian for the massless Weyl field,

$$\mathcal{L} = i\psi^\dagger \bar{\sigma}^\mu \partial_\mu \psi, \tag{4.2.6}$$

from which the Weyl equations can be derived.

As we stated before, the left- and right- handed Weyl fields are the $(\frac{1}{2}, 0)$ and $(0, \frac{1}{2})$ irreducible representations, respectively, of the Lorentz algebra. From this, it is natural to see that we can construct other spin-1/2 representations with the Weyl fields as their building blocks. Mathematically, other spin-1/2 representations can be constructed as the direct sum of the $(\frac{1}{2}, 0)$ and $(0, \frac{1}{2})$ field representations to create the $(\frac{1}{2}, 0) \oplus (0, \frac{1}{2})$ field representation called the *bispinor*. Its Lorentz transformation matrix will be a direct sum of the matrices for both handed spinors [15]

$$\begin{aligned} D(\Lambda) &= D_L(\Lambda)_a{}^c \oplus D_R(\Lambda)^{\dot{a}}{}_{\dot{c}} \\ &= \begin{pmatrix} D_L(\Lambda)_a{}^c & 0 \\ 0 & D_R(\Lambda)^{\dot{a}}{}_{\dot{c}} \end{pmatrix}. \end{aligned} \tag{4.2.7}$$

The generator of the bispinor representation is then given as [5]

$$S^{\mu\nu} = \begin{pmatrix} (+S_L^{\mu\nu})_a{}^c & 0 \\ 0 & (-S_R^{\mu\nu})^{\dot{a}}{}_{\dot{c}} \end{pmatrix}, \tag{4.2.8}$$

where

$$(S_L^{\mu\nu})_a{}^c = +\frac{i}{4}(\sigma^\mu \bar{\sigma}^\nu - \sigma^\nu \bar{\sigma}^\mu)_a{}^c, \tag{4.2.9}$$

$$(S_R^{\mu\nu})^{\dot{a}}{}_{\dot{c}} = -\frac{i}{4}(\bar{\sigma}^\mu \sigma^\nu - \bar{\sigma}^\nu \sigma^\mu)^{\dot{a}}{}_{\dot{c}}. \tag{4.2.10}$$

Now, the equations of motion (4.2.2) and (4.2.3) can be rewritten in matrix notation

as [5]

$$\begin{pmatrix} m\delta_a{}^c & -i\sigma^\mu_{a\dot{c}}\partial_\mu \\ -i\bar{\sigma}^{\mu a\dot{c}}\partial_\mu & m\delta^{\dot{a}}{}_{\dot{c}} \end{pmatrix} \begin{pmatrix} \psi_c \\ \psi^{\dagger\dot{c}} \end{pmatrix} = 0. \tag{4.2.11}$$

To make equation (4.2.11) more compact, we introduce the *gamma matrices*

$$\gamma^\mu = \begin{pmatrix} 0 & \sigma^\mu_{a\dot{c}} \\ \bar{\sigma}^{\mu a\dot{c}} & 0 \end{pmatrix}, \tag{4.2.12}$$

which follows the anticommutator algebra called the *Clifford algebra* [15]

$$\{\gamma^\nu, \gamma^\mu\} = -2g^{\mu\nu}. \tag{4.2.13}$$

Proof. Let γ^μ be defined as above in equation (4.2.12). Then

$$\begin{aligned} \{\gamma^\nu, \gamma^\mu\} &= \gamma^\mu\gamma^\nu + \gamma^\nu\gamma^\mu \\ &= \begin{pmatrix} 0 & \sigma^\mu \\ \bar{\sigma}^\mu & 0 \end{pmatrix} \begin{pmatrix} 0 & \sigma^\nu \\ \bar{\sigma}^\nu & 0 \end{pmatrix} + \begin{pmatrix} 0 & \sigma^\nu \\ \bar{\sigma}^\nu & 0 \end{pmatrix} \begin{pmatrix} 0 & \sigma^\mu \\ \bar{\sigma}^\mu & 0 \end{pmatrix} \\ &= \begin{pmatrix} \{\sigma^\mu, \bar{\sigma}^\nu\} & 0 \\ 0 & \{\sigma^\mu, \bar{\sigma}^\nu\} \end{pmatrix} \\ &= \begin{pmatrix} -2g^{\mu\nu} & 0 \\ 0 & -2g^{\mu\nu} \end{pmatrix} \\ &= -2g^{\mu\nu}, \end{aligned}$$

where we have used the relation $\{\sigma^\mu, \bar{\sigma}^\nu\} = -2g^{\mu\nu}$. □

With the gamma matrices, we can write matrix (4.2.8) as [5]

$$S^{\mu\nu} = \frac{i}{4}[\gamma^\mu, \gamma^\nu], \tag{4.2.14}$$

then the spinor representation is

$$D(\Lambda) = e^{-\frac{i}{2}\omega_{\mu\nu}S^{\mu\nu}}. \tag{4.2.15}$$

Now we introduce a new field [5].

Definition 4.2.1. The $(\frac{1}{2}, 0) \oplus (0, \frac{1}{2})$ field representation of the Lorentz algebra that obeys the *Majorana condition* $\boldsymbol{\chi}^C = \boldsymbol{\chi}$ is the four-component bispinor called the **Majorana field**, or **Majorana spinor**, $\boldsymbol{\chi}$

$$\boldsymbol{\chi} = \begin{pmatrix} \psi_c \\ \psi^{\dagger\dot{c}} \end{pmatrix}. \tag{4.2.16}$$

Under a Lorentz transformation, the Majorana field transforms as

$$U(\Lambda)^{-1}\boldsymbol{\chi}(x)U(\Lambda) = D(\Lambda)\boldsymbol{\chi}(\Lambda^{-1}x),$$

where

$$D(\Lambda) = e^{-\frac{i}{2}\omega_{\mu\nu}S^{\mu\nu}},$$
$$S^{\mu\nu} = \frac{i}{4}[\gamma^\mu, \gamma^\nu].$$

With the Majorana field and gamma matrices, equation (4.2.11) elegantly simplifies to

$$(-i\gamma^\mu\partial_\mu + m)\boldsymbol{\chi} = 0, \tag{4.2.17}$$

which is Dirac equation from the introduction. Amazingly, we see that the Dirac equation naturally arises when we consider quantum fields rather than wave functions. With this new field, we seek to construct a proper Lagrangian that will produce the Dirac equation (4.2.17) as its equation of motion. To do this, some technical nuance is needed because the Majorana field obeys a unique symmetry condition.

Recall from the chapter on symmetries that the charge conjugation matrix $\mathcal{C}$ was introduced as

$$\mathcal{C} = \begin{pmatrix} -\varepsilon^{ac} & 0 \\ 0 & -\varepsilon_{\dot{a}\dot{c}} \end{pmatrix},$$

with the property

$$\mathcal{C}^\dagger = -\mathcal{C}.$$

It is obvious that $\boldsymbol{\chi}^C = \boldsymbol{\chi}$ for the Majorana field, as was stated in the definition, thus the Majorana field is its own charge conjugate [5]. If we are to construct a proper Lagrangian for the Majorana field, it must obey this condition as well. With this key feature in mind, we seek to rearrange the Lagrangian (4.2.1) to construct the Lagrangian for the Majorana field. To do this, we first define the matrix

$$\beta = \begin{pmatrix} 0 & \delta^{\dot{a}}{}_{c} \\ \delta_{a}{}^{c} & 0 \end{pmatrix}, \tag{4.2.18}$$

and then introduce

$$\bar{\boldsymbol{\chi}} = \boldsymbol{\chi}^\dagger \beta = (\psi^a, \psi^\dagger_{\dot{a}}), \tag{4.2.19}$$

where $\boldsymbol{\chi}^\dagger = (\psi^\dagger_{\dot{a}}, \psi^a)$ [5]. Then

$$\begin{aligned} \boldsymbol{\chi}\gamma^\mu\partial_\mu\boldsymbol{\chi} &= \psi^a\bar{\sigma}^\mu\partial_\mu\psi^\dagger_{\dot{a}} + \psi^\dagger_{\dot{a}}\bar{\sigma}^\mu\partial_\mu\psi^a \\ &= -(\partial_\mu\psi^a)\bar{\sigma}\psi^\dagger_{\dot{a}} + \partial_\mu(\psi^a\bar{\sigma}^\mu\psi^\dagger_{\dot{a}}) + \psi^\dagger_{\dot{a}}\bar{\sigma}^\mu\partial_\mu\psi^a, \end{aligned}$$

where we have used the identity $A\partial B = -(\partial A) + \partial(AB)$ [5]. By using the anticommutation of the fields and index manipulation, the first term can be rewritten as

$$-(\partial_\mu\psi^a)\bar{\sigma}\psi^\dagger_{\dot{a}} = \psi^\dagger_{\dot{a}}\bar{\sigma}^\mu\partial_\mu\psi^a,$$

which then gives us

$$\boldsymbol{\chi}\gamma^\mu\partial_\mu\boldsymbol{\chi} = 2\psi^\dagger_{\dot{a}}\bar{\sigma}^\mu\partial_\mu\psi^a + \partial_\mu(\psi^a\bar{\sigma}^\mu\psi^\dagger_{\dot{a}}). \tag{4.2.20}$$

The last term is a total divergence which is removed when we find the action $\mathcal{S}$ so we discard the second term as it is irrelevant [5]. We also see that the mass terms of equation (4.2.1) can rewritten as

$$\frac{1}{2}m\bar{\boldsymbol{\chi}}\boldsymbol{\chi} = \frac{1}{2}m(\psi\psi + \psi^\dagger\psi^\dagger). \tag{4.2.21}$$

Although it now seems like we can combine our results to construct the Majorana Lagrangian, we must incorporate the condition $\boldsymbol{\chi}^{\mathrm{C}} = \boldsymbol{\chi}$. To do this, we use the fact that $\boldsymbol{\chi}^{\mathrm{C}} = \mathcal{C}\bar{\boldsymbol{\chi}} = \boldsymbol{\chi}$ to write $\bar{\boldsymbol{\chi}} = \boldsymbol{\chi}^{\mathrm{T}}\mathcal{C}$ [5]. With this final piece of the puzzle, we now have the correct Lagrangian for the Majorana field

$$\mathcal{L} = \boldsymbol{\chi}^{\mathrm{T}}\Big(\frac{i}{2}\mathcal{C}\gamma^\mu\partial_\mu - \frac{1}{2}m\mathcal{C}\Big)\boldsymbol{\chi}, \tag{4.2.22}$$

which yields the correct equation of motion (4.2.17) [5].

It is quite remarkable how we were able to construct a new fermion field by just using the Lagrangian for a single left-handed spinor field to build a $(\frac{1}{2}, 0) \oplus (0, \frac{1}{2})$ representation. Let us now consider extending the Lagrangian (4.2.1) for two left handed Weyl fields

$$\mathcal{L} = i\psi_i^\dagger\bar{\sigma}^\mu\partial_\mu\psi_i - \frac{1}{2}m\psi_i\psi_i - \frac{1}{2}m\psi_i^\dagger\psi_i^\dagger, \tag{4.2.23}$$

where as usual the repeated indices are to be implicitly summed for $i = 1, 2$ [15]. Lagrangian (4.2.23) exhibits SO(2) symmetry which we now prove [5].

Proof. Let $\mathcal{L}$ be the Lagrangian defined as in (4.2.23) in matrix notation

$$\mathcal{L} = i(\boldsymbol{\psi}^\dagger\bar{\sigma}^\mu)^{\mathrm{T}}\partial_\mu\boldsymbol{\psi} - \frac{1}{2}m\boldsymbol{\psi}^{\mathrm{T}}\boldsymbol{\psi} - \frac{1}{2}m(\boldsymbol{\psi}^\dagger)^{\mathrm{T}}\boldsymbol{\psi}^\dagger,$$

where $\boldsymbol{\psi} = (\psi_1, \psi_2)^{\mathrm{T}}$ and $\boldsymbol{\psi}^\dagger = (\psi_1^\dagger, \psi_2^\dagger)^{\mathrm{T}}$ and consider the SO(2) transformation

$$\boldsymbol{\psi} \longrightarrow \boldsymbol{\psi}' = G\boldsymbol{\psi}$$
$$\begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix} \longrightarrow \begin{pmatrix} \psi_1' \\ \psi_2' \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} \psi_1 \\ \psi_2 \end{pmatrix}$$

where $G \in \mathrm{SO}(2)$. Then the Lagrangian under the SO(2) transformation is

$$\begin{aligned}
\mathcal{L}' &= i(\boldsymbol{\psi}^{\dagger\prime}\bar{\sigma}^\mu)^{\mathrm{T}}\partial_\mu\boldsymbol{\psi}' - \frac{1}{2}m\boldsymbol{\psi}'^{\mathrm{T}}\boldsymbol{\psi}' - \frac{1}{2}m(\boldsymbol{\psi}^{\dagger\prime})^{\mathrm{T}}\boldsymbol{\psi}^{\dagger\prime} \\
&= i(G\boldsymbol{\psi}^\dagger\bar{\sigma}^\mu)^{\mathrm{T}}\partial_\mu(G\boldsymbol{\psi}) - \frac{1}{2}m(G\boldsymbol{\psi})^{\mathrm{T}}G\boldsymbol{\psi} - \frac{1}{2}m(G\boldsymbol{\psi}^\dagger)^{\mathrm{T}}G\boldsymbol{\psi}^\dagger \\
&= i(\boldsymbol{\psi}^\dagger\bar{\sigma}^\mu)^{\mathrm{T}}G^{\mathrm{T}}G\partial_\mu\boldsymbol{\psi} - \frac{1}{2}m\boldsymbol{\psi}^{\mathrm{T}}G^{\mathrm{T}}G\boldsymbol{\psi} - \frac{1}{2}m(\boldsymbol{\psi}^\dagger)^{\mathrm{T}}G^{\mathrm{T}}G\boldsymbol{\psi}^\dagger \\
&= i(\boldsymbol{\psi}^\dagger\bar{\sigma}^\mu)^{\mathrm{T}}\partial_\mu\boldsymbol{\psi} - \frac{1}{2}m\boldsymbol{\psi}^{\mathrm{T}}\boldsymbol{\psi} - \frac{1}{2}m(\boldsymbol{\psi}^\dagger)^{\mathrm{T}}\boldsymbol{\psi}^\dagger \\
&= \mathcal{L},
\end{aligned}$$

where we have used the property $G^{\mathrm{T}}G = I$. Thus the Lagrangian $\mathcal{L}$ exhibits SO(2) symmetry. □

Since the Lie group SO(2) is isomorphic to the Lie group U(1), the Lagrangian (4.2.23) is also invariant under U(1) transformations and this can be directly observed if we make the substitution [5]

$$\chi = \frac{1}{\sqrt{2}}(\psi_1 + i\psi_2),$$
$$\xi = \frac{1}{\sqrt{2}}(\psi_1 - i\psi_2),$$

which then turns the Lagrangian into

$$\mathcal{L} = i\chi^\dagger\bar{\sigma}^\mu\partial_\mu\chi + i\xi^\dagger\bar{\sigma}^\mu\partial_\mu\xi - m\chi\xi - m\xi^\dagger\chi^\dagger. \tag{4.2.24}$$

Before we prove the U(1) symmetry of the above Lagrangian, let us first see if we can reduce the Lagrangian into a simpler form, much like how we derived the

Majorana field. Upon close inspection, Lagrangian (4.2.24) almost resembles a quadratic form and a quadratic form can be expressed as the inner product of two vectors. However, the last two terms of the Lagrangian and $\bar{\sigma}^{\mu}$ mix up the terms. Nevertheless, we were able to deal with this problem when deriving the Majorana Lagrangian by introducing the gamma matrices γ^{μ}. Along with this, we were guided by the notion of combining the irreducible representations to form reducible representations of spin-1/2 which resulted in the Majorana field. These reasons give us a strong reason to consider combining ξ and χ into a new spinor. To see if this is correct, we vary the action with respect to ξ and $\chi^{\dagger}$ to get the equations of motion

$$-\frac{\delta \mathcal{S}}{\delta \xi} = m\chi - i\bar{\sigma}^{\mu}\partial_{\mu}\xi^{\dagger} = 0, \tag{4.2.25}$$

$$-\frac{\delta \mathcal{S}}{\delta \chi^{\dagger}} = m\xi^{\dagger} - i\bar{\sigma}^{\mu}\partial_{\mu}\chi = 0. \tag{4.2.26}$$

Then we express the equations of motion above in matrix form

$$\begin{pmatrix} m\delta_{a}{}^{c} & -i\sigma^{\mu}_{a\dot{c}}\partial_{\mu} \\ -i\bar{\sigma}^{\mu\dot{a}c}\partial_{\mu} & m\delta^{\dot{a}}{}_{\dot{c}} \end{pmatrix} \begin{pmatrix} \chi \\ \xi^{\dagger} \end{pmatrix} = 0, \tag{4.2.27}$$

which confirms our assumption and we are left with a new field [15].

Definition 4.2.2. The $(\frac{1}{2}, 0) \oplus (0, \frac{1}{2})$ field representation of the Lorentz algebra is the four-component bispinor called the **Dirac field**, or **Dirac spinor**, $\mathbf{\Psi}$

$$\mathbf{\Psi} = \begin{pmatrix} \chi_{c} \\ \xi^{\dagger\dot{c}} \end{pmatrix}, \tag{4.2.28}$$

where χ_{c} and $\xi^{\dagger\dot{c}}$ are left- and right-handed Weyl fields, respectively. Under a Lorentz transformation, the field transforms as

$$U(\Lambda)^{-1}\mathbf{\Psi}U(\Lambda) = D(\Lambda)\mathbf{\Psi}(\Lambda^{-1}x),$$

where

$$D(\Lambda) = e^{-\frac{i}{2}\omega_{\mu\nu}S^{\mu\nu}},$$
$$S^{\mu\nu} = \frac{i}{4}[\gamma^{\mu}, \gamma^{\nu}].$$

The matrix in equation (4.2.27) is the same as the matrix in equation (4.2.11) so like before, we introduce the gamma matrices γ^{μ} and the equations of motion (4.2.27) once again elegantly reduces to

$$(-i\gamma^{\mu}\partial_{\mu} + m)\mathbf{\Psi} = 0, \tag{4.2.29}$$

which is the Dirac equation. Unsurprisingly, we see that the Dirac field's equation of motion is the Dirac equation. The Lagrangian for the Dirac field is derived in the same manner as the Majorana Lagrangian with the substitution $\psi \to \chi$ and $\psi \to \xi$ which yields

$$\mathcal{L} = \overline{\mathbf{\Psi}}(i\gamma^{\mu}\partial_{\mu} - m)\mathbf{\Psi}. \tag{4.2.30}$$

where $\overline{\mathbf{\Psi}} = \mathbf{\Psi}^{\dagger}\beta = (\xi, \chi^{\dagger})$ [5]. Now we finally prove the U(1) symmetry of the Dirac Lagrangian.

Proof. Let $\mathcal{L}$ be the Dirac Lagrangian as defined in equation (4.2.30) and consider the U(1) transformation

$$\mathbf{\Psi} \to \mathbf{\Psi}' = e^{-i\alpha}\mathbf{\Psi},$$
$$\overline{\mathbf{\Psi}} \to \overline{\mathbf{\Psi}}' = e^{+i\alpha}\overline{\mathbf{\Psi}},$$

where $\alpha \in \mathbb{R}$. Then the Dirac Lagrangian transforms as

$$\begin{aligned}\mathcal{L} \to \mathcal{L}' &= \overline{\mathbf{\Psi}}'(i\gamma^\mu \partial_\mu - m)\mathbf{\Psi}' \\ &= e^{+i\alpha}\overline{\mathbf{\Psi}}(i\gamma^\mu \partial_\mu - m)e^{-i\alpha}\mathbf{\Psi} \\ &= \overline{\mathbf{\Psi}}(i\gamma^\mu \partial_\mu - m)\mathbf{\Psi} \\ &= \mathcal{L}.\end{aligned}$$

Thus, the Dirac Lagrangian exhibits U(1) symmetry. □

With the Dirac and Majorana fields, it is sometimes advantageous to recover their individual components, namely, the Weyl fields. To do this, we define a new gamma matrix

$$\gamma_5 = i\gamma^0\gamma^1\gamma^2\gamma^3 = \begin{pmatrix} -\delta_a{}^c & 0 \\ 0 & \delta^{\dot{a}}{}_{\dot{c}} \end{pmatrix}, \tag{4.2.31}$$

then define left and right projection matrices [5]

$$P_L \equiv \frac{1}{2}(1-\gamma_5) = \begin{pmatrix} \delta_a{}^c & 0 \\ 0 & 0 \end{pmatrix}, \tag{4.2.32}$$

$$P_R \equiv \frac{1}{2}(1+\gamma_5) = \begin{pmatrix} 0 & 0 \\ 0 & \delta^{\dot{a}}{}_{\dot{c}} \end{pmatrix}. \tag{4.2.33}$$

Operating both projection matrices on a Dirac field yields

$$P_L\mathbf{\Psi} = \begin{pmatrix} \chi_c \\ 0 \end{pmatrix},$$

$$P_R\mathbf{\Psi} = \begin{pmatrix} 0 \\ \xi^{\dagger\dot{c}} \end{pmatrix},$$

and similarly for a Majorana field yields

$$P_L\boldsymbol{\chi} = \begin{pmatrix} \psi_a \\ 0 \end{pmatrix},$$
$$P_R\boldsymbol{\chi} = \begin{pmatrix} 0 \\ \psi_{\dot{a}}^\dagger \end{pmatrix}.$$

From this point until the end of this section, we will use various identities and formulas that will be derived in the next chapter with formal proofs. We will present every identity and formula used so as to explain each step of our derivation.

Before we quantize our three fields, let us carry out further analysis into our fields so that the quantization procedure yields deeper insights. Now it is evident that both the Majorana and Dirac fields have the same equation of motion which is the Dirac equation,

$$(-i\partial\!\!\!/ + m)\boldsymbol{\Psi} = 0, \tag{4.2.34}$$
$$(-i\partial\!\!\!/ + m)\boldsymbol{\chi} = 0, \tag{4.2.35}$$

where we have introduced the *Feynman slash* notation $a\!\!\!/ = a_\mu\gamma^\mu$ [5]. Multiplying on the left by the operator $(i\partial\!\!\!/ + m)$ for either of the equations gives

$$\begin{aligned} 0 &= (i\partial\!\!\!/ + m)(-i\partial\!\!\!/ + m)\boldsymbol{\Psi} \\ &= (i\partial\!\!\!/\partial\!\!\!/ + m^2)\boldsymbol{\Psi} \\ &= (-\partial^2 + m^2)\boldsymbol{\Psi}, \end{aligned}$$

where we have used the identity $a\!\!\!/a\!\!\!/ = -a^2$ in the third line [5]. Unsurprisingly, we see that both the Majorana and Dirac fields satisfy the Klein-Gordon equation which makes sense given that in nonrelativistic quantum mechanics, any wave function that satisfies the Dirac equation also satifies the Klein-Gordon equation

[11]. Thus, the Dirac equation admits plane-wave solutions of the form,

$$\mathbf{\Psi}(x) = u(\mathbf{p})e^{ipx} + v(\mathbf{p})e^{-ipx}, \tag{4.2.36}$$

where $e^{\pm ipx} = e^{\pm i(\mathbf{p}\cdot\mathbf{x})\mp i\omega t}$ and $u(\mathbf{p})$ and $v(\mathbf{p})e^{-ipx}$ are constant, four-component spinors. Recall that $\mathbf{\Psi}(x) = \mathbf{\Psi}(\mathbf{x}, t)$ since we are working with four-vectors. Now applying the Dirac equation on the Dirac field above yields

$$\begin{aligned} 0 &= (-i\not{\partial} + m)\mathbf{\Psi}(x) \\ &= (-i\not{\partial} + m)u(\mathbf{p})e^{ipx} + (-i\not{\partial} + m)v(\mathbf{p})e^{-ipx} \\ &= (\not{p} + m)u(\mathbf{p})e^{ipx} + (-\not{p} + m)v(\mathbf{p})e^{-ipx}, \end{aligned}$$

which implies that [15]

$$(\not{p} + m)u(\mathbf{p}) = 0, \tag{4.2.37}$$

$$(-\not{p} + m)v(\mathbf{p}) = 0. \tag{4.2.38}$$

The above equations each admit two linearly independent solutions $u_{\pm}(\mathbf{p})$ and $v_{\pm}(\mathbf{p})$ thus the general solution for the Dirac field is [15]

$$\mathbf{\Psi}(x) = \sum_{s=\pm} \int d^3p_L \left[b_s(\mathbf{p})u_s(\mathbf{p})e^{ipx} + d_s^{\dagger}(\mathbf{p})v_s(\mathbf{p})e^{-ipx} \right]. \tag{4.2.39}$$

For the Majorana field, some technical nuance is needed for deriving the correct general solution. The Majorana field has the condition that $\boldsymbol{\chi} = \mathcal{C}\bar{\boldsymbol{\chi}}^{\mathrm{T}}$ so if a Majorana field has the form

$$\boldsymbol{\chi}(x) = \sum_{s=\pm} \int d^3p_L \left[b_s(\mathbf{p})u_s(\mathbf{p})e^{ipx} + d_s^{\dagger}(\mathbf{p})v_s(\mathbf{p})e^{-ipx} \right], \tag{4.2.40}$$

then

$$\begin{aligned}\bar{\boldsymbol{\chi}}(x) &= \boldsymbol{\chi}^\dagger\beta \\ &= \sum_{s=\pm}\int d^3p_L\left[b_s^\dagger(\mathbf{p})u_s^\dagger(\mathbf{p})\beta e^{-ipx} + d_s(\mathbf{p})v_s^\dagger(\mathbf{p})\beta e^{ipx}\right] \\ &= \sum_{s=\pm}\int d^3p_L\left[b_s^\dagger(\mathbf{p})\bar{u}_s(\mathbf{p})e^{-ipx} + d_s(\mathbf{p})\bar{v}_s(\mathbf{p})e^{ipx}\right],\end{aligned}$$

where $\bar{u}_s(\mathbf{p}) = u_s^\dagger(\mathbf{p})\beta$ and $\bar{v}_s(\mathbf{p}) = v_s^\dagger(\mathbf{p})\beta$ [5]. Then we use the identities

$$\mathcal{C}\bar{u_s}(\mathbf{p})^{\mathrm{T}} = v_s(\mathbf{p}), \tag{4.2.41}$$

$$\mathcal{C}\bar{v_s}(\mathbf{p})^{\mathrm{T}} = u_s(\mathbf{p}), \tag{4.2.42}$$

to get

$$\begin{aligned}\mathcal{C}\bar{\boldsymbol{\chi}}^{\mathrm{T}}(x) &= \sum_{s=\pm}\int d^3p_L\left[b_s^\dagger(\mathbf{p})\mathcal{C}\bar{u_s}^{\mathrm{T}}(\mathbf{p})e^{-ipx} + d_s(\mathbf{p})\mathcal{C}\bar{v_s}^{\mathrm{T}}(\mathbf{p})e^{ipx}\right] \\ &= \sum_{s=\pm}\int d^3p_L\left[b_s^\dagger(\mathbf{p})v_s(\mathbf{p})e^{-ipx} + d_s(\mathbf{p})u_s(\mathbf{p})e^{ipx}\right].\end{aligned}$$

Comparing this result and equation (4.2.40) leads the conclusion that

$$d_s(\mathbf{p}) = b_s(\mathbf{p}), \tag{4.2.43}$$

thus we arrive at the general solutions for the Dirac and Majorana fields [5]

$$\boldsymbol{\Psi}(x) = \sum_{s=\pm}\int d^3p_L\left[b_s(\mathbf{p})u_s(\mathbf{p})e^{ipx} + d_s^\dagger(\mathbf{p})v_s(\mathbf{p})e^{-ipx}\right], \tag{4.2.44}$$

$$\boldsymbol{\chi}(x) = \sum_{s=\pm}\int d^3p_L\left[b_s(\mathbf{p})u_s(\mathbf{p})e^{ipx} + b_s^\dagger(\mathbf{p})v_s(\mathbf{p})e^{-ipx}\right]. \tag{4.2.45}$$

When we quantize the Dirac and Majorana fields, we anticipate that the Fourier coefficients $b_s(\mathbf{p})$ and $d_s(\mathbf{p})$ will be operators since this is a consequence of quantizing a field that obeys the Klein-Gordon equation. For the sake of future brevity and computational ease, let us follow the same procedure of the real scalar field and

rewrite the Fourier coefficients $b_s(\mathbf{p})$ and $d_s(\mathbf{p})$ in terms of $\boldsymbol{\Psi}$ ($\boldsymbol{\chi}$) and $\overline{\boldsymbol{\Psi}}$ ($\bar{\boldsymbol{\chi}}$) and write the Hamiltonian in terms of these Fourier coefficients. Since it is obvious that the Majorana field is a specific case of the Dirac field, we will start with the Dirac field.

Starting with the general solution of the Dirac field,

$$\boldsymbol{\Psi}(x) = \sum_{s'=\pm} \int d^3p'_L \left[b_{s'}(\mathbf{p}')u_{s'}(\mathbf{p}')e^{ip'x} + d^\dagger_{s'}(\mathbf{p}')v_{s'}(\mathbf{p}')e^{-ip'x} \right], \tag{4.2.46}$$

we multiply on the left by e^{-ipx} then integrate with respect to x [5]

$$\begin{aligned}
\int d^3x\, e^{-ipx}\, \boldsymbol{\Psi}(x) &= \int d^3x \sum_{s'=\pm} \int d^3p'_L \left[b_{s'}(\mathbf{p}')u_{s'}(\mathbf{p}')e^{ip'x} \right. \\
&\quad \left. + d^\dagger_{s'}(\mathbf{p}')v_{s'}(\mathbf{p}')e^{-ip'x} \right] e^{-ipx} \\
\int d^3x\, e^{-ipx}\, \boldsymbol{\Psi}(x) &= \sum_{s'=\pm} \int d^3p'_L\, d^3x \left[b_{s'}(\mathbf{p}')u_{s'}(\mathbf{p}')e^{i(p'-p)x} \right. \\
&\quad \left. + d^\dagger_{s'}(\mathbf{p}')v_{s'}(\mathbf{p}')e^{-i(p'+p)x} \right] \\
\int d^3x\, e^{-ipx}\, \boldsymbol{\Psi}(x) &= \sum_{s'=\pm} \int d^3p' \frac{1}{2\omega'} \left[b_{s'}(\mathbf{p}')u_{s'}(\mathbf{p}')e^{i(\omega'-\omega)t}\delta^3(\mathbf{p}'-\mathbf{p}) \right. \\
&\quad \left. + d^\dagger_{s'}(\mathbf{p}')v_{s'}(\mathbf{p}')e^{i(\omega'+\omega)t}\delta^3(\mathbf{p}'+\mathbf{p}) \right] \\
\int d^3x\, e^{-ipx}\, \boldsymbol{\Psi}(x) &= \sum_{s'=\pm} \frac{1}{2\omega} \left[b_{s'}(\mathbf{p})u_{s'}(\mathbf{p}) + d^\dagger_{s'}(-\mathbf{p})v_{s'}(-\mathbf{p})e^{2i\omega t} \right].
\end{aligned}$$

Applying the same operations except with e^{ipx} yields

$$\int d^3x\, e^{ipx}\, \boldsymbol{\Psi}(x) = \sum_{s'=\pm} \frac{1}{2\omega} \left[b_{s'}(-\mathbf{p})u_{s'}(-\mathbf{p})e^{-2i\omega t} + d^\dagger_{s'}(\mathbf{p})v_{s'}(\mathbf{p}) \right],$$

thus we have the following equations,

$$\int d^3x\, e^{-ipx}\, \boldsymbol{\Psi}(x) = \sum_{s'=\pm} \frac{1}{2\omega} \left[b_{s'}(\mathbf{p})u_{s'}(\mathbf{p}) + d^\dagger_{s'}(-\mathbf{p})v_{s'}(-\mathbf{p})e^{2i\omega t} \right], \tag{4.2.47}$$

$$\int d^3x\, e^{ipx}\, \boldsymbol{\Psi}(x) = \sum_{s'=\pm} \frac{1}{2\omega} \left[b_{s'}(-\mathbf{p})u_{s'}(-\mathbf{p})e^{-2i\omega t} + d^\dagger_{s'}(\mathbf{p})v_{s'}(\mathbf{p}) \right]. \tag{4.2.48}$$

We multiply on the left $\bar{u}_s(\mathbf{p})\gamma^0$ and $\bar{v}_s(\mathbf{p})\gamma^0$ on equations (4.2.47) and (4.2.48), respectively, to get

$$\begin{aligned}\int d^3x\, e^{-ipx}\, \bar{u}_s(\mathbf{p})\gamma^0\mathbf{\Psi}(x) &= \sum_{s'=\pm}\frac{1}{2\omega}\Big[b_{s'}(\mathbf{p})\bar{u}_s(\mathbf{p})\gamma^0 u_{s'}(\mathbf{p}) \\ &\quad + d^\dagger_{s'}(-\mathbf{p})\bar{u}_s(\mathbf{p})\gamma^0 v_{s'}(-\mathbf{p})e^{2i\omega t}\Big], \\ \int d^3x\, e^{ipx}\, \bar{v}_s(\mathbf{p})\gamma^0\mathbf{\Psi}(x) &= \sum_{s'=\pm}\frac{1}{2\omega}\Big[b_{s'}(-\mathbf{p})\bar{v}_s(\mathbf{p})\gamma^0 u_{s'}(-\mathbf{p})e^{-2i\omega t} \\ &\quad + d^\dagger_{s'}(\mathbf{p})\bar{v}_s(\mathbf{p})\gamma^0 v_{s'}(\mathbf{p})\Big].\end{aligned}$$

We use the following identities,

$$\begin{aligned}\bar{u}_{s'}(\mathbf{p})\gamma^0 u_s(\mathbf{p}) &= 2\omega\delta_{s's}, \\ \bar{v}_{s'}(\mathbf{p})\gamma^0 v_s(\mathbf{p}) &= 2\omega\delta_{s's}, \\ \bar{u}_{s'}(\mathbf{p})\gamma^0 v_s(-\mathbf{p}) &= 0, \\ \bar{v}_{s'}(\mathbf{p})\gamma^0 u_s(-\mathbf{p}) &= 0,\end{aligned} \tag{4.2.49}$$

which reduces the above equations to finally yield [5]

$$b_s(\mathbf{p}) = \int d^3x\, e^{-ipx}\bar{u}_s(\mathbf{p})\gamma^0\mathbf{\Psi}(x), \tag{4.2.50}$$

$$d^\dagger_s(\mathbf{p}) = \int d^3x\, e^{ipx}\bar{v}_s(\mathbf{p})\gamma^0\mathbf{\Psi}(x). \tag{4.2.51}$$

Hermitian conjugating the above coefficients gives us the last two coefficients, namely $b^\dagger_s(\mathbf{p})$ and $d_s(\mathbf{p})$, thus we finally have the desired coefficients in terms of $\mathbf{\Psi}$

and $\overline{\mathbf{\Psi}}$,

$$b_s(\mathbf{p}) = \int d^3x\, e^{-ipx} \bar{u}_s(\mathbf{p}) \gamma^0 \mathbf{\Psi}(x), \tag{4.2.52}$$

$$b_s^\dagger(\mathbf{p}) = \int d^3x\, e^{ipx} \overline{\mathbf{\Psi}}(x) \gamma^0 u_s(\mathbf{p}), \tag{4.2.53}$$

$$d_s(\mathbf{p}) = \int d^3x\, e^{-ipx} \overline{\mathbf{\Psi}}(x) \gamma^0 v_s(\mathbf{p}), \tag{4.2.54}$$

$$d_s^\dagger(\mathbf{p}) = \int d^3x\, e^{ipx} \bar{v}_s(\mathbf{p}) \gamma^0 \mathbf{\Psi}(x), \tag{4.2.55}$$

where we remind ourselves that $s = \pm$ [5]. At last, we are finally in a position to express the Hamiltonian of the Dirac (Majorana) field in terms of the Fourier coefficients. First, we identify the conjugate momentum of the Dirac field

$$\Pi(x) = \frac{\partial \mathcal{L}}{\partial(\partial_0 \mathbf{\Psi})} = -i\overline{\mathbf{\Psi}}(x)\gamma^0, \tag{4.2.56}$$

then the Hamiltonian of the Dirac field is

$$\begin{aligned} H_D &= \int d^3x \left(\Pi(\partial_0 \mathbf{\Psi}) - \mathcal{L} \right) \\ &= \int d^3x \left(-i\overline{\mathbf{\Psi}}\gamma^0(\partial_0 \mathbf{\Psi}) + i\overline{\mathbf{\Psi}}\gamma^0(\partial_0 \mathbf{\Psi}) - i\overline{\mathbf{\Psi}}\gamma^i(\partial_i \mathbf{\Psi}) + m\overline{\mathbf{\Psi}}\mathbf{\Psi} \right) \\ &= \int d^3x\, \overline{\mathbf{\Psi}}(-i\gamma^i\partial_i + m)\mathbf{\Psi}. \end{aligned} \tag{4.2.57}$$

To ease further computations and preserve clarity in our derivations, we will proceed with simplifying the Dirac Hamiltonian one part at a time. Now note that

$$\begin{aligned} (-i\gamma^i\partial_i + m)\mathbf{\Psi} &= \sum_{s=\pm} \int d^3p_L \Big[b_s(\mathbf{p})(-i\gamma^i\partial_i + m)u_s(\mathbf{p})e^{ipx} \\ &\quad + d_s^\dagger(\mathbf{p})(-i\gamma^i\partial_i + m)v_s(\mathbf{p})e^{-ipx} \Big] \\ &= \sum_{s=\pm} \int d^3p_L \Big[b_s(\mathbf{p})(\gamma^i p_i + m)u_s(\mathbf{p})e^{ipx} \\ &\quad + d_s^\dagger(\mathbf{p})(-\gamma^i p_i + m)v_s(\mathbf{p})e^{-ipx} \Big] \\ &= \sum_{s=\pm} \int d^3p_L \Big[b_s(\mathbf{p})(\gamma^0\omega)u_s(\mathbf{p})e^{ipx} + d_s^\dagger(\mathbf{p})(-\gamma^0\omega)v_s(\mathbf{p})e^{-ipx} \Big] \end{aligned}$$

where in the third line, we used [5]

$$(\gamma^0\omega)u_s(\mathbf{p}) = (\gamma^i p_i + m)u_s(\mathbf{p}),$$
$$(-\gamma^0\omega)v_s(\mathbf{p}) = (-\gamma^i p_i + m)v_s(\mathbf{p}).$$

Then the Dirac Hamiltonian becomes

$$\begin{aligned}
H_D &= \int d^3x\, \overline{\mathbf{\Psi}}(-i\gamma^i\partial_i + m)\mathbf{\Psi} \\
&= \sum_{s',s=\pm}\int d^3p_L\, d^3p'_L\, d^3x \left[b^\dagger_{s'}(\mathbf{p}')\bar{u}_{s'}(\mathbf{p}')e^{-ip'x} + d_{s'}(\mathbf{p}')\bar{v}_{s'}(\mathbf{p}')e^{ip'x}\right] \\
&\times \left[b_s(\mathbf{p})(\gamma^0\omega)u_s(\mathbf{p})e^{ipx} + d^\dagger_s(\mathbf{p})(-\gamma^0\omega)v_s(\mathbf{p})e^{-ipx}\right] \\
&= \sum_{s',s=\pm}\int d^3p_L\, d^3p'_L\, d^3x \left[b^\dagger_{s'}(\mathbf{p}')b_s(\mathbf{p})\bar{u}_{s'}(\mathbf{p}')(\gamma^0\omega)u_s(\mathbf{p})e^{i(p-p')x}\right. \\
&+ b^\dagger_{s'}(\mathbf{p}')d^\dagger_s(\mathbf{p})\bar{u}_{s'}(\mathbf{p}')(-\gamma^0\omega)v_s(\mathbf{p})e^{-i(p+p')x} \\
&+ d_{s'}(\mathbf{p}')b_s(\mathbf{p})\bar{v}_{s'}(\mathbf{p}')(\gamma^0\omega)u_s(\mathbf{p})e^{i(p+p')x} \\
&\left.+ d_{s'}(\mathbf{p}')d^\dagger_s(\mathbf{p})\bar{v}_{s'}(\mathbf{p}')(-\gamma^0\omega)v_s(\mathbf{p})e^{-i(p'-p)x}\right] \\
&= \sum_{s',s=\pm}\int d^3p_L\, d^3p'\, \frac{1}{2\omega'}\left[b^\dagger_{s'}(\mathbf{p}')b_s(\mathbf{p})\bar{u}_{s'}(\mathbf{p}')(\gamma^0\omega)u_s(\mathbf{p})\delta^3(p'-p)\right. \\
&+ b^\dagger_{s'}(\mathbf{p}')d^\dagger_s(\mathbf{p})\bar{u}_{s'}(\mathbf{p}')(-\gamma^0\omega)v_s(\mathbf{p})\delta^3(p'+p) \\
&+ d_{s'}(\mathbf{p}')b_s(\mathbf{p})\bar{v}_{s'}(\mathbf{p}')(\gamma^0\omega)u_s(\mathbf{p})\delta^3(p'+p) \\
&\left.+ d_{s'}(\mathbf{p}')d^\dagger_s(\mathbf{p})\bar{v}_{s'}(\mathbf{p}')(-\gamma^0\omega)v_s(\mathbf{p})\delta^3(p'-p)\right] \\
&= \sum_{s',s=\pm}\int d^3p_L\, \frac{1}{2\omega}\left[b^\dagger_{s'}(\mathbf{p})b_s(\mathbf{p})\bar{u}_{s'}(\mathbf{p})(\gamma^0\omega)u_s(\mathbf{p})\right. \\
&+ b^\dagger_{s'}(-\mathbf{p})d^\dagger_s(\mathbf{p})\bar{u}_{s'}(-\mathbf{p})(-\gamma^0\omega)v_s(\mathbf{p}) \\
&+ d_{s'}(-\mathbf{p})b_s(\mathbf{p})\bar{v}_{s'}(-\mathbf{p})(\gamma^0\omega)u_s(\mathbf{p}) \\
&\left.+ d_{s'}(\mathbf{p})d^\dagger_s(\mathbf{p})\bar{v}_{s'}(\mathbf{p})(-\gamma^0\omega)v_s(\mathbf{p})\right] \\
&= \sum_{s',s=\pm}\int d^3p_L\, \omega\Big[b^\dagger_{s'}(\mathbf{p})b_s(\mathbf{p}) - d_{s'}(\mathbf{p})d^\dagger_s(\mathbf{p})\Big]\delta_{s's} \\
&= \sum_{s=\pm}\int d^3p_L\, \omega\Big[b^\dagger_s(\mathbf{p})b_s(\mathbf{p}) - d_s(\mathbf{p})d^\dagger_s(\mathbf{p})\Big],
\end{aligned}$$

hence

$$H_D = \sum_{s=\pm} \int d^3p_L \, \omega \Big[b_s^\dagger(\mathbf{p}) b_s(\mathbf{p}) - d_s(\mathbf{p}) d_s^\dagger(\mathbf{p}) \Big], \tag{4.2.58}$$

where we have carefully preserved the order of each term in anticipation for quantization [5]. Imposing the Majorana condition $\boldsymbol{\chi} = \mathcal{C}\bar{\boldsymbol{\chi}}^{\mathrm{T}}$ implies that $d_s(\mathbf{p}) = b_s(\mathbf{p})$ so the Majorana Hamiltonian is

$$H_M = \frac{1}{2} \sum_{s=\pm} \int d^3p_L \, \omega \Big[b_s^\dagger(\mathbf{p}) b_s(\mathbf{p}) - b_s(\mathbf{p}) b_s^\dagger(\mathbf{p}) \Big]. \tag{4.2.59}$$

Now the conjugate momentum $\pi^a(x)$ for the Weyl field is

$$\begin{aligned} \pi^a(x) &= \frac{\partial \mathcal{L}}{\partial(\partial_0 \psi_a)} \\ &= i\psi_a^\dagger(x) \bar{\sigma}^{0a\dot{a}}, \end{aligned}$$

so the Hamiltonian for the Weyl field is

$$\begin{aligned} H_W &= \int d^3x \, \mathcal{H} \\ &= \int d^3x \, \Big[-i\psi^\dagger \bar{\sigma}^i \partial_i \psi + \frac{1}{2} m(\psi\psi + \psi^\dagger \psi^\dagger) \Big]. \end{aligned}$$

Finally, we now have our fermion Hamiltonians

$$H_D = \sum_{s=\pm} \int d^3p_L \, \omega \Big[b_s^\dagger(\mathbf{p}) b_s(\mathbf{p}) - d_s(\mathbf{p}) d_s^\dagger(\mathbf{p}) \Big], \tag{4.2.60}$$

$$H_M = \frac{1}{2} \sum_{s=\pm} \int d^3p_L \, \omega \Big[b_s^\dagger(\mathbf{p}) b_s(\mathbf{p}) - b_s(\mathbf{p}) b_s^\dagger(\mathbf{p}) \Big], \tag{4.2.61}$$

$$H_W = \int d^3x \, \Big[-i\psi^\dagger \bar{\sigma}^i \partial_i \psi + \frac{1}{2} m(\psi\psi + \psi^\dagger \psi^\dagger) \Big], \tag{4.2.62}$$

where H_D, H_M, and H_W are the Hamiltonians for the Dirac, Majorana, and Weyl fields, respectively [5]. With this, we conclude our derivation and analysis of classical spinor fields and we proceed onward to quantize our classical theory into a quantum theory.

Quantizing the Classical Fermion Fields

It is rather obvious at this point that the Weyl, Dirac, and Majorana fields possess spin-1/2. The astute reader familiar with nonrelativistic quantum mechanics will know that quantizing half-integer fields and particles requires that we use *anticommutators* rather than commutators since spin-1/2 fields and particles are *fermions*. However for reasons of pedagogy and insight into the underlying physics of fermions, we will present the quantization of the spin-1/2 fields in the incorrect fashion by using commutators to show why this approach leads to incorrect results. Since the Dirac and Majorana fields are built using Weyl fields, we promote the left-handed Weyl field $\psi_a(x)$ and its conjugate momentum $\pi^a(x)$ to operators $\hat{\psi}_a(x)$ and $\hat{\pi}^a(x)$, respectively, then (naively) impose the canonical commutation relations on the operators,

$$[\hat{\psi}_a(\mathbf{x},t),\hat{\psi}_c(\mathbf{y},t)] = 0, \tag{4.2.63}$$

$$[\hat{\psi}_a(\mathbf{x},t),\hat{\pi}^c(\mathbf{y},t)] = i\delta_a{}^c\delta^3(\mathbf{x}-\mathbf{y}). \tag{4.2.64}$$

Relation (4.2.64) can be further reduced by inserting the definition of $\hat{\pi}^c(\mathbf{y},t)$ which yields

$$\begin{aligned}
[\hat{\psi}_a(\mathbf{x},t),\hat{\pi}^c(\mathbf{y},t)] &= i\delta_a{}^c\delta^3(\mathbf{x}-\mathbf{y}) \\
[\hat{\psi}_a(\mathbf{x},t), i\hat{\psi}^\dagger_{\dot{c}}(\mathbf{y},t)\bar{\sigma}^{0\dot{c}c}] &= i\delta_a{}^c\delta^3(\mathbf{x}-\mathbf{y}) \\
[\hat{\psi}_a(\mathbf{x},t),\hat{\psi}^\dagger_{\dot{c}}(\mathbf{y},t)]\bar{\sigma}^{0\dot{c}c} &= \delta_a{}^c\delta^3(\mathbf{x}-\mathbf{y}).
\end{aligned}$$

We multiply on the right by $\bar{\sigma}^0_{\dot{c}c}$ and we get

$$[\hat{\psi}_a(\mathbf{x},t),\hat{\psi}^\dagger_{\dot{c}}(\mathbf{y},t)] = \bar{\sigma}^0_{a\dot{c}}\delta^3(\mathbf{x}-\mathbf{y}), \tag{4.2.65}$$

$$[\hat{\psi}^a(\mathbf{x},t),\hat{\psi}^{\dagger\dot{c}}(\mathbf{y},t)] = \bar{\sigma}^{0\dot{c}a}\delta^3(\mathbf{x}-\mathbf{y}), \tag{4.2.66}$$

where the second relation comes from lowering and switching the order of the indices. The Weyl field commutation relations implies that for the Dirac field

$$[\hat{\mathbf{\Psi}}_\alpha(\mathbf{x},t), \hat{\mathbf{\Psi}}_\beta(\mathbf{y},t)] = 0, \tag{4.2.67}$$

$$[\hat{\mathbf{\Psi}}_\alpha(\mathbf{x},t), \hat{\overline{\mathbf{\Psi}}}_\beta(\mathbf{y},t)] = (\gamma^0)_{\alpha\beta}\, \delta^3(\mathbf{x}-\mathbf{y}). \tag{4.2.68}$$

where α and β are four-component spinor indices.

Proof. Let $\hat{\mathbf{\Psi}}$ and $\hat{\overline{\mathbf{\Psi}}}$ be defined as

$$\hat{\mathbf{\Psi}} = \begin{pmatrix} \hat{\chi}_c \\ \hat{\xi}^{\dagger\dot{c}} \end{pmatrix},$$
$$\hat{\overline{\mathbf{\Psi}}} = (\hat{\xi}^a, \hat{\chi}^\dagger_{\dot{a}}).$$

The relation (4.2.67) immediately follows from relation (4.2.63). For (4.2.68),

$$\begin{aligned}
[\hat{\mathbf{\Psi}}_\alpha(\mathbf{x},t), \hat{\overline{\mathbf{\Psi}}}_\beta(\mathbf{y},t)] &= \begin{pmatrix} [\hat{\chi}_c, \hat{\xi}^a] & [\hat{\chi}_c, \hat{\chi}^\dagger_{\dot{a}}] \\ [\hat{\xi}^{\dagger\dot{c}}, \hat{\xi}^a] & [\hat{\xi}^{\dagger\dot{c}}, \hat{\chi}^\dagger_{\dot{a}}] \end{pmatrix} \\
&= \begin{pmatrix} 0 & \sigma^0_{c\dot{a}} \\ \bar{\sigma}^0_{c\dot{a}} & 0 \end{pmatrix} \delta^3(\mathbf{x}-\mathbf{y}) \\
&= \gamma^0 \delta^3(\mathbf{x}-\mathbf{y}),
\end{aligned}$$

where we used relations (4.2.65) and (4.2.66) in the second line and the definition of the gamma matrix γ^0 in the last line. □

Since the Dirac field is now an operator, we see that the Fourier coefficients turn into operators with the following commutation relations following immediately

from the Dirac commutation relations,

$$
\begin{aligned}
[\hat{b}_s(\mathbf{p}), \hat{b}_{s'}(\mathbf{p}')] &= 0, \\
[\hat{d}_s(\mathbf{p}), \hat{d}_{s'}(\mathbf{p}')] &= 0, \\
[\hat{b}_s(\mathbf{p}), \hat{d}^\dagger_{s'}(\mathbf{p}')] &= 0.
\end{aligned} \tag{4.2.69}
$$

Hermitian conjugating the above trivially yields

$$
\begin{aligned}
[\hat{b}^\dagger_s(\mathbf{p}), \hat{b}^\dagger_{s'}(\mathbf{p}')] &= 0, \\
[\hat{d}^\dagger_s(\mathbf{p}), \hat{d}^\dagger_{s'}(\mathbf{p}')] &= 0, \\
[\hat{b}^\dagger_s(\mathbf{p}), \hat{d}_{s'}(\mathbf{p}')] &= 0.
\end{aligned} \tag{4.2.70}
$$

Observe that,

$$
\begin{aligned}
[\hat{b}_s(\mathbf{p}), \hat{b}^\dagger_{s'}(\mathbf{p}')] &= \int d^3x\, d^3y\, y e^{-i(px-p'y)} \bar{u}_s(\mathbf{p}) \gamma^0 [\hat{\mathbf{\Psi}}(x), \hat{\overline{\mathbf{\Psi}}}(y)] \gamma^0 u_{s'}(\mathbf{p}') \\
&= \int d^3x\, d^3y\, e^{-i(px-p'y)} \bar{u}_s(\mathbf{p}) \gamma^0 \gamma^0 \delta^3(\mathbf{x}-\mathbf{y}) \gamma^0 u_{s'}(\mathbf{p}') \\
&= \int d^3x\, e^{-i(p-p')x} \bar{u}_s(\mathbf{p}) \gamma^0 u_{s'}(\mathbf{p}') \\
&= (2\omega)(2\pi)^3 \delta_{ss'} \delta^3(\mathbf{p}-\mathbf{p}'), \\
[\hat{d}^\dagger_s(\mathbf{p}), \hat{d}_{s'}(\mathbf{p}')] &= \int d^3x\, d^3y\, y e^{i(px-p'y)} \bar{v}_s(\mathbf{p}) \gamma^0 [\hat{\mathbf{\Psi}}(x), \hat{\overline{\mathbf{\Psi}}}(y)] \gamma^0 v_{s'}(\mathbf{p}') \\
&= \int d^3x\, d^3y\, e^{i(px-p'y)} \bar{v}_s(\mathbf{p}) \gamma^0 \gamma^0 \delta^3(\mathbf{x}-\mathbf{y}) \gamma^0 v_{s'}(\mathbf{p}') \\
&= \int d^3x\, e^{i(p-p')x} \bar{v}_s(\mathbf{p}) \gamma^0 v_{s'}(\mathbf{p}') \\
&= (2\omega)(2\pi)^3 \delta_{ss'} \delta^3(\mathbf{p}-\mathbf{p}'), \\
[\hat{b}_s(\mathbf{p}), \hat{d}^\dagger_{s'}(\mathbf{p}')] &= \int d^3x\, d^3y\, y e^{-i(px+p'y)} \bar{u}_s(\mathbf{p}) \gamma^0 [\hat{\mathbf{\Psi}}(x), \hat{\overline{\mathbf{\Psi}}}(y)] \gamma^0 v_{s'}(\mathbf{p}') \\
&= \int d^3x\, d^3y\, e^{-i(px+p'y)} \bar{u}_s(\mathbf{p}) \gamma^0 \gamma^0 \delta^3(\mathbf{x}-\mathbf{y}) \gamma^0 v_{s'}(\mathbf{p}') \\
&= \int d^3x\, e^{-i(p+p')x} \bar{u}_s(\mathbf{p}) \gamma^0 v_{s'}(\mathbf{p}') \\
&= 0,
\end{aligned}
$$

where we have used the identities

$$(\gamma^0)^2 = 1, \tag{4.2.71}$$

$$\bar{u}_s(\mathbf{p})\gamma^0 u_{s'}(\mathbf{p}') = (2\omega)\delta_{ss'}, \tag{4.2.72}$$

$$\bar{v}_s(\mathbf{p})\gamma^0 v_{s'}(\mathbf{p}') = (2\omega)\delta_{ss'}, \tag{4.2.73}$$

$$\bar{u}_s(\mathbf{p})\gamma^0 v_{s'}(\mathbf{p}') = 0. \tag{4.2.74}$$

Hence we have

$$[\hat{b}_s(\mathbf{p}), \hat{b}^\dagger_{s'}(\mathbf{p}')] = (2\omega)(2\pi)^3\delta_{ss'}\delta^3(\mathbf{p}-\mathbf{p}'), \tag{4.2.75}$$

$$[\hat{d}^\dagger_s(\mathbf{p}), \hat{d}_{s'}(\mathbf{p}')] = (2\omega)(2\pi)^3\delta_{ss'}\delta^3(\mathbf{p}-\mathbf{p}'), \tag{4.2.76}$$

$$[\hat{b}_s(\mathbf{p}), \hat{d}^\dagger_{s'}(\mathbf{p}')] = 0. \tag{4.2.77}$$

Along this, we also see that

$$[\hat{H}, \hat{b}^\dagger_s(\mathbf{p})] = \omega\hat{b}^\dagger_s(\mathbf{p}), \qquad [\hat{H}, \hat{b}_s(\mathbf{p})] = -\omega\hat{b}_s(\mathbf{p}), \tag{4.2.78}$$

$$[\hat{H}, \hat{d}^\dagger_s(\mathbf{p})] = \omega\hat{d}^\dagger_s(\mathbf{p}), \qquad [\hat{H}, \hat{d}_s(\mathbf{p})] = -\omega\hat{d}_s(\mathbf{p}). \tag{4.2.79}$$

With these commutation relations, we see that the operators $\hat{b}$ and $\hat{d}$ each satisfy separate harmonic oscillator algebras and they are decoupled which follows from (4.2.77). Thus, we can identify $\hat{b}^\dagger_s(\mathbf{p})$ and $\hat{d}^\dagger_s(\mathbf{p})$ as creation operators and $\hat{b}_s(\mathbf{p})$ and $\hat{d}_s(\mathbf{p})$ as destruction operators. Using these relations, the Dirac Hamiltonian can be rewritten as

$$\begin{aligned}\hat{H} &= \sum_{s=\pm}\int d^3p_L\,\omega\Big[\hat{b}^\dagger_s(\mathbf{p})\hat{b}_s(\mathbf{p}) - \hat{d}_s(\mathbf{p})\hat{d}^\dagger_s(\mathbf{p})\Big] \\ &= \sum_{s=\pm}\int d^3p_L\,\omega\Big[\hat{b}^\dagger_s(\mathbf{p})\hat{b}_s(\mathbf{p}) - \hat{d}^\dagger_s(\mathbf{p})\hat{d}_s(\mathbf{p}) + (2\omega)(2\pi)^3\delta^3(0)\Big].\end{aligned} \tag{4.2.80}$$

We have reached several serious issues. First, we have violated one of the most important axioms of quantum *physics*, namely, unitarity. To see this, we first note

that the coommutation relation (4.2.76) can alternatively be written as

$$[\hat{d}_s(\mathbf{p}), \hat{d}^\dagger_{s'}(\mathbf{p}')] = -(2\omega)(2\pi)^3\delta_{ss'}\delta^3(\mathbf{p}-\mathbf{p}'), \tag{4.2.81}$$

where we have multiplied (4.2.76) by -1. Now like the real scalar field, we then define a vacuum state $|0\rangle$ to be the state with no particles with the normalization $\langle 0|0\rangle = 1$. Using the commutation (4.2.81) with $s = s'$ and $\mathbf{p} = \mathbf{p}'$, we observe that

$$\begin{aligned}
||\hat{d}^\dagger_s(\mathbf{p})|0\rangle||^2 &= \langle 0|\hat{d}_s(\mathbf{p})\hat{d}^\dagger_s(\mathbf{p})|0\rangle \\
&= \langle 0|\hat{d}_s(\mathbf{p})\hat{d}^\dagger_s(\mathbf{p}) - 0|0\rangle \\
&= \langle 0|\hat{d}_s(\mathbf{p})\hat{d}^\dagger_s(\mathbf{p}) - \hat{d}^\dagger_s(\mathbf{p})\hat{d}_s(\mathbf{p})|0\rangle \\
&= \langle 0|[\hat{d}_s(\mathbf{p}), \hat{d}^\dagger_s(\mathbf{p})]|0\rangle \\
&= -(2\omega)(2\pi)^3\delta^3(0) \\
&< 0.
\end{aligned}$$

Another issue is that the Dirac Hamiltonian in its current form is unbounded from below. To see this, the second term in (4.2.80), $\hat{d}^\dagger_s(\mathbf{p})\hat{d}_s(\mathbf{p})$, is the Hermitian number operator which can have values $n_d = 0, 1, 2, ...$ depending on the state $|n_d\rangle$. Thus for arbitrarily higher excited states, the energy spectrum of the Dirac Hamiltonian is negative and can continue to lower indefinitely. Physically, this would mean that since there is no ground state, a system can continue to create an infinite amount of d-type particles to lower its energy, which is highly absurd. Similar results also occur for the Majorana Hamiltonian as well. By now, it is rather clear that the flaws of the Dirac Hamiltonian come from the troublesome commutation relation,

$$[\hat{d}_s(\mathbf{p}), \hat{d}^\dagger_{s'}(\mathbf{p}')] = -(2\omega)(2\pi)^3\delta_{ss'}\delta^3(\mathbf{p}-\mathbf{p}'),$$

which was the result when we naively chose to quantize the spin-1/2 fields using commutators. Thus, our quantization procedure using commutators must be

abandoned for spin-1/2 fields since they lead to absurd physical results.

To fix this issue, we must look into the nature of the field that we are trying, and failing, to quantize. Recall that when we quantized the Klein-Gordon field, we found that the quantum Klein-Gordon field and its excitations were spin zero thus they obeyed Bose-Einstein statistics hence the Klein-Gordon field is a bosonic field with excitations called bosons. Our mistake was that we treated the spinor field and its excitations as bosons whose quantum state is symmetric under the interchange of any two particles. As mentioned in the beginning of this section, the spin-1/2 fields and particles are fermions hence they obey *Fermi-Dirac statistics*, namely, their quantum state state is *antisymmetric* under the exchange of any two particles. Quantizing fermions requires that we use *anticommutators* which are defined as $\{A, B\} = AB + BA$. Thus the spinor fields, must be quantized using anti-commutators. So we proceed with quantizing the Weyl field by imposing the canonical *anticommutation* relations,

$$\{\hat{\psi}_a(\mathbf{x}, t), \hat{\psi}_c(\mathbf{y}, t)\} = 0, \tag{4.2.82}$$

$$\{\hat{\psi}_a(\mathbf{x}, t), \hat{\pi}^c(\mathbf{y}, t)\} = i\delta_a{}^c\delta^3(\mathbf{x} - \mathbf{y}). \tag{4.2.83}$$

With this, all the other commutation relations that we derived earlier follow since all the relation came from the Weyl field relations so all we have to do is simply replace all the commutators with anti-commutators. Thus we have the canonical anti-commutation relations for the Dirac field,

$$\{\hat{\mathbf{\Psi}}_\alpha(\mathbf{x}, t), \hat{\mathbf{\Psi}}_\beta(\mathbf{y}, t)\} = 0, \tag{4.2.84}$$

$$\{\hat{\mathbf{\Psi}}_\alpha(\mathbf{x}, t), \hat{\overline{\mathbf{\Psi}}}_\beta(\mathbf{y}, t)\} = (\gamma^0)_{\alpha\beta}\,\delta^3(\mathbf{x} - \mathbf{y}). \tag{4.2.85}$$

and the anticommutation relations for the Majorana field,

$$\{\hat{\boldsymbol{\chi}}_\alpha(\mathbf{x},t),\hat{\boldsymbol{\chi}}_\beta(\mathbf{y},t)\} = (\mathcal{C}\gamma^0)_{\alpha\beta}\delta^3(\mathbf{x}-\mathbf{y}), \tag{4.2.86}$$

$$\{\hat{\boldsymbol{\chi}}_\alpha(\mathbf{x},t),\hat{\overline{\boldsymbol{\chi}}}_\beta(\mathbf{y},t)\} = (\gamma^0)_{\alpha\beta}\,\delta^3(\mathbf{x}-\mathbf{y}). \tag{4.2.87}$$

Similarly, the anticommutation relations of the Fourier coefficients of the Dirac and Majorana field are

$$\begin{aligned}
&\{\hat{b}_s(\mathbf{p}),\hat{b}_{s'}(\mathbf{p}')\} = 0, && \{\hat{b}_s^\dagger(\mathbf{p}),\hat{b}_{s'}^\dagger(\mathbf{p}')\} = 0,\\
&\{\hat{d}_s(\mathbf{p}),\hat{d}_{s'}(\mathbf{p}')\} = 0, && \{\hat{d}_s^\dagger(\mathbf{p}),\hat{d}_{s'}^\dagger(\mathbf{p}')\} = 0,\\
&\{\hat{b}_s(\mathbf{p}),\hat{d}_{s'}^\dagger(\mathbf{p}')\} = 0, && \{\hat{b}_s^\dagger(\mathbf{p}),\hat{d}_{s'}(\mathbf{p}')\} = 0,
\end{aligned} \tag{4.2.88}$$

along with

$$\begin{aligned}
\{\hat{b}_s(\mathbf{p}),\hat{b}_{s'}^\dagger(\mathbf{p}')\} &= (2\omega)(2\pi)^3\delta_{ss'}\delta^3(\mathbf{p}-\mathbf{p}'),\\
\{\hat{d}_s^\dagger(\mathbf{p}),\hat{d}_{s'}(\mathbf{p}')\} &= (2\omega)(2\pi)^3\delta_{ss'}\delta^3(\mathbf{p}-\mathbf{p}'),\\
\{\hat{b}_s(\mathbf{p}),\hat{d}_{s'}^\dagger(\mathbf{p}')\} &= 0.
\end{aligned} \tag{4.2.89}$$

Then the Dirac Hamiltonian becomes [5]

$$\hat{H} = \sum_{s=\pm}\int d^3p_L\,\omega\Big[\hat{b}_s^\dagger(\mathbf{p})\hat{b}_s(\mathbf{p}) - \hat{d}_s(\mathbf{p})\hat{d}_s^\dagger(\mathbf{p})\Big]. \tag{4.2.90}$$

Even though we are using anticommutation relations, we find that the Fourier coefficient operators still obey the following commutation relations with the Dirac Hamiltonian,

$$[\hat{H},\hat{b}_s^\dagger(\mathbf{p})] = \omega\hat{b}_s^\dagger(\mathbf{p}), \qquad [\hat{H},\hat{b}_s(\mathbf{p})] = -\omega\hat{b}_s(\mathbf{p}), \tag{4.2.91}$$

$$[\hat{H},\hat{d}_s^\dagger(\mathbf{p})] = \omega\hat{d}_s^\dagger(\mathbf{p}), \qquad [\hat{H},\hat{d}_s(\mathbf{p})] = -\omega\hat{d}_s(\mathbf{p}). \tag{4.2.92}$$

Proof. Since the d-type operators follow the same anti-commutation and commu-

tation relations as the b-type operators, we prove, without loss of generality, the commutation relations (4.2.91). Observe that

$$\begin{aligned}
[\hat{H}, \hat{b}_s(\mathbf{p})] &= \sum_{s'=\pm} \int d^3p'_L\, \omega' \Big([\hat{b}^\dagger_{s'}(\mathbf{p}')\hat{b}_{s'}(\mathbf{p}'), \hat{b}_s(\mathbf{p})] - [\hat{d}_{s'}(\mathbf{p}')\hat{d}^\dagger_{s'}(\mathbf{p}'), \hat{b}_s(\mathbf{p})] \Big) \\
&= \sum_{s'=\pm} \int d^3p'_L\, \omega' \Big(\big(\hat{b}^\dagger_{s'}(\mathbf{p}')\{\hat{b}_s(\mathbf{p}), \hat{b}_{s'}(\mathbf{p}')\} - \{\hat{b}_s(\mathbf{p}), \hat{b}^\dagger_{s'}(\mathbf{p}')\}\hat{b}_{s'}(\mathbf{p}')\big) \\
&- \big(\hat{d}_{s'}(\mathbf{p}')\{\hat{b}_s(\mathbf{p}), \hat{d}'_{s'}(\mathbf{p}')\} - \{\hat{b}_s(\mathbf{p}), \hat{d}_{s'}(\mathbf{p}')\}\hat{d}^\dagger_{s'}(\mathbf{p}')\big) \Big) \\
&= \sum_{s'=\pm} \int d^3p'_L\, \omega' \Big(- (2\omega)(2\pi)^3 \delta_{ss'} \delta^3(\mathbf{p} - \mathbf{p}') \hat{b}_{s'}(\mathbf{p}') \Big) \\
&= -\omega \hat{b}_s(\mathbf{p}),
\end{aligned}$$

where in the second line we have used the commutator identity

$$[AB, C] = A\{C, B\} - \{C, A\}B. \tag{4.2.93}$$

Complex conjugating the above yields the final relation, $[\hat{H}, \hat{b}^\dagger_s(\mathbf{p})] = \omega \hat{b}^\dagger_s(\mathbf{p})$. $\square$

The commutation relations (4.2.91) and (4.2.92) once again allow us to associate $\hat{b}^\dagger_s(\mathbf{p})$ and $\hat{d}^\dagger_s(\mathbf{p})$ as creation operators and $\hat{b}_s(\mathbf{p})$ and $\hat{d}_s(\mathbf{p})$ as annihilation operators. We can use our anticommutation relations to rewrite the Hamiltonian as

$$\begin{aligned}
\hat{H}_D &= \sum_{s=\pm} \int d^3p_L\, \omega \Big[\hat{b}^\dagger_s(\mathbf{p})\hat{b}_s(\mathbf{p}) + \hat{d}^\dagger_s(\mathbf{p})\hat{d}_s(\mathbf{p}) - (2\omega)(2\pi)^3\delta^3(0) \Big] \\
&= \sum_{s=\pm} \int d^3p_L\, \omega \Big[\hat{b}^\dagger_s(\mathbf{p})\hat{b}_s(\mathbf{p}) + \hat{d}^\dagger_s(\mathbf{p})\hat{d}_s(\mathbf{p}) \Big] - 4E_0.
\end{aligned} \tag{4.2.94}$$

The infinite zero-point energy value

$$E_0 = \frac{1}{2} \int d^3p\, \omega\, \delta^3(0),$$

appears once again in the Dirac Hamiltonian. Unlike the Klein-Gordon field, the energy is negative and is larger than the Klein-Gordon zero-point energy by four times. This is because we have four particles that arise from the quantized Dirac

field [5]. We can omit this infinity from our calculations by including the term $\Omega = -4E_0$ in the Dirac Lagrangian to cancel out this value in the same manner as the Klein-Gordon field. In the real world, we measure energy differences thus the infinite energy value does not concern us [8]. Now the quantized Majorana Hamiltonian can be written as

$$\begin{aligned}\hat{H}_M &= \frac{1}{2}\sum_{s=\pm}\int d^3p_L\,\omega\Big[\hat{b}_s^\dagger(\mathbf{p})\hat{b}_s(\mathbf{p}) - \hat{b}_s(\mathbf{p})\hat{b}_s^\dagger(\mathbf{p})\Big] \\ &= \frac{1}{2}\sum_{s=\pm}\int d^3p_L\,\omega\left[\hat{b}_s^\dagger(\mathbf{p})\hat{b}_s(\mathbf{p})\right] - 2E_0,\end{aligned}$$

hence we have the quantized Dirac and Majorana Hamiltonians [5]

$$\hat{H}_D = \sum_{s=\pm}\int d^3p_L\,\omega\Big[\hat{b}_s^\dagger(\mathbf{p})\hat{b}_s(\mathbf{p}) + \hat{d}_s^\dagger(\mathbf{p})\hat{d}_s(\mathbf{p})\Big], \tag{4.2.95}$$

$$\hat{H}_M = \frac{1}{2}\sum_{s=\pm}\int d^3p_L\,\omega\left[\hat{b}_s^\dagger(\mathbf{p})\hat{b}_s(\mathbf{p})\right], \tag{4.2.96}$$

where we have removed the infinite zero point energies. To definitively prove that the spinor fields and its excitations have spin-1/2, we promote the spinor field-theoretic total angular momentum from the symmetries chapter into a self-adjoint operator which yields

$$\hat{\mathbf{J}} = \hat{\mathbf{L}} + \hat{\mathbf{S}}. \tag{4.2.97}$$

Since we only desire to find the spin value, we move to the rest frame thus we can neglect the orbital angular momentum operator component of the total angular momentum operator. Then we have

$$\begin{aligned}\hat{\mathbf{J}} &= \hat{\mathbf{S}} \\ &= -\int d^3x\, i\hat{\Pi}\, S^{\mu\nu}\hat{\mathbf{\Psi}} \\ &= \frac{1}{2}\int d^3x\, i\hat{\mathbf{\Psi}}^\dagger\mathbf{\Sigma}\hat{\mathbf{\Psi}},\end{aligned} \tag{4.2.98}$$

where

$$\boldsymbol{\Sigma} = \begin{pmatrix} \boldsymbol{\sigma} & 0 \\ 0 & \boldsymbol{\sigma} \end{pmatrix}. \tag{4.2.99}$$

Operating with the z-direction angular momentum $\hat{J}_z$ on the states created by $\hat{b}^\dagger_{s=1}(\mathbf{0})$ and $\hat{b}^\dagger_{s=2}(\mathbf{0})$ yields $j_z = +1/2$ and $j_z = -1/2$, respectively [15]. Operating on the states created by $\hat{d}^\dagger_{s=1}(\mathbf{0})$ and $\hat{d}^\dagger_{s=2}(\mathbf{0})$ yields $j_z = -1/2$ and $j_z = +1/2$, respectively. A boost in the z-direction leaves the spin values invariant so both the b- and d-type particles possess spin-1/2 as expected [15]. Like in nonrelativistic quantum mechanics, $\hat{\mathbf{J}}$ commutes with both the Dirac and Majorana Hamiltonians thus we can use the spin degree of freedom s to label our states and operators [7].

Using anticommutators has proven to be our salvation in that our Hamiltonian (4.2.95) is bounded from below and unitarity is preserved. Thus, we conclude that anticommutators must be used for quantizing spin-1/2 fields. In fact, all fields (and their excitations) with half-integer spin, namely, a spin value that is a multiple of 1/2 are to be quantized using anticommutators. This is a consequence of the **spin-statistics theorem** which states that integer spin particles obey Bose-Einstein statistics and half-integer spin particles obey Fermi-Dirac statistics. Like the quantum Klein-Gordon field, the quantum Dirac and Majorana fields are, in an abstract sense, an ensemble of harmonic oscillators whose modes are identified as excitations, or particles, of the fields. Namely, the excitations of the Dirac, Majorana, and Weyl fields are spin-1/2 particles known as *Dirac fermions*, *Majorana fermions*, and *Weyl fermions*, respectively. Because of the antisymmetry of the operators, we have for a two particle state

$$\hat{b}^\dagger_{s_i}\hat{b}^\dagger_{s_j}|0\rangle = -\hat{b}^\dagger_{s_j}\hat{b}^\dagger_{s_i}|0\rangle \tag{4.2.100}$$

$$|p_i, s_i, a_i; p_j, s_j, a_j\rangle = -|p_j, s_j, a_j; p_i, s_i, a_i\rangle, \tag{4.2.101}$$

where $a_{i,j}$ are any other conserved quantum numbers that may be used to label

the state. If $i = j$ then we have

$$\hat{b}^{\dagger}_{s_i}\hat{b}^{\dagger}_{s_i}|0\rangle = |p_i, s_i, a_i; p_i, s_i, a_i\rangle = 0, \tag{4.2.102}$$

which is the famous *Pauli Exclusion Principle* which states that no two or more fermions can have the same quantum numbers in a quantum system.

The quantum Dirac field is very interesting in the fact that we have two *different* creation (annihilation) operators, namely, $\hat{b}^{\dagger}_s(\mathbf{p})$ ($\hat{b}_s(\mathbf{p})$) and $\hat{d}^{\dagger}_s(\mathbf{p})$ ($\hat{d}_s(\mathbf{p})$). This means that $\hat{b}^{\dagger}_s(\mathbf{p})$ creates b-type particles and $\hat{d}^{\dagger}_s(\mathbf{p})$ creates d-type particles while their respective annihilation operators annihilates them [5]. This physical and mathematical outcome results from the U(1) symmetry of the Dirac Lagrangian. By Noether's theorem, the Noether charge corresponding to the U(1) symmetry is,

$$\begin{aligned}\hat{Q} &= \int d^3x\, j^0 \\ &= \int d^3x\, \hat{\overline{\mathbf{\Psi}}}\gamma^0\hat{\mathbf{\Psi}} \\ &= \sum_{s=\pm}\int d^3p_L \left[\hat{b}^{\dagger}_s(\mathbf{p})\hat{b}_s(\mathbf{p}) - \hat{d}^{\dagger}_s(\mathbf{p})\hat{d}_s(\mathbf{p})\right],\end{aligned} \tag{4.2.103}$$

where we identify $\hat{b}^{\dagger}_s(\mathbf{p})\hat{b}_s(\mathbf{p})$ and $\hat{d}^{\dagger}_s(\mathbf{p})\hat{d}_s(\mathbf{p})$ as the Hermitian number operators for the b-type particles and d-type particles, respectively, and the Noether charge has now become a Hermitian operator. By using the commutation relations (4.2.92), it is trivial to see that $[\hat{H}, \hat{Q}] = 0$ thus the charge operator is a conserved value of the system and we can use the eigenvalues q of $\hat{Q}$ to label the states of the Dirac fermions. The Noether charge operator $\hat{Q}$ then states that the total number of b-type particles minus the total number of d-type particles is conserved. To find the eigenvalues of $\hat{Q}$, let us first construct the Fock space by defining $|0\rangle$ to be the vacuum state with the usual normalization $\langle 0|0\rangle = 1$ where

$$\hat{b}_s(\mathbf{p})|0\rangle = \hat{d}_s(\mathbf{p})|0\rangle = 0. \tag{4.2.104}$$

Then we define the one particle states for both types of creation operators acting on the vacuum as

$$\hat{b}_s^\dagger(\mathbf{p})|0\rangle = (2\omega)^{1/2}\hat{b}_s^\dagger(\mathbf{p})|0\rangle = |p,s,q\rangle, \quad \hat{d}_s^\dagger(\mathbf{p})|0\rangle = (2\omega)^{1/2}\hat{b}_s^\dagger(\mathbf{p})|0\rangle = |p,s,q\rangle, \tag{4.2.105}$$

with the normalization

$$\langle p,s,q|p',s',q'\rangle = (2\pi)^3(2\omega)\delta^3(\mathbf{p}-\mathbf{p}')\delta_{ss'}\delta_{qq'}, \tag{4.2.106}$$

where $\omega = (\mathbf{p}^2+m^2)^{1/2}$, p is the four-momentum, $s = 1,2$ is the spin degree of freedom, and q is the eigenvalue of the U(1) charge. Then the multiparticle state is

$$|p_1,s_1,q_1;\ldots;p_n,s_n,q_n\rangle = \prod_{k=1}^{n}(2\omega_{p_k})^{1/2}\,\hat{b}_{s_1}^\dagger(\mathbf{p}_1)\ldots\hat{b}_{s_n}^\dagger(\mathbf{p}_n)|0\rangle. \tag{4.2.107}$$

If we use the anticommutation rules of the creation and annihilation operators for both types of particles along with the commutator identity

$$[A,BC] = [A,B]_\pm C \mp B[A,C]_\pm,$$

we get the following commutation relations

$$[\hat{Q},\hat{b}_s^\dagger(\mathbf{p})] = +\hat{b}_s^\dagger(\mathbf{p}), \qquad [\hat{Q},\hat{b}_s(\mathbf{p})] = -\hat{b}_s(\mathbf{p}), \tag{4.2.108}$$

$$[\hat{Q},\hat{d}_s^\dagger(\mathbf{p})] = -\hat{d}_s^\dagger(\mathbf{p}), \qquad [\hat{Q},\hat{d}_s(\mathbf{p})] = +\hat{d}_s(\mathbf{p}). \tag{4.2.109}$$

Then we see that

$$\hat{Q}\hat{b}_s^\dagger(\mathbf{p})|0\rangle = +\hat{b}_s^\dagger(\mathbf{p})|0\rangle = |p,s,+1\rangle, \tag{4.2.110}$$

$$\hat{Q}\hat{d}_s^\dagger(\mathbf{p})|0\rangle = -\hat{d}_s^\dagger(\mathbf{p})|0\rangle = |p,s,-1\rangle, \tag{4.2.111}$$

thus the b- and d-type particles create particles with $+1$ and -1, respectively. So we see that the b- and d-type particles are distinguished from each other by their

charge. Since charge conjugation is also a symmetry of the Dirac Lagrangian and implies that both particles have the same mass, we call b-type particles *fermions* and d-type particles *antifermions* [5]. Thus we stumble upon the existence of antimatter in quantum field theory. It is rather remarkable that our derivation of the quantum Dirac field naturally led to the concept of antimatter when compared to the case of nonrelativistic quantum mechanics where Paul Dirac had to postulate the existence of antimatter in order to explain the negative energy states of the wave function that satisfied the Dirac equation [1]. The fact that the concept of antimatter is effectively built into quantum field theory serves to showcase one of the many triumphs of quantum field theory over nonrelativistic quantum mechanics and implicitly shows the fundamental importance of the field over the particle.

The excitations of the quantum Majorana field does not admit a separate antifermion as can be seen by the appearance of only one type of creation and annihilation operator in its quantum Hamiltonian, due to the Majorana condition $\hat{\boldsymbol{\chi}} = \hat{\boldsymbol{\chi}}^{\mathrm{C}}$. This is because the Majorana Lagrangian does not possess a U(1) symmetry so there is only one type of particle, namely, the Majorana fermion. Since the Majorana condition states that the Majorana field is its own charge conjugate, the Majorana fermion is its own antifermion thus the Majorana fermion possesses neutral charge [5]. Recall that the real scalar bosons did not possess charge conjugation as well so the spin-0 real scalar bosons are their own charge conjugates as well thus they have neutral charge. Current experimental data shows that most, perhaps all, of the fermions of the Standard Model are Dirac fermions [1]. The one possible exception to this could be the *neutrino*. In recent decades, theoretical and experimental studies on the neutrino have revealed a plethora of deep insights into the limitations of the Standard Model and the possibility of discovering new physics beyond the Standard Model. One of these insights is that the neutrino actually possesses a very small mass which was not predicted theoretically [11]. This dispelled the possibility that the neutrino was potentially a Weyl fermion and further studies are being conducted on whether the neutrino is

definitively a Majorana fermion or a Dirac fermion.

4.3 Electromagnetic Field

Classical Electromagnetic Theory

In this section, we will briefly review classical electrodynamics and present the procedure for quantizing the classical electromagnetic field. It is assumed that the reader may have little to no experience with classical electromagnetic theory so we will briefly reintroduce SI units (inserting $\hbar$ and c whenever it is needed) to aid in comprehension. We will resort back to using natural units once we have covered the necessary materials from classical electrodynamics.

From classical electromagnetic theory, Maxwell's equations are contained in an elegant, compact format known as the *field tensor* (or *Faraday tensor*)

$$F^{\mu\upsilon} = \begin{pmatrix} 0 & E_x/c & E_y/c & E_z/c \\ -E_x/c & 0 & B_z & -B_y \\ -E_y/c & -B_z & 0 & B_x \\ -E_z/c & B_y & -B_x & 0 \end{pmatrix}, \tag{4.3.1}$$

where $\mathbf{E}$ and $\mathbf{B}$ are the electric and magnetic fields, respectively [2]. Now in potential theory, the electric and magnetic fields can be written as

$$\mathbf{E} = -\nabla\phi - \dot{\mathbf{A}}, \tag{4.3.2}$$

$$\mathbf{B} = \nabla \times \mathbf{A}, \tag{4.3.3}$$

where ϕ and $\mathbf{A}$ are the scalar and vector potentials, respectively [2]. Then we define the *four-vector potential*, or *gauge field*, as

$$A^{\mu} = (\phi/c, \mathbf{A}), \tag{4.3.4}$$

which then allows the field tensor to be written as

$$F^{\mu\upsilon} = \partial^{\mu} A^{\upsilon} - \partial^{\upsilon} A^{\mu}, \tag{4.3.5}$$

which in this form is called the *field strength*. Next, we construct the *charge density four-vector*

$$J^{\mu} = (c\rho, \mathbf{J}), \tag{4.3.6}$$

where ρ and $\mathbf{J}$ are the charge density and the current density, respectively [2]. With these definitions, we start to see how Maxwell's equations are encoded in the field tensor $F^{\mu\upsilon}$. First, we take the four-divergence of J^{μ} to get

$$\partial_{\mu} J^{\mu} = \dot{\rho} + \nabla \cdot \mathbf{J} = 0, \tag{4.3.7}$$

which is the continuity equation for the electromagnetic current which basically states that the electromagnetic current must be conserved [2]. Now from the components of $F^{\mu\upsilon}$ in equation (4.3.1), we get

$$F^{0i} = E_i, \tag{4.3.8}$$

$$F^{ij} = \varepsilon^{ijk} B_k, \tag{4.3.9}$$

and furthermore,

$$\partial_{\upsilon} F^{\mu\upsilon} = \mu_0 J^{\upsilon}. \tag{4.3.10}$$

Then,

$$\partial_{\upsilon} F^{0\upsilon} = \nabla \cdot \mathbf{E} = \mu_0 J^0 = \mu_0 c\rho = \frac{1}{\epsilon_0}\rho,$$

which is Gauss's law [2]

$$\nabla \cdot \mathbf{E} = \frac{1}{\epsilon_0}\rho, \tag{4.3.11}$$

and analogously by combining $\mu = 1, 2, 3$, we get

$$\partial_\upsilon F^{\mu\upsilon} = -\mu_0\epsilon_0\dot{\mathbf{E}} + \nabla \times \mathbf{B} = \mu_0\mathbf{J},$$

which is Maxwell's correction of Ampère's law

$$\nabla \times \mathbf{B} = \mu_0\mathbf{J} + \mu_0\epsilon_0\dot{\mathbf{E}}. \tag{4.3.12}$$

The other two Maxwell's equations can be written as

$$\varepsilon_{\mu\nu\rho\sigma}\partial^\rho F^{\mu\nu} = 0, \tag{4.3.13}$$

which gives

$$\nabla \times \mathbf{E} + \dot{\mathbf{B}} = 0, \tag{4.3.14}$$

$$\nabla \cdot \mathbf{B} = 0, \tag{4.3.15}$$

which are Faraday's law and Gauss' law for magnetism, respectively [5].

The field tensor $F^{\mu\nu}$ is also invariant under gauge transformations.

Proof. Consider the gauge transformation

$$A^{\mu\prime} = A^\mu - \partial^\mu\Gamma,$$

where $\Gamma(x)$ is an arbitrary function of spacetime. Then the field tensor transforms as

$$\begin{aligned} F^{\mu\nu} &\to F^{\mu\nu\prime} \\ &= \partial^\mu(A^\nu - \partial^\nu\Gamma) - \partial^\nu(A^\mu - \partial^\mu\Gamma) \\ &= (\partial^\mu A^\nu - \partial^\nu A^\mu) - (\partial^\mu\partial^\nu - \partial^\nu\partial^\mu)\Gamma) \\ &= F^{\mu\nu}. \end{aligned}$$

Since the choice of $\Gamma(x)$ was an arbitrary, the field tensor is gauge invariant. $\square$

The Lagrangian density is then written as

$$\mathcal{L} = -\frac{1}{4}F^{\mu\upsilon}F_{\mu\upsilon} + J^{\mu}A_{\mu} \tag{4.3.16}$$

where the current J^{μ} is treated as the external source [2]. Since the field tensor is gauge invariant, it is trivial to see that $F^{\mu\upsilon}F_{\mu\upsilon}$ is gauge invariant as well. It is also trivial to see that the field tensor invariant product $F^{\mu\upsilon}F_{\mu\upsilon}$ is Lorentz invariant.

The Lagrangian density (4.3.16) exhibits the desired gauge, Lorentz, parity, and time reversal invariance required for a classical field to be compatible with relativity and quantum mechanics once we quantize it [5]. In order to see that our Lagrangian density produces Maxwell's equations as its equations of motion, we vary the action $\mathcal{S}$ with proper boundary conditions which yields the equations of motion

$$(g^{\mu\upsilon}\partial^2 - \partial^{\mu}\partial^{\upsilon})A_{\upsilon} + J^{\mu} = 0, \tag{4.3.17}$$

which are Maxwell's equations in tensor format [5]. Thus we have our desired Lagrangian density.

Classical Electromagnetic Field

We now follow the same procedure of the previous sections of deriving the classical Hamiltonian and other necessary components of the electromagnetic field. Before we start our derivation, we must choose a gauge to work with so that we may eliminate the gauge freedom. We will choose to work in the *Coulomb gauge*

$$\nabla \cdot \mathbf{A}(x) = 0. \tag{4.3.18}$$

We will impose the stated conditions one by one for a smooth transition in calculations.

Consider the source free Lagrangian with finite boundary conditions at spatial

and temporal infinity [7]

$$\begin{aligned}\mathcal{L} &= -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} \\ &= \frac{1}{2}\left[\left(\frac{\partial \mathbf{A}}{\partial t}\right)^2 - |\nabla \times \mathbf{A}|^2\right] \\ &= \frac{1}{2}\left[\dot{\mathbf{A}}^2 - |\nabla \times \mathbf{A}|^2\right]. \end{aligned} \tag{4.3.19}$$

Then the conjugate momentum is

$$\mathbf{\Pi}(\mathbf{x}, t) = \frac{\partial \mathcal{L}}{\partial \dot{\mathbf{A}}} = \dot{\mathbf{A}}. \tag{4.3.20}$$

Recall that Maxwell's equations are stated as

$$\nabla \cdot \mathbf{E} = \rho, \tag{4.3.21}$$

$$\nabla \cdot \mathbf{B} = 0, \tag{4.3.22}$$

$$\nabla \times \mathbf{E} = -\dot{\mathbf{B}}, \tag{4.3.23}$$

$$\nabla \times \mathbf{B} = \mathbf{J} + \dot{\mathbf{E}}. \tag{4.3.24}$$

From equation (4.3.22), we see that by Helmholtz's theorem, there exists a vector potential $\mathbf{A}(\mathbf{x}, t)$ such that $\mathbf{B} = \nabla \times \mathbf{A}$ and inserting this vector potential into equation (4.3.23) gives us [2]

$$\begin{aligned}\nabla \times \mathbf{E} + \frac{\partial}{\partial t}\left[\nabla \times \mathbf{A}\right] &= 0 \\ \nabla \times \left[\mathbf{E} + \dot{\mathbf{A}}\right] &= 0,\end{aligned}$$

which implies that there exists scalar potential $\phi(\mathbf{x}, t)$ such that

$$\begin{aligned}\mathbf{E} + \dot{\mathbf{A}} &= -\nabla\phi \\ \mathbf{E} &= -\dot{\mathbf{A}} - \nabla\phi. \end{aligned} \tag{4.3.25}$$

Then equation (4.3.24) becomes

$$\begin{aligned}
\nabla \times \nabla \times \mathbf{A} - \frac{\partial}{\partial t}\Big[- \dot{\mathbf{A}} - \nabla\phi\Big] &= \mathbf{J} \\
\nabla(\nabla \cdot \mathbf{A}) - \nabla^2\mathbf{A} + \frac{\partial}{\partial t}\Big[\dot{\mathbf{A}} + \nabla\phi\Big] &= \mathbf{J} \\
\Big(\ddot{\mathbf{A}} - \nabla^2\mathbf{A}\Big) + \nabla\Big(\nabla \cdot \mathbf{A} + \dot{\phi}\Big) &= \mathbf{J}.
\end{aligned}$$

Finally, inserting equation (4.3.25) into (4.3.21) yields

$$\begin{aligned}
\nabla \cdot \Big(- \dot{\mathbf{A}} - \nabla\phi\Big) &= \rho \\
-\nabla^2\phi - \frac{\partial}{\partial t}\Big(\nabla \cdot \mathbf{A}\Big) &= \rho, \qquad (4.3.26)
\end{aligned}$$

which gives the potential formulation of classical electrodynamics

$$\Big(\ddot{\mathbf{A}} - \nabla^2\mathbf{A}\Big) + \nabla\Big(\nabla \cdot \mathbf{A} + \dot{\phi}\Big) = \mathbf{J}, \qquad (4.3.27)$$

$$-\nabla^2\phi - \frac{\partial}{\partial t}\Big(\nabla \cdot \mathbf{A}\Big) = \rho. \qquad (4.3.28)$$

Now we impose the Coulomb gauge $\nabla \cdot \mathbf{A} = 0$ and equation (4.3.28) becomes

$$\begin{aligned}
-\nabla^2\phi &= \rho \\
\phi(\mathbf{x}, t) &= \int d^3x' \frac{\rho(\mathbf{x}', t)}{4\pi\,|\mathbf{x} - \mathbf{x}'|}. \qquad (4.3.29)
\end{aligned}$$

We now impose the free field condition $\mathbf{J} = \rho = 0$ which implies that $\phi = 0$ and using equations (4.3.21) and (4.3.25), we get that [7]

$$\nabla \cdot \dot{\mathbf{A}} = \nabla \cdot \mathbf{\Pi} = 0. \qquad (4.3.30)$$

Then we are left with

$$\ddot{\mathbf{A}} - \nabla^2 \mathbf{A} = 0, \tag{4.3.31}$$

$$\nabla \cdot \mathbf{A} = 0, \tag{4.3.32}$$

$$\nabla \cdot \dot{\mathbf{A}} = 0. \tag{4.3.33}$$

In analogy with the Klein-Gordon equation, equation (4.3.31) admits plane wave solutions [7]

$$\mathbf{A}(\mathbf{x}, t) = \int d^3 k_L \Big[\mathbf{a}(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{x} - i\omega t} + \mathbf{b}(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{x} + i\omega t} \Big].$$

Complex conjugating the equation yields

$$\begin{aligned} \mathbf{A}^*(\mathbf{x}, t) &= \int d^3 k_L \Big[\mathbf{a}^*(\mathbf{k}) e^{-i\mathbf{k}\cdot\mathbf{x} + i\omega t} + \mathbf{b}^*(\mathbf{k}) e^{-i\mathbf{k}\cdot\mathbf{x} - i\omega t} \Big] \\ &= \int d^3 k_L \Big[\mathbf{a}^*(\mathbf{k}) e^{-i\mathbf{k}\cdot\mathbf{x} + i\omega t} + \mathbf{b}^*(-\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{x} - i\omega t} \Big]. \end{aligned}$$

Since the vector potential $\mathbf{A}$ is real, we impose the condition $\mathbf{A} = \mathbf{A}^*$ then comparing both sides of the equations yields the relation $\mathbf{a}(\mathbf{k}) = \mathbf{b}^*(-\mathbf{k})$. Now we are left with

$$\begin{aligned} \mathbf{A}(\mathbf{x}, t) &= \int d^3 k_L \Big[\mathbf{a}(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{x} - i\omega t} + \mathbf{a}^*(-\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{x} + i\omega t} \Big] \\ &= \int d^3 k_L \Big[\mathbf{a}(\mathbf{k}) e^{i\mathbf{k}\cdot\mathbf{x} - i\omega t} + \mathbf{a}^*(\mathbf{k}) e^{-i\mathbf{k}\cdot\mathbf{x} + i\omega t} \Big] \\ &= \int d^3 k_L \Big[\mathbf{a}(\mathbf{k}) e^{ikx} + \mathbf{a}^*(\mathbf{k}) e^{-ikx} \Big] \end{aligned}$$

where $k^0 = \omega = |\mathbf{k}|$ [5]. Taking the differential with respect to time on $\mathbf{A}$ gives us the conjugate momentum hence we arrive at

$$\mathbf{A}(x) = \int d^3 k_L \, \Big[\mathbf{a}(\mathbf{k}) e^{ikx} + \mathbf{a}^*(\mathbf{k}) e^{-ikx} \Big], \tag{4.3.34}$$

$$\mathbf{\Pi}(x) = \int d^3 k_L \, (-i\omega) \Big[\mathbf{a}(\mathbf{k}) e^{ikx} - \mathbf{a}^*(\mathbf{k}) e^{-ikx} \Big], \tag{4.3.35}$$

where it is implicit that both functions are functions of spacetime, namely,

$\mathbf{A}(x) = \mathbf{A}(\mathbf{x}, t)$ and $\mathbf{\Pi}(x) = \mathbf{\Pi}(\mathbf{x}, t)$ [5]. The Coulomb gauge $\nabla \cdot \mathbf{A} = 0$ implies that

$$\nabla \cdot \mathbf{\Pi} = 0, \tag{4.3.36}$$

so imposing the Coulomb gauge and condition (4.3.36) on the equations (4.3.34) and (4.3.35), respectively, yields the following relations

$$\mathbf{k} \cdot \left[\mathbf{a}(\mathbf{k}) + \mathbf{a}^*(-\mathbf{k})\right] = 0,$$
$$\mathbf{k} \cdot \left[\mathbf{a}(\mathbf{k}) - \mathbf{a}^*(-\mathbf{k})\right] = 0,$$

which implies that [7]

$$\mathbf{k} \cdot \mathbf{a}(\mathbf{k}) = 0. \tag{4.3.37}$$

Condition (4.3.37) says that $\mathbf{A}(x)$ is perpendicular to the propagation direction $\mathbf{k}$ hence it is a transverse wave (this is why the Coulomb gauge is often referred to as the "transverse gauge") [7]. Since $\mathbf{a}(\mathbf{k})$ and $\mathbf{a}^*(\mathbf{k})$ are vectors with three components, we can expand them at each point $\mathbf{k}$ using some set of basis vectors. Due to the transverse condition of our wave in question, we strategically choose our basis vectors to be the circular polarization vectors

$$\varepsilon_+(\mathbf{k}) = \frac{1}{\sqrt{2}}(1, -i, 0), \tag{4.3.38}$$
$$\varepsilon_-(\mathbf{k}) = \frac{1}{\sqrt{2}}(1, +i, 0), \tag{4.3.39}$$

with which we get the following relations

$$\varepsilon_\lambda(\mathbf{k}) \cdot \varepsilon^*_{\lambda'}(\mathbf{k}) = \delta_{\lambda\lambda'}, \tag{4.3.40}$$
$$\mathbf{k} \cdot \varepsilon_\lambda(\mathbf{k}) = 0, \tag{4.3.41}$$
$$\sum_{\lambda=\pm} \varepsilon^*_{i\lambda}(\mathbf{k})\varepsilon_{j\lambda}(\mathbf{k}) = \delta_{ij} - \frac{k_i k_j}{\mathbf{k}^2}, \tag{4.3.42}$$

where it is evident that our chosen basis is perpendicular to the propagation vector $\mathbf{k}$ [15]. Then we expand equations (4.3.34) and (4.3.35) in terms of the bases to get

$$\mathbf{A}(x) = \sum_{\lambda=\pm} \int d^3k_L \left[\varepsilon^*_\lambda(\mathbf{k}) a_\lambda(\mathbf{k}) e^{ikx} + \varepsilon_\lambda(\mathbf{k}) a^*_\lambda(\mathbf{k}) e^{-ikx}\right], \tag{4.3.43}$$

$$\mathbf{\Pi}(x) = \sum_{\lambda=\pm} \int d^3k_L \, (-i\omega)\left[\varepsilon^*_\lambda(\mathbf{k}) a_\lambda(\mathbf{k}) e^{ikx} - \varepsilon_\lambda(\mathbf{k}) a^*_\lambda(\mathbf{k}) e^{-ikx}\right]. \tag{4.3.44}$$

Now other than the fact that we have vector coefficients, it is clear that equations (4.3.34) and (4.3.35) have the same form as the Klein-Gordon field $\varphi(x)$ and its conjugate momentum $\Pi(x)$ so we apply the same operations used on the Klein-Gordon field to get the coefficients $\mathbf{a}(\mathbf{k})$ and $\mathbf{a}^*(\mathbf{k})$ in terms of $\mathbf{A}$,

$$\begin{aligned}
\mathbf{a}(\mathbf{k}) &= i \int d^3x \, e^{-ikx} \overleftrightarrow{\partial_0} \mathbf{A}(x) \\
\sum_{\lambda=\pm} \varepsilon^*_\lambda a_\lambda(\mathbf{k}) &= i \int d^3x \, e^{-ikx} \overleftrightarrow{\partial_0} \mathbf{A}(x) \\
\left(\sum_{\lambda=\pm} \varepsilon^*_\lambda a_\lambda(\mathbf{k}) = i \int d^3x \, e^{-ikx} \overleftrightarrow{\partial_0} \mathbf{A}(x)\right) &\varepsilon_{\lambda'}(\mathbf{k}) \\
a_\lambda(\mathbf{k}) &= i\varepsilon_{\lambda'}(\mathbf{k}) \int d^3x \, e^{-ikx} \overleftrightarrow{\partial_0} \mathbf{A}(x).
\end{aligned}$$

Complex conjugating $a_\lambda(\mathbf{k})$ then gives us the other coefficient and we are left with [5]

$$a_\lambda(\mathbf{k}) = i\varepsilon_{\lambda'}(\mathbf{k}) \int d^3x \, e^{-ikx} \overleftrightarrow{\partial_0} \mathbf{A}(x), \tag{4.3.45}$$

$$a^*_\lambda(\mathbf{k}) = -i\varepsilon^*_{\lambda'}(\mathbf{k}) \int d^3x \, e^{ikx} \overleftrightarrow{\partial_0} \mathbf{A}(x). \tag{4.3.46}$$

Now we can write the Hamiltonian in terms of the Fourier coefficients $a_\lambda(\mathbf{k})$ and

$a^*_\lambda(\mathbf{k})$ to get

$$\begin{aligned} H &= \int d^3x\, \mathcal{H} \\ &= \frac{1}{2}\int d^3x \left[\mathbf{\Pi}^2 - |\nabla \times \mathbf{A}|\right] \\ &= \frac{1}{2}\int d^3x \left[\mathbf{\Pi}^2 - \mathbf{A}\cdot\nabla^2\mathbf{A}\right] \\ &= \frac{1}{2}\sum_{\lambda=\pm}\int d^3k_L\, \omega\Big[a^*_\lambda(\mathbf{k})a_\lambda(\mathbf{k}) + a_\lambda(\mathbf{k})a^*_\lambda(\mathbf{k})\Big], \end{aligned} \tag{4.3.47}$$

where we have carefully arranged our coefficients in a noncommutative manner so as to preserve the structure of the equation in anticipation for the quantization procedure that is to follow [5].

Quantizing the Classical Electromagnetic Field

Instead of promoting $\mathbf{A}(\mathbf{x},t)$ and $\mathbf{\Pi}(\mathbf{x},t)$ to operators $\hat{\mathbf{A}}(\mathbf{x},t)$ and $\hat{\mathbf{\Pi}}(\mathbf{x},t)$, respectively, and imposing the canonical commutation relations, we will instead promote the Fourier coefficients to operators and impose the harmonic oscillator commutation relations first then find the commutation relations of $\hat{\mathbf{A}}(\mathbf{x},t)$ and $\hat{\mathbf{\Pi}}(\mathbf{x},t)$ [15]. Thus we have

$$\begin{aligned} &[\hat{a}_\lambda(\mathbf{k}), \hat{a}_{\lambda'}(\mathbf{k}')] = 0, \\ &[\hat{a}^\dagger_\lambda(\mathbf{k}), \hat{a}^\dagger_{\lambda'}(\mathbf{k}')] = 0, \\ &[\hat{a}_\lambda(\mathbf{k}), \hat{a}^\dagger_{\lambda'}(\mathbf{k}')] = (2\pi)^3(2\omega)\delta_{\lambda\lambda'}\delta^3(\mathbf{k}-\mathbf{k}'). \end{aligned} \tag{4.3.48}$$

Then we use equations (4.3.43) and (4.3.44) along with equation (4.3.42) to get the correct commutation relations for the field operators

$$\begin{aligned}
&[\hat{A}_i(\mathbf{x},t),\hat{A}_j(\mathbf{x}',t)] = 0,\\
&[\hat{\Pi}_i(\mathbf{x},t),\hat{\Pi}_j(\mathbf{x}',t)] = 0,\\
&[\hat{A}_i(\mathbf{x},t),\hat{\Pi}_j(\mathbf{x}',t)] = i\Big(\delta_{ij} - \frac{\nabla_i\nabla_j}{\nabla^2}\delta^3(\mathbf{x}-\mathbf{x}')\Big)\\
&\qquad\qquad = i\int \frac{d^3k}{(2\pi)^3}\, e^{i\mathbf{k}\cdot(\mathbf{x}-\mathbf{x}')}\Big(\delta_{ij} - \frac{k_ik_j}{\mathbf{k}^2}\Big).
\end{aligned} \tag{4.3.49}$$

To show why these commutation relations are correct, we impose the Coulomb gauge on the third line

$$\begin{aligned}
[\partial_i\hat{A}_i(\mathbf{x},t),\hat{\Pi}_j(\mathbf{x}',t)] &= i\int \frac{d^3k}{(2\pi)^3}\, e^{i\mathbf{k}\cdot(\mathbf{x}-\mathbf{x}')}ik_i\Big(\delta_{ij} - \frac{k_ik_j}{\mathbf{k}^2}\Big)\\
&= 0,
\end{aligned} \tag{4.3.50}$$

which satisfies the Coulomb gauge. Had we first imposed the standard commutation relations on the field operators in the usual manner

$$[\hat{A}_i(\mathbf{x},t),\hat{\Pi}_j(\mathbf{x}',t)] = i\delta_{ij}\delta^3(\mathbf{x}-\mathbf{x}'), \tag{4.3.51}$$

then we would have

$$\begin{aligned}
[\nabla\cdot\hat{A}_i(\mathbf{x},t),\nabla\cdot\hat{\Pi}_j(\mathbf{x}',t)] &= i\nabla^2\delta^3(\mathbf{x}-\mathbf{x}')\\
&\neq 0,
\end{aligned} \tag{4.3.52}$$

which is not consistent with our choice of gauge.

Since we imposed the harmonic oscillator algebra on the Fourier coefficients, we appropriately designate $\hat{a}^\dagger_\lambda(\mathbf{k})$ and $\hat{a}_\lambda(\mathbf{k})$ as the creation and annihilation operators, respectively. Indeed, our quantum Hamiltonian mathematically resembles the continuous, field analogue of the quantum harmonic oscillator which can be seen

in the following:

$$\begin{aligned}
\hat{H} &= \frac{1}{2}\sum_{\lambda=\pm}\int d^3k_L\,\omega\Big[\hat{a}^\dagger_\lambda(\mathbf{k})\hat{a}_\lambda(\mathbf{k}) + \hat{a}_\lambda(\mathbf{k})\hat{a}^\dagger_\lambda(\mathbf{k})\Big] \\
&= \frac{1}{2}\sum_{\lambda=\pm}\int d^3k_L\,\omega\Big[\hat{a}^\dagger_\lambda(\mathbf{k})\hat{a}_\lambda(\mathbf{k}) + \hat{a}_\lambda(\mathbf{k})\hat{a}^\dagger_\lambda(\mathbf{k})\Big] \qquad (4.3.53) \\
&= \frac{1}{2}\sum_{\lambda=\pm}\int d^3k_L\,\omega\Big[2\hat{a}^\dagger_\lambda(\mathbf{k})\hat{a}_\lambda(\mathbf{k}) + (2\pi)^3(2\omega)\delta^3(0)\Big] \qquad (4.3.54) \\
&= \sum_{\lambda=\pm}\int d^3k_L\,\omega\Big[\hat{a}^\dagger_\lambda(\mathbf{k})\hat{a}_\lambda(\mathbf{k}) + E_0\Big], \qquad (4.3.55)
\end{aligned}$$

where the infinite term

$$E_0 = \frac{1}{2}\int d^3k\,\omega\delta^3(0), \qquad (4.3.56)$$

is energy of the vacuum, namely, the sum of the zero point energies of all the oscillators. Like the Klein-Gordon field, the vacuum energy cannot be measured and while the electromagnetic vacuum energy does in fact have some physical consequences, such as the *Casimir effect*, we will ignore the vacuum energy for our calculations unless it is explicitly needed [1, 11].

From the form of the Hamiltonian along with the fact that we have used commutators, it is clear that the wave function is a three component vector describing a symmetric state thus it can be easily deduced that the excitation, or particle, of the electromagnetic field is a spin-1 particle thus this particle obeys Bose-Einstein statistics [7]. Due to the transverse condition (4.3.37), these particles can only have spin values of $\pm\hbar$ parallel to its momentum vector [7]. Thus we say that a particle has *helicity* ±1 which is the spin component parallel to its momentum vector [5]. In the formalism of quantum field theory, the creation and annihilation operators $\hat{a}^\dagger_\lambda(\mathbf{k})$ and $\hat{a}_\lambda(\mathbf{k})$ create excitations, or particles, of the quantum electromagnetic field of definite helicity $+1$ (right-circularly polarized) and helicity -1 (left-circularly polarized), respectively. To identify what these particles are, let us first define the ground, or vacuum, state state $|0\rangle$ as the state with no particles with the usual normalization $\langle 0|0\rangle = 1$. Similarly, we define the

state $|k,\lambda\rangle$ as the state with one particle with energy ω and helicity λ with the normalization [5]

$$\langle k,\lambda|k',\lambda'\rangle = (2\pi)^3(2\omega)\delta_{\lambda,\lambda'}\delta^3(\mathbf{k}-\mathbf{k}'). \tag{4.3.57}$$

Now from classical electrodynamics, the momentum of the electromagnetic field is

$$\mathbf{P}(\mathbf{x},t) = \int d^3x\,\mathbf{S} = \int d^3x\,\big(\mathbf{E}+\mathbf{B}\big),$$

where $\mathbf{S}$ is the *Poynting vector* [2]. Using the potential formulation in the Coulomb gauge, the momentum can be written as

$$\begin{aligned}\mathbf{P}(\mathbf{x},t) &= \int d^3x\,\big(\mathbf{E}+\mathbf{B}\big)\\ &= \int d^3x\,(-\dot{\mathbf{A}})\times(\nabla\times\mathbf{A})\\ &= \int d^3x\,(-\mathbf{\Pi})\times(\nabla\times\mathbf{A})\\ &= \int d^3x\,\big(\mathbf{\Pi}\cdot\nabla\big)\mathbf{A}-\mathbf{\Pi}\cdot\nabla\mathbf{A},\end{aligned}$$

which when quantized yields

$$\hat{\mathbf{P}} = \sum_{\lambda=\pm}\int \frac{d^3k}{(2\pi)^3}\,\omega\,\hat{a}^\dagger_\lambda(\mathbf{k})\hat{a}_\lambda(\mathbf{k}), \tag{4.3.58}$$

where $k^0=\omega=|\mathbf{k}|$ [7]. Operator (4.3.58) is the total spatial momentum operator and by applying the operator on some state $|k,\lambda\rangle$, we get [7]

$$\hat{\mathbf{P}}|k,\lambda\rangle = \omega|k,\lambda\rangle = \mathbf{k}|k,\lambda\rangle.$$

By reintroducing $\hbar$ and c into the correct places, we have that the energy of a single particle is $E=\hbar\omega$, which is the *Planck relation*, and the total momentum is $\mathbf{p}=\hbar\mathbf{k}$, which is one of the *De Broglie relations* [7]. From the total momentum, it is clear that an *equivalent* energy formula of the *same* particle in question is $E=\mathbf{p}c=\hbar\mathbf{k}c$. Then by rearranging the relativistic energy formula $E^2=(\mathbf{p}c)^2+(mc^2)^2$, we get

the following result

$$
\begin{aligned}
(mc^2)^2 &= E^2 - (\mathbf{p}c)^2 \\
(mc^2)^2 &= (\hbar\omega)^2 - (\mathbf{k}\hbar c)^2 \\
(mc^2)^2 &= 0 \\
m &= 0,
\end{aligned}
$$

hence the particle created by the quantized electromagnetic field is massless. According to relativity, a massless particle does not have a rest frame which means that it must always travel at the speed of light c (in a vacuum) [4]. By now, it is very obvious from our thorough analysis that the excitation, or particle, of the electromagnetic field is the *photon* which satisfies all of the results we have derived above.

From the mathematical form of the quantum electromagnetic field, the electromagnetic field can be viewed as an infinite set of decoupled harmonic oscillators at every point in space. In fact, the notion of a "vacuum" loses its meaning in the traditional sense when we make the transition to quantum physics. Classically, a vacuum is typically idealized as the state of a system with no matter or energy thus there should be no dynamics or processes occurring. But in quantum physics, the ground state of the electromagnetic field, which has no photons, has energy therefore the "vacuum" is not "empty" [7].

Chapter 5

Path Integrals of the Quantum Fields

For this section we will apply the path integral formulation introduced in the background chapter to the quantum fields we have constructed. As mentioned before, the real scalar field will serve as a pedagogical tool and as the prototype for which the other path integrals of other fields will be modeled on.

5.1 Real Scalar Field

Using the path integral formulation of the quantum harmonic oscillator, the path integral of the *free field* theory of the real scalar field is constructed as

$$\begin{aligned} Z_0(J) \equiv \langle 0|0\rangle_J &= \int \mathcal{D}\varphi \, e^{i\mathcal{S}} \\ &= \int \mathcal{D}\varphi \, e^{i\int d^4x \, [\mathcal{L}_0 + J\varphi]}, \end{aligned} \tag{5.1.1}$$

where $J(x)$ is the source term, $\mathcal{D}\varphi \propto \prod_x d\varphi(x)$ is the functional measure, and

$$\mathcal{L}_0 = -\frac{1}{2}\partial^\mu\varphi\partial_\mu\varphi - \frac{1}{2}m^2\varphi^2, \tag{5.1.2}$$

is the free field real scalar Lagrangian [5]. Note that we have implicitly multiplied the Hamiltonian density by $(1-i\epsilon)$ so the mass term m^2 is actually $m^2-i\epsilon$ but we will suppress this in our derivations. Also note that equations of motion, namely, the Klein-Gordon equation now becomes $(\partial^2+m^2-i\epsilon)\hat{\varphi}(x)=0$ [5]. In accordance with the path integral formulation and the field formalism, the path integral now represents the path in the space of field configurations. Recall that when we applied the path integral formulation to the simple harmonic oscillator, we were able to derive the Green's function for the oscillator which led to the path integral to be simplified as

$$\langle 0|0\rangle_f = \exp\Big[\frac{i}{2}\int_{-\infty}^{+\infty} dt\,dt'\,f(t)G(t-t')f(t')\Big], \tag{5.1.3}$$

where $G(t-t')=i\langle 0|\Gamma\hat{X}(t)\hat{X}(t')|0\rangle$ was the propagator for the harmonic oscillator. This form of the path integral allowed us to easily compute the vacuum expectation values using functional derivatives so we seek to derive a similar expression for the real scalar field path integral. We first start by expanding the action $\mathcal{S}$ in the exponential of equation (5.1.1),

$$\begin{aligned}\mathcal{S} &= \int d^4x\,[\mathcal{L}_0 + J\varphi(x)]\\ &= \int d^4x\,\Big[-\frac{1}{2}\partial^\mu\varphi\partial_\mu\varphi - \frac{1}{2}(m^2-i\epsilon)\varphi^2 + J\varphi\Big]\\ &= \int d^4x\,\Big[-\frac{1}{2}\varphi(\partial^2+m^2-i\epsilon)\varphi + J\varphi\Big],\end{aligned}$$

where we integrated by parts twice and imposed the boundary condition $\varphi\to 0$ at infinity to get to the third line [5]. Then we treat the field φ as the integration variable and shift the field $\varphi'=\varphi+\Gamma$ where the field Γ is arbitrary. Since we are essentially shifting by a constant, the functional measure remains the same, namely, $\mathcal{D}\varphi'=\mathcal{D}\varphi$ thus

$$\begin{aligned}Z_0(J) &= \int \mathcal{D}\varphi'\exp\Big[i\int d^4x\,\Big(-\frac{1}{2}\varphi'(\partial^2+m^2-i\epsilon)\varphi' + J\varphi'\Big)\Big]\\ &= \int \mathcal{D}\varphi\exp\Big[i\int d^4x\,\Big(-\frac{1}{2}(\varphi+\Gamma)(\partial^2+m^2-i\epsilon)(\varphi+\Gamma) + J(\varphi+\Gamma)\Big)\Big].\end{aligned}$$

Before we continue, note that

$$\int d^4x\,\varphi\,\partial^2\,\Gamma = \int d^4x\,\Gamma\,\partial^2\,\varphi, \tag{5.1.4}$$

which is obtained by integrating by parts twice the left expression and imposing the boundary condition for the field φ. With this in mind, we continue to simplify the above path integral,

$$\begin{aligned} Z_0(J) &= \int \mathcal{D}\varphi \exp\Big[i\int d^4x\,\Big(-\frac{1}{2}(\varphi+\Gamma)(\partial^2+m^2-i\epsilon)(\varphi+\Gamma)+J(\varphi+\Gamma)\Big)\Big] \\ &= \int \mathcal{D}\varphi \exp\Big[i\int d^4x\,\frac{1}{2}\Big(-\varphi(-\partial^2+m^2-i\epsilon)\varphi-\Gamma(-\partial^2+m^2-i\epsilon)\Gamma\Big) \\ &\quad -\varphi(-\partial^2+m^2-i\epsilon)\Gamma+J(\varphi+\Gamma)\Big]. \end{aligned} \tag{5.1.5}$$

To get equation (5.1.5) into the same mathematical form as equation (5.1.3), recall that the field Γ was arbitrarily chosen so we impose the condition that

$$(-\partial^2+m^2-i\epsilon)\Gamma(x) = J(x). \tag{5.1.6}$$

Then we introduce the *Feynman propagator* which satisfies the condition

$$(-\partial^2+m^2-i\epsilon)\Delta_F(x-x') = \delta^4(x-x'), \tag{5.1.7}$$

thus the Feynman propagator is the Green's function of the (modified) Klein-Gordon equation. With these two equations, we solve for $\Gamma(x)$ explicitly,

$$\begin{gathered} \Big((-\partial^2+m^2-i\epsilon)\Gamma(x) = J(x)\Big)\Delta_F(x-x') \\ \int d^4x\,\Big((-\partial^2+m^2-i\epsilon)\Gamma(x)\Delta_F(x-x') = J(x)\Delta_F(x-x')\Big) \\ \Gamma(x) = \int d^4x'\,\Delta_F(x-x')J(x'), \end{gathered} \tag{5.1.8}$$

where we have switched the variables $x \leftrightarrow x'$ to get the last line. Now the path integral becomes

$$\begin{aligned}
Z_0(J) &= \int \mathcal{D}\varphi \exp\Big[i \int d^4x \, \frac{1}{2}\Big(-\varphi(-\partial^2 + m^2 - i\epsilon)\varphi - \Gamma(-\partial^2 + m^2 - i\epsilon)\Gamma\Big) \\
&\quad - \varphi(-\partial^2 + m^2 - i\epsilon)\Gamma + J(\varphi + \Gamma)\Big] \\
&= \int \mathcal{D}\varphi \exp\Big[\frac{i}{2} \int d^4x \, \Big(-\varphi(-\partial^2 + m^2 - i\epsilon)\varphi + J\Gamma\Big)\Big] \\
&= \int \mathcal{D}\varphi \exp\Big[\frac{i}{2} \int d^4x \, \Big(-\varphi(-\partial^2 + m^2 - i\epsilon)\varphi \\
&\quad + \int d^4x' \, J(x)\Delta_F(x - x')J(x')\Big)\Big] \\
&= \exp\Big[\frac{i}{2} \int d^4x \, d^4x' \, J(x)\Delta_F(x - x')J(x')\Big] \\
&\quad \times \int \mathcal{D}\varphi \exp\Big[-\frac{i}{2} \int d^4x \, \varphi(-\partial^2 + m^2 - i\epsilon)\varphi\Big].
\end{aligned}$$

Now recall that we must have $Z_0(0) = \langle 0|0\rangle_{J=0} = 1$ so the second term must equal 1 thus we finally get our desired path integral for the free Klein-Gordon field,

$$Z_0(J) = \exp\Big[\frac{i}{2} \int d^4x \, d^4x' \, J(x)\Delta_F(x - x')J(x')\Big]. \tag{5.1.9}$$

To get an expression for the Feynman propagator $\Delta_F(x - x')$, we cleverly use the Fourier transform of the Feynman propagator [5]

$$\Delta_F(x) = \int \frac{d^4p}{(2\pi)^4} \, e^{ipx} \, \tilde{\Delta}_F(p), \quad \tilde{\Delta}_F(p) = \int d^4x \, e^{-ipx} \, \Delta_F(x), \tag{5.1.10}$$

then apply the operator $(-\partial^2 + m^2 - i\epsilon)$ to get

$$\begin{aligned}
(-\partial^2 + m^2 - i\epsilon)\Big(\Delta_F(x) = \int \frac{d^4x}{(2\pi)^4} \, e^{ipx} \, \tilde{\Delta}_F(p)\Big) \\
\delta^4(x) = \int \frac{d^4p}{(2\pi)^4} \, (p^2 + m^2 - i\epsilon)\tilde{\Delta}_F(p) \, e^{ipx} \\
\int \frac{d^4p}{(2\pi)^4} \, (p^2 + m^2 - i\epsilon)\tilde{\Delta}_F(p) \, e^{ipx} = \int \frac{d^4p}{(2\pi)^4} \, e^{ipx},
\end{aligned}$$

which implies that

$$(p^2 + m^2 - i\epsilon)\tilde{\Delta}_F(p) = 1$$
$$\tilde{\Delta}_F(p) = \frac{1}{(p^2 + m^2 - i\epsilon)}. \tag{5.1.11}$$

Thus the Feynman propagator for the Klein-Gordon field is

$$\Delta_F(x - x') = \int \frac{d^4p}{(2\pi)^4} \frac{e^{ip(x-x')}}{p^2 + m^2 - i\epsilon}. \tag{5.1.12}$$

Applying the Klein-Gordon equation $(-\partial_x^2 + m^2)$ on the Feynman propagator (5.1.12) and taking the limit $\epsilon \to 0$ yields the the four-dimensional delta function which confirms that equation (5.1.12) is indeed the Green's function of the Klein-Gordon equation. We can also evaluate the Feynman propagator by contour integration by integrating in the complex p^0 plane to get the explicit form of the propagator [5]. To do this, we can expand equation (5.1.12)

$$\begin{aligned}
\Delta_F(x - x') &= \int \frac{d^4p}{(2\pi)^4} \frac{e^{ip(x-x')}}{p^2 + m^2 - i\epsilon} \\
&= \int \frac{d^3p}{(2\pi)^3} e^{i\mathbf{p}\cdot(\mathbf{x}-\mathbf{x}')} \int \frac{dp^0}{2\pi} \frac{e^{-i\omega(t-t')}}{p^2 + m^2 - i\epsilon} \\
&= \int \frac{d^3p}{(2\pi)^3} e^{i\mathbf{p}\cdot(\mathbf{x}-\mathbf{x}')} \int \frac{dp^0}{2\pi} f(p^0),
\end{aligned}$$

where,

$$f(p^0) = \frac{-e^{-ip^0\tau}}{(p^0 - \sqrt{\mathbf{p}^2 + m^2 - i\epsilon})(p^0 + \sqrt{\mathbf{p}^2 + m^2 - i\epsilon})}, \tag{5.1.13}$$

where $\tau = (t - t')$ and we have used the relation from special relativity $p^2 = -(p^0)^2 + \mathbf{p}^2$. For the case when $\tau > 0$, we integrate in the lower half plane where the closed contour encircles the pole at $p^0 = \sqrt{\mathbf{p}^2 + m^2 - i\epsilon}$ in a clockwise direction. Then

by the residue theorem,

$$\begin{aligned}\int \frac{dp^0}{2\pi} f(p^0) &= \frac{1}{2\pi}(-2\pi i)\sum_k \text{Res}(f, a_k) \\ &= \frac{ie^{-i\sqrt{\mathbf{p}^2+m^2-i\epsilon}\tau}}{2\sqrt{\mathbf{p}^2+m^2-i\epsilon}} \\ &\xrightarrow[\epsilon\to 0]{} \frac{ie^{-i\omega\tau}}{2\omega},\end{aligned}$$

where we multiplied by $-2\pi i$ since we integrating in the clockwise direction. If $\tau < 0$, we integrate in the upper half plane where the closed contour encircles the pole at $p^0 = -\sqrt{\mathbf{p}^2 + m^2 - i\epsilon}$. Then by the residue theorem,

$$\begin{aligned}\int \frac{dp^0}{2\pi} f(p^0) &= \frac{1}{2\pi}(2\pi i)\sum_k \text{Res}(f, a_k) \\ &= \frac{ie^{i\sqrt{\mathbf{p}^2+m^2-i\epsilon}\tau}}{2\sqrt{\mathbf{p}^2+m^2-i\epsilon}} \\ &\xrightarrow[\epsilon\to 0]{} \frac{ie^{i\omega\tau}}{2\omega}.\end{aligned}$$

Combining our results yields

$$\begin{aligned}\int \frac{dp^0}{2\pi} f(p^0) &= \frac{ie^{-i\omega|\tau|}}{2\omega} \\ &= \frac{ie^{-i\omega|t-t'|}}{2\omega}.\end{aligned} \tag{5.1.14}$$

Thus the Feynman propagator is

$$\begin{aligned}\Delta_F(x-x') &= i\int d^3p_L\, e^{i\mathbf{p}\cdot(\mathbf{x}-\mathbf{x}')-i\omega|t-t'|} \\ &= i\theta(t-t')\Delta(x-x') + i\theta(t'-t)\Delta(x'-x) \\ &= i\langle 0|\mathrm{T}\hat{\varphi}(x)\hat{\varphi}(x')|0\rangle,\end{aligned} \tag{5.1.15}$$

where $\Delta(x-x') = \langle 0|\hat{\varphi}(x)\hat{\varphi}(x')|0\rangle = \int d^3p_L\, e^{ip(x-x')}$ is the amplitude for a particle to propagate from position x' to position x [5]. Recall that we first introduced Wick's Theorem when we applied the path integral formalism to the quantum

harmonic oscillator. Before we prove the theorem for fields, we will first introduce some concepts that will aid us. For any free field $\hat{\varphi}(x)$, we can decompose the field into its positive and negative frequency parts [8]:

$$\hat{\varphi}(x) = \hat{\varphi}^+(x) + \hat{\varphi}^-(x), \tag{5.1.16}$$

where

$$\hat{\varphi}^+(x) = \int d^3p_L \hat{a}_p e^{-i\mathbf{x}\cdot\mathbf{p}},$$
$$\hat{\varphi}^-(x) = \int d^3p_L \hat{a}_p^\dagger e^{i\mathbf{x}\cdot\mathbf{p}},$$

then we have

$$\hat{\varphi}^+(x)|0\rangle = 0,$$
$$\langle 0|\hat{\varphi}^-(x) = 0.$$

We define a product of creation and annihilation operators written with all the creation operators to left and all the annihilation operators to the right, such as $\hat{a}_{p_1}^\dagger \hat{a}_{p_2}^\dagger \hat{a}_{p_3} \hat{a}_{p_4}$, as to be in *normal order* and will have a vanishing vacuum expectation value [8]. We also define the normal ordering operation $N()$ that places operators in normal order

$$N(\hat{a}_{p_1} \hat{a}_{p_2}^\dagger \hat{a}_{p_3}) = \hat{a}_{p_2}^\dagger \hat{a}_{p_1} \hat{a}_{p_3}. \tag{5.1.17}$$

The order of the annihilation operators does not matter since they commute. We define a *contraction* to be

$$\overbrace{\hat{\varphi}(x)\hat{\varphi}}(x) = \Delta_F(x-y), \tag{5.1.18}$$

which is the Feynman propagator. Then for two fields

$$T\{\hat{\varphi}(x)\hat{\varphi}(y)\} = N\{\hat{\varphi}(x)\hat{\varphi}(y) + \overbrace{\hat{\varphi}(x)\hat{\varphi}}(x)\}.$$

When two non-adjacent operators are contracted, we will still assign the propagator

$$N\{\widehat{\varphi}_1\widehat{\varphi}_2\widehat{\varphi}_3\widehat{\varphi}_4\} \equiv \Delta_F(x_1 - x_3) \cdot N\{\widehat{\varphi}_2\widehat{\varphi}_4\}.$$

Theorem 5.1.1 (**Wick's Theorem**). *Let $\hat{\varphi}(x)$ be a field operator with the propagator $\Delta_F(x-y)$. Then*

$$T\{\hat{\varphi}_1 \dots \hat{\varphi}_n\} = N\{\hat{\varphi}_1 \dots \hat{\varphi}_n + \text{all possible contractions}\},$$

or equivalently,

$$\begin{aligned}\langle 0|\mathrm{T}\hat{\varphi}(x_1)\hat{\varphi}(x_2)\dots\hat{\varphi}(x_{2n})|0\rangle &= \frac{1}{i}\frac{\delta}{\delta J(x_1)}\frac{1}{i}\frac{\delta}{\delta J(x_2)}\dots\frac{1}{i}\frac{\delta}{\delta J(x_{2n})}Z_0(J)\Big|_{J=0} \\ &= \frac{1}{i^n}\sum_{\text{pairings}}\Delta_F(x_{i_1}-x_{i_2})\dots\Delta_F(x_{i_{2n-1}}-x_{i_{2n}}).\end{aligned} \tag{5.1.19}$$

Proof. We follow the proof according to Peskin and Schroeder [8]. We proceed by induction. The base case $n = 2$ was proved earlier so suppose the theorem is true for $n-1$ fields. Without loss of generality, we consider the case $x_1^0 \geq x_2^0 \geq \dots x_n^0$. We use Wick's theorem on the fields $\hat{\varphi}_2 \dots \hat{\varphi}_n$ to get

$$\begin{aligned}T\{\hat{\varphi}_1 \dots \hat{\varphi}_n\} &= \hat{\varphi}_1 \dots \hat{\varphi}_n \\ &= \hat{\varphi}_1 N\Big\{\hat{\varphi}_2 \dots \hat{\varphi}_n + \Big(\text{all contractions not involving } \hat{\varphi}_1\Big)\Big\} \\ &= (\hat{\varphi}_1^+ + \hat{\varphi}_1^-)N\Big\{\hat{\varphi}_2 \dots \hat{\varphi}_n + \Big(\text{all contractions not involving } \hat{\varphi}_1\Big)\Big\}.\end{aligned} \tag{5.1.20}$$

We can move $\hat{\varphi}_1^-$ inside $N\{\}$ since it is in normal order. For $\hat{\varphi}_1^+$, observe that

$$\begin{aligned}\hat{\varphi}_1^+ N(\hat{\varphi}_2 \dots \hat{\varphi}_n) &= N(\hat{\varphi}_2 \dots \hat{\varphi}_n)\hat{\varphi}_1^+ + [\hat{\varphi}_1^+], N(\hat{\varphi}_2 \dots \hat{\varphi}_n)] \\ &= N(\hat{\varphi}_1^+ \hat{\varphi}_2 \dots \hat{\varphi}_n) + N([\hat{\varphi}_1^+, \hat{\varphi}_2^-]\hat{\varphi}_3 \dots \hat{\varphi}_n \\ &+ \hat{\varphi}_2[\hat{\varphi}_1^+, \hat{\varphi}_3^-]\hat{\varphi}_4 \dots \hat{\varphi}_n + \dots) \\ &= N(\hat{\varphi}_1^+ \hat{\varphi}_2 \dots \hat{\varphi}_n + \overbrace{\hat{\varphi}_1 \hat{\varphi}_2}\hat{\varphi}_3 \dots \hat{\varphi}_n + \overbrace{\hat{\varphi}_1 \hat{\varphi}_2 \hat{\varphi}_3} \cdots + \dots). \quad (5.1.21)\end{aligned}$$

The first term in the last line of (5.1.21) combines with the $\hat{\varphi}_1^-$ term in equation (5.1.20) to form $N\{\hat{\varphi}_1 \hat{\varphi}_2 \dots \hat{\varphi}_n\}$ and other terms involving the contraction of $\hat{\varphi}_1$ with another field. We also see that a contraction with one term in equation (5.1.20) will yield all possible terms involving both that particular contraction and a contraction of $\hat{\varphi}_1$ with one of the other fields. Continuing this process gives all possible contractions of all the fields including $\hat{\varphi}_1$ thus the induction step is established therefore Wick's theorem holds. □

Wick's theorem will be used frequently when we construct Feynman diagrams and for computing vacuum expectation values.

5.2 Spinor and Gamma Matrix Technology

Before we move onto the construction of the path integral of the free fermion field and the derivation of its Feynman propagator, we will finally present and prove the various identities and computational shortcuts for spinors that was first introduced in the fields chapter. These identities and shortcuts will be used ubiquitously when doing calculations involving spinors.

Spinor Technology

We begin with computations involving spinors specifically Dirac spinors since the Majorana spinor is a specific case of the Dirac spinor. Recall that the general

solution to the Dirac equation for the Dirac field was

$$\hat{\mathbf{\Psi}}(x) = \sum_{s=\pm} \int d^3p_L \left[\hat{b}_s(\mathbf{p})u_s(\mathbf{p})e^{ipx} + \hat{d}_s^\dagger(\mathbf{p})v_s(\mathbf{p})e^{-ipx}\right], \tag{5.2.1}$$

where the spinors $u_s(\mathbf{p})$ and $v_s(\mathbf{p})$ obey

$$\begin{aligned}(\not{p} + m)u_s(\mathbf{p}) = 0,\\ (-\not{p} + m)v_s(\mathbf{p}) = 0.\end{aligned} \tag{5.2.2}$$

and similarly,

$$\begin{aligned}\bar{u}_s(\mathbf{p})(\not{p} + m) = 0,\\ \bar{v}_s(\mathbf{p})(-\not{p} + m) = 0.\end{aligned} \tag{5.2.3}$$

Note that each of the spinor equations have two solutions based on their spin values. To find the explicit solutions for the spinors, recall from special relativity that a particle with $m \neq 0$ possesses a rest frame where $\mathbf{p} = 0$ [4]. Let us move to the rest frame where $\not{p} = p_\mu\gamma^\mu$ then becomes $\not{p} = p_0\gamma^0 = -m\gamma^0$ [5]. Then equation (5.2.2) becomes

$$\begin{aligned}(-m\gamma^0 + m)u_s(\mathbf{0}) = 0,\\ (m\gamma^0 + m)v_s(\mathbf{0}) = 0,\end{aligned} \tag{5.2.4}$$

which yields

$$\begin{aligned} u_+(\mathbf{0}) &= \sqrt{m}\begin{pmatrix} +|\uparrow\rangle \\ +|\uparrow\rangle \end{pmatrix}, & u_-(\mathbf{0}) &= \sqrt{m}\begin{pmatrix} +|\downarrow\rangle \\ +|\downarrow\rangle \end{pmatrix},\\ v_+(\mathbf{0}) &= \sqrt{m}\begin{pmatrix} +|\downarrow\rangle \\ -|\downarrow\rangle \end{pmatrix}, & v_-(\mathbf{0}) &= \sqrt{m}\begin{pmatrix} -|\uparrow\rangle \\ +|\uparrow\rangle \end{pmatrix}, \end{aligned} \tag{5.2.5}$$

and similarly we also get

$$\begin{aligned}
\bar{u}_+(\mathbf{0}) &= \sqrt{m}\,(+\langle\uparrow|, +\langle\uparrow|), \\
\bar{u}_-(\mathbf{0}) &= \sqrt{m}\,(+\langle\downarrow|, +\langle\downarrow|), \\
\bar{v}_+(\mathbf{0}) &= \sqrt{m}\,(-\langle\downarrow|, +\langle\downarrow|), \\
\bar{v}_-(\mathbf{0}) &= \sqrt{m}\,(+\langle\uparrow|, -\langle\uparrow|),
\end{aligned} \tag{5.2.6}$$

where

$$|\uparrow\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \tag{5.2.7}$$

$$|\downarrow\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \tag{5.2.8}$$

are the spin-up and spin-down eigenkets of the Pauli matrices, respectively [11]. With (5.2.5), it is a simple matter of finding the spinors with an arbitrary three-momentum $\mathbf{p}$. To do this, we follow the rules of quantum mechanics and construct the necessary one-parameter unitary operator $\hat{U}(\Lambda)$, whose generator is the boost matrix $\mathbf{K}$, parametrized by the rapidity $\eta = \sinh^{-1}(|\mathbf{p}|/m)$ in the direction of the unit vector $\hat{\mathbf{p}}$ which is [5]

$$\hat{U}(\Lambda) = \exp(i\eta\,\hat{\mathbf{p}}\cdot\mathbf{K}). \tag{5.2.9}$$

Then we apply the operator $\hat{U}(\Lambda)$ to equations (5.2.5) and (5.2.6) to get

$$\begin{aligned}
u_s(\mathbf{p}) &= \exp(i\eta\,\hat{\mathbf{p}}\cdot\mathbf{K})u_s(\mathbf{0}), \\
v_s(\mathbf{p}) &= \exp(i\eta\,\hat{\mathbf{p}}\cdot\mathbf{K})v_s(\mathbf{0}).
\end{aligned} \tag{5.2.10}$$

Now we define the operation [5]

$$\overline{A} \equiv \beta A^{\dagger} \beta, \tag{5.2.11}$$

and it is trivial to see that

$$\begin{aligned}
\overline{\gamma^{\mu}} &= \gamma^{\mu}, \\
\overline{S^{\mu\nu}} &= S^{\mu\nu}, \\
\overline{i\gamma_5} &= i\gamma_5, \\
\overline{\gamma^{\mu}\gamma_5} &= \gamma^{\mu}\gamma_5, \\
\overline{\gamma_5 S^{\mu\nu}} &= \gamma_5 S^{\mu\nu}.
\end{aligned} \tag{5.2.12}$$

Recall that the boost matrix components are formed by

$$K^j = \frac{i}{4}[\gamma^j, \gamma^0] = \frac{i}{2}\gamma^j\gamma^0, \tag{5.2.13}$$

so it is trivial to see that $\overline{K^j} = K^j$ thus we get [8]

$$\begin{aligned}
\bar{u}_s(\mathbf{p}) &= \bar{u}_s(\mathbf{0})\exp(-i\eta\,\hat{\mathbf{p}}\cdot\mathbf{K}), \\
\bar{v}_s(\mathbf{p}) &= \bar{v}_s(\mathbf{0})\exp(-i\eta\,\hat{\mathbf{p}}\cdot\mathbf{K}).
\end{aligned} \tag{5.2.14}$$

Now it is trivial to derive

$$\begin{aligned}
\bar{u}_{s'}(\mathbf{p})u_s(\mathbf{p}) &= +2m\delta_{s's}, \\
\bar{v}_{s'}(\mathbf{p})v_s(\mathbf{p}) &= -2m\delta_{s's}, \\
\bar{u}_{s'}(\mathbf{p})v_s(\mathbf{p}) &= 0, \\
\bar{v}_{s'}(\mathbf{p})u_s(\mathbf{p}) &= 0,
\end{aligned} \tag{5.2.15}$$

from equations (5.2.10) and (5.2.14) and the orthonormality of the spin kets and bras. From here, we are in a position to derive the *Gordon identities* that we used

in the fields chapter (we used the case $\mu = 0$) [5]

$$2m\bar{u}_{s'}(\mathbf{p}')\gamma^\mu u_s(\mathbf{p}) = \bar{u}_{s'}(\mathbf{p}')\Big[(p'+p)^\mu - 2iS^{\mu\nu}(p'-p)_\nu\Big]u_s(\mathbf{p}),$$
$$2m\bar{v}_{s'}(\mathbf{p}')\gamma^\mu v_s(\mathbf{p}) = \bar{v}_{s'}(\mathbf{p}')\Big[(p'+p)^\mu - 2iS^{\mu\nu}(p'-p)_\nu\Big]v_s(\mathbf{p}). \tag{5.2.16}$$

Proof. Consider the following identities

$$\gamma^\mu \not{p} = \frac{1}{2}\{\gamma^\mu, \not{p}\} + \frac{1}{2}[\gamma^\mu, \not{p}] = -p^\mu - 2iS^{\mu\nu}p_\nu, \tag{5.2.17}$$
$$\not{p}'\gamma^\mu = \frac{1}{2}\{\gamma^\mu, \not{p}'\} + \frac{1}{2}[\gamma^\mu, \not{p}'] = -p'^\mu + 2iS^{\mu\nu}p'_\nu. \tag{5.2.18}$$

We add the above identities to get

$$\gamma^\mu \not{p} + \not{p}'\gamma^\mu = -\Big[(p'+p)^\mu - 2iS^{\mu\nu}(p'-p)_\nu\Big]. \tag{5.2.19}$$

To get the first Gordon identity, we multiply the above expression by $\bar{u}_{s'}(\mathbf{p}')$ and $u_s(\mathbf{p})$ on the left and right, respectively, then use equations (5.2.2) and (5.2.3). Following the same procedure except with $\bar{v}_{s'}(\mathbf{p}')$ and $v_s(\mathbf{p})$ yields the second Gordon identity. □

A similar identity that can be derived is

$$2m\bar{u}_{s'}(\mathbf{p}')\gamma^\mu v_s(\mathbf{p}) = \bar{u}_{s'}(\mathbf{p}')\Big[(p'-p)^\mu - 2iS^{\mu\nu}(p'+p)_\nu\Big]v_s(\mathbf{p}),$$
$$-2m\bar{v}_{s'}(\mathbf{p}')\gamma^\mu u_s(\mathbf{p}) = \bar{v}_{s'}(\mathbf{p}')\Big[(p'-p)^\mu - 2iS^{\mu\nu}(p'+p)_\nu\Big]u_s(\mathbf{p}). \tag{5.2.20}$$

Proof. The proof is similar to the Gordon identity. Subtract (5.2.18) from (5.2.17) to get

$$\gamma^\mu \not{p} - \not{p}'\gamma^\mu = \Big[(p'-p)^\mu - 2iS^{\mu\nu}(p'+p)_\nu\Big]. \tag{5.2.21}$$

Then multiply the above by $\bar{u}_{s'}(\mathbf{p}')$ and $v_s(\mathbf{p})$ on the left and right, respectively, and use equations (5.2.2) and (5.2.3). Using the same procedure with $\bar{v}_{s'}(\mathbf{p}')$ and $u_s(\mathbf{p})$ yields the second identity. □

When $p' = p$, the Gordon identities reduce to

$$\begin{aligned} \bar{u}_{s'}(\mathbf{p})\gamma^{\mu}u_s(\mathbf{p}) &= 2p^{\mu}\delta_{s's}, \\ \bar{v}_{s'}(\mathbf{p})\gamma^{\mu}v_s(\mathbf{p}) &= 2p^{\mu}\delta_{s's}. \end{aligned} \tag{5.2.22}$$

Similarly, when $p' = -p$ and $\mu = 0$ in (5.2.20), we get

$$\begin{aligned} \bar{u}_{s'}(\mathbf{p})\gamma^{0}v_s(-\mathbf{p}) &= 0, \\ \bar{v}_{s'}(\mathbf{p})\gamma^{\mu}u_s(-\mathbf{p}) &= 0. \end{aligned} \tag{5.2.23}$$

An incredibly useful identity that we will use for calculations in scattering cross sections for fermions are the spin-sum identities

$$\begin{aligned} \sum_{s=\pm} u_s(\mathbf{p})\bar{u}_s(\mathbf{p}) &= -\not{p} + m, \\ \sum_{s=\pm} v_s(\mathbf{p})\bar{v}_s(\mathbf{p}) &= -\not{p} - m, \end{aligned} \tag{5.2.24}$$

which is derived from the fact that in the rest frame

$$\begin{aligned} \sum_{s=\pm} u_s(\mathbf{0})\bar{u}_s(\mathbf{0}) &= m\gamma^{0} + m, \\ \sum_{s=\pm} v_s(\mathbf{0})\bar{v}_s(\mathbf{0}) &= m\gamma^{0} - m. \end{aligned}$$

Summing over the eigenstates of S_z allows one to rewrite the result in terms of the four-momentum and gamma matrices with contracted vector indices so generalizing the rest frame spin-sum to a frame with $\mathbf{p} \neq 0$ yields the spin-sum identities above [5].

Gamma Matrix Technology

There many kinds of identities that can be derived using the gamma matrices along with the Feynman slash notation for variables. Some identities will, of course,

prove to be more useful than others so we will derive the most commonly used gamma matrix identities. To begin our derivations, we start with the following relations of the gamma matrices as our "axioms"

$$\{\gamma^\mu, \gamma^\nu\} = -2g^{\mu\nu}, \tag{5.2.25}$$

$$\gamma_5^2 = 1, \tag{5.2.26}$$

$$\{\gamma^\mu, \gamma_5\} = 0, \tag{5.2.27}$$

$$\mathrm{Tr}\, 1 = 4, \tag{5.2.28}$$

where $\gamma_5 = i\gamma^0\gamma^1\gamma^2\gamma^3$ [5]. We claim that

$$\mathrm{Tr}\,[\text{product of odd number of } \gamma^\mu] = 0, \tag{5.2.29}$$

$$\mathrm{Tr}\,[\gamma_5(\text{product of odd number of } \gamma^\mu)] = 0. \tag{5.2.30}$$

Proof. Using the cyclic property of the trace and the above axioms, observe that [5]

$$\begin{aligned}
\mathrm{Tr}\,[\gamma^{\mu_1}\dots\gamma^{\mu_n}] &= \mathrm{Tr}\,[\gamma_5^2\gamma^{\mu_1}\gamma_5^2\dots\gamma_5^2\gamma^{\mu_n}] \\
&= \mathrm{Tr}\,[(-\gamma_5^2\gamma^{\mu_1})\dots(-\gamma_5^2\gamma^{\mu_n})] \\
&= (-1)^n\mathrm{Tr}\,[\gamma^{\mu_1}\dots\gamma^{\mu_n}],
\end{aligned}$$

and similarly,

$$\begin{aligned}
\mathrm{Tr}\,[\gamma_5(\gamma^{\mu_1}\dots\gamma^{\mu_n})] &= \mathrm{Tr}\,[\gamma_5(\gamma_5^2\gamma^{\mu_1}\gamma_5^2\dots\gamma_5^2\gamma^{\mu_n})] \\
&= \mathrm{Tr}\,[\gamma_5((-\gamma_5^2\gamma^{\mu_1})\dots(-\gamma_5^2\gamma^{\mu_n}))] \\
&= (-1)^n\mathrm{Tr}\,[\gamma_5(\gamma^{\mu_1}\dots\gamma^{\mu_n})],
\end{aligned}$$

hence,

$$\begin{aligned}\mathrm{Tr}\left[\gamma^{\mu_1}\ldots\gamma^{\mu_n}\right] &= (-1)^n\mathrm{Tr}\left[\gamma^{\mu_1}\ldots\gamma^{\mu_n}\right],\\ \mathrm{Tr}\left[\gamma_5(\gamma^{\mu_1}\ldots\gamma^{\mu_n})\right] &= (-1)^n\mathrm{Tr}\left[\gamma_5(\gamma^{\mu_1}\ldots\gamma^{\mu_n})\right].\end{aligned} \tag{5.2.31}$$

If n is odd then (5.2.31) equals its negative value which means that $\mathrm{Tr}\left[\gamma^{\mu_1}\ldots\gamma^{\mu_n}\right] = 0$ and $\mathrm{Tr}\left[\gamma_5(\gamma^{\mu_1}\ldots\gamma^{\mu_n})\right] = 0$. □

Now

$$\mathrm{Tr}\left[\gamma^\mu\gamma^\nu\right] = -4g^{\mu\nu}. \tag{5.2.32}$$

Proof. Observe that

$$\begin{aligned}\mathrm{Tr}\left[\gamma^\mu\gamma^\nu\right] &= \mathrm{Tr}\left[\gamma^\nu\gamma^\mu\right]\\ &= \mathrm{Tr}\left[-\gamma^\mu\gamma^\nu - 2g^{\mu\nu}\right]\\ &= -\mathrm{Tr}\left[\gamma^\mu\gamma^\nu\right] - 2g^{\mu\nu}\mathrm{Tr}\,1\\ &= -\mathrm{Tr}\left[\gamma^\mu\gamma^\nu\right] - 8g^{\mu\nu},\end{aligned}$$

then rearranging the terms gives us our desired result

$$\begin{aligned}2\mathrm{Tr}\left[\gamma^\mu\gamma^\nu\right] &= -8g^{\mu\nu}\\ \mathrm{Tr}\left[\gamma^\mu\gamma^\nu\right] &= -4g^{\mu\nu}.\end{aligned}$$

□

It is trivial to see that

$$\mathrm{Tr}\left[\gamma_5\right] = 0, \tag{5.2.33}$$

since,

$$\gamma_5 = \begin{pmatrix} -I & 0\\ 0 & I\end{pmatrix}. \tag{5.2.34}$$

With this, we claim that

$$\mathrm{Tr}\left[\gamma_5\gamma^\mu\gamma^\nu\right] = 0, \tag{5.2.35}$$

and more generally,

$$\mathrm{Tr}\left[\gamma_5\gamma^\mu\gamma^\nu\gamma^\rho\gamma^\sigma\right] = -4i\varepsilon^{\mu\nu\rho\sigma} \tag{5.2.36}$$

Proof. Observe that

$$\begin{aligned}
\mathrm{Tr}\left[\gamma_5\gamma^\mu\gamma^\nu\right] &= \mathrm{Tr}\left[\gamma_5(-\gamma^\nu\gamma^\mu - 2g^{\mu\nu})\right] \\
&= -\mathrm{Tr}\left[\gamma_5\gamma^\nu\gamma^\mu\right] - 2g^{\mu\nu}\mathrm{Tr}\left[\gamma_5\right] \\
&= -\mathrm{Tr}\left[\gamma_5\gamma^\mu\gamma^\nu\right],
\end{aligned}$$

which implies that $\mathrm{Tr}\left[\gamma_5\gamma^\mu\gamma^\nu\right] = 0$.

Now consider

$$\mathrm{Tr}\left[\gamma_5\gamma^\mu\gamma^\nu\gamma^\rho\gamma^\sigma\right].$$

If any of the indices have the same value, then the expression will always be zero. Suppose all the indices are distinct. Because of $\{\gamma^\mu, \gamma_5\} = 0$, we see that switching any two of the indices will change the sign of the trace and since we showed that the trace equaled zero if any of the indices were the same, the trace must therefore by proportional to $\varepsilon^{\mu\nu\rho\sigma}$. The proportionality constant is found by considering the case $\varepsilon^{0123} = \varepsilon^{3210} = +1$ which is

$$\begin{aligned}
\mathrm{Tr}\left[(\gamma_5)\gamma^0\gamma^1\gamma^2\gamma^3\right] &= -i\mathrm{Tr}\left[(\gamma_5)(-i\gamma_5)\right] \\
&= -i\mathrm{Tr}\left[1\right] \\
&= -4i.
\end{aligned}$$

Therefore

$$\mathrm{Tr}\left[\gamma_5\gamma^\mu\gamma^\nu\gamma^\rho\gamma^\sigma\right] = -4i\varepsilon^{\mu\nu\rho\sigma} = -4i\varepsilon^{\mu\nu\rho\sigma}.$$

□

Now the definition of the Feynman slash notation was given as

$$\not{a} = a_\mu \gamma^\mu. \tag{5.2.37}$$

The generalization of this notation to two or more quantities is [5]

$$\not{a}_1 \not{a}_2 \ldots \not{a}_n = a_{\mu_1} a_{\mu_2} \ldots a_{\mu_n} \gamma^{\mu_1} \gamma^{\mu_2} \ldots \gamma^{\mu_n}. \tag{5.2.38}$$

It will prove fruitful for future derivations to evaluate the anticommutator of two Feynman slashed variables which is

$$\begin{aligned} \{\not{a}, \not{b}\} &= \{a_\mu \gamma^\mu, b_\nu \gamma^\nu\} \\ &= a_\mu b_\nu \{\gamma^\mu, \gamma^\nu\} \\ &= -2 a_\mu b_\nu g^{\mu\nu} \\ &= -2(ab), \end{aligned} \tag{5.2.39}$$

where $(ab) = a^\mu b_\mu$. This implies that [5]

$$\not{a}\not{b} = -\not{b}\not{a} - 2(ab). \tag{5.2.40}$$

With equations (5.2.40) and (5.2.32), we are able to write

$$\mathrm{Tr}\,[\not{a}\not{b}] = -4(ab). \tag{5.2.41}$$

Now we claim that

$$\mathrm{Tr}\,[\not{a}\not{b}\not{c}\not{d}] = 4\big[(ab)(bc) - (ac)(bd) + (ab)(cd)\big]. \tag{5.2.42}$$

Proof. Observe that

$$\begin{aligned}
\mathrm{Tr}\,[\not{a}\not{b}\not{c}\not{d}] &= \mathrm{Tr}\,[(-\not{b}\not{a} - 2(ab))\not{c}\not{d}] \\
&= -\mathrm{Tr}\,[\not{b}\not{a}\not{c}\not{d}] - 2(ab)\mathrm{Tr}\,[\not{c}\not{d}] \\
&= -\mathrm{Tr}\,[\not{b}(-\not{c}\not{a} - 2(ac))\not{d}] - 2(ab)\mathrm{Tr}\,[\not{c}\not{d}] \\
&= \mathrm{Tr}\,[\not{b}\not{c}\not{a}\not{d}] + 2(ac)\mathrm{Tr}\,[\not{b}\not{d}] - 2(ab)\mathrm{Tr}\,[\not{c}\not{d}] \\
&= \mathrm{Tr}\,[\not{b}\not{c}(-\not{d}\not{a} - 2(ad))] + 2(ac)\mathrm{Tr}\,[\not{b}\not{d}] - 2(ab)\mathrm{Tr}\,[\not{c}\not{d}] \\
&= -\mathrm{Tr}\,[\not{b}\not{c}\not{d}\not{a}] - 2(ad)\mathrm{Tr}\,[\not{b}\not{c}] + 2(ac)\mathrm{Tr}\,[\not{b}\not{d}] - 2(ab)\mathrm{Tr}\,[\not{c}\not{d}] \\
&= -\mathrm{Tr}\,[\not{a}\not{b}\not{c}\not{d}] - 2(ad)\mathrm{Tr}\,[\not{b}\not{c}] + 2(ac)\mathrm{Tr}\,[\not{b}\not{d}] - 2(ab)\mathrm{Tr}\,[\not{c}\not{d}],
\end{aligned}$$

where we have used the cyclic property of trace and equation (5.2.40) [5]. This implies that

$$\begin{aligned}
2\mathrm{Tr}\,[\not{a}\not{b}\not{c}\not{d}] &= -2(ad)\mathrm{Tr}\,[\not{b}\not{c}] + 2(ac)\mathrm{Tr}\,[\not{b}\not{d}] - 2(ab)\mathrm{Tr}\,[\not{c}\not{d}] \\
\mathrm{Tr}\,[\not{a}\not{b}\not{c}\not{d}] &= -(ad)\mathrm{Tr}\,[\not{b}\not{c}] + (ac)\mathrm{Tr}\,[\not{b}\not{d}] - (ab)\mathrm{Tr}\,[\not{c}\not{d}] \\
\mathrm{Tr}\,[\not{a}\not{b}\not{c}\not{d}] &= 4\big[(ab)(bc) - (ac)(bd) + (ab)(cd)\big],
\end{aligned}$$

where we have applied (5.2.41) [5]. □

Throughout the course of this paper, we have considered four spacetime dimensions for the metric $g^{\mu\nu}$. Suppose we generalized our metric to d-dimensions. Then the inner product of two gamma matrices is [5]

$$\begin{aligned}
\gamma^\mu\gamma_\mu &= g_{\mu\nu}\gamma^\mu\gamma^\nu \\
&= g_{\mu\nu}(-\gamma^\nu\gamma^\mu - 2g^{\mu\nu}) \\
&= -\gamma_\mu\gamma^\mu - 2g_{\mu\nu}g^{\mu\nu},
\end{aligned}$$

which after rearranging yields

$$\gamma^\mu\gamma_\mu = -g_{\mu\nu}g^{\mu\nu} = -d. \tag{5.2.43}$$

Then we have

$$\gamma^\mu \not{a} \gamma_\mu = (d-2)\not{a}, \tag{5.2.44}$$

$$\gamma^\mu \not{a} \not{b} \gamma_\mu = 4(ab) - (d-4)\not{a}\not{b}, \tag{5.2.45}$$

$$\gamma^\mu \not{a} \not{b} \not{c} \gamma_\mu = 2\not{c}\not{b}\not{a} + (d-4)\not{a}\not{b}\not{c}. \tag{5.2.46}$$

Proof. We note that

$$\{\not{a}, \gamma^\mu\} = -2a_\mu, \tag{5.2.47}$$

then [5]

$$\begin{aligned}
\gamma^\mu \not{a} \gamma_\mu &= \gamma^\mu(-\gamma_\mu \not{a} - 2a_\mu) \\
&= -\gamma^\mu \gamma_\mu \not{a} - 2\gamma^\mu a_\mu \\
&= (d-2)\not{a}.
\end{aligned}$$

Now observe

$$\begin{aligned}
\gamma^\mu \not{a} \not{b} \gamma_\mu &= \gamma^\mu \not{a}(-\gamma_\mu \not{b} - 2b_\mu \\
&= -\gamma^\mu \not{a} \gamma_\mu \not{b} - 2\gamma^\mu \not{a} b_\mu \\
&= -(d-2)\not{a}\not{b} - 2(-\not{a}\gamma^\mu - 2a^\mu)b_\mu \\
&= -(d-2)\not{a}\not{b} + 2\not{a}\not{b} + 4(ab) \\
&= 4(ab) - (d-4)\not{a}\not{b}.
\end{aligned}$$

Lastly,

$$\begin{aligned}
\gamma^\mu \not{a}\not{b}\not{c}\gamma_\mu &= \gamma^\mu \not{a}\not{b}(-\gamma_\mu \not{c} - 2c_\mu) \\
&= -\gamma^\mu \not{a}\not{b}\gamma_\mu \not{c} - 2\gamma^\mu \not{a}\not{b} c_\mu \\
&= -(4(ab) - (d-4)\not{a}\not{b})\not{c} - 2\gamma^\mu(-\not{b}\not{a} - 2(ab))c_\mu \\
&= (d-4)\not{a}\not{b}\not{c} - 4(ab)\not{c} + 2\gamma^\mu \not{b}\not{a} c_\mu + 4\gamma^\mu (ab) c_\mu \\
&= (d-4)\not{a}\not{b}\not{c} - 4(ab)\not{c} + 2\not{c}\not{b}\not{a} + 4(ab)\not{c} \\
&= 2\not{c}\not{b}\not{a} + (d-4)\not{a}\not{b}\not{c}.
\end{aligned}$$

□

5.3 Fermion Fields

To construct the path integrals of the free fermion fields, we must account for the antisymmetric nature of the fields along with the fact that the fields are complex valued. We anticipate that the Dirac field, with its U(1) symmetry, will need two complex valued, anticommuting source terms in its path integral [5]. We introduce the *anticommuting* source field $\eta(x)$ and its Hermitian conjugate equivalent $\bar{\eta}(x) = \eta^\dagger(x)\beta$ thus the path integral for the free Dirac field is [5]

$$\begin{aligned}
Z_0(\eta, \bar{\eta}) &= \int \mathcal{D}\mathbf{\Psi}\, \mathcal{D}\overline{\mathbf{\Psi}} \exp\Big[i \int d^4x \, (\mathcal{L}_0 + \bar{\eta}\mathbf{\Psi} + \overline{\mathbf{\Psi}}\eta)\Big] \\
&= \int \mathcal{D}\mathbf{\Psi}\, \mathcal{D}\overline{\mathbf{\Psi}} \exp\Big[i \int d^4x \, (-\overline{\mathbf{\Psi}}(-i\not{\partial} + m)\mathbf{\Psi} + \bar{\eta}\mathbf{\Psi} + \overline{\mathbf{\Psi}}\eta)\Big]. \qquad (5.3.1)
\end{aligned}$$

By using the equivalent methods in the derivation of the real scalar free field path integral (5.1.9), the path integral for the free Dirac field is

$$Z_0(\eta, \bar{\eta}) = \exp\Big[i \int d^4x \, d^4y \, \bar{\eta}(x) S_F(x-y) \eta(y)\Big], \qquad (5.3.2)$$

where $S_F(x-y)$ is the Feynman propagator for the free fermion field [5]. The Majorana Lagrangian does not possess U(1) symmetry so we expect its path

integral to be analogous to the real scalar path integral. We must also incorporate the Majorana condition $\hat{\overline{\boldsymbol{\chi}}} = \hat{\boldsymbol{\chi}}^{\mathrm{T}}\mathcal{C}$ in its path integral as well. With this, the path integral of the free Majorana field is [5]

$$\begin{aligned} Z_0(\eta) &= \int \mathcal{D}\boldsymbol{\chi} \exp\Big[i \int d^4x \, (\mathcal{L}_0 + \eta^T \boldsymbol{\chi})\Big] \\ &= \exp\Big[-\frac{i}{2} \int d^4x \, d^4y \, \eta^{\mathrm{T}}(x) S_F(x-y) \mathcal{C}^{-1} \eta(y) \Big]. \end{aligned} \tag{5.3.3}$$

Much like the Feynman propagator for the real scalar field, the free fermion propagator is defined as

$$\begin{aligned} S_F(x-y)_{\alpha\beta} &\equiv i\langle 0|\mathrm{T}\hat{\boldsymbol{\Psi}}_\alpha(x)\hat{\overline{\boldsymbol{\Psi}}}_\beta(y)|0\rangle \\ &= i\theta(x^0-y^0)\langle 0|\hat{\boldsymbol{\Psi}}_\alpha(x)\hat{\overline{\boldsymbol{\Psi}}}_\beta(y)|0\rangle - i\theta(y^0-x^0)\langle 0|\hat{\overline{\boldsymbol{\Psi}}}_\beta(y)\hat{\boldsymbol{\Psi}}_\alpha(x)|0\rangle, \end{aligned} \tag{5.3.4}$$

where T is the time ordered product and α and β are four component spinor indices [5]. Note that unlike the Feynman propagator for the real scalar field, we have a negative sign in equation (5.3.4) since the fields must anticommute at spacelike separations [5]. Along with this, it is important to note that the free fermion propagator $S_F(x-y)_{\alpha\beta}$ is a 4×4 matrix and is also the Green's function for the Dirac equation [5]

$$(-i\partial\!\!\!/ + m)_{\alpha\beta} S_F(x-y)_{\alpha\gamma} = \delta^4(x-y)\delta_{\alpha\gamma}. \tag{5.3.5}$$

It will prove fruitful in future calculations to find the more compact, Fourier representation of the Feynman propagator of the free fermion field. To find the free fermion propagator's Fourier representation, we can compute the amplitudes in equation (5.3.4) by using the plane wave solution (4.2.44) of the Dirac field and combine the results to form the Fourier representation. This is a valid approach albeit a very algebraically tedious method. Let us instead avoid this approach and analyze how we were able to derive the general plane wave solution of the Dirac

field. Recall that multiplying the complex conjugate Dirac operator $(i\partial\!\!\!/ + m)$ to the Dirac equation yielded

$$\begin{aligned}(i\partial\!\!\!/ + m)\Big((-i\partial\!\!\!/ + m)\hat{\mathbf{\Psi}} = 0\Big)\\ (\partial\!\!\!/\partial\!\!\!/ + m^2)\hat{\mathbf{\Psi}} = 0\\ (-\partial^2 + m^2)\hat{\mathbf{\Psi}} = 0,\end{aligned}$$

namely that the Dirac field satisfied the Klein-Gordon equation [5]. Now the Feynman propagator $\Delta_F(x-y)$ for the real scalar field is the Green's function for the Klein-Gordon field thus it obeys the condition

$$(-\partial^2 + m^2)\Delta_F(x-y) = \delta^4(x-y). \tag{5.3.6}$$

Now observe that,

$$\begin{aligned}(-\partial^2 + m^2)\Delta_F(x-y) &= \delta^4(x-y)\\ (-i\partial\!\!\!/ + m)(i\partial\!\!\!/ + m)\Delta_F(x-y) &= \delta^4(x-y)\\ (-i\partial\!\!\!/ + m)\underbrace{\Big((i\partial\!\!\!/ + m)\Delta_F(x-y)\Big)}_{S_F(x-y)} &= \delta^4(x-y),\end{aligned}$$

thus the Feynman propagator for the free fermion field is

$$\begin{aligned}S_F(x-y) &= (i\partial\!\!\!/ + m)\Delta_F(x-y)\\ &= (i\partial\!\!\!/ + m)\int \frac{d^4p}{(2\pi)^4}\,\frac{e^{ip(x-y)}}{p^2+m^2-i\epsilon}\\ &= \int \frac{d^4p}{(2\pi)^4}\,\frac{(-p\!\!\!/ + m)}{p^2+m^2-i\epsilon}\,e^{ip(x-y)},\end{aligned} \tag{5.3.7}$$

or in component form, [5]

$$S_F(x-y)_{\alpha\beta} = \int \frac{d^4p}{(2\pi)^4}\,\frac{(-p\!\!\!/ + m)_{\alpha\beta}}{p^2+m^2-i\epsilon}\,e^{ip(x-y)}. \tag{5.3.8}$$

Notice that we have been referring to equation (5.3.8) as the free *fermion* propagator rather than the free Dirac propagator. As the reader may have deduced by now from our derivation, equation (5.3.8) is indeed also the Feynman propagator for the free Majorana field since the Majorana field, whose equation of motion is the Dirac equation as well, also satisfies the Klein-Gordon equation so

$$i\langle 0|\mathrm{T}\hat{\boldsymbol{\chi}}_\alpha(x)\hat{\overline{\boldsymbol{\chi}}}_\beta(y)|0\rangle = S_F(x-y)_{\alpha\beta} \tag{5.3.9}$$

$$= \int \frac{d^4p}{(2\pi)^4}\,\frac{(-\not{p}+m)_{\alpha\beta}}{p^2+m^2-i\epsilon}\,e^{ip(x-y)}. \tag{5.3.10}$$

When computing the time-ordered amplitude of other field operators, there are some subtle differences between the Dirac and Majorana fields. For example, we have that for the Dirac field,

$$\langle 0|\mathrm{T}\hat{\boldsymbol{\Psi}}_\alpha(x)\hat{\boldsymbol{\Psi}}_\beta(y)|0\rangle = 0, \tag{5.3.11}$$

$$\langle 0|\mathrm{T}\hat{\overline{\boldsymbol{\Psi}}}_\alpha(x)\hat{\overline{\boldsymbol{\Psi}}}_\beta(y)|0\rangle = 0, \tag{5.3.12}$$

which can be easily proved by using the plane wave solution of the Dirac field into equation (5.3.4) and noting that the ordering of the Fourier coefficient operators cancel with respect to the vacuum kets and bras [5]. For the Majorana field, the above equations are not correct. To see this, recall that the Majorana field obeys the Majorana condition $\hat{\overline{\boldsymbol{\chi}}} = \hat{\boldsymbol{\chi}}^{\mathrm{T}}\mathcal{C}$. This condition can be rewritten as $\hat{\boldsymbol{\chi}}^{\mathrm{T}} = \hat{\overline{\boldsymbol{\chi}}}\mathcal{C}^{-1}$ or $\hat{\overline{\boldsymbol{\chi}}}^{\mathrm{T}} = \mathcal{C}^{-1}\hat{\boldsymbol{\chi}}$ where we have used the property $\mathcal{C}^{-1} = \mathcal{C}^T$ [5]. Thus, equations (5.3.11) and (5.3.12) become [5]

$$\langle 0|\mathrm{T}\hat{\boldsymbol{\chi}}_\alpha(x)\hat{\boldsymbol{\chi}}_\beta(y)|0\rangle = \langle 0|\mathrm{T}\hat{\boldsymbol{\chi}}_\alpha(x)\hat{\overline{\boldsymbol{\chi}}}_\gamma(y)|0\rangle(\mathcal{C}^{-1})_{\gamma\beta}$$

$$= \frac{1}{i}[S_F(x-y)\mathcal{C}^{-1}]_{\alpha\beta}, \tag{5.3.13}$$

$$\langle 0|\mathrm{T}\hat{\overline{\boldsymbol{\chi}}}_\alpha(x)\hat{\overline{\boldsymbol{\chi}}}_\beta(y)|0\rangle = (\mathcal{C}^{-1})_{\alpha\gamma}\langle 0|\mathrm{T}\hat{\boldsymbol{\chi}}_\gamma(x)\hat{\overline{\boldsymbol{\chi}}}_\beta(y)|0\rangle$$

$$= \frac{1}{i}[\mathcal{C}^{-1}S_F(x-y)]_{\alpha\beta}. \tag{5.3.14}$$

Now we must impose more rules on the functional differentiation of anticommuting variables before we move on. To analyze this further, it is clear that the algebraic structure of anticommuting numbers (and variables) will have different properties. To elaborate further, given any arbitrary anticommuting numbers ψ_i that obey the condition

$$\{\psi_i, \psi_j\} = 0, \tag{5.3.15}$$

it is clear that

$$\psi_i \psi_j = -\psi_j \psi_i, \tag{5.3.16}$$

namely, switching the order of the numbers imparts a change in sign [5]. Even more interestingly, if we let $\psi_j = \psi_i$, then we have that $\psi^2 = 0$. Although we have only just scratched the surface in the metaphorical sense, numbers that obey the condition (5.3.15) clearly have very different properties than the usual commuting numbers that most people are familiar with. Indeed, numbers that obey the condition (5.3.15) are mathematical objects known as *Grassmann variables* [8]. With this notion in mind, we hypothesize that functionally differentiating the anticommuting terms in the fermion path integrals will yield a negative sign. To test this hypothesize, observe that

$$\begin{aligned}
\frac{\delta}{\delta\eta(x)} \int d^4y \left[\bar{\eta}\mathbf{\Psi}(y) + \overline{\mathbf{\Psi}}(y)\eta(y)\right] &= \int d^4y \, \frac{\delta}{\delta\eta(x)} \left[\overline{\mathbf{\Psi}}(y)\eta(y)\right] \\
&= \int d^4y \, \frac{\delta}{\delta\eta(x)} \left[-\eta(y)\overline{\mathbf{\Psi}}(y)\right] \\
&= \int d^4y - \delta^4(x-y)\mathbf{\Psi}(y) \\
&= -\mathbf{\Psi}(x),
\end{aligned}$$

and similarly,

$$\frac{\delta}{\delta\bar{\eta}(x)} \int d^4y \left[\bar{\eta}\mathbf{\Psi}(y) + \overline{\mathbf{\Psi}}(y)\eta(y)\right] = +\mathbf{\Psi}(x).$$

We see that our hypothesis is true thus whenever we functionally differentiate

Grassmann variables, we must apply the appropriate sign changes [5].

5.4 Electromagnetic Field

In this section, we derive the path integral of the electromagnetic field as well as the photon propagator. The path integral formalism shows its strength here because deriving the photon propagator using the interaction picture is quite difficult in practice but the path integral formulation allows us to derive the propagator rather easily.

Following what we have done in previous sections, the path integral for the electromagnetic field is

$$\begin{aligned} Z_0(J) \equiv \langle 0|0\rangle_J &= \int \mathcal{D}A\, e^{i\mathcal{S}_0} \\ &= \int \mathcal{D}A \exp\Big[i\int d^4x \Big(-\frac{1}{4}F^{\mu\nu}F_{\mu\nu} + J^\mu A_\mu\Big)\Big], \end{aligned} \tag{5.4.1}$$

where the functional measure $\mathcal{D}A \equiv \mathcal{D}A^0\mathcal{D}A^1\mathcal{D}A^2\mathcal{D}A^3$ [5]. Integrating by parts and imposing vanishing boundary conditions to eliminate the total divergence simplifies the path integral to [5]

$$Z_0(J) = \int \mathcal{D}A \exp\Big[i\int d^4x \Big(-\frac{1}{2}A_\mu(g^{\mu\nu}\partial^2 - \partial^\mu\partial^\nu)A_\nu + J^\mu A_\mu\Big)\Big]. \tag{5.4.2}$$

Let us focus on the action integral $\mathcal{S}_0$ and expand it in momentum-space

$$\begin{aligned} \mathcal{S}_0 = \frac{1}{2}\int \frac{d^4k}{(2\pi)^4}\Big[&-\tilde{A}_\mu(k)\big(k^2 g^{\mu\nu} - k^\mu k^\nu\big)\tilde{A}_\nu(-k) \\ &+ \tilde{J}^\mu(k)\tilde{A}_\mu(-k) + \tilde{J}^\mu(-k)\tilde{A}_\mu(k)\Big]. \end{aligned} \tag{5.4.3}$$

Ideally we would like to follow the same procedure for the derivation of the real scalar path integral and shift the field $\tilde{A}$ to complete the square and solve for the propagator as well. However, this is not possible because this requires that we

invert the matrix $\left(k^2 g^{\mu\nu} - k^\mu k^\nu\right)$ which is singular [5]. To see this, let

$$P^{\mu\nu}(k) = g^{\mu\nu} - \frac{k^\mu k^\nu}{k^2}, \tag{5.4.4}$$

then note that

$$\begin{aligned} P^{\mu\nu}k_\nu &= \big(g^{\mu\nu} - \frac{k^\mu k^\nu}{k^2}\big)k_\nu \\ &= k^\mu - k^\mu \\ &= 0. \end{aligned}$$

This means that one of the eigenvalues of the matrix $P^{\mu\nu}(k)$ is zero and this implies that det(P)=0 [5]. Actually, the matrix $P^{\mu\nu}(k)$ is a projection matrix which we will prove.

Proof. Let $P^{\mu\nu}(k)$ be defined as in equation (5.4.4). Then observe that [5]

$$\begin{aligned} P^{\mu\nu}(k)P_\nu{}^\lambda(k) &= \Big(g^{\mu\nu} - \frac{k^\mu k^\nu}{k^2}\Big)\Big(g_\nu{}^\lambda - \frac{k_\nu k^\lambda}{k^2}\Big) \\ &= g^{\mu\lambda} - g^{\mu\nu} - g^{\mu\nu}\frac{k_\nu k^\lambda}{k^2} - \frac{k^\mu k^\nu}{k^2}g_\nu{}^\lambda + \frac{k^\mu k^\nu}{k^2}\frac{k_\nu k^\lambda}{k^2} \\ &= g^{\mu\lambda} - \frac{k^\mu k^\lambda}{k^2} \\ &= P^{\mu\lambda}(k). \end{aligned}$$

Thus $P^{\mu\nu}(k)$ is a projection matrix. □

From elementary linear algebra, projection matrices have eigenvalues 0 and 1 and we can see this by taking the trace of (5.4.4),

$$\begin{aligned} g_{\mu\nu}P^{\mu\nu} &= g_{\mu\nu}g^{\mu\nu} - g_{\mu\nu}\frac{k^\mu k^\nu}{k^2} \\ &= 4 - 1 \\ &= 3, \end{aligned}$$

so the other three eigenvalues of $P^{\mu\nu}(k)$ are 1 [5]. Since the matrix $\left(k^2 g^{\mu\nu} - k^\mu k^\nu\right)$

is singular, the photon propagator $D_F^{\nu\rho}(x-y)$ has no solution because we must invert the operator

$$\left(k^2 g^{\mu\nu} - k^\mu k^\nu\right)\tilde{D}_F^{\nu\rho}(k) = \delta_\mu{}^\rho, \tag{5.4.5}$$

which is not possible thus it cannot be solved [5].

The root of these problems ultimately comes from gauge invariance. To fix this problem, we will utilize a clever trick. We follow the derivation according to Peskin and Schroeder [8]. We introduce a *gauge-fixing function* $G(A^\alpha)$ defined as

$$G(A^\alpha) = \partial^\mu A_\mu^\alpha(x) - \omega^\alpha(x), \tag{5.4.6}$$

where $\omega^\alpha(x)$ is an arbitrary scalar function and $A_\mu^a(x)$ is the gauge transformed gauge field defined as

$$A_\mu^a(x) = A_\mu(x) - \partial_\mu \alpha(x). \tag{5.4.7}$$

Before we continue, recall the property of the delta function,

$$1 = \int dy\, \delta(y). \tag{5.4.8}$$

This result does not change if we shift the argument of the delta function so let us shift the argument by $y \to y - f(x)$ [8]. Along with this, let $g(x,y)$ be an arbitrary function then we use the delta function identity

$$\delta(G(x,y)) = \frac{\delta(y-f(x))}{|\partial g/\partial y|}, \tag{5.4.9}$$

to turn equation (5.4.8) into

$$1 = \int dy\, \frac{\partial g}{\partial y}\, \delta(g(x,y)), \tag{5.4.10}$$

where we have assumed that $\partial g/\partial y$ is positive when $y = f(x)$. The generalization

of equation (5.4.10) for n-dimensions is then

$$1 = \int d^n y \det\Big(\frac{\partial g_i}{\partial y_j}\Big) \prod_i \delta(g_i). \tag{5.4.11}$$

With this rather complicated form of the identity property of the delta function, we generalize it to the functional form for $G(A^\alpha)$, [8]

$$1 = \int \mathcal{D}\alpha \det\Big(\frac{\delta G(A^\alpha)}{\delta \alpha}\Big) \delta(G(A^\alpha)). \tag{5.4.12}$$

We insert this functional identity into the path integral for the photon

$$\begin{aligned} Z_0(J) &= \int \mathcal{D}A\, e^{i\mathcal{S}_0} \\ &= \int \mathcal{D}\alpha \int \mathcal{D}A \det\Big(\frac{\delta G(A^\alpha)}{\delta \alpha}\Big) \delta(G(A^\alpha))\, e^{i\mathcal{S}_0} \\ &= \det\Big(\frac{\delta G(A^\alpha)}{\delta \alpha}\Big)\Big(\int \mathcal{D}\alpha\Big) \int \mathcal{D}A\, \delta(G(A^\alpha))\, e^{i\mathcal{S}_0}, \end{aligned}$$

where the functional determinant is pulled out since it is a constant [5]. Note that the term $\int \mathcal{D}\alpha$ is a constant as well. Let us treat A_μ as the integration variable and change variables $A_\mu \to A^\alpha_\mu$ which leads to $\mathcal{D}A \to \mathcal{D}A^\alpha$ and $\mathcal{S}_0[A] \to \mathcal{S}_0[A^\alpha]$. Recall that A^α is a gauge transformed field so in terms in functional calculus, we have merely shifted the original field A_μ by a constant thus the functional measure is invariant under this change of integration variables, namely, $\mathcal{D}A^\alpha = \mathcal{D}A$. Along with this, we constructed the electromagnetic Lagrangian density $\mathcal{S}_0$ to be gauge invariant so $\mathcal{S}_0[A^\alpha] = \mathcal{S}_0[A]$. Since A^α is arbitrary, we relabel it as A then the path integral is [8]

$$\begin{aligned} Z_0(J) &= \det\Big(\frac{\delta G(A^\alpha)}{\delta \alpha}\Big)\Big(\int \mathcal{D}\alpha\Big) \int \mathcal{D}A\, \delta(G(A))\, e^{i\mathcal{S}_0} \\ &= \det\Big(\frac{\delta G(A^\alpha)}{\delta \alpha}\Big)\Big(\int \mathcal{D}\alpha\Big) \int \mathcal{D}A\, \delta(\partial^\mu A_\mu - \omega(x))\, e^{i\mathcal{S}_0}. \end{aligned}$$

The above equality holds for any arbitrary scalar function $\omega(x)$ and since the path

integral is independent of $\omega(x)$, we multiply the path integral by

$$N(\xi)\exp\Big[-i\int d^4x\,\frac{\omega^2}{2\xi}\Big], \tag{5.4.13}$$

which is the functional generalization of the Gaussian function centered at $\omega = 0$ and $N(\xi)$ is the normalization constant [8]. Then we integrate over the functional measure $\mathcal{D}\omega$,

$$\begin{aligned} Z_0(J) &= \int \mathcal{D}\omega\, N(\xi)\exp\Big[-i\int d^4x\,\frac{\omega^2}{2\xi}\Big] \\ &\quad\times \det\Big(\frac{\delta G(A^\alpha)}{\delta\alpha}\Big)\Big(\int\mathcal{D}\alpha\Big)\int\mathcal{D}A\,\delta(\partial^\mu A_\mu-\omega(x))\,e^{i\mathcal{S}_0} \\ &\propto \int\mathcal{D}A\exp\Big[-i\int d^4x\,\frac{(\partial^\mu A_\mu)^2}{2\xi}\Big]e^{i\mathcal{S}_0} \\ &\propto \int\mathcal{D}A\,e^{i\mathcal{S}_0+i\mathcal{S}_{gf}}, \end{aligned}$$

where

$$\mathcal{S}_{gf} = \int d^4x\,\mathcal{L}_{gf}, \tag{5.4.14}$$

$$\mathcal{L}_{gf} = -\frac{1}{2\xi}\big(\partial^\mu A_\mu\big)^2, \tag{5.4.15}$$

are the *gauge-fixed action* and *gauge-fixed Lagrangian*, respectively [5]. Note that the proportionality factor is due to multiplying the Gaussian function to Z_0 and integrating it which simply multiplies a constant to the original path integral $Z_0(J)$. This constant is removed when we normalize $Z_0(J)$. After properly normalizing the path integral, we get

$$\begin{aligned} Z_0(J) &= \int\mathcal{D}A\,e^{i\mathcal{S}_0+i\mathcal{S}_{gf}} \\ &= \int\mathcal{D}A\exp\Big[i\int d^4x\,\Big(-\frac{1}{4}F^{\mu\nu}F_{\mu\nu}+J^\mu A_\mu-\frac{1}{2\xi}\big(\partial^\mu A_\mu\big)^2\Big)\Big], \end{aligned} \tag{5.4.16}$$

and in momentum-space the action term is [8]

$$\begin{aligned}\mathcal{S} &= \int d^4x\ -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} - \frac{1}{2\xi}\big(\partial^\mu A_\mu\big)^2 \\ &= \frac{1}{2}\int \frac{d^4k}{(2\pi)^4}\Big[-\tilde{A}_\mu(k)\Big(k^2 g_{\mu\nu} - \big(1-\frac{1}{\xi}\big)\frac{k_\mu k_\nu}{k^2}\Big)\tilde{A}_\mu(-k) \\ &+ \tilde{J}^\mu(k)\tilde{A}_\mu(-k) + \tilde{J}^\mu(-k)\tilde{A}_\mu(k)\Big]. \end{aligned} \tag{5.4.17}$$

Finally the matrix $\Big(k^2 g_{\mu\nu} - \big(1-\frac{1}{\xi}\big)\frac{k_\mu k_\nu}{k^2}\Big)$ is nonsingular thus we can solve for the propagator

$$\Big(k^2 g_{\mu\nu} - \big(1-\frac{1}{\xi}\big)\frac{k_\mu k_\nu}{k^2}\Big)\tilde{D}^{\nu\rho}_{F,\xi}(k) = \delta_\mu{}^\rho,$$

which yields [8]

$$\tilde{D}^{\mu\nu}_{F,\xi}(k) = \frac{1}{k^2 - i\epsilon}\Big(g^{\mu\nu} - \big(1-\xi\big)\frac{k^\mu k^\nu}{k^2}\Big). \tag{5.4.18}$$

The free photon propagator $\tilde{D}^{\mu\nu}_{F,\xi}(k)$ is said to be in the *generalized Feynman gauge* or R_ξ *gauge* [5]. Choosing $\xi = 0$ leads to the *Lorenz gauge* (or *Landau gauge*) and choosing $\xi = 1$ leads to the *Feynman gauge* [8]. With the photon propagator, we can now shift the field by a constant and follow the same procedure as the real scalar field then the path integral for the photon is [5]

$$Z_0(J) = \exp\Big[\frac{i}{2}\int d^4x\, d^4y\, J_\mu(x) D^{\mu\nu}_{F,\xi}(x-y) J_\nu(y)\Big], \tag{5.4.19}$$

where

$$D^{\mu\nu}_{F,\xi}(x-y) = \int \frac{d^4k}{(2\pi)^4}\, e^{ik(x-y)}\ \frac{g^{\mu\nu} - \big(1-\xi\big)\frac{k^\mu k^\nu}{k^2}}{k^2 - i\epsilon}. \tag{5.4.20}$$

The choice of gauge will depend on the problem and we will see later on which gauge to choose. Notice that this form of the propagator is rather useful due to its generality since we did not have to quantize the electromagnetic field again in the Lorenz gauge to determine the Lorenz gauge propagator.

Chapter 6

Scattering

6.1 LSZ Reduction Formula

Let us use the real scalar field to calculate the probability for a scattering event to occur. We follow the derivation according to Srednicki [5]. In the Schrödinger picture, let $|i\rangle \equiv |i, \text{in}\rangle = |\mathbf{p}_1 \mathbf{p}_2 \dots \mathbf{p}_n; t_i\rangle$ be the initial state ket of the incoming particles with and let $|f\rangle \equiv |f, \text{out}\rangle = |\mathbf{p}_{1'} \mathbf{p}_{2'} \dots \mathbf{p}_{n'}; t_f\rangle$ be the final state ket of the outgoing particles where $\mathbf{p}_i \neq \mathbf{p}_j$ when $i \neq j$. If we consider the distant, or asymptotic, past $t \to -\infty$, the interacting theory is expected to reduce to the free theory which makes sense physically because the incoming particles will be far apart from each other so that they cannot interact [11]. We expect the incoming particles in the distant past to then behave as free particles that are localized near $\mathbf{p}$ and the origin in momentum- and position-space, respectively [5]. So the creation operator $\hat{a}^\dagger(\mathbf{p})$ is expected to be time-independent in the distant past and is expected to create a particle with a wave packet wavefunction. Now recall from the fields chapter that the creation and annihilation operators can be written as

$$\hat{a}^\dagger(\mathbf{p}) = -i \int d^3x \, e^{ipx} \overleftrightarrow{\partial_0} \hat{\varphi}(x), \tag{6.1.1}$$

$$\hat{a}(\mathbf{p}) = i \int d^3x \, e^{-ipx} \overleftrightarrow{\partial_0} \hat{\varphi}(x), \tag{6.1.2}$$

respectively. Let us move to the Heisenberg picture and make our operators time-dependent and localized in momentum-space

$$\hat{a}_i^\dagger(t) = \int d^3p\, f_i(\mathbf{p})\hat{a}^\dagger(\mathbf{p}), \tag{6.1.3}$$

$$f_i(\mathbf{p}_i) \propto \exp\Big[\frac{-(\mathbf{p}-\mathbf{p}_i)^2}{4\sigma^2}\Big], \tag{6.1.4}$$

where $f_i(\mathbf{p}_i)$ is the Gaussian wave packet with its width σ in momentum-space [5]. Now in the distant, or asymptotic, past $t \to -\infty$, the initial state can be described as a wave packet with a well defined momentum

$$|i\rangle = \lim_{t\to-\infty} \hat{a}_1^\dagger(t)\hat{a}_2^\dagger(t)\ldots\hat{a}_n'(t)|0\rangle, \tag{6.1.5}$$

and similarly in the far, or asymptotic, future $t \to +\infty$, the final state is

$$|f\rangle = \lim_{t\to+\infty} \hat{a}_1^\dagger(t)\hat{a}_2^\dagger(t)\ldots\hat{a}_n'(t)|0\rangle. \tag{6.1.6}$$

Using the postulates of quantum mechanics and the fact that we can build multi-particle states from the vacuum state $|0\rangle$ by acting on the vacuum state with the appropriate creation operators $\hat{a}_i(\mathbf{p}_i)$, the *scattering amplitude* for the scattering event is

$$\langle f|i\rangle \equiv \langle \mathbf{p}_{1'}\mathbf{p}_{2'}\ldots\mathbf{p}_{n'}|\mathbf{p}_1\mathbf{p}_2\ldots\mathbf{p}_n\rangle, \tag{6.1.7}$$

where p_i and p_i' represent the initial and final momenta of the n particle states. The probability for the scattering process to occur is

$$|\langle f|i\rangle|^2 = |\langle p_{1'}p_{2'}\ldots p_{n'}|p_1p_2\ldots p_n\rangle|^2\,. \tag{6.1.8}$$

This makes physical sense because in a real scattering experiment, the atoms are prepared in the momentum eigenstate which is clearly measured beforehand. Then the particles are sped up and eventually they collide with one another and the detector measures the cross section and the final momenta of the particles which

means that the final state of the particles is once again a momentum eigenstate [8]. The particles in the initial and final states essentially behave as free particles localized in momentum-space since at these two asymptotic times, the particles are not interacting with each other or with any potential that might influence each particle's wavefunction. We know that between these two asymptotic times, the particles will interact with each other, namely, scatter off each other which results in their wavepackets overlapping [5]. Now observe that for an arbitrary creation operator $\hat{a}_i^\dagger(t)$

$$
\begin{aligned}
\big(\lim_{t\to+\infty} - \lim_{t\to-\infty}\big)\hat{a}_i^\dagger(t) &= \int_{-\infty}^{+\infty} dt\,\partial_0\,\hat{a}_i^\dagger(t) \\
&= \int_{-\infty}^{+\infty} dt\, f_i(\mathbf{p}_i)\int d^3p\,\partial_0\big(\hat{a}^\dagger(\mathbf{p})\big) \\
&= -i\int d^3p\, f_i(\mathbf{p}_i)\int d^4x\,\partial_0\big(e^{ipx}\overleftrightarrow{\partial_0}\,\hat{\varphi}(x)\big) \\
&= -i\int d^3p\, f_i(\mathbf{p}_i)\int d^4x\, e^{ipx}(\partial_0^2+\omega^2)\hat{\varphi}(x) \\
&= -i\int d^3p\, f_i(\mathbf{p}_i)\int d^4x\, e^{ipx}(\partial_0^2+\mathbf{p}^2+m^2)\hat{\varphi}(x) \\
&= -i\int d^3p\, f_i(\mathbf{p}_i)\int d^4x\, e^{ipx}(\partial_0^2+\overleftarrow{\nabla}^2+m^2)\hat{\varphi}(x) \\
&= -i\int d^3p\, f_i(\mathbf{p}_i)\int d^4x\, e^{ipx}(\partial_0^2+\overrightarrow{\nabla}^2+m^2)\hat{\varphi}(x) \\
&= -i\int d^3p\, f_i(\mathbf{p}_i)\int d^4x\, e^{ipx}(-\partial^2+m^2)\hat{\varphi}(x), \qquad (6.1.9)
\end{aligned}
$$

where we used the expansion of the Fourier coefficient operator in terms of the field operator which was derived in the fields chapter and integrated by parts to shift the Laplacian to act on the right [5]. In a free theory, the last term above equals zero since the equation of motion is the Klein-Gordon equation but we want an interacting theory so the Klein-Gordon equation will equal some interaction term from the interaction part of the Lagrangian which makes the above expression

nonzero. Then we have

$$\lim_{t\to-\infty}\hat{a}_i^\dagger(t)=\lim_{t\to+\infty}\hat{a}_i^\dagger(t)+i\int d^3k\,f_i(\mathbf{p}_i)\int d^4x\,e^{ikx}(-\partial^2+m^2)\hat{\varphi}(x), \tag{6.1.10}$$

$$\lim_{t\to+\infty}\hat{a}_i^\dagger(t)=\lim_{t\to-\infty}\hat{a}_i^\dagger(t)+i\int d^3k\,f_i(\mathbf{p}_i)\int d^4x\,e^{-ikx},(-\partial^2+m^2)\hat{\varphi}(x). \tag{6.1.11}$$

and the scattering amplitude becomes

$$\langle f|i\rangle=\lim_{\substack{t_i\to-\infty\\ t_f\to+\infty}}\langle 0|\mathrm{T}\hat{a}_{1'}(t_f)\dots\hat{a}_{n'}(t_f)\hat{a}'_1(t_i)\dots\hat{a}'_n(t_i)|0\rangle. \tag{6.1.12}$$

When the n particles come within vicinity of each other so that their wavefunctions can interact, the wavepackets lose their significance since the state of each particle changes so we let $\sigma\to\infty$ which then results in the Gaussian wavepacket to become delta function normalized $f_i(\mathbf{p})=\delta^3(\mathbf{p}-\mathbf{p}_i)$ [5]. Thus we finally get

$$\begin{aligned}\langle f|i\rangle=i^{n+n'}\int d^4x_1\,e^{ip_1x_1}(-\partial_1^2+m^2)\dots\\ d^4x'_1e^{-ip'_1x_1}(-\partial_{1'}^2+m^2)\dots\\ \times\langle 0|\mathrm{T}\hat{\varphi}(x_1)\dots\hat{\varphi}(x'_1)\dots|0\rangle,\end{aligned} \tag{6.1.13}$$

which is called the *Lehmann-Symanzik-Zimmermann reduction formula*, or LSZ formula [5]. Although we derived the LSZ formula under the assumption the free field operators would work in a noninteracting theory, the LSZ formula that we derived is valid for the interacting theory as well provided that the field operator obeys the following normalization conditions

$$\langle 0|\hat{\varphi}(x)|0\rangle=0, \tag{6.1.14}$$

$$\langle p|\hat{\varphi}(x)|0\rangle=e^{-ipx}, \tag{6.1.15}$$

since these conditions ensure that the field operator creates mostly one particle states and the contributions from multiparticle states can be made as small as we like by waiting long enough for the multiparticle wave packet to go to zero [5].

We can use the exact same arguments to derive the LSZ reduction formula for the fermion and electromagnetic fields as well. For fermion fields, both the Dirac field and Majorana field share the same LSZ reduction formula which is given by [5]

$$\begin{aligned}\langle f|i\rangle = i^{n+n'} \int d^4x_1\, d^4x_2 \ldots d^4x_{1'}\, d^4x_{2'} \ldots \\ \times e^{-ip'_1x'_1}[\bar{u}_{s_{1'}}(\mathbf{p}_{1'})(-i\partial\!\!\!/_{1'} + m)]_{\alpha_{1'}} \\ \times e^{-ip'_2x'_2}[\bar{u}_{s_{2'}}(\mathbf{p}_{2'})(-i\partial\!\!\!/_{2'} + m)]_{\alpha_{2'}} \ldots \\ \times \langle 0|\mathrm{T}\hat{\mathbf{\Psi}}_{\alpha_{2'}}(x_{2'})\hat{\mathbf{\Psi}}_{\alpha_{1'}}(x_{1'}) \ldots \hat{\overline{\mathbf{\Psi}}}_{\alpha_1}(x_1)\hat{\overline{\mathbf{\Psi}}}_{\alpha_2}(x_2) \ldots |0\rangle \\ \times [(i\overleftarrow{\partial}_1 + m)u_{s_1}(\mathbf{p}_1)]_{\alpha_1} e^{ip_1x_1} \\ \times [(i\overleftarrow{\partial}_2 + m)u_{s_2}(\mathbf{p}_2)]_{\alpha_2} e^{ip_2x_2} \ldots . \end{aligned} \tag{6.1.16}$$

The LSZ reduction formula will hold for a Dirac field given that we impose the following conditions

$$\begin{aligned}\langle 0|\hat{\mathbf{\Psi}}(x)|0\rangle &= 0, \\ \langle p,s,+|\hat{\mathbf{\Psi}}(x)|0\rangle &= 0, \\ \langle p,s,-|\hat{\mathbf{\Psi}}(x)|0\rangle &= v_s(\mathbf{p})e^{-ipx}, \\ \langle p,s,+|\hat{\overline{\mathbf{\Psi}}}(x)|0\rangle &= \bar{u}_s(\mathbf{p})e^{-ipx}, \\ \langle p,s,-|\hat{\overline{\mathbf{\Psi}}}(x)|0\rangle &= 0, \end{aligned} \tag{6.1.17}$$

while for a Majorana field,

$$\begin{aligned}\langle 0|\hat{\mathbf{\chi}}(x)|0\rangle &= 0, \\ \langle p,s|\hat{\mathbf{\chi}}(x)|0\rangle &= v_s(p)e^{-ipx}, \\ \langle p,s|\hat{\overline{\mathbf{\chi}}}(x)|0\rangle &= \bar{u}_s(p)e^{-ipx}, \end{aligned} \tag{6.1.18}$$

and as usual the single particle states obey their Lorentz invariant normalization stated in the fields chapter and $\langle 0|0\rangle = 1$ [5].

For the electromagnetic field, to work directly with four-vectors instead of the three-vectors we derived for the creation and annihilation operators, we can rewrite the incoming photon creation operator and outgoing photon annihilation operator as

$$\hat{a}_\lambda^\dagger(\mathbf{k})_{\text{in}} = i\varepsilon_\lambda^{\mu *}(\mathbf{k}) \int d^4x\, e^{ikx}(-\partial^2)\hat{A}_\mu(x), \tag{6.1.19}$$

$$\hat{a}_\lambda(\mathbf{k})_{\text{out}} = i\varepsilon_\lambda^{\mu}(\mathbf{k}) \int d^4x\, e^{-ikx}(-\partial^2)\hat{A}_\mu(x), \tag{6.1.20}$$

where $\varepsilon_\lambda^0(\mathbf{k}) = 0$ [5]. Then the LSZ reduction formula for photons is [5]

$$\begin{aligned}\langle f|i\rangle = {} & i^{n+n'}\varepsilon_\lambda^{\mu_1 *}(\mathbf{k})\varepsilon_\lambda^{\mu_2 *}(\mathbf{k})\ldots\varepsilon_\lambda^{\mu_1'}(\mathbf{k})\varepsilon_\lambda^{\mu_2'}(\mathbf{k})\ldots \\ & \int d^4x_1\, d^4x_2 \ldots d^4x_{1'}\, d^4x_{2'}\ldots \\ & \times e^{-ik_1'x_1'}(-\partial_{1'}^2)e^{-ik_2'x_2'}(-\partial_{2'}^2)\ldots \\ & \times e^{-ik_1x_1}(-\partial_1^2)e^{-ik_2x_2}(-\partial_2^2)\ldots \\ & \times \langle 0|\mathrm{T}\hat{A}_{\mu_1'}(x_1')\hat{A}_{\mu_2'}(x_2')\ldots\hat{A}_{\mu_2}(x_2)\hat{A}_{\mu_2}(x_2)\ldots|0\rangle.\end{aligned} \tag{6.1.21}$$

The LSZ reduction formula is valid provided that we impose the following conditions

$$\begin{aligned}\langle 0|\hat{A}^i(x)|0\rangle &= 0, \\ \langle k,\lambda|\hat{A}^i(x)|0\rangle &= \varepsilon_\lambda^i(\mathbf{k})e^{-ikx},\end{aligned} \tag{6.1.22}$$

along with the Lorentz invariant single photon state normalization [5].

Now from nonrelativistic quantum mechanics, we know that the scattering amplitude will give us a general term of the form

$$\langle f|i\rangle = (2\pi)^3\delta^3(\mathbf{p}_{\text{in}} - \mathbf{p}_{\text{out}})i\mathcal{T}, \tag{6.1.23}$$

where the three-dimensional delta function represents momentum conservation and $\mathcal{T}$ is the scattering matrix [11]. We expect this to be the case for quantum field

theory as well and indeed later on, we will see that this is the case. One could also note that up to now, we have the necessary tools and knowledge needed to compute the vacuum expectation values and in doing so will yield a result similar in appearance to equation (6.1.23). Since we are working in Minkowski space, we anticipate that the scattering amplitude in quantum field theory will be of the form

$$\langle f|i\rangle = (2\pi)^4\delta^4(\mathbf{p}_{\text{in}} - \mathbf{p}_{\text{out}})i\mathcal{T}, \tag{6.1.24}$$

where the four-dimensional delta function represents the conservation of four-momentum. We now wish to relate the scattering matrix $\mathcal{T}$ to the differential cross section so that we are able to derive results that pertain to real life results.

6.2 Differential Cross Section

We now turn to more concrete ideas and seek to apply everything we have developed up to now to real life applications. In particle physics experiments, particle accelerators are used to speed up atoms to ultrarelativsitc speeds and make them collide with each other which results in the atoms exploding into their elementary components for a fraction of a second [1]. The quantities that are measured in these experiments are the cross sections σ and decay rates Γ. Let us derive the differential cross section formula [8].

We follow the derivation according to Srednicki [8]. In scattering experiments, the two most commonly used frames are the *center-of-mass frame*, or CM frame, and the *fixed target frame* (*lab frame*) or FT frame [8]. Let us consider the case of two incoming particles and two outgoing particles. In the CM frame, the total momentum of the incoming particles is zero $\mathbf{p}_1 + \mathbf{p}_2 = 0$ with $\mathbf{p}_1$ chosen in the z-direction. Then we have that the total energy of the incoming particles in the CM frame is just $E_1 + E_2 = -(p_1 + p_2)^2$. For notational brevity, we let $s = -(p_1 + p_2)^2$ which is a Lorentz scalar that is equivalent to $s = (E_1 + E_2)^2$ in the CM frame [5]. By the conservation of momentum and energy, the total momentum and energy of

the outgoing particles are $\mathbf{p}_1' + \mathbf{p}_2' = 0$ and $(E_1' + E_2')^2 = s$. To specify the angle θ between $\mathbf{p}_1$ and $\mathbf{p}_1'$, let us define the Lorentz scalar $t = -(p_1 - p_1')^2$ which is equal to the frame independent [5]

$$t = m_1^2 + m_{1'}^2 - 2E_1E_1' + 2\,|\mathbf{p}_1|\,|\mathbf{p}_1'|\cos\theta. \tag{6.2.1}$$

The Lorentz scalars

$$s = -(p_1 + p_2)^2 = -(p_1' + p_2')^2, \tag{6.2.2}$$

$$t = -(p_1 - p_1')^2 = -(p_2 - p_2')^2, \tag{6.2.3}$$

$$u = -(p_1 - p_2')^2 = -(p_2 - p_1')^2, \tag{6.2.4}$$

are the *Mandelstam variables* that are related by the relation [5]

$$s + t + u = m_1^2 + m_2^2 + m_{1'}^2 + m_{2'}^2. \tag{6.2.5}$$

In the CM frame, we see that

$$|\mathbf{p}_1| = \frac{1}{2\sqrt{s}}\sqrt{s^2 - 2(m_1^2 + m_2^2)s + (m_1^2 - m_2^2)^2}, \tag{6.2.6}$$

$$|\mathbf{p}_1'| = \frac{1}{2\sqrt{s}}\sqrt{s^2 - 2(m_{1'}^2 + m_{2'}^2)s + (m_{1'}^2 - m_{2'}^2)^2}. \tag{6.2.7}$$

In the FT frame, a particle, say particle 2, is set at rest which means that $\mathbf{p}_2 = 0$ then we see that

$$|\mathbf{p}_1| = \frac{1}{2m_2}\sqrt{s^2 - 2(m_1^2 + m_2^2)s + (m_1^2 - m_2^2)^2} \tag{6.2.8}$$

and [5]

$$m_2\,|\mathbf{p}_1|_{\mathrm{FT}} = \sqrt{s}\,|\mathbf{p}_1|_{\mathrm{CM}}\,. \tag{6.2.9}$$

The Mandelstam variables and these identities will prove to be extremely helpful in reducing cross section formulas and transitioning between frames as we will see

in the subsequent sections and chapters.

From the previous section, we found that the scattering amplitude was given as, in its most general form,

$$\langle f|i\rangle = (2\pi)^4\delta^4(p_{\text{in}} - p_{\text{out}})i\mathcal{T}. \tag{6.2.10}$$

Using the postulates of quantum mechanics, the probability for the initial state $|i\rangle$ to transition to the final state $|f\rangle$ is

$$P(i \to f) = \frac{|\langle f|i\rangle|^2}{\langle f|f\rangle\langle i|i\rangle}. \tag{6.2.11}$$

Much like scattering theory in nonrelativistic quantum mechanics, we wish to compute the probability rate $\dot{P}$ [11]. We will assume that the experiment takes place in a big box of volume V for a large time T [5]. Then the volume of the box is

$$(2\pi)^3\delta^3(0) = V,$$

and we see that

$$(2\pi)^4\delta^4(0) = VT.$$

Then the norm of the single particle states implies that the inner product of the initial and final states are

$$\langle i|i\rangle = 4E_1E_2V^2, \tag{6.2.12}$$

$$\langle f|f\rangle = \prod_{j=1}^{n'} 2p_j'^0 V, \tag{6.2.13}$$

where n' is the number of outgoing particles [5]. Then

$$\begin{aligned}\langle f|i\rangle &= \left[(2\pi)^4\delta^4(p_{\text{in}}-p_{\text{out}})\right]^2|\mathcal{T}|^2\\ &= (2\pi)^4\delta^4(p_{\text{in}}-p_{\text{out}})(2\pi)^4\delta^4(0)\\ &= (2\pi)^4\delta^4(p_{\text{in}}-p_{\text{out}})VT. \end{aligned} \tag{6.2.14}$$

Then the probability rate is

$$\dot{P} = \frac{P}{T} = \frac{(2\pi)^4\delta^4(p_{\text{in}}-p_{\text{out}})V|\mathcal{T}|^2}{4E_1E_2V^2\prod_{j=1}^{n'}2p_j'^0V}. \tag{6.2.15}$$

We sum this expression over all possible discrete momenta of the final particles and in the limit of large V, we get

$$\sum_{\mathbf{n}_j'} \to \frac{V}{(2\pi)^3}\int d^3p_j', \tag{6.2.16}$$

thus

$$\dot{P} = \frac{(2\pi)^4\delta^4(p_{\text{in}}-p_{\text{out}})|\mathcal{T}|^2}{4E_1E_2V}\prod_{j=1}^{n'}d^3p_{L\,j}'. \tag{6.2.17}$$

We divide $\dot{P}$ by the incident flux to get the differential cross section $d\sigma$ which in the FT frame is $|\mathbf{p}_1|/E_1V$ [5]. In the FT frame, we also have that $E_2 = m_2$ and dividing by the incident flux will yield the term $|p_1|\,m_2$ in the denominator which we can convert to $\sqrt{s}\,|\mathbf{p}_1|_{\text{CM}}$ so we get

$$\begin{aligned} d\sigma &= \frac{\dot{P}}{|\mathbf{p}_1|/E_1V}\\ &= \frac{(2\pi)^4\delta^4(p_{\text{in}}-p_{\text{out}})|\mathcal{T}|^2}{4E_2|\mathbf{p}_1|}\prod_{j=1}^{n'}d^3p_{L\,j}'\\ &= \frac{(2\pi)^4\delta^4(p_{\text{in}}-p_{\text{out}})|\mathcal{T}|^2}{4|\mathbf{p}_1|_{\text{CM}}\sqrt{s}}\prod_{j=1}^{n'}d^3p_{L\,j}'\\ &= \frac{1}{4|\mathbf{p}_1|_{\text{CM}}\sqrt{s}}|\mathcal{T}|^2\,d\text{LIPS}_{n'}(p_1+p_2), \end{aligned} \tag{6.2.18}$$

where $d\text{LIPS}_{n'}(p_1+p_2)$ is the n'-body Lorentz invariant phase space measure

$$d\text{LIPS}_{n'}(p) = (2\pi)^4\delta^4\Big(p - \sum_{i=1}^{n'} p_i'\Big)\prod_{j=1}^{n'} d^3p_{L\,j}', \tag{6.2.19}$$

where $p = \sum_{i=1}^{n} p_i$ [5]. Let us consider two outgoing particles so we must evaluate

$$d\text{LIPS}_2(p) = (2\pi)^4\delta^4(p - p_1' - p_2')\, d^3p_{L\,1}'\, d^3p_{L\,2}', \tag{6.2.20}$$

where $p = p_1 + p_2$. In the CM frame, $\mathbf{p} = \mathbf{p}_1 + \mathbf{p}_2 = \mathbf{0}$ and $p^0 = \sqrt{s}$ thus

$$\begin{aligned} d\text{LIPS}_2(p) &= (2\pi)^4\delta^4(p - p_1' - p_2')\, d^3p_{L\,1}'\, d^3p_{L\,2}' \\ &= \frac{1}{16\pi^2 E_1' E_2'}\delta^3(\mathbf{p}_1' + \mathbf{p}_2')\delta(E_1' + E_2' - \sqrt{s})\, d^3p_1'\, d^3p_2' \\ &= \frac{1}{16\pi^2 E_1' E_2'}\delta(E_1' + E_2' - \sqrt{s})\, d^3p_1', \end{aligned} \tag{6.2.21}$$

where we have integrated over the spatial delta function which yields $E_1' = \sqrt{\mathbf{p}_1^2 + m_{1'}^2}$ and $E_2' = \sqrt{\mathbf{p}_1^2 + m_{2'}^2}$ [5]. We now write $d^3p_1' = |\mathbf{p}_1'|^2\, d\,|\mathbf{p}_1'|\, d\Omega_{\text{CM}}$ where

$$d\Omega_{\text{CM}} = \sin\theta d\theta d\phi,$$

is the differential solid angle with the angle θ being between $\mathbf{p}_1'$ and $\mathbf{p}_1$ [5]. We use the delta function identity

$$\int dx\, \delta(f(x)) = \sum_i |f'(x_i)|^{-1}, \tag{6.2.22}$$

where x_i obeys $f(x_i) = 0$ to integrate equation (6.2.21) over the last differential to finally get

$$d\text{LIPS}_2(k) = \frac{|\mathbf{k}_1'|}{16\pi^2\sqrt{s}\, d\Omega_{\text{CM}}}, \tag{6.2.23}$$

which yields the differential cross section in the CM frame

$$\frac{d\sigma}{d\Omega_{\text{CM}}} = \frac{1}{64\pi^2\sqrt{s}}\frac{|\mathbf{p}_1'|}{|\mathbf{p}_1|}\,|\mathcal{T}|^2\,. \tag{6.2.24}$$

If we take the differential of the Mandelstam variable t at fixed s, we get

$$\begin{aligned} dt &= 2\left|\mathbf{p}_1\right|\left|\mathbf{p}_1'\right| d\cos\theta \\ &= 2\left|\mathbf{p}_1\right|\left|\mathbf{p}_1'\right| \frac{d\Omega_{\mathrm{CM}}}{2\pi}, \end{aligned} \tag{6.2.25}$$

then we get a frame-independent differential cross section [5]

$$\frac{d\sigma}{dt} = \frac{1}{64\pi s\left|\mathbf{p}_1\right|^2}\left|\mathcal{T}\right|^2. \tag{6.2.26}$$

To get the differential decay rate $d\Gamma$ for n' outgoing particles, we repeat the same derivations for the differential cross section except with $\langle i|i\rangle = 2E_1V$ which yields

$$d\Gamma = \frac{1}{2E_1}\left|\mathcal{T}\right|^2 d\mathrm{LIPS}_{n'}(p_1), \tag{6.2.27}$$

where $s = -p_1^2 = m_1^2$ [5]. To compute the total cross section and total decay rate, we compute

$$\sigma = \frac{1}{S}\int d\sigma, \tag{6.2.28}$$

$$\Gamma = \frac{1}{S}\int d\Gamma, \tag{6.2.29}$$

where for n_i' outgoing identical particles of type i

$$S = \prod_i n_i'!, \tag{6.2.30}$$

which is the *symmetry factor* [5]. This concept will become clear when we introduce the Feynman diagrams and we will see why we must divide by the symmetry factor.

Chapter 7

Quantum Symmetries

Recall from the symmetries chapter that we presented Noether's theorem along with its main results which were the Noether current and Noether charge. Despite the fact that these concepts are rooted in classical physics, we liberally applied its results to our quantum fields. It is natural to consider a quantum equivalent version of Noether's theorem so that we may have the theoretical justification to apply these results to quantum field theory. Thus we seek to generalize Noether's theorem to quantum physics. Based on the derivation of Noether's theorem in classical field theory, the natural place to start for our quantum Noether's theorem will be the path integral [8].

7.1 Schwinger-Dyson Equations

To begin our discussion on the symmetries of a quantum field theory, let us see how the quantum equations of motion can be derived from the path integral. Consider a general set of fields $\varphi_a(x)$ with the path integral

$$Z(J) = \int \mathcal{D}\varphi \, e^{i[\mathcal{S} + \int d^4y \, J_a \varphi_a]}. \tag{7.1.1}$$

In analogy with classical mechanics, let us consider a infinitesimal variation of the fields $\varphi_a(x) \to \varphi'_a(x) = \varphi_a(x) + \delta\varphi_a(x)$ [8]. This is merely a shift by a constant so

the functional measure remains invariant $\mathcal{D}\varphi' = \mathcal{D}\varphi$ under this change. Then we have

$$\delta Z = 0$$
$$i\int \mathcal{D}\varphi\, e^{i[\mathcal{S}+\int d^4y\, J_b\varphi_b]} \int d^4x \left(\frac{\delta\mathcal{S}}{\delta\varphi_a(x)} + J_a(x)\right)\delta\varphi_a(x) = 0. \tag{7.1.2}$$

If we take n functional derivatives with respect to $J_{a_k}(x_k)$ and set $J = 0$, we get the following [8]

$$\int \mathcal{D}\varphi\, e^{i\mathcal{S}} \int d^4x \left[i\frac{\delta\mathcal{S}}{\delta\varphi_a(x)}\hat{\varphi}_{a_1}(x_1)\ldots\hat{\varphi}_{a_n}(x_n) + \sum_{k=1}^{n} \hat{\varphi}_{a_1}(x_1)\ldots\delta_{aa_k}\delta^4(x-x_k)\ldots\hat{\varphi}_{a_n}(x_n)\right]\delta\varphi_a(x) = 0. \tag{7.1.3}$$

The variation $\delta\varphi_a(x)$ was arbitrary thus the term in the brackets must equal zero but since the functional derivatives compute the vacuum expectation value when used on the partition function, we get

$$i\langle 0|\mathrm{T}\frac{\delta\mathcal{S}}{\delta\varphi_a(x)}\hat{\varphi}_{a_1}(x_1)\ldots\hat{\varphi}_{a_n}(x_n)|0\rangle + \sum_{k=1}^{n} \langle 0|\mathrm{T}\hat{\varphi}_{a_1}(x_1)\ldots\delta_{aa_k}\delta^4(x-x_k)\ldots\hat{\varphi}_{a_n}(x_n)|0\rangle = 0, \tag{7.1.4}$$

which are called the *Schwinger-Dyson equations* [5]. To understand the meaning of these equations, consider the free field theory of a single real scalar field. From our derivation of the real scalar field in the fields chapter, we have

$$\frac{\delta\mathcal{S}}{\delta\varphi(x)} = (\partial_x^2 - m^2)\hat{\varphi}(x),$$

so for $n = 1$, the Schwinger-Dyson equations reduce to [5]

$$(-\partial_x^2 + m^2)i\langle 0|\mathrm{T}\hat{\varphi}(x)\hat{\varphi}(x_1)|0\rangle = 0. \tag{7.1.5}$$

Recall that $i\langle 0|\mathrm{T}\hat{\varphi}(x)\hat{\varphi}(x_1)|0\rangle = \Delta_F(x - x_1)$ which is the Feynman propagator of the Klein-Gordon equation thus we have

$$(-\partial_x^2 + m^2)\Delta_F(x - x_1) = 0. \tag{7.1.6}$$

Then for n real scalar fields, we get the more general form [5]

$$(-\partial^2 + m^2)\langle 0|\mathrm{T}\frac{\delta\mathcal{S}}{\delta\varphi_a(x)}\hat{\varphi}_{a_1}(x_1)\dots\hat{\varphi}_{a_n}(x_n)|0\rangle = 0 \quad \text{for } x \neq x_{1,\dots n}. \tag{7.1.7}$$

Thus we see that the classical equations of motion for classical real scalar field are also obeyed by the quantum real scalar field inside a correlation function as long as the spacetime arguments of the fields are different from each other [5]. Otherwise, we get *contact terms*. In general, as long as the functional measure is invariant under an infinitesimal change of variables, this process can be repeated for any of the fields we have considered and will yield the appropriate equations of motion for the propagators of the corresponding quantum field theory [8]. In this respect, the Schwinger-Dyson equations are the quantum analogue of the Euler-Lagrange equations.

7.2 Ward-Takahashi Identity

Now consider a theory that possesses a continuous symmetry with a corresponding Noether current. We repeat the derivation of the Schwinger-Dyson equations except with an infinitesimal variation in $\varphi_a(x)$ that results in $\delta\mathcal{L}(x) = 0$ [5]. Then we sum over the index a and use

$$\partial_\mu j^\mu(x) = \delta\mathcal{L}(x) - \frac{\delta\mathcal{S}}{\delta\varphi_a(x)}\delta\varphi_a(x),$$

to get

$$\partial_\mu \langle 0|\mathrm{T} j^\mu(x)\hat{\varphi}_{a_1}(x_1)\dots\hat{\varphi}_{a_n}(x_n)|0\rangle + i\sum_{k=1}^{n}\langle 0|\mathrm{T}\hat{\varphi}_{a_1}(x_1)\dots\delta\hat{\varphi}_{a_k}(x)\delta^4(x-x_k)\dots\hat{\varphi}_{a_n}(x_n)|0\rangle = 0, \tag{7.2.1}$$

which is the *Ward-Takahashi identity* [5]. Thus we see that Noether's theorem and consequently the conservation of the Noether currents hold in quantum physics as well up to contact terms that depend on the form of the infinitesimal variation that leaves the Lagrangian invariant [5]. We see that the Ward-Takahashi identity is the quantum analogue of Noether's theorem for continuous symmetries.

Ward Identity

We will apply our results to derive a crucial identity that will aid us in computing scattering amplitudes for quantum electrodynamics. We will state and introduce results from the next chapter to facilitate our discussion.

The scattering amplitude for quantum electrodynamics is given as

$$\mathcal{T} = \varepsilon^\mu \mathcal{M}_\mu, \tag{7.2.2}$$

where ε^μ is the external polarization vector and $\mathcal{M}_\mu$ is a matrix [5]. We expect that the scattering amplitude should be gauge invariant under gauge transformations of the external polarization vectors of the form

$$\varepsilon^\mu \to \varepsilon^\mu + ck^\mu, \tag{7.2.3}$$

where k^μ is the four-momentum of the photon and c is an arbitrary constant [5]. Thus if we replace ε^μ with k^μ, we expect to have

$$k^\mu \mathcal{M}_\mu = 0, \tag{7.2.4}$$

which is called the *Ward identity*. We now prove the Ward identity.

Proof. Recall that the general form of the scattering amplitude was

$$\langle f|i\rangle = (2\pi)^4\delta^4\big(\sum_i k_i\big) f(k_i^2, k_i\cdot k_j),$$

where $f(k_i^2, k_i\cdot k_j)$ is a function of the Lorentz scalar k_i^2 and $k_i\cdot k_j$. If f is of the form

$$f(k_i^2, k_i\cdot k_j) \propto \frac{i\mathcal{T}}{(k_1^2+m^2)\dots(k_n^2+m^2)},$$

then f has contributions to the scattering amplitude $\mathcal{T}$. Otherwise, f will not contribute to $\mathcal{T}$ [5].

The Schwinger-Dyson equations were derived as

$$\begin{aligned} &i\langle 0|\mathrm{T}\frac{\delta\mathcal{S}}{\delta\varphi_a(x)}\hat{\varphi}_{a_1}(x_1)\dots\hat{\varphi}_{a_n}(x_n)|0\rangle \\ &\quad+\sum_{k=1}^{n}\langle 0|\mathrm{T}\hat{\varphi}_{a_1}(x_1)\dots\delta_{aa_k}\delta^4(x-x_k)\dots\hat{\varphi}_{a_n}(x_n)|0\rangle = 0. \end{aligned}$$

If we Fourier transform one of the contact terms, we get

$$\begin{aligned} &\int d^4x\, d^4x_k\, e^{i(kx+k_kx_k)}\,\delta^4(x-x_k)f(x,x_k) \\ &= \int d^4x\, e^{i(kx+k_k)x}\, f(x, x=x_k) \\ &= \tilde{f}(k+k_k), \end{aligned}$$

so $\tilde{f}(k+k_k)$ is a function of $k+k_k$ hence it is independent of $k-k_k$ thus it cannot contribute to the scattering amplitude. Therefore, contact terms do not contribute to the scattering amplitude $\mathcal{T}$ [5].

Recall that the LSZ reduction formula for outgoing photons in the Lorenz gauge

$$\langle f|i\rangle = i\varepsilon^\mu\int d^4x\, e^{ikx}(-\partial^2)\dots\langle 0|\mathrm{T}\hat{A}_\mu\dots|0\rangle,$$

which yields the equations of motion (from the *renormalized* QED Lagrangian)

$$-Z_3\partial^2 A_\mu = Z_1 j^\mu, \tag{7.2.5}$$

where Z_1 and Z_3 are the renormalization factors and j^μ is the electromagnetic current [5]. Then we use the Schwinger-Dyson equations along the fact that the contact terms do not contribute to the scattering amplitude to get

$$\begin{aligned}
\langle f|i\rangle &= i\varepsilon^\mu \int d^4x\, e^{-ikx}(-\partial^2)\ldots\langle 0|\mathrm{T}\hat{A}_\mu\ldots|0\rangle \\
&= i\frac{Z_1}{Z_3}\varepsilon^\mu \int d^4x\, e^{-ikx}\ldots\langle 0|\mathrm{T}j_\mu(x)\ldots|0\rangle \\
&= \varepsilon^\mu \mathcal{M}_\mu.
\end{aligned}$$

Then

$$\begin{aligned}
k^\mu\mathcal{M}_\mu &= \frac{Z_1}{Z_3}\int d^4x\, ik^\mu\, e^{-ikx}\ldots\langle 0|\mathrm{T}j_\mu(x)\ldots|0\rangle \\
&= \frac{Z_1}{Z_3}\int d^4x\,(-\partial^\mu)e^{-ikx}\ldots\langle 0|\mathrm{T}j_\mu(x)\ldots|0\rangle \\
&= -\frac{Z_1}{Z_3}\int d^4x\, e^{-ikx}\ldots\langle 0|\mathrm{T}\partial^\mu j_\mu(x)\ldots|0\rangle \\
&= 0,
\end{aligned}$$

where we integrated by parts to get the four-gradient inside the correlation function in the third equality and used the conservation of the Noether current $\partial^\mu j_\mu = 0$ to get the last equality. Thus the Ward identity holds. □

With the gauge invariance of the scattering amplitude firmly established, we can now compute the useful identity that will aid us in the next chapter which is the summation of the polarization four-vectors

$$\sum_{\lambda=\pm}\varepsilon_\lambda^{\mu*}(\mathbf{k})\varepsilon_\lambda^\nu(\mathbf{k}). \tag{7.2.6}$$

Recall the identity from fields chapter for the electromagnetic field

$$\sum_{\lambda=\pm} \varepsilon_\lambda^{i*}(\mathbf{k})\varepsilon_\lambda^j(\mathbf{k}) = \delta_{ij} - \frac{k_i k_j}{\mathbf{k}^2}. \tag{7.2.7}$$

To generalize this to the four-vector case, let us introduce the unit four-vector in the time direction [5]

$$\hat{t}^\mu = (1, \mathbf{0}). \tag{7.2.8}$$

Then we see that $\hat{t}^\mu k_\mu = -k^0$ and

$$k^\mu + (\hat{t}\cdot k_\mu)\hat{t}^\mu = (0, \mathbf{k}), \tag{7.2.9}$$

which leads to

$$\mathbf{k}^2 = k^2 + (\hat{t}^\mu k_\mu)^2. \tag{7.2.10}$$

Now we define a new unit vector in the $\mathbf{k}$ direction [5]

$$\hat{z}^\mu = \frac{k^\mu + (\hat{t}\cdot k)\hat{t}^\mu}{[k^2 + (\hat{t}\cdot k)^2]^{1/2}}. \tag{7.2.11}$$

Then we can extend the summation to the four-vector case which leads to

$$\begin{aligned}\sum_{\lambda=\pm} \varepsilon_\lambda^{\mu*}(\mathbf{k})\varepsilon_\lambda^\nu(\mathbf{k}) &= g^{\mu\nu} + \hat{t}^\mu\hat{t}^\nu - \frac{(k^\mu + (t\cdot k)\hat{t}^\mu)(k^\rho + (t\cdot k)\hat{t}^\rho)}{k^2 + (\hat{t}\cdot k)^2} \\ &= g^{\mu\nu} + \hat{t}^\mu\hat{t}^\nu - \hat{z}^\mu\hat{z}^\nu. \end{aligned} \tag{7.2.12}$$

Now since the scattering amplitude is gauge invariant due to the Ward identity, when we sum over the photon polarizations to calculate the scattering amplitude we can drop the k^μ and k^ρ terms since the photons will couple to a conserved current [5]. Thus

$$\sum_{\lambda=\pm} \varepsilon_\lambda^{\mu*}(\mathbf{k})\varepsilon_\lambda^\nu(\mathbf{k}) = g^{\mu\nu} + \hat{t}^\mu\hat{t}^\nu - \frac{(t\cdot k)^2}{k^2 + (\hat{t}\cdot k)^2}\hat{t}^\mu\hat{t}^\rho. \tag{7.2.13}$$

We will see later that for *external* photons, we have $k^2 = 0$ which makes sense since photons are massless so we reach another simplification [5]

$$\sum_{\lambda=\pm} \varepsilon_\lambda^{\mu*}(\mathbf{k})\varepsilon_\lambda^{\nu}(\mathbf{k}) = g^{\mu\nu}. \tag{7.2.14}$$

This result will be extremely useful for calculating scattering amplitudes in quantum electrodynamics.

Chapter 8

Interactions and Feynman Diagrams

8.1 Constructing Interaction Lagrangians

We explained earlier in the chapter of symmetries on how to build appropriate Lagrangians for quantum field theory. The criteria was that all Lagrangians must be Lorentz invariant and possess CPT symmetry. When constructing Lagrangians that model interactions, we must impose more conditions so that the interaction Lagrangian is consistent with the framework of quantum field theory and does not produce absurd results. We typically add to the free field Lagrangian $\mathcal{L}_0$ an interaction term $\mathcal{L}_{\text{int}}$ which, of course, must obey Lorentz and CPT symmetry so we are able to eliminate a fair amount of the number of possible terms we could add. However, even with these conditions there are still many ways to construct the interaction Lagrangian and clearly all of them cannot be physically correct. One more condition must be imposed and it is that the theory must be *renormalizable* [8]. Renormalization is the generic term for systemically cancelling out divergent integrals that frequently appear in higher order perturbations in a mathematically well defined way [1]. Not all theories are renormalizable thus there are certain criteria that must be met for a theory to be renormalizable. We will present only

the simplest criteria since we will not delve into renormalization techniques in this thesis.

Recall from the section on natural units that the action $\mathcal{S}$ must be dimensionless and $[\mathcal{L}] = d$ in d spacetime dimensions. Now the interaction term will have a *coupling constant* g and assuming that free field Lagrangian has the required Lorentz and CPT symmetry along with $[\mathcal{L}_0] = 0$, the dimension of the coupling term determines whether a theory is renormalizable [8]. The scattering amplitude that involves the coupling g must be parameterized as a ratio of the particle mass m in the low energy regime or the Mandelstam variable s in the high energy regime [5]. The parameter is $gs^{-[g]/2}$ [5]. If the dimension of the coupling is positive or dimensionless, the perturbation expansion does not break down at high energies and the theory is renormalizable [8]. Generally, we desire that the coupling be dimensionless. If the dimension is negative, the perturbation expansion blows up at high energies and the theory is nonrenormalizable. Thus when we construct interaction terms, we seek dimensionless couplings so that our theory is normalizable, although an exception will be given for the φ^3 theory for pedagogical reasons.

8.2 φ^3 and φ^4 Theory

We present φ^3 and φ^4 theory of the real scalar field. These theories are almost always presented for pedagogical purposes yet φ^4 actually has physical relevance in the fact that the self-interaction of the Higgs boson, a spin-0 particle, can be described using the φ^4 theory. We will begin with the φ^4 theory and then present the φ^3 theory.

The Lagrangian for the φ^4 theory is

$$\mathcal{L} = -\frac{1}{2}\partial^\mu\varphi\partial_\mu\varphi - \frac{1}{2}m^2\varphi^2 + \frac{\lambda}{4!}\varphi^4, \tag{8.2.1}$$

where λ is the coupling constant [8]. Using dimensional analysis, we see that

$[\varphi] = 1$ and it is trivial to prove that terms of the form φ^n are always Lorentz scalars [5] but are not allowed for $n > 4$ since it would result in $[\lambda] = 4 - n$ [8]. Using dimensional analysis reveals that $[\lambda] = 0$, namely, the coupling constant λ is dimensionless thus the φ^4 theory is renormalizable. Now the path integral for this theory is given as

$$Z(J) \equiv \langle 0|0\rangle_J = \int \mathcal{D}\varphi \, e^{i\int d^4x\,[\mathcal{L}_0+\mathcal{L}_{int}+J\varphi]}, \tag{8.2.2}$$

where $\mathcal{L}_0$ is the free field real scalar Lagrangian and $\mathcal{L}_{int} = (\lambda/4!)\varphi^4$ is the interaction term. Once again, we assume the following normalizations

$$\begin{aligned} \langle 0|\hat{\varphi}(x)|0\rangle &= 0, \\ \langle p|\hat{\varphi}(x)|0\rangle &= e^{-ipx}, \\ \langle p'|p\rangle &= (2\pi^3)(2\omega)\delta^3(\mathbf{p}' - \mathbf{p}), \end{aligned} \tag{8.2.3}$$

for our interacting theory. Recall that the expression for the four-point vacuum expectation value for the real scalar fields was given by Wick's theorem,

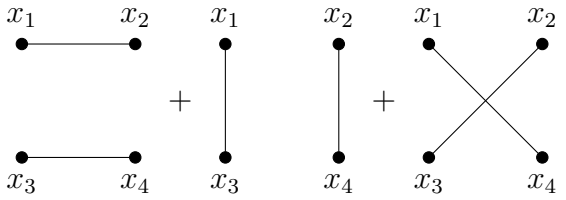
$$\begin{aligned} \langle 0|\mathrm{T}\hat{\varphi}(x_1)\hat{\varphi}(x_2)\hat{\varphi}(x_3)\hat{\varphi}(x_4)|0\rangle = \frac{1}{i^2}\Big[&\Delta_F(x_1 - x_2)\Delta_F(x_3 - x_4) \\ &+ \Delta_F(x_1 - x_3)\Delta_F(x_2 - x_4) \\ &+ \Delta_F(x_1 - x_4)\Delta_F(x_2 - x_3)\Big]. \end{aligned} \tag{8.2.4}$$

We can represent this expression diagrammatically by drawing a line with vertices labeled with the propagator's spacetime positions to represent an individual propagator in the following manner:

$$\frac{1}{i}\Delta_F(x - y) = \qquad x \bullet\!\!\!-\!\!\!-\!\!\!-\!\!\!-\!\!\!-\!\!\!-\!\!\!-\!\!\!-\!\!\!\bullet\, y \tag{8.2.5}$$

Then the four point vacuum expectation value can be drawn as [8]

$$\langle 0|\mathrm{T}\hat{\varphi}(x_1)\hat{\varphi}(x_2)\hat{\varphi}(x_3)\hat{\varphi}(x_4)|0\rangle =$$



(8.2.6)

Thus we see that the four point vacuum expectation value is the sum of these three diagrams which are called *Feynman diagrams*. The diagrams provide a rather interesting interpretation of the underlying physics involved. For each term in (8.2.6), two particles are created at the spacetime points x_1 and x_2 and each particle propagates to the their corresponding spacetime points x_3 and x_4 where they are annihilated [8]. Each term in the expansion above represents the different ways in which the particles can do this process. Along with this, each term in the expansion represents the amplitude for the process to occur in the particular fashion specified by the Feynman diagram [8]. The sum of all three of these diagrams ultimately represents the total amplitude for this process. The Feynman diagrams we have drawn for the four-point correlation function can easily be generalized for any n-point correlation function as long as the number of fields $\hat{\varphi}$ is even since the vacuum expectation value vanishes for an odd number of fields $\hat{\varphi}$ [5]. Although it seems as though the Feynman diagrams seem applicable only to the propagators, we have just scratched the surface in terms of the utility and applications of the diagrams. We will now derive the *Feynman rules* for the φ^4 theory and how it will aid us greatly in calculations and interpreting the underlying physics.

Following the procedure for the path integral of the harmonic oscillator, we can write the path integral as

$$\begin{aligned} Z(J) &= e^{i\int d^4x\,\mathcal{L}_{int}\left(\frac{1}{i}\frac{\delta}{\delta J(x)}\right)^4}\int \mathcal{D}\varphi e^{i\int d^4x\,[\mathcal{L}_0+J\varphi]} \\ &\propto e^{i\int d^4x\,\mathcal{L}_{int}\left(\frac{1}{i}\frac{\delta}{\delta J(x)}\right)^4} Z_0(J), \end{aligned} \tag{8.2.7}$$

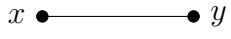
where $Z_0(J)$ is the path integral for the free real scalar field [8]. The second line is proportional to some normalization constant that is fixed by the normalization of the path integral $Z(0) = 1$ which we impose by hand [5]. Assuming that the coupling constant λ is small, we make a dual Taylor expansion in powers of λ and J [5]

$$Z(J) \propto \sum_{V=0}^{\infty} \frac{1}{V!}\Big[\frac{i\lambda}{4!}\int d^4x \left(\frac{1}{i}\frac{\delta}{\delta J(x)}\right)^4\Big]^V \times \sum_{P=0}^{\infty} \frac{1}{P!}\Big[\frac{i}{2}\int d^4y\, d^4z\, J(y)\Delta_F(y-z)J(z)\Big]^P. \tag{8.2.8}$$

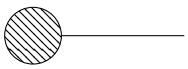
To keep things simple and less cluttered, let us focus on the first few terms of the expansion for $V = P = 1$,

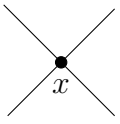
$$\begin{aligned}&\frac{i\lambda}{4!}\Big(\frac{1}{i}\Big)^4\Big[i^2\Delta_F(x_1-x_2)\Delta_F(x_3-x_4)\\ &+ i\Delta_F(x_1-x_2)\Big(\frac{i}{2}\Big(\int d^4z\,\Delta_F(x_3-z)J(z) + \int d^4y\, J(y)\Delta_F(y-x_3)\Big)\Big)\\ &\times\Big(\frac{i}{2}\Big(\int d^4z\,\Delta_F(x_4-z)J(z) + \int d^4y\, J(y)\Delta_F(y-x_4)\Big)\Big) + \ldots\Big]\, Z_0.\end{aligned}$$

To organize these terms, we will draw a blob with a line to represent the term $i\int d^4x\, J(x)$, a vertex joining four line segments to represent $i\lambda\int d^4x$, and of course a line with two endpoints to represent the propagator $\frac{1}{i}\Delta_F(x-y)$ [8]:

1. Propagators : 

$= \frac{1}{i}\Delta_F(x-y)$;

2. External lines :  $= i\int d^4x\, J(x)$;

3. Vertices : 

$= i\lambda\int d^4x$.

To explain our specific choice of labels, E stands for external lines (for now it will refer to source terms, namely, blobs), P stands for propagators, and V stands for vertices. Hence, the labels E, V, and P count the number of external lines (sources), vertices, and propagators, respectively and their relationship is given by

the graph equation

$$E = 2P - 4V, \tag{8.2.9}$$

which also gives the number of nonvanishing sources. The graph equation provides an easy way to draw and check if a diagram has all the necessary components given some restrictions and/or parameters on the values of E, V, and P. The overall phase factor is

$$i^V (1/i)^{4V} i^P = i^{V+E-P}, \tag{8.2.10}$$

where the exponential term of the overall phase factor is the famous *Euler characteristic*. Now the $4V$ functional derivatives can act on the $2P$ propagators in $(2P!)/(2P - 4V)!$ different ways but it is clear that we will have many identical terms in the expansion [5]. This means that if we were to draw every single diagram for each term in the dual Taylor expansion, there would actually be a small number of truly different configured diagrams that would yield different answers. However, each of these diagrams would have a large multiplicity, namely, different ways of rearranging parts of the diagram without changing the answer. Thus we see that we only have to draw a few diagrams to count these identical terms and each of these distinct diagrams will be multiplied by some constant [5]. Finding these constants is essentially a combinatorics problem of finding all the possible ways we can rearrange the vertices, propagators, and sources. The vertices can be interchanged in $V!$ ways and the propagators can be rearranged in $P!$ ways. Along with this, the functional derivatives for each vertex can be rearranged in 4! ways without changing the diagram. The two sources at the ends of a propagator can be interchanged in 2! ways [5]. Thus, each vertex and propagator has a counting factor of $4!V!$ and $2!P!$, respectively. When added onto the diagrams, the counting factors cancel the numbers from the Taylor expanded path integral (8.2.8) [5]. If the reader was curious about the inclusion of the 4! factor in the interaction term, it is now clear that the inclusion of the 4! cancels the counting factor of the functional derivatives that come from contracting the φ^4 terms [8]. This is

essentially the reason why the vertex term is $i\lambda \int d^4x$ instead of $(i\lambda/4!) \int d^4x$.

Despite the cancellation of the numbers in the expansion, the overall constant of each diagram is too large because we have over counted the number of identical terms. This is due to the fact some rearrangement of the derivatives give the same configuration of sources as the rearrangement of sources. This occurs when there is an underlying symmetry associated with a diagram so the factor by which we have over-counted is called the *symmetry factor* [5]. Thus another rule that we must add is that after assigning the appropriate factors to a diagram, we divide by the symmetry factor of the diagram, provided it has one.

Even with all these rules, there are still quite a bit of diagrams we must draw and calculate. Along with this, consider the following diagrams [8]

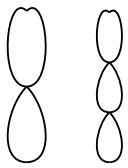


Figure 8.1: Disconnected diagrams

Diagrams such as the ones in fig. 8.1 are *disconnected* diagrams meaning that the diagram is not connected to an external point [8]. A *connected* diagram is a diagram where a path can be traced from any two points on the diagram. From equation (8.2.6), we see that some of the terms in the expansion consists of a product connected diagrams that are disconnected from each other. In fact, the most general diagram will consist of a product of connected diagrams. Let D be an arbitrary general diagram then

$$D = \frac{1}{S_D} \prod_I (C_I)^{n_I}, \tag{8.2.11}$$

where C_I is a connected diagram with its specific symmetry factor included, n_I is an integer that counts the number of C_I in D, and S_D is the additional symmetry factor for D not accounted for by the symmetry factors that are already present in

the connected diagrams [5]. This unaccounted symmetry factor S_D will then involve interchanges of propagators and vertices between different connected diagrams [5]. Exchanges made among different but identical connected diagrams and exchanges that involve all of the propagators and vertices in a connected diagram leave the overall diagram D unchanged [5]. Since there are n_I factors of C_I then there are $n_I!$ ways for these types of rearrangements. Thus the value of S_D is

$$S_D = \prod_I \frac{1}{n_I!}. \tag{8.2.12}$$

Since the path integral of the φ^4 theory is diagrammatically the sum of all general diagrams D, we have [5]

$$\begin{aligned} Z(J) &\propto \sum_{\{n_I\}} D \\ &\propto \sum_{\{n_I\}} \prod_I \frac{1}{n_I!} (C_I)^{n_I} \\ &\propto \prod_I \sum_{\{n_I\}} \frac{(C_I)^{n_I}}{n_I!} \\ &\propto \prod_I \exp\Big[C_I\Big] \\ &\propto \exp\Big[\sum_I C_I\Big]. \end{aligned} \tag{8.2.13}$$

With this, we impose the normalization condition $Z(0) = 1$ which means that the value of *vacuum diagrams* in (8.2.13), such as fig. 8.1, are omitted thus we have the elegant result

$$Z(J) = \exp\Big[iW(J)\Big], \tag{8.2.14}$$

where

$$iW(J) \equiv \sum_{I \neq \{0\}} C_I, \tag{8.2.15}$$

and $I \neq \{0\}$ means to omit vacuum diagrams such that $W(0) = 0$ [5]. This result holds for all other theories that we will consider throughout the subsequent sections.

With this, we see that the propagator becomes

$$\begin{aligned}\langle 0|\mathrm{T}\hat{\varphi}(x_1)\hat{\varphi}(x_2)|0\rangle &= \delta_1\delta_2\, iW(J)\Big|_{J=0} \\ &= \frac{1}{i}\Delta_F(x_1 - x_2),\end{aligned} \tag{8.2.16}$$

and more generally, Wick's theorem becomes

$$\begin{aligned}\langle 0\mathrm{T}\hat{\varphi}(x_1)\hat{\varphi}(x_2)\dots\hat{\varphi}(x_1)\hat{\varphi}(x_{2n})|0\rangle &= \delta_1\delta_2\dots\delta_{2n}\, Z(J)\Big|_{J=0} \\ &= \delta_1\delta_2\dots\delta_{2n}\, iW(J)\Big|_{J=0} \\ &= \frac{1}{i^n}\sum_{\text{pairings}}\Delta_F(x_{i_1} - x_{i_2})\dots\Delta_F(x_{i_{2n-1}} - x_{i_{2n}}).\end{aligned} \tag{8.2.17}$$

where $\delta_k \equiv \frac{1}{i}\frac{\delta}{\delta J(x_k)}$ [5].

Now we will shift our attention from φ^4 theory to φ^3 theory so that we can derive additional properties of $iW(J)$ and other necessary properties. We will use the φ^3 theory to derive the Feynman rules so that we can introduce key ideas that arise from the φ^3 theory that will appear frequently in the other theories we will cover. The Lagrangian for φ^3 theory is

$$\mathcal{L} = -\frac{1}{2}Z_\varphi\partial^\mu\varphi\partial_\mu\varphi - \frac{1}{2}Z_m m^2\varphi^2 + \frac{\lambda}{3!}Z_\lambda\varphi^3 + Y\varphi, \tag{8.2.18}$$

where it is expected that $Y \to 0$ and $Z_{\lambda,\varphi,m} \to 1$ as $\lambda \to 0$ [5]. The reason for the inclusion of Y and $Z_{\lambda,\varphi,m}$ will be clear shortly. Now the path integral is

$$Z(J) \propto e^{i\int d^4x\mathcal{L}_1\left(\frac{1}{i}\frac{\delta}{\delta J(x)}\right)}Z_0(J), \tag{8.2.19}$$

where

$$\mathcal{L}_1 = \frac{\lambda}{3!} Z_\lambda \varphi^3 + \mathcal{L}_{ct}, \tag{8.2.20}$$

$$\mathcal{L}_{ct} = -\frac{1}{2}(Z_\varphi - 1)\partial^\mu \varphi \partial_\mu \varphi - \frac{1}{2}(Z_m - 1)m^2\varphi^2 + Y\varphi. \tag{8.2.21}$$

The second term above $\mathcal{L}_{ct}$ is called the *counterterm* Lagrangian and it will play an important role in keeping our normalization consistent. If we were to momentarily ignore the counterterm Lagrangian, the diagrammatic rules for φ^3 theory would remain the same as φ^4 theory except that the graph equation would be

$$E = 2P - 3V, \tag{8.2.22}$$

and the vertex would join three line segments for a vertex factor of $iZ_\lambda \lambda \int d^4x$ [5]. Repeating the derivations from the previous analysis will lead to the same result $Z(J) = \exp\Big[iW(J)\Big]$. Now consider the following expression

$$\begin{aligned} \langle 0|\hat{\varphi}(x)|0\rangle &= \frac{1}{i}\frac{\delta}{\delta J(x)} Z(J)\Big|_{J=0} \\ &= \frac{\delta}{\delta J(x)} W(J)\Big|_{J=0}, \end{aligned}$$

which is the vacuum expectation value of the field $\hat{\varphi}(x)$. Diagrammatically, this expression is sum of diagrams such as

Figure 8.2: Tadpole diagrams

with the source removed [5]. In φ^4 theory, the vacuum expectation of a single field would be zero since from the graph equation $E = 2P - 4V$, there is no way to draw a diagram with a single source which agrees with our normalization condition $\langle 0|\hat{\varphi}(x)|0\rangle = 0$. Let us see what happens for the φ^3 theory. Ignoring the

counterterm Lagrangian, we get

$$\langle 0|\hat{\varphi}(x)|0\rangle = \frac{1}{2}i\lambda \int d^4x \, \frac{1}{i}\Delta_F(x-y)\frac{1}{i}\Delta_F(y-y) + O(\lambda^3),$$

where the $1/2$ comes from the symmetry factor and we have set $Z_\lambda = 1$ since at first order $Z_\lambda = 1 + O(\lambda^2)$ [5]. This is a problem because we must have $\langle 0|\hat{\varphi}(x)|0\rangle = 0$. So, we add back the counterterm Lagrangian which introduces a new vertex term $iY \int d^4x$ where a single line segment ends. The Feynman diagrams for this are shown in fig. 8.3.

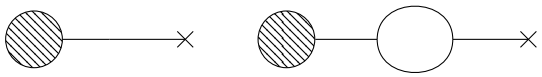


Figure 8.3: Tadpole diagrams with one source removed

Up to order $O(\lambda)$, the first diagram above yields the vacuum expectation value of a single field

$$\langle 0|\hat{\varphi}(x)|0\rangle = \left(iY + \frac{1}{2}(i\lambda)\frac{1}{i}\Delta_F(0)\right) \int d^4x \, \frac{1}{i}\Delta_F(x-y) + O(\lambda^3).$$

We now take advantage of the coefficients to correct and fix our normalization. It is clear that we should choose

$$Y = \frac{1}{2}i\lambda\Delta_F(0) + O(\lambda^3), \tag{8.2.23}$$

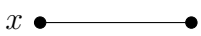
for first order $O(\lambda)$. For higher orders, we employ the exact same scheme and simply choose the proper value of Y that corrects the vacuum expectation value. It is obvious by now that $\Delta_F(0)$ is divergent for large p hence Y is infinite but parameters like these will cancel out when calculating cross sections [5]. Along with this, infinite parameters such as Y are not directly measurable hence they have no physical relevance. Since we are constantly fixing our Y value so that $\langle 0|\hat{\varphi}(x)|0\rangle = 0$ for every order, we see that the sum of of all connected diagrams with a single source equals zero. Like the vacuum diagrams, we can ignore any

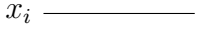
diagrams where when a single line segment is cut into two pieces, one of the two pieces has no sources. These types of diagrams like in fig. 8.3 are called *tadpole* diagrams, or *tadpoles* for short [5]. Thus $iW(J)$ is the sum of connected diagrams with no sources.

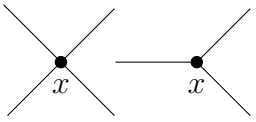
We see that from equations (8.2.16) and (8.2.17), a functional derivative δ_k strips off a source and labels it with the spacetime position coordinate x_k which brings us back to the original diagrammatic four-point correlation function (8.2.6) we introduced earlier. We see that we must modify our rule for external lines. Recall that the rule for external lines for the φ^3 and φ^4 theories was to draw a line segment with a blob at one end which was then assigned a factor of $i\int d^4x\, J(x)$. The functional derivatives δ_k strip off the blobs and relabel them with a spacetime coordinate x_k. Mathematically, we see then that the source factor becomes

$$\begin{aligned}\frac{1}{i}\frac{\delta}{\delta J(x_k)} i\int d^4y\, J(y) &= \int d^4y\, \delta^4(x_k - y)\\ &= 1. \end{aligned} \tag{8.2.24}$$

Thus the new rule for external lines for the φ^3 and φ^4 theories is to draw a line segment with one end labeled with a spacetime coordinate x_i [5]. The rules that we have outlined up to now for the φ^3 and φ^4 theories are called the *position-space Feynman rules* which are given as [8]:

1. Propagators : 

$= \frac{1}{i}\Delta_F(x-y);$
2. External lines : 

$= 1;$
3. Vertices : 

$= i\lambda\int d^4x;$
4. Divide by symmetry factor.

In practice however, the *momentum-space Feynman rules* are used to calculate the Feynman diagrams because of computational ease and since in real life scattering

experiments, it is the momenta of the particles that are measured rather than the positions [8]. To derive the momentum-space Feynman rules, we could simply replace the position-space values with their Fourier transformed counterparts in momentum-space. However, since we will ultimately be computing cross sections from scattering theory, we should consider any *additional* rules that would arise from the context of scattering so that we can create a full set of Feynman rules, in momentum-space, that gives us all the necessary information to draw and compute scattering diagrams.

Let us start with the simplest case of scattering which is the scattering of two (real) scalar particles written as $\varphi\varphi \to \varphi\varphi$. As we outlined in the previous chapter, we wish to compute the LSZ formula for scalar fields

$$\begin{aligned}\langle f|i\rangle = i^4 \int d^4x_1\, d^4x_2\, d^4x_1'\, d^4x_2'\, e^{i(p_1x_1+p_2x_2-p_1'x_1'-p_2'x_2')}(-\partial_1^2+m^2) \\ \times (-\partial_2^2+m^2)(-\partial_{1'}^2+m^2)(-\partial_{2'}^2+m^2)\langle 0|\mathrm{T}\hat{\varphi}(x_1)\hat{\varphi}(x_2)\hat{\varphi}(x_1')\hat{\varphi}(x_2')|0\rangle,\end{aligned} \tag{8.2.25}$$

where $\hat{\varphi}(x_1)$ and $\hat{\varphi}(x_2)$ represent incoming particles and $\hat{\varphi}(x_1')$ and $\hat{\varphi}(x_2')$ represent outgoing particles [5]. By Wick's theorem, the vacuum expectation value is diagrammatically [5]

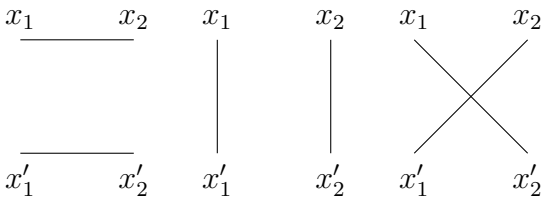
$$\begin{aligned}\langle 0|\mathrm{T}\hat{\varphi}(x_1)\hat{\varphi}(x_2)\hat{\varphi}(x_1')\hat{\varphi}(x_2')|0\rangle &= \delta_1\delta_2\delta_{1'}\delta_{2'}\, Z(J)\Big|_{J=0} \\ &= \delta_1\delta_2\delta_{1'}\delta_{2'}\, iW(J)\Big|_{J=0} \\ &= \Big[(\delta_1\delta_2\delta_{1'}\delta_{2'}\, iW) \\ &\quad + (\delta_1\delta_2 iW)(\delta_{1'}\delta_{2'} iW) \\ &\quad + (\delta_1\delta_{1'} iW)(\delta_2\delta_{2'} iW) \\ &\quad + (\delta_1\delta_{2'} iW)(\delta_2\delta_{1'} iW)\Big|_{J=0}.\end{aligned} \tag{8.2.26}$$

In practice, one would first draw the necessary Feynman diagrams but there is a key subtlety that we have not considered with our current rules. To avoid clutter,

let us insert the second line of the last equality of equation (8.2.26) into the LSZ formula and use $\langle 0|\mathrm{T}\hat{\varphi}(x)\hat{\varphi}(y)|0\rangle = \frac{1}{i}\Delta_F(x-y)$ to get

$$\begin{aligned}\langle f|i\rangle &= i^4 \int d^4x_1\, d^4x_2\, d^4x_1'\, d^4x_2'\, e^{i(p_1x_1+p_2x_2-p_1'x_1'-p_2'x_2')}(-\partial_1^2+m^2)\\ &\quad\times(-\partial_2^2+m^2)(-\partial_{1'}^2+m^2)(-\partial_{2'}^2+m^2)\Big(\frac{1}{i}\Big)^2\Delta_F(x_1-x_2)\Delta_F(x_1'-x_2')\\ &= \int d^4x_1\, d^4x_2\, d^4x_1'\, d^4x_2'\, e^{i(p_1x_1+p_2x_2-p_1'x_1'-p_2'x_2')}\, F(x_{11'})F(x_{22'})\\ &= (2\pi)^4\delta^4(p_1-p_1')(2\pi)^4\delta^4(p_2-p_2')\tilde{F}(\bar{p}_{11'})\tilde{F}(\bar{p}_{22'}),\end{aligned} \tag{8.2.27}$$

where $F(x_{ij}) = (-\partial_i^2+m^2)(-\partial_j^2+m^2)\Delta_F(x_{ij})$, $\tilde{F}(p)$ is its Fourier conjugate, $x_{ij'} = x_i - x_j'$, and $\bar{p}_{ij'} = (p_i+p_j')/2$ [5]. Of course, this result along with the other terms could also have been deduced by drawing the position-space Feynman diagrams for the three terms



(8.2.28)

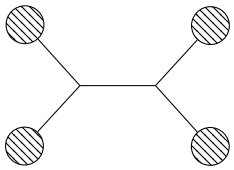
corresponding to $(\delta_1\delta_2 iW)(\delta_{1'}\delta_{2'}iW)$, $(\delta_1\delta_{1'}iW)(\delta_2\delta_{2'}iW)$, and $(\delta_1\delta_{2'}iW)(\delta_2\delta_{1'}iW)$, respectively. The first diagram represents two particles created at spacetime positions x_1 and x_1', propagating to spacetime positions x_2 and x_2', then annihilating. The second diagram shows two particles created at spacetime positions x_1 and x_2, propagating to spacetime positions x_1' and x_2', then annihilating. The third diagram shows two particles created at spacetime positions x_1 and x_2, propagating to spacetime positions x_2' and x_1', then annihilating. In summary, the particles do not interact at all which is not what we want to calculate. So, we omit these diagrams from our calculations [5].

The only term in equation (8.2.26) left to consider is $(\delta_1\delta_2\delta_{1'}\delta_{2'}\, iW)$ which represents a diagram that is *fully connected* meaning that a path can be traced on the diagram from any endpoint to another endpoint on the diagram [5]. Since

we are only considering one term of (8.2.26), we define a new vacuum expectation value called the *connected vacuum expectation value* which is

$$\langle 0|\mathrm{T}\hat{\varphi}(x_1)\hat{\varphi}(x_2)\hat{\varphi}(x_1')\hat{\varphi}(x_2')|0\rangle_{\mathrm{C}} = \delta_1\delta_2\delta_{1'}\delta_{2'}\, iW\Big|_{J=0}. \tag{8.2.29}$$

From the graph formula $E = 2P - 3V$, diagrams with one vertex $V = 1$ which is a first order $O(\lambda)$ approximation, would result in tadpole diagrams which we concluded earlier would be omitted from our calculations. Thus the lowest order approximation is the second order $O(\lambda^2)$ approximation which are diagrams with two vertices so $V = 2$. Now from equation (8.2.29), the four functional derivatives indicate that we will need a diagram with four sources which will be stripped off and relabeled with the position labels so $E = 4$. This leaves us with $P = 5$ thus the lowest order diagram will need five propagators. The corresponding diagram that matches these parameters is given in fig. 8.4.



$S = 8$

Figure 8.4: Graph with E=4, P=5, and V=2

Now when the functional derivatives strip off the sources and label them with the position coordinates, we have 24 position-space Feynman diagrams since the four functional derivatives can act on the four sources in 4! ways [5]. As mentioned previously, we luckily do not have to draw and compute all 24 diagrams because many of the diagrams are identical to each other. The 24 diagrams can be categorized evenly into three distinct groups of diagrams which are [5]

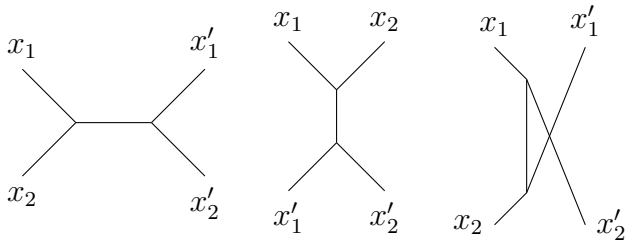


Figure 8.5: Position-space graphs of fig. (8.4)

Each of the three diagrams are assigned a factor of eight thus the symmetry factor of each diagram is neatly cancelled. Diagrams such as these are called *tree diagrams* and they do not contain any closed loops and have a symmetry factor of one [5]. Many of the Feynman diagrams that are actually computed in quantum field theory are tree diagrams and they often serve as the basis for which further corrections are considered. Reading the above position-space diagrams from left to right and using the position space Feynman rules, we get [5]

$$\begin{aligned}\langle 0|\mathrm{T}\hat{\varphi}(x_1)\hat{\varphi}(x_2)\hat{\varphi}(x_1')\hat{\varphi}(x_2')|0\rangle_{\mathrm{C}} &= (ig)^2\Big(\frac{1}{i}\Big)^5\int d^4y\, d^4z\, \Delta_F(y-z)\\ &\times\Big[\Delta_F(x_1-y)\Delta_F(x_2-y)\Delta_F(x_1'-z)\Delta_F(x_2'-z)\\ &+\Delta_F(x_1-y)\Delta_F(x_1'-y)\Delta_F(x_2-z)\Delta_F(x_2'-z)\\ &+\Delta_F(x_1-y)\Delta_F(x_2'-y)\Delta_F(x_2-z)\Delta_F(x_1'-z)\Big]+O(\lambda^4).\end{aligned}\tag{8.2.30}$$

Then the LSZ formula with the connected four-point correlation function is

$$\begin{aligned}
\langle f|i\rangle &= i^4 \int d^4x_1\, d^4x_2\, d^4x_1'\, d^4x_2'\, e^{i(p_1x_1+p_2x_2-p_1'x_1'-p_2'x_2')}(-\partial_1^2+m^2) \\
&\times (-\partial_2^2+m^2)(-\partial_{1'}^2+m^2)(-\partial_{2'}^2+m^2)\langle 0|\mathrm{T}\hat{\varphi}(x_1)\hat{\varphi}(x_2)\hat{\varphi}(x_1')\hat{\varphi}(x_2')|0\rangle_{\mathrm{C}} \\
&= i\lambda^2 \int d^4y\, d^4z\, d^4x_1\, d^4x_2\, d^4x_1'\, d^4x_2'\, \Delta_F(y-z)e^{i(p_1x_1+p_2x_2-p_1'x_1'-p_2'x_2')} \\
&\times (-\partial_1^2+m^2)(-\partial_2^2+m^2)(-\partial_{1'}^2+m^2)(-\partial_{2'}^2+m^2) \\
&\times \Big[\Delta_F(x_1-y)\Delta_F(x_2-y)\Delta_F(x_1'-z)\Delta_F(x_2'-z) \\
&+ \Delta_F(x_1-y)\Delta_F(x_1'-y)\Delta_F(x_2-z)\Delta_F(x_2'-z) \\
&+ \Delta_F(x_1-y)\Delta_F(x_2'-y)\Delta_F(x_2-z)\Delta_F(x_1'-z)\Big] + O(\lambda^4) \\
&= i\lambda^2 \int d^4y\, d^4z\, \Delta_F(y-z)\Big[e^{i(p_1y+p_2y-p_1'z-p_2'z)} + e^{i(p_1y+p_2z-p_1'y-p_2'z)} \\
&+ e^{i(p_1y+p_2z-p_1'z-p_2'y)}\Big] + O(\lambda^4), \qquad (8.2.31)
\end{aligned}$$

where we used the property of the Feynman propagator

$$(-\partial_i^2+m^2)\Delta_F(x_i-y) = \delta^4(x_i-y),$$

to get the last equality [5]. To derive an expression for the LSZ formula in momentum-space, we use the Fourier conjugate of the propagator

$$\Delta_F(y-z) = \int \frac{d^4p}{(2\pi)^4}\,\frac{e^{ip(y-z)}}{p^2+m^2-i\epsilon}.$$

When we integrate over the last two spacetime variables y and z, we get

$$\begin{aligned}
\langle f|i\rangle &= i\lambda^2(2\pi)^8 \int \frac{d^4p}{(2\pi)^4}\,\frac{1}{p^2+m^2-i\epsilon}\Big[\delta^4(p_1+p_2+p)\delta^4(p_1'+p_2'+p) \\
&+ \delta^4(p_1-p_1'+p)\delta^4(p_2'+p_2+p) + \delta^4(p_1-p_2'+p)\delta^4(p_1'+p_2+p)\Big] + O(\lambda^4) \\
&= (2\pi)^4\delta^4(p_1+p_2-p_1'-p_2')i\mathcal{T} + O(\lambda^4), \qquad (8.2.32)
\end{aligned}$$

where

$$\mathcal{T} = \lambda^2\Big[\frac{1}{(p_1+p_2)^2+m^2} + \frac{1}{(p_1-p_1')^2+m^2} + \frac{1}{(p_1-p_2')^2+m^2}\Big], \tag{8.2.33}$$

where we have suppressed $-i\epsilon$ into m^2 [5]. The delta function $\delta^4(p_1+p_2-p_1'-p_2')$ says that the four-momentum is conserved during the scattering process which is in full agreement with the conservation of the stress energy tensor from the spacetime translational symmetry of the Lagrangian. Since the unprimed and primed four-momenta represent the momenta of the incoming and outgoing particles, respectively, we can abstract equation (8.2.32) to

$$\langle f|i\rangle = (2\pi)^4\delta^4(p_{in}-p_{out})i\mathcal{T}, \tag{8.2.34}$$

where p_{in} and p_{out} represent the total four-momenta of the incoming and outgoing particles, respectively [5].

Let us analyze what we have done thus far so that we can extract patterns that will lead to the momentum-space Feynman rules. In deriving equations (8.2.32) and (8.2.33), the coupling constant appears as $(i\lambda)^2$ which makes sense given that all the diagrams in fig. 8.5 have two vertices. Since we have been using $Z_\lambda = 1$, we will return back to using Z_λ thus each vertex will be assigned a factor of $iZ_\lambda\lambda$. Along with this, we see that the momentum-space Feynman propagator appears in equation (8.2.33) except that the momentum component is replaced with the combined momenta of the particles in a way so that momenta is conserved. To explain this further, let us focus on the physics of just the first term of equation (8.2.33) which corresponds to the first diagram in fig. 8.5. The four-momentum components p_1 and p_2 correspond to the incoming particles at the spacetime positions x_1 and x_2, respectively. Since we already labeled unprimed and primed variables as incoming and outgoing particles, respectively, we will read the diagram from left to right. If we imagine two incoming scalar particles with four-momentum components p_1 and p_2 that collide at the first vertex and travel as one composite

scalar particle, this composite scalar particle will have four-momentum $p_1 + p_2$, which appears in the propagator of the first term. The composite particle would break up and separate into two outgoing scalar particles with four-momentum components p_1' and p_2' at the second vertex. Diagrammatically, we can draw this as

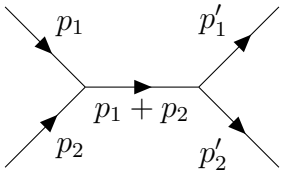


Figure 8.6: Tree-level s-channel diagram

If we were to imagine the momenta as rivers, we see that the incoming momenta rivers "flow" into the first vertex, combining their channel strength, then separating the channel at the second vertex as two outgoing momenta rivers. From our reasoning, we see that we should label all incoming particles as solid lines with arrows pointing *towards* the vertex with unprimed momentum labels p_i and all outgoing particles as solid lines with arrows pointing *away* from the vertex with primed momentum labels p_i'. These lines will be the momentum-space external lines and will be assigned a factor of 1 for each line. By following the arrows, the four-momenta will be conserved at each vertex and the *internal lines* will be labeled with the combined four-momentum that corresponds to the contribution of the four-momenta at the first vertex. The internal line will be assigned a factor of $-i/(p^2 + m^2 - i\epsilon)$ where of course the four-momentum p will be replaced with combined four-momentum labeling of the internal line. Now the momentum-space diagram for the second term in equation (8.2.33) is given in fig. 8.7.

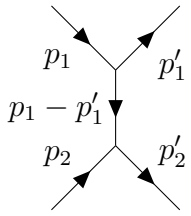


Figure 8.7: Tree-level t-channel diagram

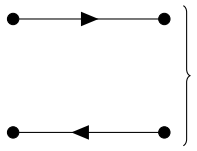
Here the combined four-momentum on the internal line is $p_1 - p_1'$ which is just the conservation of momentum. The four-momentum labeling on the internal line matches the four-momentum of the propagator of the second term as well thus our rules are consistent. We should point out that when drawing the necessary Feynman diagrams, one should fix the ordering of the labels of the four-momenta on the external lines for all of the contributing diagrams. This will standardize the process of labeling the four-momentum of the internal line because there will be no ambiguity in deciding which vertex will contribute to the four-momentum of the internal line. We will use the labeling scheme that we have been using for every diagram thus far. Depending on the configuration of the diagram, the first vertex will refer to the very first vertex when reading the diagram either from left to right or from top to bottom. We now present the momentum-space Feynman rules for the φ^3 and φ^4 theories according to Srednicki [5].

Feynman Rules for φ^3 Theory (Real)

1. The external lines for incoming scalar particles are solid lines with an arrow pointing towards the vertex and are labeled with its four-momentum p_i. The assigned factor is 1.

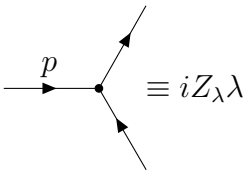
2. The external lines for outgoing scalar particles are solid lines with an arrow pointing away from the vertex and are labeled with its four-momentum p_i'. The assigned factor is 1.

3. For internal lines, draw a solid line with an arrow in an arbitrary direction with an assigned factor of $-i/(p^2 + m^2 - i\epsilon)$.



$$\equiv \frac{1}{i}\frac{1}{p^2+m^2-i\epsilon}$$

4. Construct each vertex from three solid lines. One of the solid lines must have an arrow pointing towards the vertex while the other must have an arrow pointing away from the vertex. The arrow on the third line can point in any direction. Draw all topologically inequivalent diagrams using the vertex, external lines, and any extra internal lines when needed. The assigned factor is $iZ_\lambda\lambda$.



$$\equiv iZ_\lambda\lambda$$

5. Arrange every diagram so that the ordering of their four-momenta labels are the same for each diagram. Depending on the configuration of the diagram, read the vertices from left to right or from top to bottom.
6. Conserve four-momentum at each vertex by following the arrows of each external line. This fixes the momenta on all the internal lines for tree diagrams.
7. A diagram with L closed loops will have L internal momenta not fixed according to rule (6). These loops will be labeled as ℓ_i and will be integrated with measure $d^4\ell_i/(2\pi)^4$. If a looped diagram has symmetry factors associated with the exchange of internal propagators and vertices that leave the diagram unchanged, divide the value of the diagram by this symmetry factor.
8. Include diagrams with counterterm vertices that connects two propagators, each with four-momentum p. The counterterm vertex will be assigned a factor of $-i(Ap^2 + Bm^2)$ where the $O(\lambda^2)$ coefficients are $A = Z_\varphi - 1$ and $B = Z_m - 1$.

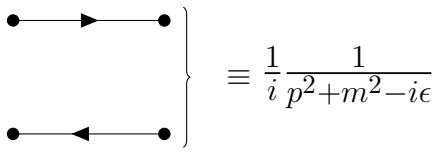
9. Sum all the values from the contributing diagrams to get the value of $i\mathcal{T}$.

Feynman Rules for φ^4 Theory (Real)

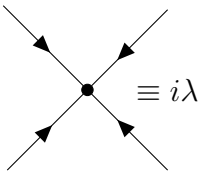
1. The external lines for incoming scalar particles are solid lines with an arrow pointing towards the vertex and are labeled with its four-momentum p_i. The assigned factor is 1.

2. The external lines for outgoing scalar particles are solid lines with an arrow pointing away from the vertex and are labeled with its four-momentum p_i'. The assigned factor is 1.

3. For internal lines, draw a solid line with an arrow in an arbitrary direction with an assigned factor of $-i/(p^2 + m^2 - i\epsilon)$.

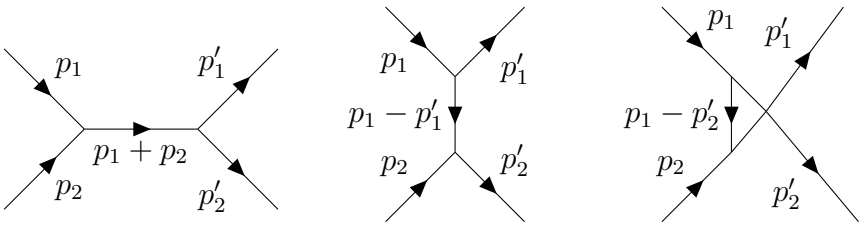


4. Construct each vertex from four solid lines. Two of the solid lines must have arrows pointing towards the vertex while the other two must have arrows pointing away from the vertex. The arrow on the third line can point in any direction. Draw all topologically inequivalent diagrams using the vertex, external lines, and any extra internal lines when needed. The assigned factor is $i\lambda$.



5. Arrange every diagram so that the ordering of their four-momenta labels are the same for each diagram. Depending on the configuration of the diagram, read the vertices from left to right or from top to bottom.
6. Conserve four-momentum at each vertex by following the arrows of each external line. This fixes the momenta on all the internal lines for tree diagrams.
7. A diagram with L closed loops will have L internal momenta labeled as ℓ_i. Integrate each loop with measure $d^4\ell_i/(2\pi)^4$. If a looped diagram has symmetry factors associated with the exchange of internal propagators and vertices that leave the diagram unchanged, divide the value of the diagram by this symmetry factor.
8. Sum all the values from the contributing diagrams to get the value of $i\mathcal{T}$.

Now the momentum-space diagrams of fig. 8.5 are



(a) s-channel (b) t-channel (c) u-channel

From left to right, the three tree diagrams are called the s-, t-, and u-channel

diagrams which is in reference to the Mandelstam variables [5]

$$\begin{aligned} s &= -(p_1 + p_2)^2 = -(p_1' + p_2')^2, \\ t &= -(p_1 - p_1')^2 = -(p_2 - p_2')^2, \\ u &= -(p_1 - p_2')^2 = -(p_2 - p_1')^2. \end{aligned} \tag{8.2.35}$$

which satisfy the relation

$$s + t + u = m_1^2 + m_2^2 + m_{1'}^2 + m_{2'}^2. \tag{8.2.36}$$

At last we finally have our first set of Feynman rules to draw our first Feynman diagrams. Note that the rules we have outlined will refer to tree diagrams. The Feynman rules that we will derive later on will also refer to the tree diagrams of the corresponding theory as well. However, tree diagrams are the most commonly calculated diagrams with loop corrections only being needed in cases where extreme accuracy is paramount. In fact, the results from a tree diagram can provide extremely accurate results on its own. Indeed, we will see in subsequent sections how to compute the scattering and decay processes from tree diagrams and the results will be extremely accurate when compared to experiments.

We will finish this section by computing the differential cross section of the scattering amplitude [5]

$$\begin{aligned} \mathcal{T} &= \lambda^2 \Big[\frac{1}{(p_1 + p_2)^2 + m^2} + \frac{1}{(p_1 - p_1')^2 + m^2} + \frac{1}{(p_1 - p_2')^2 + m^2} \Big] \\ &= \lambda^2 \Big[\frac{1}{-s + m^2} + \frac{1}{-t + m^2} + \frac{1}{-u + m^2} \Big]. \end{aligned} \tag{8.2.37}$$

We then have

$$\begin{aligned} |\mathcal{T}|^2 = \lambda^4 \Big[&\frac{1}{(m^2 - s)^2} + \frac{1}{(m^2 - t)^2} + \frac{1}{(m^2 - u)^2} \\ &+ \frac{2}{(m^2 - s)(m^2 - t)} + \frac{2}{(m^2 - s)(m^2 - u)} + \frac{2}{(m^2 - u)(m^2 - t)} \Big]. \end{aligned} \tag{8.2.38}$$

After integrating over $t_{\min} = -(s-4m^2)$ and $t_{\max} = 0$ with $S = 2$, we get the total cross section [5]

$$\begin{aligned}\sigma &= \frac{\lambda^4}{32\pi s(s-4m^2)}\Big[\frac{2}{m^2} + \frac{s-4m^2}{(s-m^2)^2} - \frac{2}{s-3m^2} \\ &= \frac{4m^2}{(s-m^2)(s-2m^2)}\ln\Big(\frac{s-3m^2}{m^2}\Big)\Big] + O(\lambda^6). \end{aligned} \tag{8.2.39}$$

In the ultrarelativistic limit $s \gg m^2$, the total cross section becomes

$$\sigma = \frac{\lambda^4}{16\pi m^2 s^2}\Big[1 + \frac{7}{2}\frac{m^2}{s} + \dots\Big] + O(\lambda^6), \tag{8.2.40}$$

while in the nonrelativistic limit $s - 4m^2 \ll m^2$, we get [5]

$$\sigma = \frac{25\lambda^4}{1152\pi m^6}\Big[1 - \frac{79}{60}\Big(\frac{s-4m^2}{m^2}\Big) + \dots\Big] + O(\lambda^6). \tag{8.2.41}$$

8.3 Yukawa Theory (Dirac)

In this section we will explore and derive the Feynman rules for *Yukawa theory* in four spacetime dimensions which is the interaction theory of a Dirac field and a real scalar field given by the Lagrangian

$$\begin{aligned}\mathcal{L} &= \mathcal{L}_{KG} + \mathcal{L}_D + \mathcal{L}_{int} \\ &= -\frac{1}{2}\partial^\mu\varphi\partial_\mu\varphi - \frac{1}{2}M^2\varphi^2 + \overline{\mathbf{\Psi}}(i\partial\!\!\!/ - m)\mathbf{\Psi} + g\varphi\overline{\mathbf{\Psi}}\mathbf{\Psi}, \end{aligned} \tag{8.3.1}$$

where g is the dimensionless coupling constant and M and m are the masses of the real scalar particle and Dirac fermion, respectively [5]. To differentiate between the momenta of the scalar particles from the Dirac fermions, we will use k and p to denote the momenta of the scalar particle and the Dirac fermion, respectively. Now the partition function of (8.3.1) is

$$Z(\bar\eta, \eta, J) \propto \exp\Big[ig\int d^4x \Big(\frac{1}{i}\frac{\delta}{\delta J(x)}\Big)\Big(\frac{1}{i}\frac{\delta}{\delta\eta_\alpha(x)}\Big)\Big(\frac{1}{i}\frac{\delta}{\delta\bar\eta_\alpha(x)}\Big)\Big] Z_0(\bar\eta, \eta, J), \tag{8.3.2}$$

where

$$Z_0(\bar{\eta},\eta,J) = \exp\Big[i\int d^4x\, d^4y\, \bar{\eta}(x)S_F(x-y)\eta(y)\Big] \times \exp\Big[\frac{i}{2}\int d^4x\, d^4y\, J(x)\Delta_F(x-y)J(y)\Big], \tag{8.3.3}$$

is normalized by the condition $Z_0(0,0,0) = 1$ [5]. Upon inspection, we see that the interaction Lagrangian $\mathcal{L}_{int}$ is invariant under a U(1) transformation, namely, $\boldsymbol{\Psi} \to e^{-i\alpha}\boldsymbol{\Psi}$. Thus Lagrangian (8.3.1) possesses U(1) symmetry so by Noether's theorem, the Noether current of the Dirac field $j^\mu = \overline{\boldsymbol{\Psi}}\gamma^\mu\boldsymbol{\Psi}$, and hence the Noether charge Q, is conserved [5]. Recall that the Noether charge for the Dirac field was interpreted as the total number of b-type particles (matter) minus the total number of d-type particles (antimatter). Thus when we construct the Feynman diagrams of Yukawa theory, we must be mindful in keeping the Noether charge Q conserved for fermions. Recall that real scalar field does not possess U(1) symmetry thus the scalar particle possesses a neutral charge (it is its own antiparticle).

Yukawa theory was originally formulated by Hideki Yukawa to describe the nuclear force between nucleons mediated by spin-0 particles called *pions* [1]. However, Yukawa theory can apply to other types of interactions between fermions and other spin-0 particles as well so for this section, we will consider the Dirac field $\boldsymbol{\Psi}$ as electrons. Then the Noether charge Q becomes the conservation of electric charge thus the b- and d-type particles will be electrons e^- and positrons e^+, respectively [5]. Under this interpretation, the scalar particle will be electrically neutral.

To begin our derivation of the position-space Feynman rules for Yukawa theory, we first note that since we have Dirac fermions and real scalar particles, we distinguish the two by drawing solid lines for fermions and dashed lines for the real scalar particles [5]. Following from this, the fermion propagator $\frac{1}{i}S_F(x-y)$ will be represented by solid internal lines while the real scalar propagator $\frac{1}{i}\Delta_F(x-y)$ will be represented by dashed internal lines. We must also differentiate between the $\boldsymbol{\Psi}$ source $i\int d^4x\,\bar{\eta}(x)$ and the $\overline{\boldsymbol{\Psi}}$ source $i\int d^4x\,\eta(x)$. To do this, an arrow drawn

on a solid line pointing *towards* a source is $i\int d^4x\,\bar{\eta}(x)$ while an arrow drawn on a solid line pointing *away* from a source is $i\int d^4x\,\eta(x)$. The interaction term $\mathcal{L}_{int}$ contains one $\mathbf{\Psi}$, one $\overline{\mathbf{\Psi}}$, and one φ so the vertex must be constructed with two solid lines and one dashed line. We draw one solid line with an arrow pointing *towards* the vertex and the other solid line with an arrow pointing *away* from the vertex [5]. The vertex will contribute a factor of $ig\int d^4x$. It is clear that many of the Feynman rules for the real scalar field will carry over to Yukawa theory but we must also derive the rules for interaction of the scalar particle with electrons and positrons. To do this, we shall consider the scattering processes $e^-\varphi\to e^-\varphi$ and $e^+e^-\to e^+e^-$ so that we can generalize and abstract the position-space Feynman rules to derive the momentum-space Feynman rules for Yukawa theory.

Let us begin with the scattering process $e^-\varphi\to e^-\varphi$ which is the scattering of an electron and a scalar particle. The source diagram to lowest order in g is given by [5]

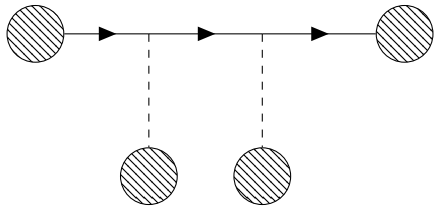


Figure 8.9: Source diagram for $e^-\varphi\to e^-\varphi$.

The scattering amplitude that we wish to calculate is

$$\begin{aligned}\langle f|i\rangle &= \langle 0|\mathrm{T}\hat{a}(\mathbf{p}')\hat{b}_{s'}(\mathbf{p}')\hat{b}_s^\dagger(\mathbf{p})\hat{a}^\dagger(\mathbf{p})|0\rangle\\ &= (i)^4\int d^4x\,d^4y\,d^4z_1\,d^4z_2\,e^{-ik'z_2}(-\partial^2+m^2)\,e^{-ip'x}[\overline{u}_{s'}(\mathbf{p}')(-i\not{\partial}+m)]_\alpha\\ &\times e^{ikz_1}(-\partial^2+m^2)\langle 0|\mathrm{T}\hat{\mathbf{\Psi}}_\alpha(x)\hat{\overline{\mathbf{\Psi}}}_\beta(y)\hat{\varphi}(z_1)\hat{\varphi}(z_2)|0\rangle[(i\overleftarrow{\not{\partial}}+m)u_s(\mathbf{p})e^{ipy}]_\beta.\end{aligned}\tag{8.3.4}$$

We use the functional derivatives to strip off the sources and label them with the spacetime coordinates to produce [5]

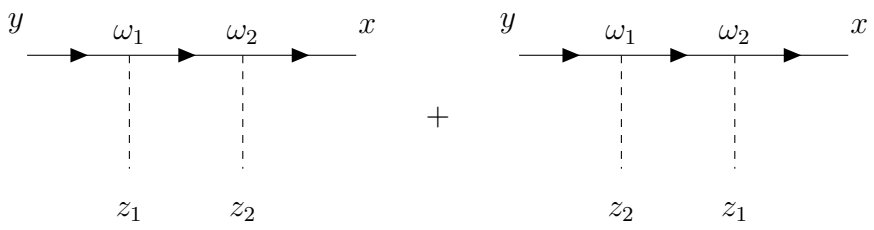


Figure 8.10: Position-space diagram of fig. 8.9

Then we use fig. 8.10 along with our position space rules to calculate

$$\langle 0|\mathrm{T}\hat{\mathbf{\Psi}}_\alpha(x)\hat{\bar{\mathbf{\Psi}}}_\beta(y)\hat{\varphi}(z_1)\hat{\varphi}(z_2)|0\rangle_{\mathrm{C}} = \left(\frac{1}{i}\right)^5 (ig)^2 \int d^4\omega_1\, d^4\omega_2\, [S_F(x-\omega_2) \times S_F(\omega_2-\omega_1)S_F(\omega_2-y)]_{\alpha\beta} \Big(\Delta_F(z_1-\omega_1)\Delta_F(z_2-\omega_2) + \Delta_F(z_2-\omega_1)\Delta_F(z_1-\omega_2)\Big). \tag{8.3.5}$$

Then we insert equation (8.3.5) into (8.3.4) and follow the same computations that was done for the scalar fields to get the scattering matrix

$$i\mathcal{T}_{e^-\varphi\to e^-\varphi} = \frac{1}{i}(ig)^2 \bar{u}_{s'}(\mathbf{p}')\Big[\frac{-\not{p}-\not{k}+m}{-s+m^2} + \frac{-\not{p}+\not{k}'+m}{-u+m^2}\Big]u_s(\mathbf{p}), \tag{8.3.6}$$

where $s=-(p+k)^2$ and $u=-(p-k')^2$ are the Mandelstam variables [5]. With our scattering matrix, the spacetime diagram is now relabeled with its momentum components to form the momentum-space Feynman diagrams for the scattering process $e^-\varphi \to e^-\varphi$ [5]

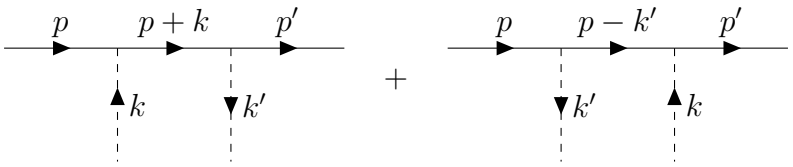


Figure 8.11: Momentum-space diagrams for $e^-\varphi \to e^-\varphi$

Now we consider the scattering of an electron and positron mediated by a scalar particle $e^+e^- \to e^+e^-$. The source diagram to lowest order for this process is

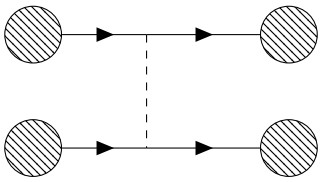


Figure 8.12: Source diagram for $e^+e^- \to e^+e^-$.

The scattering amplitude for this process is [5]

$$\begin{aligned}
\langle f|i\rangle &= \langle 0|\mathrm{T}\hat{b}_{s'}(\mathbf{p}')\hat{d}_{s'}(\mathbf{p}')\hat{b}^\dagger_s(\mathbf{p})\hat{d}^\dagger_s(\mathbf{p})|0\rangle \\
&= (i)^4 \int d^4x_1\, d^4x_2\, d^4y_1\, d^4y_2 \\
&\times e^{-ikp'x_1}[\bar{u}_{s'}(\mathbf{p}')(-\not{\partial} + m)]_{\alpha_1} e^{-ipx_2}[\overline{v}_s(\mathbf{p})(-i\not{\partial} + m)]_{\alpha_2} \\
&\times \langle 0|\mathrm{T}\hat{\mathbf{\Psi}}_{\alpha_1}(x_1)\hat{\overline{\mathbf{\Psi}}}_{\beta_1}(y_1)\hat{\mathbf{\Psi}}_{\alpha_2}(x_2)\hat{\mathbf{\Psi}}(y_2)|0\rangle \\
&\times [(i\overleftarrow{\not{\partial}} + m)v_{s'}(\mathbf{p}')]_{\beta_1} e^{ip'y_1}[(i\overleftarrow{\not{\partial}} + m)u_s(\mathbf{p})]_{\beta_2} e^{ipy_2}. \qquad (8.3.7)
\end{aligned}$$

We use functional derivatives to strip off the sources in fig. 8.12 and label them with spacetime coordinates to get the position-space diagrams [5]

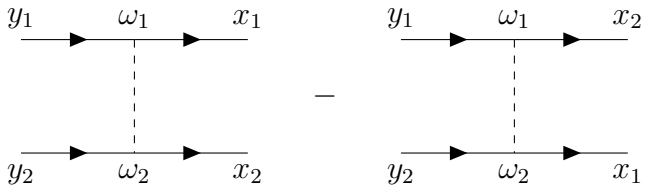


Figure 8.13: Position space diagram of fig. 8.12

Now we use the position-space diagrams to calculate the vacuum expectation value

$$\begin{aligned}
\langle 0|\mathrm{T}\hat{\mathbf{\Psi}}_{\alpha_1}(x_1)\hat{\overline{\mathbf{\Psi}}}_{\beta_1}(y_1)\hat{\mathbf{\Psi}}_{\alpha_2}(x_2)\hat{\mathbf{\Psi}}(y_2)|0\rangle &= \left(\frac{1}{i}\right)^5 (ig)^2 \int d^4\omega_1\, d^4\omega_2\, \Delta_F(\omega_1 - \omega_2) \\
&\times \Big([S_F(x_1 - \omega_1)S_F(\omega_1 - y_1)]_{\alpha_1\beta_1}[S_F(x_2 - \omega_2)S_F(\omega_2 - y_2)]_{\alpha_2\beta_2} \\
&- [S_F(x_2 - \omega_1)S_F(\omega_1 - y_1)]_{\alpha_1\beta_2}[S_F(x_1 - \omega_2)S_F(\omega_2 - y_2)]_{\alpha_2\beta_1}\Big). \qquad (8.3.8)
\end{aligned}$$

Note the negative sign that appears is due to applying the functional derivative with respect to $\bar{\eta}$ which is a Grassmann variable [5]. Then we insert equation

(8.3.8) into equation (8.3.7) and after simplification, we get the matrix

$$i\mathcal{T}_{e^+e^-\to e^+e^-} = \frac{1}{i}(ig)^2\Big[\frac{(\bar{u}_1' u_1)(\bar{v}_2 v_2')}{-t+M^2} - \frac{(\bar{v}_2 u_1)(\bar{u}_1' v_2')}{-s+M^2}\Big], \tag{8.3.9}$$

where we have suppressed the momentum argument and $s = -(p_1+p_2)^2$ and $u = -(p_1-p_1')^2$ [5]. Now we relabel the spacetime diagrams to get the following momentum diagrams for the scattering process $e^+e^- \to e^+e^-$

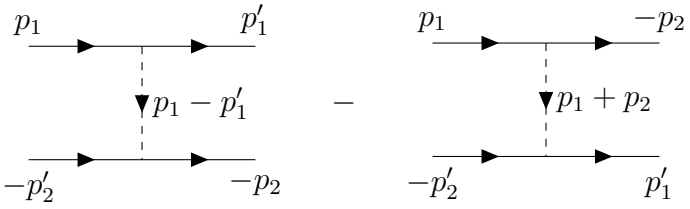
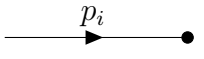


Figure 8.14: Momentum-space diagrams of fig. 8.13

From our calculations, we now recognize the patterns that allow us to formulate the momentum-space Feynman rules for Yukawa theory with Dirac fermions. We now present these rules according to Srednicki [5].

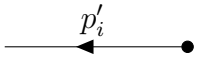
Feynman Rules for Yukawa Theory (Dirac Fermion)

1. Incoming electrons are solid lines with an arrow pointing towards the vertex and are labeled with its four-momentum p_i. The assigned factor is $u_{s_i}(\mathbf{p}_i)$.



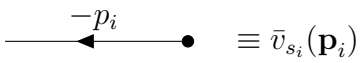
$\equiv u_{s_i}(\mathbf{p}_i)$

2. Outgoing electrons are solid lines with an arrow pointing away from the vertex and are labeled with its four-momentum p_i'. The assigned factor is $\bar{u}_{s_i'}(\mathbf{p}_i')$.



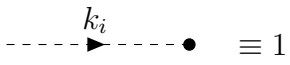
$\equiv \bar{u}_{s_i'}(\mathbf{p}_i')$

3. Incoming positrons are solid lines with an arrow pointing away from the vertex and are labeled with its four-momentum $-p_i$. The assigned factor is $\bar{v}_{s_i}(\mathbf{p}_i)$.



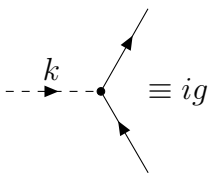
4. Outgoing positrons are solid lines with an arrow pointing towards the vertex and are labeled with its four-momentum $-p_i'$. The assigned factor is $v_{s_i'}(\mathbf{p}_i')$.

5. Incoming scalars are dashed lines with an arrow pointing towards the vertex and are labeled with its four-momentum k_i. The assigned factor is 1.



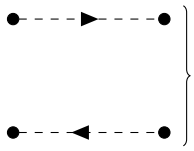
6. Outgoing scalars are dashed lines with an arrow pointing away from the vertex and are labeled with its four-momentum k_i'. The assigned factor is 1.

7. Construct each vertex from two solid lines and one dashed line. One of the solid lines must have an arrow pointing towards the vertex while the other must have an arrow pointing away from the vertex. The arrow on the dashed line can point in any direction. Draw all topologically inequivalent diagrams using the vertex, external lines, and any extra internal lines when needed. The assigned factor is ig.



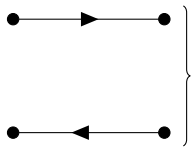
8. Conserve four-momentum at each vertex by following the arrows of each external line.
9. Assign the appropriate propagators to each internal line:

- Internal scalar:



$$\equiv \frac{1}{i}\frac{1}{p^2+m^2-i\epsilon}$$

- Internal fermion:



$$\equiv \frac{1}{i}\frac{-\not{p}+m}{p^2+m^2-i\epsilon}$$

10. Arrange the diagrams in standard form: all fermion lines are horizontal with their arrows pointing from left to right and label the left endpoints of the contributing diagrams with the same, arbitrarily fixed order from top to bottom. Contract the spinor indices by starting at the fermion line with the arrow pointing away from the vertex, follow the complete fermion line backwards, and write down from left to right the assigned factors that correspond to the vertices and propagators. The first factor should be either $\bar{u}$ or $\bar{v}$ and the last factor should be either u or v. Repeat for other fermion lines if necessary.
11. To determine the sign of a tree diagram, draw all the contributing diagrams in standard form. If the ordering of the right endpoints are an even (odd) permutation of the fixed ordering then the sign is positive (negative). This method can be reversed to so that we fix the right endpoints and observe the left endpoints instead.
12. Sum all the values from the contributing diagrams to get the value of $i\mathcal{T}$.

Let us continue further with the two scattering processes we have constructed. For

the $e^{-}\varphi \to e^{-}\varphi$ process, we calculated the matrix

$$\mathcal{T}_{e^{-}\varphi\to e^{-}\varphi} = g^2 \bar{u}_{s'}(\mathbf{p}')\Big[\frac{-\not{p} - \not{k} + m}{-s + m^2} + \frac{-\not{p} + \not{k}' + m}{-u + m^2}\Big] u_s(\mathbf{p}). \tag{8.3.10}$$

We now wish to compute $|\mathcal{T}|^2$ so that we can insert it into the differential cross section formula. It is a bit cumbersome to directly calculate $\mathcal{T}\mathcal{T}^*$ so we will apply a clever trick. First, observe that we can immediately apply the spinor equation $(\not{p} + m)u_s(\mathbf{p})$ to get [5]

$$\mathcal{T}_{e^{-}\varphi\to e^{-}\varphi} = g^2 \bar{u}_{s'}(\mathbf{p}')\Big[\frac{-\not{k} + 2m}{m^2 - s} + \frac{\not{k}' + 2m}{m^2 - u}\Big] u_s(\mathbf{p}). \tag{8.3.11}$$

Now upon close inspection, equation (8.3.11) is essentially, by reading from left to right, a row vector times a 4×4 matrix times a column vector. Let us abstract this idea and let [5]

$$A = g^2\Big[\frac{-\not{k} + 2m}{m^2 - s} + \frac{\not{k}' + 2m}{m^2 - u}\Big], \tag{8.3.12}$$

then suppress the spin state and momentum argument to get the matrix form of equation (8.3.11)

$$\mathcal{T} = \bar{u}' A u. \tag{8.3.13}$$

Now we can easily calculate $\mathcal{T}^*$ which is

$$\begin{aligned} \mathcal{T}^* &= \overline{\mathcal{T}} \\ &= \overline{\bar{u}' A u} \\ &= \bar{u} A u', \end{aligned} \tag{8.3.14}$$

where in the last equality we have used the identity $\bar{A} = A$. Then we have

$$\begin{aligned}
|\mathcal{T}|^2 &= \mathcal{T}\overline{\mathcal{T}} \\
&= (\bar{u}' A u)(\bar{u} A u') \\
&= \sum_{\alpha\beta\gamma\delta} \bar{u}'|\alpha\rangle\langle\alpha|A|\beta\rangle\langle\beta|u\bar{u}|\gamma\rangle\langle\gamma|A|\delta\rangle\langle\delta|u' \\
&= \sum_{\alpha\beta\gamma\delta} \langle\delta|u'\bar{u}'|\alpha\rangle\langle\alpha|A|\beta\rangle\langle\beta|u\bar{u}|\gamma\rangle\langle\gamma|A|\delta\rangle \\
&= \sum_{\delta} \langle\delta|u'\bar{u}' A u\bar{u} A|\delta\rangle \\
&= \mathrm{Tr}\,[(u'\bar{u}')A(u\bar{u})A]. \qquad (8.3.15)
\end{aligned}$$

From here, we could use the spinor and gamma matrix identities for the spin state but usually this is not done because knowing the spin states of the scattered particles is not the main focus. In fact, it can be extremely difficult to measure experimentally or to prepare particles in a certain spin state [5]. Since we will know the momentum of the particles after scattering but not the spin states, we will sum $\mathcal{T}^2$ over the final two spin states and average over the initial spin states [8]. Mathematically, we will calculate

$$\begin{aligned}
\langle|\mathcal{T}^2|\rangle &= \frac{1}{2}\sum_{s,s'} \mathrm{Tr}\,[(u'\bar{u}')A(u\bar{u})A] \\
&= \frac{1}{2}\mathrm{Tr}\,[(-\not{p}' + m)A(-\not{p} + m)A]. \qquad (8.3.16)
\end{aligned}$$

Now observe

$$\begin{aligned}
\mathrm{Tr}\,[(-\not{p}' + m)A(-\not{p} + m)A] &= g^4 \mathrm{Tr}\Big[(-\not{p}' + m)\Big(\frac{-\not{k} + 2m}{m^2 - s} + \frac{\not{k}' + 2m}{m^2 - u}\Big)(-\not{p} + m) \\
&\quad \times (-\not{p}' + m)\Big(\frac{-\not{k} + 2m}{m^2 - s} + \frac{\not{k}' + 2m}{m^2 - u}\Big)\Big] \\
&= g^4\Big[\frac{\langle\Phi_{ss}\rangle}{(m^2 - s)^2} + \frac{\langle\Phi_{su}\rangle + \langle\Phi_{us}\rangle}{(m^2 - s)(m^2 - u)} + \frac{\langle\Phi_{uu}\rangle}{(m^2 - u)^2}\Big], \\
& \qquad (8.3.17)
\end{aligned}$$

where [5]

$$\begin{aligned}
\langle\Phi_{ss}\rangle &= \frac{1}{2}\mathrm{Tr}\left[(-\not{p}' + m)(-\not{k} + 2m)(-\not{p} + m)(-\not{k} + 2m)\right],\\
\langle\Phi_{su}\rangle &= \frac{1}{2}\mathrm{Tr}\left[(-\not{p}' + m)(-\not{k} + 2m)(-\not{p} + m)(\not{k}' + 2m)\right],\\
\langle\Phi_{us}\rangle &= \frac{1}{2}\mathrm{Tr}\left[(-\not{p}' + m)(\not{k}' + 2m)(-\not{p} + m)(-\not{k} + 2m)\right],\\
\langle\Phi_{uu}\rangle &= \frac{1}{2}\mathrm{Tr}\left[(-\not{p}' + m)(\not{k}' + 2m)(-\not{p} + m)(\not{k}' + 2m)\right].
\end{aligned} \tag{8.3.18}$$

To solve the above equations, we apply our knowledge of spinor and gamma matrix identities and make liberal use of the Mandelstam variable identities [5]

$$\begin{aligned}
s &= -(p+k)^2 = -(p'+k')^2,\\
t &= -(p-p')^2 = -(k-k')^2,\\
u &= -(p-k')^2 = -(k-p')^2,\\
s+t+u &= 2m^2 + 2M^2,
\end{aligned} \tag{8.3.19}$$

which also imply the following identities

$$\begin{aligned}
pk = p'k' &= -\frac{1}{2}(s - m^2 - M^2),\\
pp' &= \frac{1}{2}(t - 2m^2),\\
kk' &= \frac{1}{2}(t - 2M^2),\\
pk' = p'k &= -\frac{1}{2}(u - m^2 - M^2).
\end{aligned} \tag{8.3.20}$$

Before we hastily attempt to solve all four spin averages, note that if we make the replacement $k \to -k'$ in $\langle\Phi_{ss}\rangle$, we get the transformation $\langle\Phi_{ss}\rangle \to \langle\Phi_{uu}\rangle$. Similarly, making the same replacement $k \to -k'$ in $\langle\Phi_{su}\rangle$ results in the transformation $\langle\Phi_{su}\rangle \to \langle\Phi_{us}\rangle$. It is obvious that we can make this transformation in the reverse direction as well thus we see that $k \leftrightarrow -k'$ is analogous to $s \leftrightarrow u$ [5]. This significantly relieves our computational burden since we only have to compute $\langle\Phi_{ss}\rangle$ ($\langle\Phi_{uu}\rangle$) and $\langle\Phi_{su}\rangle$ ($\langle\Phi_{us}\rangle$) then make the appropriate transformation to get

the other terms. Thus we have

$$\begin{aligned}\langle\Phi_{ss}\rangle &= \frac{1}{2}\mathrm{Tr}\left[(-\not{p}'+m)(-\not{k}+2m)(-\not{p}+m)(-\not{k}+2m)\right]\\ &= \frac{1}{2}\Big[\mathrm{Tr}\left[\not{p}'\not{k}\not{p}\not{k}\right]+4m^2\mathrm{Tr}\left[\not{p}'\not{p}\right]+4m^2\mathrm{Tr}\left[\not{p}'\not{k}\right]+2m^2\mathrm{Tr}\left[\not{p}\not{k}\right]\\ &+ m^2\mathrm{Tr}\left[\not{k}\not{k}\right]+2m^2\mathrm{Tr}\left[\not{k}\not{p}\right]+4m^4\mathrm{Tr}\left[1\right]\Big]\\ &= -su+m^2u+9m^2s+7m^4-8M^2m^2+M^4, \end{aligned} \tag{8.3.21}$$

where in the last step, we rewrote the entire expression in terms of s and u using $t = 2m^2 + 2M^2 - s - u$ [5]. Similar computations yield

$$\langle\Phi_{su}\rangle = su+3m^2u+3m^2s+9m^4-8M^2m^2-M^4. \tag{8.3.22}$$

Then applying the Mandelstam variable change $s \leftrightarrow u$ finally yields [5]

$$\begin{aligned}\langle\Phi_{ss}\rangle &= -su+m^2u+9m^2s+7m^4-8M^2m^2+M^4,\\ \langle\Phi_{uu}\rangle &= -su+m^2s+9m^2u+7m^4-8M^2m^2+M^4,\\ \langle\Phi_{su}\rangle &= su+3m^2u+3m^2s+9m^4-8M^2m^2-M^4,\\ \langle\Phi_{us}\rangle &= su+3m^2s+3m^2u+9m^4-8M^2m^2-M^4. \end{aligned} \tag{8.3.23}$$

Now let us compute $\langle|\mathcal{T}|^2\rangle$ for the process $e^+e^- \to e^+e^-$. Recall that the $\mathcal{T}_{e^+e^-\to e^+e^-}$ matrix was

$$\mathcal{T}_{e^+e^-\to e^+e^-} = g^2\Big[\frac{(\bar{u}_1'u_1)(\bar{v}_2v_2')}{M^2-t}-\frac{(\bar{v}_2u_1)(\bar{u}_1'v_2')}{M^2-s}\Big], \tag{8.3.24}$$

where $s = -(p_1+p_2)^2$ and $t = -(p_1-p_1')^2$. Since the $\mathcal{T}$ is not in an abstract matrix form, we proceed directly with calculating $\overline{\mathcal{T}}$

$$\begin{aligned}\overline{\mathcal{T}} &= g^2\Big[\overline{\frac{(\bar{u}_1'u_1)(\bar{v}_2v_2')}{M^2-t}}-\overline{\frac{(\bar{v}_2u_1)(\bar{u}_1'v_2')}{M^2-s}}\Big]\\ &= g^2\Big[\frac{(\bar{u}_1u_1')(\bar{v}_2'v_2)}{M^2-t}-\frac{(\bar{u}_1v_2)(\bar{v}_2'u_1')}{M^2-s}\Big]. \end{aligned} \tag{8.3.25}$$

Then

$$|\mathcal{T}|^2 = g^4\Big[\frac{(\bar{u}_1'u_1)(\bar{v}_2v_2')(\bar{u}_1u_1')(\bar{v}_2'v_2)}{(M^2-t)^2} - \frac{(\bar{u}_1'u_1)(\bar{v}_2v_2')(\bar{u}_1v_2)(\bar{v}_2'u_1')}{(M^2-t)(M^2-s)} - \frac{(\bar{v}_2u_1)(\bar{u}_1'v_2')(\bar{u}_1u_1')(\bar{v}_2'v_2)}{(M^2-s)(M^2-t)} + \frac{(\bar{v}_2u_1)(\bar{u}_1'v_2')(\bar{u}_1v_2)(\bar{v}_2'u_1')}{(M^2-s)^2}\Big]. \tag{8.3.26}$$

To evaluate the above expression into something more meaningful, observe that the numerator of the first term can be rewritten as [5]

$$\begin{aligned}(\bar{u}_1'u_1)(\bar{v}_2v_2')(\bar{u}_1u_1')(\bar{v}_2'v_2) &= \sum_{\alpha\beta\gamma\delta} \bar{u}_1'|\alpha\rangle\langle\alpha|u_1\bar{v}_2|\beta\rangle\langle\beta|v_2'\bar{u}_1|\gamma\rangle\langle\gamma|u_1'\bar{v}_2'|\delta\rangle\langle\delta|v_2 \\ &= \sum_{\alpha\beta\gamma\delta} \langle\alpha|u_1\bar{u}_1|\gamma\rangle\langle\gamma|u_1'\bar{u}_1'|\alpha\rangle\langle\beta|v_2'\bar{v}_2'|\delta\rangle\langle\delta|v_2\bar{v}_2|\beta\rangle \\ &= \sum_{\alpha\beta} \langle\alpha|u_1\bar{u}_1u_1'\bar{u}_1'|\alpha\rangle\langle\beta|v_2'\bar{v}_2'v_2\bar{v}_2|\beta\rangle \\ &= \mathrm{Tr}\,[u_1\bar{u}_1u_1'\bar{u}_1']\,\mathrm{Tr}\,[v_2'\bar{v}_2'v_2\bar{v}_2],\end{aligned} \tag{8.3.27}$$

and similarly the numerator in the fourth term can be written as

$$(\bar{v}_2u_1)(\bar{u}_1'v_2')(\bar{u}_1v_2)(\bar{v}_2'u_1') = \mathrm{Tr}\,[u_1\bar{u}_1v_2\bar{v}_2]\,\mathrm{Tr}\,[v_2'\bar{v}_2'u_1'\bar{u}_1']. \tag{8.3.28}$$

The numerator in the third term is [5]

$$\begin{aligned}(\bar{v}_2u_1)(\bar{u}_1'v_2')(\bar{u}_1u_1')(\bar{v}_2'v_2) &= \sum_{\alpha\beta\gamma\delta} \bar{v}_2|\alpha\rangle\langle\alpha|u_1\bar{u}_1'|\beta\rangle\langle\beta|v_2'\bar{u}_1|\gamma\rangle\langle\gamma|u_1'\bar{v}_2'|\delta\rangle\langle\delta|v_2 \\ &= \sum_{\alpha\beta\gamma\delta} \langle\alpha|u_1\bar{u}_1|\gamma\rangle\langle\gamma|u_1'\bar{u}_1'|\beta\rangle\langle\beta|v_2'\bar{v}_2'|\delta\rangle\langle\delta|v_2\bar{v}_2|\alpha\rangle \\ &= \sum_{\alpha} \langle\alpha|u_1\bar{u}_1u_1'\bar{u}_1'v_2'\bar{v}_2'v_2\bar{v}_2|\alpha\rangle \\ &= \mathrm{Tr}\,[u_1\bar{u}_1u_1'\bar{u}_1'v_2'\bar{v}_2'v_2\bar{v}_2],\end{aligned} \tag{8.3.29}$$

and similarly the numerator in the second term is

$$(\bar{u}_1'u_1)(\bar{v}_2v_2')(\bar{u}_1v_2)(\bar{v}_2'u_1') = \mathrm{Tr}\,[u_1\bar{u}_1v_2\bar{v}_2v_2'\bar{v}_2'u_1'\bar{u}_1']. \tag{8.3.30}$$

Thus $|\mathcal{T}|^2$ becomes

$$|\mathcal{T}|^2 = g^4\Big[\frac{\Phi_{ss}}{(M^2-s)^2} - \frac{\Phi_{st}+\Phi_{ts}}{(M^2-s)(M^2-t)} + \frac{\Phi_{tt}}{(M^2-t)^2}\Big], \tag{8.3.31}$$

where [5]

$$\begin{aligned}
\Phi_{ss} &= \mathrm{Tr}\,[u_1\bar{u}_1u_1'\bar{u}_1']\,\mathrm{Tr}\,[v_2'\bar{v}_2'v_2\bar{v}_2],\\
\Phi_{st} &= \mathrm{Tr}\,[u_1\bar{u}_1u_1'\bar{u}_1'v_2'\bar{v}_2'v_2\bar{v}_2],\\
\Phi_{ts} &= \mathrm{Tr}\,[u_1\bar{u}_1v_2\bar{v}_2v_2'\bar{v}_2'u_1'\bar{u}_1'],\\
\Phi_{tt} &= \mathrm{Tr}\,[u_1\bar{u}_1v_2\bar{v}_2]\,\mathrm{Tr}\,[v_2'\bar{v}_2'u_1'\bar{u}_1'].
\end{aligned} \tag{8.3.32}$$

Now we average the initial spins and sum over the final spins

$$\begin{aligned}
\langle|\mathcal{T}|^2\rangle &= \frac{1}{4}\sum_{s_1,s_2,s_1',s_2'}|\mathcal{T}|^2\\
&= g^4\Big[\frac{\langle\Phi_{ss}\rangle}{(M^2-s)^2} - \frac{\langle\Phi_{st}\rangle+\langle\Phi_{ts}\rangle}{(M^2-s)(M^2-t)} + \frac{\langle\Phi_{tt}\rangle}{(M^2-t)^2}\Big],
\end{aligned} \tag{8.3.33}$$

where [5]

$$\begin{aligned}
\langle\Phi_{ss}\rangle &= \frac{1}{4}\mathrm{Tr}\,[(-\not{p}_1+m)(-\not{p}_2-m)]\,\mathrm{Tr}\,[(-\not{p}_2'-m)(-\not{p}_1'+m)],\\
\langle\Phi_{st}\rangle &= \frac{1}{4}\mathrm{Tr}\,[(-\not{p}_1+m)(-\not{p}_1'+m)(-\not{p}_2'-m)(-\not{p}_2-m)],\\
\langle\Phi_{ts}\rangle &= \frac{1}{4}\mathrm{Tr}\,[(-\not{p}_1+m)(-\not{p}_2-m)(-\not{p}_2'-m)(-\not{p}_1'+m)],\\
\langle\Phi_{tt}\rangle &= \frac{1}{4}\mathrm{Tr}\,[(-\not{p}_1+m)(-\not{p}_1'-m)]\,\mathrm{Tr}\,[(-\not{p}_2'-m)(-\not{p}_2-m)].
\end{aligned} \tag{8.3.34}$$

Once again we notice that making the change $p_2 \leftrightarrow -p_1'$ results in the transformation $s \leftrightarrow t$ thus we reduce our workload to calculating $\langle\Phi_{ss}\rangle$ ($\langle\Phi_{tt}\rangle$) and $\langle\Phi_{st}\rangle$

($\langle\langle\Phi_{ts}\rangle\rangle$). The Mandelstam variables for this process is

$$\begin{aligned} s &= -(p_1+p_2)^2 = -(p_1'+p_2')^2, \\ t &= -(p_1-p_1')^2 = -(p_2-p_2')^2, \\ u &= -(p_1-p_2')^2 = -(p_2-p_1')^2, \\ s+t+u &= 4m^2, \end{aligned} \tag{8.3.35}$$

which implies

$$\begin{aligned} p_1p_2 = p_1'p_2' &= -\frac{1}{2}(s-2m^2), \\ p_1p_1' = p_2p_2' &= \frac{1}{2}(t-2m^2), \\ p_1p_2' = p_1'p_2 &= \frac{1}{2}(u-2m^2). \end{aligned} \tag{8.3.36}$$

Next

$$\begin{aligned} \langle\Phi_{ss}\rangle &= \frac{1}{4}\mathrm{Tr}\left[(-\not{p}_1+m)(-\not{p}_2-m)\right]\mathrm{Tr}\left[(-\not{p}_2'-m)(-\not{p}_1'+m)\right] \\ &= \frac{1}{4}\Big[\big(\mathrm{Tr}\left[\not{p}_1\not{p}_2\right]-m^2\mathrm{Tr}\left[1\right]\big)\big(\mathrm{Tr}\left[\not{p}_2'\not{p}_1'\right]-m^2\mathrm{Tr}\left[1\right]\big)\Big] \\ &= \frac{1}{4}\Big[(-4(p_1p_2)-4m^2)(-4(p_2'p_1')-4m^2)\Big] \\ &= 4\big[(p_1p_2)^2+2m^2(p_1p_2)+m^4\big] \\ &= 4((p_1p_2)+m^2)^2 \\ &= (s-4m^2)^2, \end{aligned} \tag{8.3.37}$$

and similarly [5]

$$\begin{aligned} \langle\Phi_{st}\rangle &= \frac{1}{4}\mathrm{Tr}\left[(-\not{p}_1+m)(-\not{p}_1'+m)(-\not{p}_2'-m)(-\not{p}_2-m)\right] \\ &= \frac{1}{4}\Big[\mathrm{Tr}\left[\not{p}_1\not{p}_1'\not{p}_2'\not{p}^2\right]+m^2\mathrm{Tr}\left[\not{p}_2'\not{p}_2\right]-m^2\mathrm{Tr}\left[\not{p}_1'\not{p}_2'\right] \\ &- m^2\mathrm{Tr}\left[\not{p}_1'\not{p}_2\right]-m^2\mathrm{Tr}\left[\not{p}_1\not{p}_2'\right]-m^2\mathrm{Tr}\left[\not{p}_1\not{p}_2\right]+m^2\mathrm{Tr}\left[\not{p}_1\not{p}_1'\right]+m^4\mathrm{Tr}\left[1\right]\Big] \\ &= -\frac{1}{2}st+2m^2u. \end{aligned} \tag{8.3.38}$$

Thus we have [5]

$$\begin{aligned}
\langle \Phi_{ss} \rangle &= (s - 4m^2)^2, \\
\langle \Phi_{st} \rangle &= -\frac{1}{2} st + 2m^2 u, \\
\langle \Phi_{ts} \rangle &= -\frac{1}{2} st + 2m^2 u, \\
\langle \Phi_{tt} \rangle &= (t - 4m^2)^2.
\end{aligned} \tag{8.3.39}$$

From here, it is a straightforward application of $\langle |\mathcal{T}|^2 \rangle$ into the formula for the differential cross section to get the spin averaged cross section or, depending on the nature of the problem at hand, we can compute the decay rate. The two processes that we have covered here, namely $e^- \varphi \to e^- \varphi$ and $e^+ e^- \to e^+ e^-$, will serve as a template and guide on how to approach calculating $\langle |\mathcal{T}|^2 \rangle$ using Feynman diagrams for interactions in Yukawa theory involving Dirac fermions. We will cover one more example of a process in Yukawa theory to demonstrate the utility of the Feynman diagrams.

Yukawa Potential

Consider the scattering of *distinguishable* fermions (we will continue to use electrons) in the nonrelativistic limit $e^- e^- \to e^- e^-$. The Feynman diagram for this process is

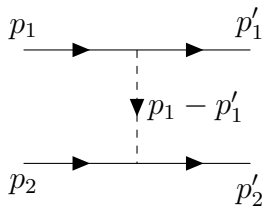


Figure 8.15: Yukawa potential $e^- e^- \to e^- e^-$ process

The scattering amplitude is

$$\begin{aligned} i\mathcal{T}_{e^-e^-\to e^-e^-} &= \bar{u}_1'(ig)u_1\frac{1}{i}\frac{1}{(p_1-p_1')^2+M^2-i\epsilon}\bar{u}_2'(ig)u_2 \\ &= \frac{1}{i}(ig)^2\Big[\frac{(\bar{u}_1'u_1)(\bar{u}_2'u_2)}{-t+M^2}\Big] \\ &= \frac{ig^2}{-t+M^2}(2m\delta_{s,s'})_1(2m\delta_{s,s'})_2 \end{aligned} \tag{8.3.40}$$

where $t=-(p_1-p_1')^2$. In the nonrelativistic limit, we have

$$\begin{aligned} p_1 &= (m,\mathbf{p}_1), & p_1' &= (m,\mathbf{p}_1'), \\ p_2 &= (m,\mathbf{p}_2), & p_2' &= (m,\mathbf{p}_2'), \end{aligned} \tag{8.3.41}$$

thus

$$t=-(p_1-p_1')^2=-\left|\mathbf{p}_1-\mathbf{p}_1'\right|^2, \tag{8.3.42}$$

and the scattering amplitude reduces to

$$\begin{aligned} i\mathcal{T} &= \frac{ig^2}{\left|\mathbf{p}_1-\mathbf{p}_1'\right|^2+M^2}(2m\delta_{s,s'})_1(2m\delta_{s,s'})_2 \\ &= \frac{ig^2}{\left|\mathbf{p}\right|^2+M^2}(2m\delta_{s,s'})_1(2m\delta_{s,s'})_2, \end{aligned} \tag{8.3.43}$$

where $\mathbf{p}=\mathbf{p}_1-\mathbf{p}_1'$ [8]. From nonrelativistic quantum mechanics, the scattering amplitude is [8]

$$\langle p'|i\mathrm{T}|p\rangle = -i\tilde{V}(\mathbf{p})(2\pi)\delta(E_{\mathbf{p}'}-E_{\mathbf{p}}). \tag{8.3.44}$$

By comparing with the Born approximation, we see that the potential for this process in momentum-space is

$$\tilde{V}(\mathbf{p})=\frac{-g^2}{\left|\mathbf{p}\right|^2+M^2}. \tag{8.3.45}$$

The potential in position-space is

$$\begin{aligned} V(\mathbf{x}) &= \int \frac{d^3p}{(2\pi)^3} \frac{-g^2}{|\mathbf{p}|^2 + M^2} e^{i\mathbf{p}\cdot\mathbf{x}} \\ &= -\frac{-g^2}{4ir\pi^2} \int dp \, \frac{pe^{ipr}}{p^2 + M^2} \\ &= -\frac{-g^2}{4\pi} \frac{1}{r} e^{-Mr}, \end{aligned}$$

hence

$$V(r) = -\frac{-g^2}{4\pi} \frac{1}{r} e^{-Mr}, \tag{8.3.46}$$

which is the *Yukawa potential*. We see that the scattering of fermions is attractive from the minus sign.

Let us now consider the scattering of a fermion and an antifermion. The Feynman diagram for this process is

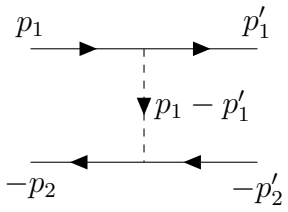


Figure 8.16: Yukawa potential $e^+e^- \to e^+e^-$ process

The scattering amplitude is

$$\begin{aligned} i\mathcal{T}_{e^+e^-\to e^+e^-} &= -\bar{u}_1'(ig)u_1 \frac{1}{i} \frac{1}{(p_1 - p_1')^2 + M^2 - i\epsilon} \bar{v}_2'(ig)\bar{v}_2 \\ &= -\frac{1}{i}(ig)^2 \Big[\frac{(\bar{u}_1' u_1)(\bar{v}_2' \bar{v}_2)}{-t + M^2} \Big] \\ &= -\frac{ig^2}{-t + M^2} (2m\delta_{s,s'})_1 (-2m\delta_{s,s'})_2 \\ &= \frac{ig^2}{-t + M^2} (2m\delta_{s,s'})_1 (2m\delta_{s,s'})_2, \end{aligned} \tag{8.3.47}$$

where $t = -(p_1 - p_1')^2$ and the minus sign comes from the exchange of antisymmetric operators. We see that the scattering amplitude is the same as the case for the

scattering of two fermions thus we get the same attractive Yukawa potential

$$V(r) = -\frac{-g^2}{4\pi}\frac{1}{r}e^{-Mr}, \tag{8.3.48}$$

with range $1/M = \hbar/Mc$, namely, the Compton wavelength of the force carrying boson [8]. If we consider the scattering of two antifermions, we will get the same scattering amplitude therefore the same attractive Yukawa potential because we are switching $\bar{u}u$ with $\bar{v}v$ which cancels the minus sign. Thus the Yukawa poential is a short range interaction that is universally attractive [15]. By studying the diagrams we have computed, we see that the interaction is mediated by a massive boson which carries the force to the other particles.

8.4 Quantum Electrodynamics (Spinor)

At last we finally arrive at the pièce de résistance, quantum electrodynamics. Specifically, we will explore the theory of the electromagnetic field interacting with Dirac spinor fields which is what is most commonly referred to when the term quantum electrodynamics is used. Thus when we invoke the term quantum electrodynamics, or QED for short, we will actually refer to quantum spinor electrodynamics. When constructing the proper Lagrangian, we impose the extra condition that the Lagrangian preserves the gauge symmetry of the electromagnetic field so that it produces the correct equations of motion [5]. We also require that the Lagrangian preserves the conservation of the electromagnetic current j^μ. Thus we let j^μ be proportional to the Noether current from the U(1) symmetry of the Dirac field, namely,

$$j^\mu = e\overline{\mathbf{\Psi}}\gamma^\mu\mathbf{\Psi}, \tag{8.4.1}$$

where e is the elementary charge which is dimensionless in natural units and makes the theory renormalizable [5]. Then the interaction part of the Lagrangian is

$$\mathcal{L}_{int} = e\overline{\mathbf{\Psi}}\gamma^\mu\mathbf{\Psi}A_\mu, \tag{8.4.2}$$

thus we have the Lagrangian for quantum electrodynamics [5]

$$\begin{aligned}\mathcal{L} &= \mathcal{L}_{EM} + \mathcal{L}_D + \mathcal{L}_{int} \\ &= -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} + \overline{\mathbf{\Psi}}(i\not{\partial} - m)\mathbf{\Psi} + e\overline{\mathbf{\Psi}}\gamma^\mu\mathbf{\Psi}A_\mu. \end{aligned} \tag{8.4.3}$$

We showed earlier that the electromagnetic field tensor and hence the Lagrangian for the electromagnetic field was gauge invariant but the Dirac field has not been accounted for in terms of gauge transformations. To see that Lagrangian (8.4.3) is gauge invariant, we must consider gauge transformations on the Dirac field of the form

$$\mathbf{\Psi}(x) \to \mathbf{\Psi}'(x) = \exp[-ie\alpha(x)]\mathbf{\Psi}(x), \tag{8.4.4}$$

$$\overline{\mathbf{\Psi}}(x) \to \overline{\mathbf{\Psi}}'(x) = \exp[ie\alpha(x)]\overline{\mathbf{\Psi}}(x), \tag{8.4.5}$$

along with the usual gauge transformation on the four-potential

$$A^\mu(x) \to A'^\mu(x) = A^\mu(x) - \partial^\mu\alpha(x), \tag{8.4.6}$$

where the phase angle $\alpha(x)$ now depends locally on spacetime [5]. Before we prove that Lagrangian (8.4.3) is invariant under these gauge transformations, we rewrite the Lagrangian in the following form

$$\mathcal{L} = -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} + \overline{\mathbf{\Psi}}(i\not{D} - m)\mathbf{\Psi}, \tag{8.4.7}$$

where $D_\mu = \partial_\mu - ieA_\mu$ is the *gauge covariant derivative* or *covariant derivative*. For those familiar with differential geometry, the vector field $A_\mu(x)$ in the covariant derivative is called a *connection* [8]. Now we prove that (8.4.7) is gauge invariant.

Proof. Let $\mathcal{L}$ be defined as in equation (8.4.7) and consider the gauge transforma-

tions

$$\begin{aligned}
\mathbf{\Psi}(x) &\to \mathbf{\Psi}'(x) = \exp[-ie\alpha(x)]\mathbf{\Psi}(x), \\
\overline{\mathbf{\Psi}}(x) &\to \overline{\mathbf{\Psi}}'(x) = \exp[ie\alpha(x)]\overline{\mathbf{\Psi}}(x), \\
A^{\mu}(x) &\to A'^{\mu}(x) = A^{\mu}(x) - \partial^{\mu}\alpha(x).
\end{aligned}$$

Under the gauge transformations, the covariant derivative transforms as

$$\begin{aligned}
D_{\mu}\mathbf{\Psi} &\to D'_{\mu}\mathbf{\Psi}' \\
&= \big(\partial_{\mu} - ie(A^{\mu} - \partial^{\mu}\alpha)\big)\big(\exp[-ie\alpha]\mathbf{\Psi}\big) \\
&= \exp[-ie\alpha]\big(\partial_{\mu} - ie\partial_{\mu}\alpha - ie(A^{\mu} - \partial^{\mu}\alpha)\big)\mathbf{\Psi} \\
&= \exp[-ie\alpha]\big(\partial_{\mu} - ieA_{\mu}\big)\mathbf{\Psi} \\
&= \exp[-ie\alpha]D_{\mu}\mathbf{\Psi}.
\end{aligned}$$

Thus

$$\begin{aligned}
&\mathcal{L} \to \mathcal{L}' \\
&= -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} + \overline{\mathbf{\Psi}}'(i\not{D}' - m)\mathbf{\Psi}' \\
&= -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} + \exp[ie\alpha]\overline{\mathbf{\Psi}}(i\exp[-ie\alpha]\not{D} - \exp[-ie\alpha]m)\mathbf{\Psi} \\
&= -\frac{1}{4}F^{\mu\nu}F_{\mu\nu} + \overline{\mathbf{\Psi}}(i\not{D} - m)\mathbf{\Psi} \\
&= \mathcal{L}.
\end{aligned}$$

Therefore the Lagrangian is gauge invariant. □

To see that our Lagrangian produces the correct equation of motion that describes the interaction of fermions and the electromagnetic field, we vary the action $\mathcal{S}$ with respect to $\overline{\mathbf{\Psi}}$ and impose the stationary action principle to get

$$(i\not{D}_{\mu} - m)\mathbf{\Psi} = 0, \tag{8.4.8}$$

which is the Dirac equation for a Dirac field coupled to the electromagnetic field. The coupling of the electromagnetic field to the Dirac field involves the replacement $\partial_\mu \to D_\mu$ which is called the *minimal coupling* prescription [15]. Now we vary the action with respect to A^μ to get

$$\begin{aligned}(g^{\mu\nu}\partial^2 - \partial^\mu\partial^\nu)A_\nu + e\overline{\mathbf{\Psi}}\gamma^\nu\mathbf{\Psi} = 0\\ \partial_\nu F^{\mu\nu} = ej^\nu,\end{aligned} \tag{8.4.9}$$

which are the inhomogeneous Maxwell's equations with current density $j^\nu = \overline{\mathbf{\Psi}}\gamma^\nu\mathbf{\Psi}$. Thus our Lagrangian produces the correct equations of motion that describe fermions coupled to the electromagnetic field.

Recall that the U(1) symmetry of the Dirac Lagrangian was an invariance under a phase rotation by a constant, arbitrary phase angle $\alpha \in \mathbb{R}$. Physically this means that if we were to shift the phase of all Dirac spinors in the universe by some phase angle α, the physics of the Dirac spinors would not change meaning that the U(1) symmetry is a global symmetry. When we imposed the gauge transformations (8.4.4) and (8.4.5), we allowed the phase angle $\alpha(x)$ to be a function of spacetime thus the value of the phase angle depends locally on its spacetime coordinates. We say that the global U(1) symmetry has been promoted to a *local* U(1) symmetry thus the gauge transformation is a local U(1) transformation [5]. Since Lagrangian (8.4.7) is invariant under this local symmetry as well, this means that physically, if the phase angle changed in some arbitrary region of spacetime, the physics of spinors would not change. Recall that a local symmetry implies an equivalent global symmetry which is the U(1) symmetry. Then by Noether's theorem, the

Noether charge operator of quantum (spinor) electrodynamics is

$$\begin{aligned}\hat{Q} &= \int d^3x\, j^0 \\ &= e\int d^3x\, \hat{\overline{\mathbf{\Psi}}}\gamma^0\hat{\mathbf{\Psi}} \\ &= e\sum_{s=\pm}\int d^3p_L\, \Big[\hat{b}_s^\dagger(\mathbf{p})\hat{b}_s(\mathbf{p}) - \hat{d}_s^\dagger(\mathbf{p})\hat{d}_s(\mathbf{p})\Big]. \end{aligned} \tag{8.4.10}$$

We see that the the inclusion of the elementary charge e makes Noether charge operator the electric charge operator which means that the electric charge is the generator of the U(1) gauge symmetry [15]. Since the Dirac field also possesses charge conjugation as an additional symmetry, which implies that the b-type particles and d-type particles have the same mass, the electric charge operator counts the number of negatively charged fermions (b-type) minus their positively charged antifermion counterparts (d-type) [5]. Hence, the Noether charge physically translates into the conservation of electric charge. Like Yukawa theory, we will treat the b-type particles as electrons and the d-type particles as positrons.

The path integral for quantum electrodynamics is given by

$$\begin{aligned}Z(\overline{\eta},\eta,J) \propto \exp\Big[ie\int d^4x\, \Big(\frac{1}{i}\frac{\delta}{\delta J^\mu(x)}\Big)\Big(i\frac{\delta}{\delta\eta_\alpha(x)}\Big)(\gamma^\mu)_{\alpha\beta}\Big(\frac{1}{i}\frac{\delta}{\delta\overline{\eta}_\beta(x)}\Big)\Big] \\ \times Z_0(\overline{\eta},\eta,J),\end{aligned} \tag{8.4.11}$$

where

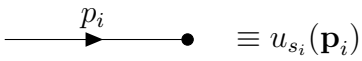
$$\begin{aligned}Z_0(\overline{\eta},\eta,J) = \exp\Big[i\int d^4x\, d^4y\, \overline{\eta}(x)S_F(x-y)\eta(y)\Big] \\ \times \exp\Big[\frac{i}{2}\int d^4x\, d^4y\, J^\mu(x)D_{\mu\nu}(x-y)J^\nu(y)\Big],\end{aligned} \tag{8.4.12}$$

where $D_{\mu\nu}(x-y) \equiv D^{\mu\nu}_{F,\xi=1}(x-y)$ is the Feynman propagator for the photon in the Feynman gauge [5]. In accordance with virtually every concept we have covered thus far, we have chosen the Feynman gauge $\xi = 1$ for the photon propagator since the photon propagator is Lorentz invariant with this choice of gauge [8]. Since we

were able to establish the Feynman rules for fermions in Yukawa theory, it is easy to see that the Feynman rules for fermions from Yukawa theory will generalize to quantum electrodynamics except that we must consider the Feynman rules for the photon. From the path integral (8.4.11), we see that the vertex term will be $ie\gamma^{\mu}$. In our derivation of the Fourier coefficient operators of the electromagnetic field and the LSZ reduction formula for photons, we see that the external lines for incoming and outgoing photons should be assigned the polarization vectors $\varepsilon_{\lambda}^{\mu *}(\mathbf{k})$ and $\varepsilon_{\lambda}^{\mu}(\mathbf{k})$, respectively [5]. The internal lines for photons will obviously be assigned the Feynman gauge photon propagator. Thus, the Feynman rules for quantum electrodynamics are in accordance with Srednicki [5]:

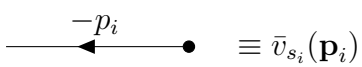
Feynman Rules for Quantum (Spinor) Electrodynamics

1. Incoming electrons are solid lines with an arrow pointing towards the vertex and are labeled with its four-momentum p_i. The assigned factor is $u_{s_i}(\mathbf{p}_i)$.



2. Outgoing electrons are solid lines with an arrow pointing away from the vertex and are labeled with its four-momentum p_i'. The assigned factor is $\bar{u}_{s_i'}(\mathbf{p}_i')$.

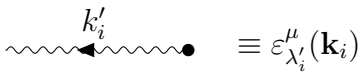
3. Incoming positrons are solid lines with an arrow pointing away from the vertex and are labeled with its four-momentum $-p_i$. The assigned factor is $\bar{v}_{s_i}(\mathbf{p}_i)$.



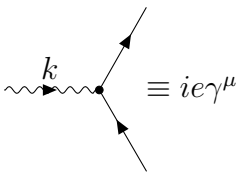
4. Outgoing positrons are solid lines with an arrow pointing towards the vertex and are labeled with its four-momentum $-p_i'$. The assigned factor is $v_{s_i'}(\mathbf{p}_i')$.

5. Incoming photons are wavy lines with an arrow pointing towards the vertex and are labeled with its four-momentum k_i. The assigned factor is $\varepsilon^{\mu *}_{\lambda_i}(\mathbf{k}_i)$.

6. Outgoing photons are wavy lines with an arrow pointing away from the vertex and are labeled with its four-momentum k_i'. The assigned factor is $\varepsilon^{\mu}_{\lambda_i'}(\mathbf{k}_i)$.

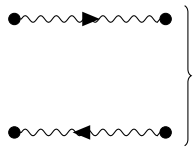


7. Construct each vertex from two solid lines and one wavy line. One of the solid lines must have an arrow pointing towards the vertex while the other must have an arrow pointing away from the vertex. The arrow on the wavy line can point in any direction. Draw all topologically inequivalent diagrams using the vertex, external lines, and any extra internal lines when needed. The assigned factor is $ie\gamma^{\mu}$.



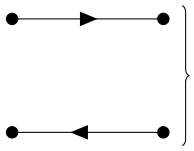
8. Conserve four-momentum at each vertex by following the arrows of each external line.
9. Assign the appropriate propagators to each internal line:

    - Internal photon:



$$\equiv \frac{1}{i}\frac{g^{\mu\nu}}{k^2 - i\epsilon}$$

- Internal fermion:



$$\equiv \frac{1}{i}\frac{-\not{p} + m}{p^2 + m^2 - i\epsilon}$$

10. Arrange the diagrams in standard form: all fermion lines are horizontal with their arrows pointing from left to right and label the left endpoints of the contributing diagrams with the same, arbitrarily fixed order from top to bottom. Contract the spinor indices by starting at the fermion line with the arrow pointing away from the vertex, follow the complete fermion line backwards, and write down from left to right the assigned factors that correspond to the vertices and propagators. The first factor should be either $\bar{u}$ or $\bar{v}$ and the last factor should be either u or v. For each vertex, the vector index is contracted on either the propagator or the photon polarization vector if the attached photon line is internal or external, respectively. Repeat for other fermion lines if necessary.
11. To determine the sign of a tree diagram, draw all the contributing diagrams in standard form. If the ordering of the right endpoints are an even (odd) permutation of the fixed ordering then the sign is positive (negative). This method can be reversed to so that we fix the right endpoints and observe the left endpoints instead.
12. Sum all the values from the contributing diagrams to get the value of $i\mathcal{T}$.

Now we will do a thorough computation of a QED process to demonstrate how

to successfully apply the Feynman rules. Let us consider the famous instance of matter-antimatter annihilation in the context of fermions, namely, electron-positron annihilation $e^+e^- \to \gamma\gamma$ called *pair annihilation*. The diagrams for this process are shown below:

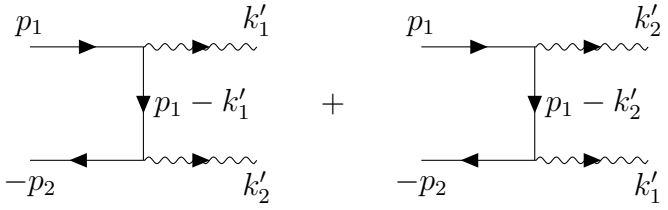


Figure 8.17: Pair annihilation $e^+e^- \to \gamma\gamma$

With our Feynman rules, the scattering matrix is

$$\begin{aligned} i\mathcal{T}_{e^+e^-\to\gamma\gamma} &= \bar{v}_2(ie\gamma_\mu)\varepsilon^\mu_{2'}\frac{1}{i}\frac{(-\not{p}_1 + \not{k}_1' + m)}{(p_1 - k_1')^2 + m^2}(ie\gamma_\nu)\varepsilon^\nu_{1'}u_1 \\ &\quad + \bar{v}_2(ie\gamma_\nu)\varepsilon^\nu_{1'}\frac{1}{i}\frac{(-\not{p}_1 + \not{k}_2' + m)}{(p_1 - k_2')^2 + m^2}(ie\gamma_\mu)\varepsilon^\mu_{2'}u_1 \\ &= ie^2\varepsilon^\mu_{1'}\varepsilon^\nu_{2'}\bar{v}_2\Big[\gamma_\mu\Big(\frac{-\not{p}_1 + \not{k}_1' + m}{-t + m^2}\Big)\gamma_\nu + \gamma_\nu\Big(\frac{-\not{p}_1 + \not{k}_2' + m}{-u + m^2}\Big)\gamma_\mu\Big]u_1, \end{aligned} \tag{8.4.13}$$

where we have suppressed some indices for brevity and suppressed $i\epsilon$ into the mass term. The Mandelstam variables for this process are

$$\begin{aligned} s &= -(p_1 + p_2)^2 = -(k_1' + k_2')^2, \\ t &= -(p_1 - k_1')^2 = -(p_2 - k_2')^2, \\ u &= -(p_1 - k_2')^2 = -(p_2 - k_1')^2, \end{aligned} \tag{8.4.14}$$

where as usual the Mandelstam variables obey the relation $s + t + u = 2m^2$ [5].

Also note that we can easily derive the following identities

$$\begin{aligned} p_1 p_2 &= -\frac{1}{2}(s-2m), \\ k_1' k_2' &= -\frac{1}{2}s, \\ p_1 k_1' = p_2 k_2' &= \frac{1}{2}(t-2m), \\ p_1 k_2' = p_2 k_1' &= \frac{1}{2}(u-2m), \end{aligned} \tag{8.4.15}$$

from the Mandelstam variables. Now let

$$A_{\mu\nu} = e^2\Big[\gamma_\mu\Big(\frac{-\not{p}_1 + \not{k}_1' + m}{-t+m^2}\Big)\gamma_\nu + \gamma_\nu\Big(\frac{-\not{p}_1 + \not{k}_2' + m}{-u+m^2}\Big)\gamma_\mu\Big], \tag{8.4.16}$$

then we have

$$\mathcal{T} = \varepsilon_{1'}^{\mu}\varepsilon_{2'}^{\nu}(\bar{v}_2 A_{\mu\nu} u_1), \tag{8.4.17}$$

and similarly

$$\overline{\mathcal{T}} = \varepsilon_{1'}^{\rho *}\varepsilon_{2'}^{\sigma *}(\bar{u}_1 \overline{A_{\rho\sigma}} v_2) = \varepsilon_{1'}^{\rho *}\varepsilon_{2'}^{\sigma *}(\bar{u}_1 A_{\sigma\rho} v_2), \tag{8.4.18}$$

where we have used our knowledge on spinor and gamma matrix identities to get $\overline{A_{\rho\sigma}} = A_{\sigma\rho}$. Then we have

$$\begin{aligned} |\mathcal{T}|^2 &= \varepsilon_{1'}^{\mu}\varepsilon_{2'}^{\nu}\varepsilon_{1'}^{\rho *}\varepsilon_{2'}^{\sigma *}(\bar{v}_2 A_{\mu\nu} u_1)(\bar{u}_1 A_{\sigma\rho} v_2) \\ &= \varepsilon_{1'}^{\mu}\varepsilon_{2'}^{\nu}\varepsilon_{1'}^{\rho *}\varepsilon_{2'}^{\sigma *}\,\mathrm{Tr}\left[A_{\mu\nu}(u_1\bar{u}_1)A_{\sigma\rho}(v_2\bar{v}_2)\right]. \end{aligned} \tag{8.4.19}$$

Now we average over the initial and final spins of the electron and positron and

sum over the final photon polarizations to get

$$\begin{aligned}
\langle|\mathcal{T}|^2\rangle &= \frac{1}{4}\sum_{\lambda=\pm}\sum_{s_1,s_2}|\mathcal{T}|^2 \\
&= \frac{1}{4}\sum_{\lambda=\pm}\sum_{s_1,s_2}\varepsilon^{\mu}_{1'}\varepsilon^{\nu}_{2'}\varepsilon^{\rho *}_{1'}\varepsilon^{\sigma *}_{2'}\,\mathrm{Tr}\left[A_{\mu\nu}(u_1\bar{u}_1)A_{\sigma\rho}(v_2\bar{v}_2)\right] \\
&= \frac{1}{4}g^{\mu\rho}g^{\nu\sigma}\,\mathrm{Tr}\left[A_{\mu\nu}(-\not{p}_1+m)A_{\sigma\rho}(-\not{p}_2-m)\right] \\
&= \frac{1}{4}\mathrm{Tr}\left[A_{\mu\nu}(-\not{p}_1+m)A^{\nu\mu}(-\not{p}_2-m)\right] \\
&= e^4\Big[\frac{\langle\Phi_{tt}\rangle}{(m^2-t)^2}+\frac{\langle\Phi_{tu}\rangle+\langle\Phi_{ut}\rangle}{(m^2-t)(m^2-u)}+\frac{\langle\Phi_{uu}\rangle}{(m^2-u)^2}\Big],
\end{aligned} \tag{8.4.20}$$

where [5]

$$\begin{aligned}
\langle\Phi_{tt}\rangle &= \frac{1}{4}\mathrm{Tr}\left[\gamma_\mu(-\not{p}_1+\not{k}'_1+m)\gamma_\nu(-\not{p}_1+m)\gamma^\nu(-\not{p}'_1+\not{k}'_1+m)\gamma^\mu(-\not{p}_2-m)\right], \\
\langle\Phi_{tu}\rangle &= \frac{1}{4}\mathrm{Tr}\left[\gamma_\mu(-\not{p}_1+\not{k}'_1+m)\gamma_\nu(-\not{p}_1+m)\gamma^\mu(-\not{p}'_1+\not{k}'_2+m)\gamma^\nu(-\not{p}_2-m)\right], \\
\langle\Phi_{ut}\rangle &= \frac{1}{4}\mathrm{Tr}\left[\gamma_\nu(-\not{p}_1+\not{k}'_2+m)\gamma_\mu(-\not{p}_1+m)\gamma^\nu(-\not{p}'_1+\not{k}'_1+m)\gamma^\mu(-\not{p}_2-m)\right], \\
\langle\Phi_{uu}\rangle &= \frac{1}{4}\mathrm{Tr}\left[\gamma_\nu(-\not{p}_1+\not{k}'_2+m)\gamma_\mu(-\not{p}_1+m)\gamma^\mu(-\not{p}'_1+\not{k}'_2+m)\gamma^\nu(-\not{p}_2-m)\right].
\end{aligned} \tag{8.4.21}$$

Similarly to how we computed the scattering processes in Yukawa theory, we make note of the fact that we only need to compute $\langle\Phi_{tt}\rangle$ ($\langle\Phi_{tt}\rangle$) and $\langle\Phi_{tu}\rangle$ ($\langle\Phi_{ut}\rangle$) since the other two terms can be gained from $t\leftrightarrow u$. We apply the spinor and gamma matrix identities along with Mandelstam variable identities to finally get

$$\begin{aligned}
\langle\Phi_{tt}\rangle &= 2[tu-m^2(3t+u)-m^4], \\
\langle\Phi_{tu}\rangle &= 2m^2(s-4m^2), \\
\langle\Phi_{uu}\rangle &= 2[tu-m^2(3u+t)-m^4], \\
\langle\Phi_{ut}\rangle &= 2m^2(s-4m^2).
\end{aligned} \tag{8.4.22}$$

To give a real application of pair annihilation, we will compute the lifetime of a exotic particle called *positronium* which is the bound system of an electron and

a positron [11]. Specifically, we will find the decay rate of positronium in its S state, namely, the ground state of the positronium. Rather than using the rather complicated expression (8.4.20) to compute the decay rate, we will instead use computational tricks on the scattering amplitude (8.4.13) using our knowledge on spinor and gamma matrix identities to derive a simpler expression to use. We will assume that the electron and positron are initially at rest thus we will work in the nonrelativistc regime. Then the spinors $u_1(\mathbf{p})$ and $\bar{v}_2(\mathbf{p})$ become $u_1(\mathbf{0})$ and $\bar{v}_2(\mathbf{0})$, respectively. We will omit the argument of the spinors to reduce computational clutter. The photons that are produced after the annihilation event will propagate in opposite directions so we have the freedom to choose the z-direction as the direction of propagation [17]. Thus we have

$$\begin{aligned} p_1 &= (m,0,0,0), & k_1' &= (m,0,0,m), \\ p_2 &= (m,0,0,0), & k_2' &= (m,0,0,-m), \end{aligned} \tag{8.4.23}$$

and the Mandelstam variables become

$$\begin{aligned} t &= -(p_1 - k_1')^2 = m^2, \\ u &= -(p_1 - k_2')^2 = m^2. \end{aligned} \tag{8.4.24}$$

Now we rewrite our scattering amplitude (8.4.13) as

$$\begin{aligned} \mathcal{T} &= e^2 \bar{v}_2 \Big[\gamma_\mu \varepsilon_{2'}^\mu \Big(\frac{-\not{p}_1 + \not{k}_1' + m}{-t + m^2} \Big) \gamma_\nu \varepsilon_{1'}^\nu + \gamma_\nu \varepsilon_{1'}^\nu \Big(\frac{-\not{p}_1 + \not{k}_2' + m}{-u + m^2} \Big) \gamma_\mu \varepsilon_{2'}^\mu \Big] u_1, \\ &= \frac{e^2}{2m^2} \bar{v}_2 \Big[\not{\varepsilon}_{2'} \big(-\not{p}_1 + \not{k}_1' + m \big) \not{\varepsilon}_{1'} + \not{\varepsilon}_{1'} \big(-\not{p}_1 + \not{k}_2' + m \big) \not{\varepsilon}_{2'} \Big] u_1. \end{aligned}$$

Now recall that the polarization vector ε^μ only has spatial components so it is trivial to see that $p_{i\,\mu} \varepsilon_{i'}^\mu = 0$ for $i, i' = 1, 2$. Along with this, we also have that polarization vector must be perpendicular the direction of propagation so $k_{i\,\mu}' \varepsilon_{i'}^\mu = 0$

for $i, i' = 1, 2$ [17]. Then observe that

$$\begin{aligned}
\not{p}_i \not{\varepsilon}_{i'} &= -\not{\varepsilon}_{i'} \not{p}_i - 2(p_i \varepsilon_{i'}) \\
&= -\not{\varepsilon}_{i'} \not{p}_i, \\
\not{k}'_i \not{\varepsilon}_{i'} &= -\not{\varepsilon}_{i'} \not{k}'_i - 2(k'_i \varepsilon_{i'}) \\
&= -\not{\varepsilon}_{i'} \not{k}'_i,
\end{aligned}$$

for $i, i' = 1, 2$. Then

$$\big(-\not{p}_1 + \not{k}'_i + m \big) \not{\varepsilon}_{i'} = \not{\varepsilon}_{i'} \big(-\not{p}_1 - \not{k}'_i + m \big), \tag{8.4.25}$$

for $i, i' = 1, 2$. We now multiply u_1 on the right side of equation (8.4.25) and use $(\not{p} + m)u = 0$ to simplify the right hand side of equation (8.4.25) to get

$$\big(-\not{p}_1 + \not{k}'_i + m \big) \not{\varepsilon}_{i'} u_1 = \not{\varepsilon}_{i'} \not{k}'_i u_1, \tag{8.4.26}$$

for $i, i' = 1, 2$ and the scattering amplitude reduces to

$$\mathcal{T} = -\frac{e^2}{2m^2} \bar{v}_2 \Big[\not{\varepsilon}_{2'} \not{\varepsilon}_{1'} \not{k}'_1 + \not{\varepsilon}_{1'} \not{\varepsilon}_{2'} \not{k}'_2 \Big] u_1. \tag{8.4.27}$$

To simplify the above amplitude even further, we use equation (8.4.23) and the definition of the Feynman slash notation to get

$$\begin{aligned}
\not{k}'_1 &= \gamma^\mu k'_{1\,\mu} \\
&= m(-\gamma^0 + \gamma^3), \\
\not{k}'_2 &= \gamma^\mu k'_{2\,\mu} \\
&= m(-\gamma^0 - \gamma^3),
\end{aligned}$$

where

$$\gamma^\mu = \begin{pmatrix} 0 & \sigma^\mu \\ \overline{\sigma}^\mu & 0 \end{pmatrix}, \tag{8.4.28}$$

where $\sigma^\mu = (I, \boldsymbol{\sigma})$ and $\overline{\sigma}^\mu = (I, -\boldsymbol{\sigma})$ [17]. Along with this, we will also use the Pauli matrix identity [7]

$$(\boldsymbol{\sigma} \cdot \boldsymbol{a})(\boldsymbol{\sigma} \cdot \boldsymbol{b}) = \boldsymbol{a} \cdot \boldsymbol{b} + i\boldsymbol{\sigma} \cdot (\boldsymbol{a} \times \boldsymbol{b}). \tag{8.4.29}$$

Then we get

$$\mathcal{T} = \frac{e^2}{2m} \bar{v}_2 \Big[\{ \not{\varepsilon}_{2'}, \not{\varepsilon}_{1'} \} \gamma^0 - [\not{\varepsilon}_{2'}, \not{\varepsilon}_{1'}] \gamma^3 \Big] u_1. \tag{8.4.30}$$

Now note that

$$\begin{aligned} \not{\varepsilon} &= \gamma^\mu \varepsilon_\mu \\ &= -\gamma^0 + \boldsymbol{\gamma} \cdot \boldsymbol{\varepsilon} \\ &= \begin{pmatrix} 0 & \boldsymbol{\sigma} \cdot \boldsymbol{\varepsilon} \\ -\boldsymbol{\sigma} \cdot \boldsymbol{\varepsilon} & 0 \end{pmatrix}, \end{aligned} \tag{8.4.31}$$

thus [17]

$$\begin{aligned} \not{\varepsilon}_{1'} \not{\varepsilon}_{2'} &= \begin{pmatrix} 0 & \boldsymbol{\sigma} \cdot \varepsilon_{1'} \\ -\boldsymbol{\sigma} \cdot \varepsilon_{1'} & 0 \end{pmatrix} \begin{pmatrix} 0 & \boldsymbol{\sigma} \cdot \varepsilon_{2'} \\ -\boldsymbol{\sigma} \cdot \varepsilon_{2'} & 0 \end{pmatrix} \\ &= - \begin{pmatrix} (\boldsymbol{\sigma} \cdot \varepsilon_{1'})(\boldsymbol{\sigma} \cdot \varepsilon_{2'}) & 0 \\ 0 & (\boldsymbol{\sigma} \cdot \varepsilon_{1'})(\boldsymbol{\sigma} \cdot \varepsilon_{2'}) \end{pmatrix} \\ &= - \begin{pmatrix} \varepsilon_{1'} \cdot \varepsilon_{2'} + i\boldsymbol{\sigma} \cdot (\varepsilon_{1'} \times \varepsilon_{2'}) & 0 \\ 0 & \varepsilon_{1'} \cdot \varepsilon_{2'} + i\boldsymbol{\sigma} \cdot (\varepsilon_{1'} \times \varepsilon_{2'}) \end{pmatrix} \\ &= -(\varepsilon_{1'} \cdot \varepsilon_{2'}) I_4 - i(\varepsilon_{1'} \times \varepsilon_{2'}) \cdot \boldsymbol{\Sigma}, \end{aligned} \tag{8.4.32}$$

where I_4 is the 4×4 identity matrix and

$$\mathbf{\Sigma} = \begin{pmatrix} \boldsymbol{\sigma} & 0 \\ 0 & \boldsymbol{\sigma} \end{pmatrix}.$$

Similarly, we also get

$$\not{\varepsilon}_{2'}\not{\varepsilon}_{1'} = -(\varepsilon_{2'} \cdot \varepsilon_{1'})I_4 - i(\varepsilon_{2'} \times \varepsilon_{1'}) \cdot \mathbf{\Sigma}. \tag{8.4.33}$$

Then we evaluate the terms in the parenthesis of equation (8.4.30) to get

$$\{\not{\varepsilon}_{2'}, \not{\varepsilon}_{1'}\} = -2\not{\varepsilon}_{2'} \cdot \not{\varepsilon}_{1'} I_4,$$
$$[\not{\varepsilon}_{2'}, \not{\varepsilon}_{1'}] = 2i(\not{\varepsilon}_{2'} \times \not{\varepsilon}_{1'}) \cdot \mathbf{\Sigma},$$

and finally our scattering amplitude becomes

$$\mathcal{T} = -\frac{e^2}{m}\bar{v}_2\Big[\varepsilon_{1'} \cdot \varepsilon_{2'}\gamma^0 - i(\varepsilon_{1'} \times \varepsilon_{2'}) \cdot \mathbf{\Sigma}\gamma^3\Big]u_1. \tag{8.4.34}$$

Before we continue further, note that since the positronium is a two particle system that consists of an electron and a positron, the Hilbert space of the positronium is a direct product space $\mathcal{H}_{e^-e^+} = \mathcal{H}_{e^-} \bigotimes \mathcal{H}_{e^+}$, namely, the direct product of the Hilbert spaces of the electron and positron form the Hilbert space of the positronium. Furthermore, recall from nonrelativistic quantum mechanics that the state vector of a system that consists of two spin-1/2 particles can exist as either a spin-0 (*singlet*) or a spin-1 particle (*triplet*) [7]. Then the Hibert space of the positronium is equivalent to $\mathcal{H}_{s=1/2} \bigotimes \mathcal{H}_{s=1/2} = \mathcal{H}_{s=1} \bigoplus \mathcal{H}_{s=0}$, namely, the direct product space of two spin-1/2 particles is the direct sum of a spin-1 and spin-0 Hilbert space. The singlet state vector in the S state is

$$|s = 0, m = 0\rangle = \frac{|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle}{\sqrt{2}}, \tag{8.4.35}$$

while the triplet state vector in the S state is

$$|s=1,m=0\rangle = \frac{|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle}{\sqrt{2}}. \tag{8.4.36}$$

From the form of the state vectors, rather than averaging the spins for the scattering amplitude, we must instead compute for the singlet state [17]

$$\mathcal{T}_{singlet} = \frac{\mathcal{T}_{\uparrow\downarrow} - \mathcal{T}_{\downarrow\uparrow}}{\sqrt{2}}, \tag{8.4.37}$$

or for the triplet state

$$\mathcal{T}_{triplet} = \frac{\mathcal{T}_{\uparrow\downarrow} + \mathcal{T}_{\downarrow\uparrow}}{\sqrt{2}}. \tag{8.4.38}$$

The S state vector of positronium expanded in terms of free particle states is

$$|Ps\rangle = \sqrt{2M}\int \frac{d^3k}{(2\pi)^3}\psi(\mathbf{k})\frac{1}{\sqrt{2m}}\frac{1}{\sqrt{2m}}|e^-\rangle|e^+\rangle, \tag{8.4.39}$$

where M and m are the masses of the positronium and electron, respectively. By using $M = 2m$ as a close approximation, we can easily identify the scattering amplitude of the singlet positronium as

$$\mathcal{T}_{Ps} = \frac{\psi_{100}(0)}{\sqrt{2m}}\mathcal{T}_{singlet}, \tag{8.4.40}$$

and similarly

$$\mathcal{T}_{Ps} = \frac{\psi_{100}(r=0)}{\sqrt{2m}}\mathcal{T}_{triplet}. \tag{8.4.41}$$

where

$$\psi_{100}(\mathbf{r}) = \left(\frac{\alpha m^3}{8\pi}\right)^{1/2} e^{-\alpha m r} \tag{8.4.42}$$

is the ground state wavefunction of the electron [17]. By using the spinor identities

$$\bar{v}_{s'}(\mathbf{0})\gamma^0 u_s(\mathbf{0}) = 0,$$

$$\bar{v}_{\pm}(\mathbf{0})\boldsymbol{\Sigma}\gamma^3 u_{\mp}(\mathbf{0}) = \mp 2m,$$

we get

$$\begin{aligned}\mathcal{T}_{\uparrow\downarrow} &= -\frac{e^2}{m}\bar{v}_{2-}\Big[\varepsilon_{1'}\cdot\varepsilon_{2'}\gamma^0 - i(\varepsilon_{1'}\times\varepsilon_{2'})\cdot\mathbf{\Sigma}\gamma^3\Big]u_{1+} \\ &= -2ie^2(\varepsilon_{1'}\times\varepsilon_{2'}) \qquad (8.4.43)\end{aligned}$$

and similarly

$$\begin{aligned}\mathcal{T}_{\downarrow\uparrow} &= -\frac{e^2}{m}\bar{v}_{2+}\Big[\varepsilon_{1'}\cdot\varepsilon_{2'}\gamma^0 - i(\varepsilon_{1'}\times\varepsilon_{2'})\cdot\mathbf{\Sigma}\gamma^3\Big]u_{1-} \\ &= 2ie^2(\varepsilon_{1'}\times\varepsilon_{2'}), \qquad (8.4.44)\end{aligned}$$

so we see that $\mathcal{T}_{\uparrow\downarrow} = -\mathcal{T}_{\downarrow\uparrow}$ [17]. The triplet scattering amplitude is then

$$\begin{aligned}\mathcal{T}_{triplet} &= \frac{\mathcal{T}_{\uparrow\downarrow} + \mathcal{T}_{\downarrow\uparrow}}{\sqrt{2}} \\ &= \frac{\mathcal{T}_{\uparrow\downarrow} - \mathcal{T}_{\uparrow\downarrow}}{\sqrt{2}} \\ &= 0, \qquad (8.4.45)\end{aligned}$$

so we see that a triplet S state positronium cannot decay into two photons [15]. For the singlet S state positronium, the scattering amplitude is

$$\begin{aligned}\mathcal{T}_{singlet} &= \frac{\mathcal{T}_{\uparrow\downarrow} - \mathcal{T}_{\downarrow\uparrow}}{\sqrt{2}} \\ &= \frac{\mathcal{T}_{\uparrow\downarrow} + \mathcal{T}_{\uparrow\downarrow}}{\sqrt{2}} \\ &= -2\sqrt{2}ie^2(\varepsilon_{1'}\times\varepsilon_{2'}),\end{aligned}$$

so the singlet S state positronium can undergo a two photon decay process [15]. To calculate the photon polarization, recall from our derivation of the quantization

of the electromagnetic field the right- and left-handed circular polarization vectors

$$\varepsilon_+(\mathbf{k}) = \frac{1}{\sqrt{2}}(1, -i, 0),$$
$$\varepsilon_-(\mathbf{k}) = \frac{1}{\sqrt{2}}(1, i, 0).$$

Using the explicit form of the polarization vectors yields $(\varepsilon_{1'} \times \varepsilon_{2'}) = i\sqrt{2}\hat{k}$ hence

$$\begin{aligned}\mathcal{T}_{singlet} &= 4e^2\hat{k} \\ &= 16\pi\alpha\hat{k}. \end{aligned} \tag{8.4.46}$$

Thus we finally have the matrix element for positronium

$$\begin{aligned}\langle|\mathcal{T}_{Ps}|^2\rangle &= \frac{|\psi_{000}(0)|}{2m}\,|\mathcal{T}_{singlet}|^2 \\ &= 16\pi\alpha^5 m^2. \end{aligned} \tag{8.4.47}$$

The decay rate is

$$\begin{aligned}\Gamma &= \frac{1}{2}\int d\Gamma \\ &= \frac{1}{2}\frac{1}{4m}\int d^3k'_{1\,L}\, d^3k'_{2\,L}\,(16\alpha^5 m^2)(2\pi)^4\delta^4(p_{Ps} - k'_1 - k'_2) \\ &= \frac{1}{2}\alpha^5 m, \end{aligned} \tag{8.4.48}$$

where the 1/2 comes from the fact that photons are identical particles [8]. Then the lifetime of positronium in the singlet S state is

$$\tau = \frac{1}{\Gamma} = \frac{2}{\alpha^5 m_e} \approx 1.25 \times 10^{-10}\,\text{sec}, \tag{8.4.49}$$

which is off from the measured value by a 0.5% discrepancy [8]. This result is one of the many instances of the extreme accuracy that is characteristic of quantum electrodynamics. Applying radiative corrects results in even more accurate values however, we should note that the Feynman diagrams that we have used are the

second order terms of the perturbation expansion for this result. The fact that a second order approximation is able to give such a remarkably precise value demonstrates the success of quantum electrodynamics and quantum field theory in general [1]. From here, we will explore other examples of QED processes.

Quantum Electrodynamics (Spinor) Applications

Coulomb Potential

As a simple application, let us consider the first diagram of the scattering process $e^-e^- \to e^-e^-$ and the first diagram of the scattering process $e^+e^- \to e^+e^-$

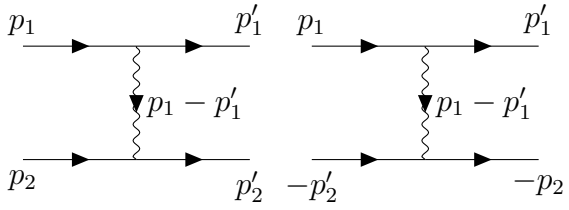


Figure 8.18: Coulomb potential diagrams

For the first diagram, the scattering matrix is

$$i\mathcal{T}_{e^-e^-\to e^-e^-} = \frac{1}{i}(ie)^2(\bar{u}_1'\gamma^\mu u_1)\frac{g^{\mu\nu}}{(p_1-p_1')^2 - i\epsilon}(\bar{u}_2'\gamma^\mu u_2). \tag{8.4.50}$$

Now in the nonrelativistic limit, we have that [8]

$$\bar{u}(\mathbf{p}')\gamma^0 u(\mathbf{p}) = u_{s'}^\dagger(\mathbf{p}')u_s(\mathbf{p}) \to 2m\delta_{s's}, \tag{8.4.51}$$

$$\bar{v}(\mathbf{p}')\gamma^0 v(\mathbf{p}) = v_{s'}^\dagger(\mathbf{p}')v_s(\mathbf{p}) \to 2m\delta_{s's}, \tag{8.4.52}$$

while $\gamma^i \to 0$ so the scattering matrix in the nonrelativistic limit becomes

$$\begin{aligned} i\mathcal{T}_{e^-e^-\to e^-e^-} &\to \frac{ie^2}{-\left|\mathbf{p}_1-\mathbf{p}_1'\right|^2}(2m\delta_{s's})_1(2m\delta_{s's})_2 \\ &= \frac{-ie^2}{\left|\mathbf{p}_1-\mathbf{p}_1'\right|^2}(2m\delta_{s's})_1(2m\delta_{s's})_2 \\ &= \frac{-ie^2}{\left|\mathbf{p}\right|^2}(2m\delta_{s's})_1(2m\delta_{s's})_2, \end{aligned} \tag{8.4.53}$$

where $\mathbf{p} = \mathbf{p}_1 - \mathbf{p}_1'$ [8]. By comparing with the Born approximation from nonrelativistic quantum mechanics, we see that the potential in momentum-space is [8]

$$V(\mathbf{p}) = \frac{e^2}{\left|\mathbf{p}\right|^2}, \tag{8.4.54}$$

and in position-space

$$\begin{aligned} V(r) &= e^2 \int \frac{d^3p}{(2\pi)^3}\,\frac{e^{i\mathbf{p}\cdot\mathbf{r}}}{\left|\mathbf{p}\right|^2} \\ &= \frac{e^2}{4\pi r}. \end{aligned} \tag{8.4.55}$$

We see that the sign of the potential for the scattering process of two fermions, which have the same negative charge, is positive which means that the potential is repulsive. If we had considered $e^+e^+ \to e^+e^+$, we would have derived the same potential with the same sign. Thus we conclude that fermions with like charges repel. Let us now do the same calculations, in the nonrelativistic limit, for the scattering process $e^+e^- \to e^+e^-$. Using the Feynman rules, we get for the second diagram

$$i\mathcal{T}_{e^+e^-\to e^+e^-} = -\frac{1}{i}(ie)^2(\bar{u}_1'\gamma^\mu u_1)\frac{g^{\mu\nu}}{(p_1-p_1')^2 - i\epsilon}(\bar{v}_2'\gamma^\mu v_2). \tag{8.4.56}$$

In the nonrelativistic limit,

$$\begin{aligned} i\mathcal{T}_{e^+e^-\to e^+e^-} &\to -\frac{ie^2}{-\left|\mathbf{p}_1-\mathbf{p}_1'\right|^2}(2m\delta_{s's})_1(2m\delta_{s's})_2 \\ &= \frac{ie^2}{\left|\mathbf{p}\right|^2}(2m\delta_{s's})_1(2m\delta_{s's})_2, \end{aligned} \tag{8.4.57}$$

where $\mathbf{p}=\mathbf{p}_1-\mathbf{p}_1'$ and the negative sign comes from the ordering permutation rule. Then the potential in momentum-space is

$$V(\mathbf{p}) = -\frac{e^2}{\left|\mathbf{p}\right|^2}(2m\delta_{s's})_1(2m\delta_{s's})_2, \tag{8.4.58}$$

and in position-space,

$$\begin{aligned} V(r) &= -e^2\int\frac{d^3p}{(2\pi)^3}\,\frac{e^{i\mathbf{p}\cdot\mathbf{r}}}{\left|\mathbf{p}\right|^2} \\ &= -\frac{e^2}{4\pi r}. \end{aligned} \tag{8.4.59}$$

We see that the negative sign of the potential for the scattering of fermions and antifermions yields an attractive potential. Since these two results apply for all fermions and antifermions interacting with the electromagnetic field, we conclude that when a vector particle (in this case, the photon) is exchanged, oppositely charged fermions attract while like charged fermions repel [15]. By now, the reader will have noticed that we have derived the *Coulomb potential* from classical electrodynamics. If we study the diagrams we have computed, we see that the interaction is mediated by the exchange of a massless boson, namely, the photon and the potential is a long-range interaction since the mediator of the interaction is massless [15]. It is now well known that there exist fermions that actually possess *fractional charges* with the most famous example being *quarks* [1]. Thus to generalize the Coulomb potential we have derived for any kind of fermion and antifermion, we make the replacement $e\to q$ where q is the charge magnitude. So $q=-1$ and $q=+1$ for electrons and positrons, respectively. Then we can combine

the repulsive and attractive Coulomb potentials into the more compact version

$$V(r) = \frac{q_1 q_2}{4\pi r}, \tag{8.4.60}$$

where q_i is the charge magnitude. If we take the gradient in spherical coordinates with no angular dependence, we get

$$\mathbf{F}(r) = \frac{q_1 q_2}{4\pi r^2}\hat{\mathbf{r}} \equiv \frac{q_1 q_2}{4\pi\varepsilon_0 r^2}\,\hat{\mathbf{r}}, \tag{8.4.61}$$

which is *Coulomb's law* [2]. This brings us back in full circle to the starting point of classical electrodynamics.

Compton Scattering

We now present the scattering process $e^-\gamma \to e^-\gamma$ called *Compton scattering*. Compton scattering is when a photon is shot at an electrically charged particle resulting in both the photon and the particle scattering inelastically [7]. This experiment was important in demonstrating the particle nature of light and its results were used in the development of quantum mechanics [1]. Since Compton scattering was first demonstrated with the electron, we will use the electron for this example. The tree-level diagrams for Compton scattering are presented below [5]

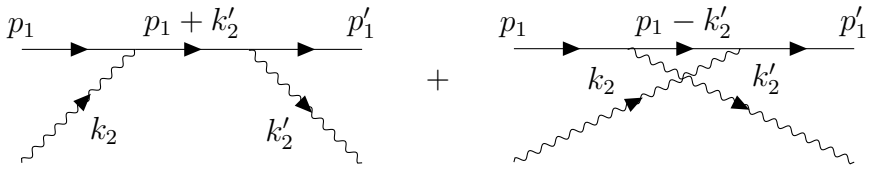


Figure 8.19: Compton scattering $e^-\gamma \to e^-\gamma$

The scattering matrix is

$$
\begin{aligned}
i\mathcal{T}_{e^-\gamma\to e^-\gamma} &= \bar{u}_1'(ie\gamma_\mu)\varepsilon_{2'}^{\mu}\frac{1}{i}\frac{(-\not{p}_1-\not{k}_2'+m)}{(p_1+k_2')^2+m^2}(ie\gamma_\nu)\varepsilon_2^{\nu*}u_1 \\
&\quad+\bar{u}_1'(ie\gamma_\nu)\varepsilon_2^{\nu*}\frac{1}{i}\frac{(-\not{p}_1+\not{k}_2'+m)}{(p_1-k_2')^2+m^2}(ie\gamma_\mu)\varepsilon_{2'}^{\mu}u_1 \\
&= ie^2\varepsilon_{2'}^{\mu}\varepsilon_2^{\nu*}\bar{u}_1'\Big[\gamma_\mu\Big(\frac{-\not{p}_1-\not{k}_1'+m}{-s+m^2}\Big)\gamma_\nu+\gamma_\nu\Big(\frac{-\not{p}_1+\not{k}_2'+m}{-u+m^2}\Big)\gamma_\mu\Big]u_1,
\end{aligned}
\tag{8.4.62}
$$

where $s=-(p_1+k_2')^2$ and $u=-(p_1-k_2')^2$. Hence

$$
\begin{aligned}
\mathcal{T} &= \varepsilon_{2'}^{\mu}\varepsilon_2^{\nu*}(\bar{u}_1'A_{\mu\nu}u_1), \\
\overline{\mathcal{T}} &= \varepsilon_{2'}^{\rho}\varepsilon_2^{\sigma*}(\bar{u}_1A_{\sigma\rho}u_1').
\end{aligned}
\tag{8.4.63}
$$

Following along the same route of calculations, we get

$$
\langle|\mathcal{T}|^2\rangle = e^4\Big[\frac{\langle\Phi_{ss}\rangle}{(m^2-s)^2}+\frac{\langle\Phi_{su}\rangle+\langle\Phi_{us}\rangle}{(m^2-s)^2(m^2-t)^2}+\frac{\langle\Phi_{uu}\rangle}{(m^2-u)^2}\Big],
\tag{8.4.64}
$$

where [5]

$$
\begin{aligned}
\langle\Phi_{ss}\rangle &= -2(su-m^2(3s+u)-m^4), \\
\langle\Phi_{su}\rangle &= -2m^2(t-4m^2), \\
\langle\Phi_{us}\rangle &= -2m^2(t-4m^2), \\
\langle\Phi_{uu}\rangle &= -2(su-m^2(3u+s)-m^4).
\end{aligned}
\tag{8.4.65}
$$

Then we can rewrite $\langle|\mathcal{T}|^2\rangle$ in terms of the dimensionless coupling $\alpha=e^2/4\pi\approx 1/137$ called the *fine-structure constant* [5]

$$
\begin{aligned}
\langle|\mathcal{T}|^2\rangle &= 32\pi^2\alpha^2\Big[\frac{m^4+m^2(3s+u)-su}{(m^2-s)^2}+\frac{m^4+m^2(3u+s)-su}{(m^2-u)^2} \\
&\quad+\frac{2m^2(s+u+2m^2)}{(m^2-s)(m^2-u)}\Big].
\end{aligned}
\tag{8.4.66}
$$

Let us go further with this example and assume that in the FT frame, we have the following

$$\begin{aligned} p &= (m, 0, 0, 0), \\ k &= (\omega, 0, 0, \omega), \\ k' &= (\omega', \omega' \sin\theta, 0, \omega' \cos\theta), \end{aligned} \tag{8.4.67}$$

where it is implied that $\theta \equiv \theta_{FT}$ and we have dropped the numbered subscripts from the four-momenta p and k [8]. Then our Mandelstam variables reduce to

$$s = -(p+k)^2 = m^2 + 2m\omega, \tag{8.4.68}$$

$$u = m^2 - 2m\omega'. \tag{8.4.69}$$

By the conservation of momentum, we have that $p' = p+k-k'$ and also $(p')^2 = -m^2$ thus

$$\begin{aligned} -m^2 &= (p + k - k')^2 \\ &= -m^2 - 2m\omega + 2m\omega' - 2\omega\omega'(\cos\theta - 1), \end{aligned}$$

which reduces to

$$1 - \cos\theta = m\Big(\frac{1}{\omega'} - \frac{1}{\omega}\Big), \tag{8.4.70}$$

which is Compton's formula for the shift in the photon wavelength [8]. Now by using equations (8.4.68), (8.4.69), and (8.4.70), our scattering matrix reduces further to

$$\langle|\mathcal{T}|^2\rangle = 32\pi^2\alpha^2\Big[-\sin^2\theta + \frac{\omega}{\omega'} + \frac{\omega'}{\omega}\Big]. \tag{8.4.71}$$

In the FT frame, we have that $s\,|\mathbf{k}_1|^2_{\mathrm{CM}} = m^2\omega^2$ so our frame independent differen-

tial cross section is

$$\begin{aligned}\frac{d\sigma}{dt} &= \frac{1}{64\pi s\,|\mathbf{k}_1|^2_{\rm CM}}\langle|\mathcal{T}|^2\rangle \\ &= \frac{1}{64\pi m^2\omega^2}\langle|\mathcal{T}|^2\rangle.\end{aligned}$$

To get the above differential cross section into the FT frame, we use equation (8.4.70) to solve for ω' which yields

$$\omega' = \frac{m\omega}{m+\omega(1-\cos\theta)}, \tag{8.4.72}$$

then we get

$$\begin{aligned}d\omega' &= \frac{m\omega^2}{(m+\omega(1-\cos\theta))^2}\,d\cos\theta \\ &= \frac{\omega'^2}{m}\,d\cos\theta.\end{aligned}$$

Since $t = 2m^2 - s - u = 2m(\omega' - \omega)$, the differential of t is [8]

$$\begin{aligned}dt &= 2m\,d\omega' \\ &= 2\omega'^2\,d\cos\theta \\ &= \frac{\omega'^2}{\pi}\,d\Omega_{\rm FT}.\end{aligned}$$

Finally, the differential cross section in the FT frame is

$$\frac{d\sigma}{d\Omega_{\rm FT}} = \frac{\alpha^2}{2m^2}\frac{\omega'^2}{\omega^2}\Big[\frac{\omega}{\omega'} + \frac{\omega'}{\omega} - \sin^2\theta_{\rm FT}\Big] \tag{8.4.73}$$

which is the *Klein-Nishina formula* [8]. If we consider the limit $\omega \to 0$, we see that $\omega'/\omega \to 1$ so the differential cross section reduces to

$$\frac{d\sigma}{d\Omega_{\rm FT}} = \frac{\alpha^2}{2m^2}(1+\cos^2\theta_{\rm FT}) \tag{8.4.74}$$

which when integrated yields

$$\sigma = \frac{8\pi\alpha^2}{3m^2}, \tag{8.4.75}$$

which is the *Thomas cross section* for *Thomas scattering* [17]. We should note that it is remarkable that we are able to glean much information from just a second order approximation of Compton scattering.

Møller Scattering

We will explore the QED process of electron-electron scattering called *Møller scattering* named after theoretical physicist Christian Møller. The simple Møller scattering serves as the basis for modelling other, more complicated interactions that involve electron repulsion [17]. The leading order diagrams for this process are given below:

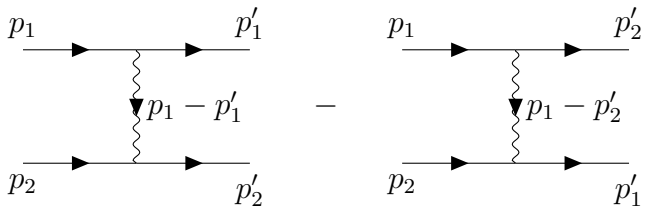


Figure 8.20: Møller scattering $e^-e^- \to e^-e^-$

In the first diagram, which is a t-channel diagram, we see that the electron with four-momentum p_1 emits a virtual photon with four-momentum $p_1 - p_1'$ to the electron with four-momentum p_2. The electron that emitted the virtual photon then scatters with four-momentum p_1' while the other electron absorbs the virtual particle and scatters with four-momentum p_2'. In the second diagram, which is a u-channel diagram, the physical process occurs except with four-momenta p_1' and p_2' switched. Thus we see that Møller scattering can occur according to the t-channel diagram or the u-channel diagram with a probability assigned to both.

By using the Feynman rules, the scattering amplitude is

$$\begin{aligned} i\mathcal{T}_{e^+e^-\to e^+e^-} &= \bar{u}_1'(ie\gamma_\nu)u_1\frac{1}{i}\frac{g^{\mu\nu}}{(p_1-p_1')^2-i\epsilon}\bar{u}_2'(ie\gamma_\mu)u_2 \\ &- \bar{u}_2'(ie\gamma_\mu)u_1\frac{1}{i}\frac{g^{\mu\nu}}{(p_1-p_2')^2-i\epsilon}\bar{u}_1'(ie\gamma_\nu)u_2 \\ &= \Big(\frac{1}{i}\Big)(ie)^2\Big[\frac{(\bar{u}_1'\gamma^\mu u_1)(\bar{u}_2'\gamma_\mu u_2)}{-t} - \frac{(\bar{u}_2'\gamma^\mu u_1)(\bar{u}_1'\gamma_\mu u_2)}{-u}\Big], \end{aligned} \tag{8.4.76}$$

and similarly

$$\overline{\mathcal{T}} = e^2\Big[\frac{(\bar{u}_1\gamma^\nu u_1')(\bar{u}_2\gamma_\nu u_2')}{-t} - \frac{(\bar{u}_1\gamma^\nu u_2)(\bar{u}_2\gamma_\nu u_1')}{-u}\Big], \tag{8.4.77}$$

where we have suppressed $i\epsilon$ and $t=-(p_1-p_1')^2$ and $u=-(p_1-p_2')^2$. Then the spin-averaged scattering amplitude is

$$\langle|\mathcal{T}|^2\rangle = e^4\Big[\frac{\langle\Phi_{tt}\rangle}{t^2} + \frac{\langle\Phi_{tu}\rangle+\langle\Phi_{ut}\rangle}{tu} + \frac{\langle\Phi_{uu}\rangle}{u^2}\Big], \tag{8.4.78}$$

where [5]

$$\begin{aligned} \langle\Phi_{tt}\rangle &= 2(t^2+2tu+2u^2-8m^2u+8m^4), \\ \langle\Phi_{tu}\rangle &= -2(s^2-8m^2s+12m^4), \\ \langle\Phi_{ut}\rangle &= -2(s^2-8m^2s+12m^4), \\ \langle\Phi_{uu}\rangle &= 2(u^2+2tu+2t^2-8m^2t+8m^4). \end{aligned} \tag{8.4.79}$$

Using the Mandelstam relation $s+t+u=4m^2$, we can rewrite $\langle\Phi_{tt}\rangle$ and $\langle\Phi_{uu}\rangle$ to include the Mandelstam variable s for uniformity so we get [5]

$$\begin{aligned} \langle\Phi_{tt}\rangle &= 2(s^2+u^2-8m^2(s+u)+24m^4), \\ \langle\Phi_{tu}\rangle &= -2(s^2-8m^2s+12m^4), \\ \langle\Phi_{ut}\rangle &= -2(s^2-8m^2s+12m^4), \\ \langle\Phi_{uu}\rangle &= 2(s^2+t^2-8m^2(s+t)+24m^4). \end{aligned} \tag{8.4.80}$$

Then the spin-averaged matrix element is [5]

$$\langle|\mathcal{T}|^2\rangle = 2e^4\Big[\frac{(s^2+u^2-8m^2(s+u)+24m^4)}{t^2} - \frac{2(s^2-8m^2s+12m^4)}{tu} + \frac{(s^2+t^2-8m^2(s+t)+24m^4)}{u^2}\Big]. \tag{8.4.81}$$

Working in the CM frame, we have that $\mathbf{k}_1' = |\mathbf{k}_1|$ and $s = (E_1+E_2)^2$ so the differential cross section is

$$\begin{aligned}\frac{d\sigma}{d\Omega_{\text{CM}}} &= \frac{1}{64\pi^2 s}\frac{|\mathbf{k}_1'|}{|\mathbf{k}_1|}\langle|\mathcal{T}|^2\rangle \\ &= \frac{e^4}{32\pi^2 s}\Big[\frac{(s^2+u^2-8m^2(s+u)+24m^4)}{t^2} - \frac{2(s^2-8m^2s+12m^4)}{tu} \\ &+ \frac{(s^2+t^2-8m^2(s+t)+24m^4)}{u^2}\Big] \\ &= \frac{\alpha^2}{2s}\Big[\frac{(s^2+u^2-8m^2(s+u)+24m^4)}{t^2} - \frac{2(s^2-8m^2s+12m^4)}{tu} \\ &+ \frac{(s^2+t^2-8m^2(s+t)+24m^4)}{u^2}\Big],\end{aligned}$$

hence

$$\frac{d\sigma}{d\Omega_{\text{CM}}} = \frac{\alpha^2}{2s}\Big[\frac{(s^2+u^2-8m^2(s+u)+24m^4)}{t^2} - \frac{2(s^2-8m^2s+12m^4)}{tu} + \frac{(s^2+t^2-8m^2(s+t)+24m^4)}{u^2}\Big], \tag{8.4.82}$$

where $\alpha = e^2/4\pi$ [5]. Suppose that the incoming electrons are prepared in a state with the same four-momentum in the $+z$ direction. Then by simple kinematics, the four-momenta of the incoming and outgoing electrons are [8]

$$\begin{aligned} p_1 &= (E,0,0,p), & p_2 &= (E,0,0,-p), \\ p_1' &= (E,p\sin\theta,0,p\cos\theta), & p_2' &= (E,-p\sin\theta,0,-p\cos\theta). \end{aligned} \tag{8.4.83}$$

Then our Mandelstam variables become

$$s = 4E^2 = E_{\text{CM}}^2, \tag{8.4.84}$$

$$t = 2p^2(\cos\theta - 1), \tag{8.4.85}$$

$$u = -2p^2(1 + \cos\theta) \tag{8.4.86}$$

and our differential cross section becomes

$$\frac{d\sigma}{d\Omega_{\text{CM}}} = \frac{\alpha^2}{E_{\text{CM}}^2\, p^4 \sin^4\theta}\Big[4(m^2 + 2p^2)^2 + (4p^4 - 3(m^2 + 2p^2)^2)\sin^2\theta + p^4\sin^4\theta\Big]. \tag{8.4.87}$$

In the nonrelativistic regime $m \gg p$, the differential cross section reduces to [8]

$$\frac{d\sigma}{d\Omega_{\text{CM}}} = \frac{m^4\alpha^2}{E_{\text{CM}}^2\, p^4 \sin^4\theta}(1 + 3\cos^2\theta). \tag{8.4.88}$$

In the ultrarelativistic regime $m \ll p$, the differential cross section reduces to [8]

$$\frac{d\sigma}{d\Omega_{\text{CM}}} = \frac{\alpha^2}{E_{\text{CM}}^2\, \sin^4\theta}(3 + \cos^2\theta)^2. \tag{8.4.89}$$

Bhabha Scattering

In this section, we will explore the electron-positron scattering process called *Bhabha scattering* named after theoretical physicist Homi Bhabha. Bhabha scattering is used in accelerator physics to measure luminosity which measures the ratio of the number of results after a fundamental interaction, which are called *events*, in a time interval to the cross section of the interaction process [1]. In summary, luminosity is a measurable quantity that quantifies the ability of a particle accelerator to produce the number of interactions and maximizing this value along with a related quantity called the *integrated luminosity* maximizes the performance of a particle accelerator. The leading order diagrams for this process are given below:

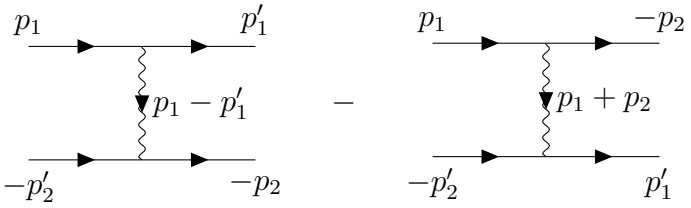


Figure 8.21: Bhabha scattering $e^+e^- \to e^+e^-$

In the first diagram, which is a t-channel diagram, the electron with four-momentum p_1 emits a virtual photon with four-momentum $p_1 - p_1'$ to the positron with four-momentum $-p_2$. The electron then scatters with four-momentum p_1' while the positron absorbs the virtual photon and scatters with four-momentum $-p_2'$. In the second diagram, which is a s-channel diagram, the electron with four-momentum p_1 and the positron with $-p_2$ annihilate to create a photon with four-momentum $p_1 + p_2$ which then recreates the electron and positron except with four-momenta p_1' and $-p_2'$, respectively. The physical interpretation of both these Feynman diagrams is that Bhabha scattering can occur according to the t-channel diagram or the s-channel diagram with a probability assigned to both. Using the Feynman rules, the scattering amplitude for this process is

$$\begin{aligned} i\mathcal{T}_{e^+e^-\to e^+e^-} &= \bar{u}_1'(ie\gamma_\nu)u_1 \frac{1}{i}\frac{g^{\mu\nu}}{(p_1 - p_1')^2 - i\epsilon}\bar{v}_2(ie\gamma_\mu)v_2' \\ &\quad - \bar{v}_2(ie\gamma_\mu)u_1 \frac{1}{i}\frac{g^{\mu\nu}}{(p_1 + p_2)^2 - i\epsilon}\bar{u}_1'(ie\gamma_\nu)v_2' \\ &= \Big(\frac{1}{i}\Big)(ie)^2\Big[\frac{(\bar{u}_1'\gamma^\mu u_1)(\bar{v}_2\gamma_\mu v_2')}{-t} - \frac{(\bar{v}_2\gamma^\mu u_1)(\bar{u}_1'\gamma_\mu v_2')}{-s}\Big], \end{aligned} \tag{8.4.90}$$

and similarly

$$\overline{\mathcal{T}} = e^2\Big[\frac{(\bar{u}_1\gamma^\nu u_1')(\bar{v}_2'\gamma_\nu v_2)}{-t} - \frac{(\bar{u}_1\gamma^\nu v_2)(\bar{v}_2'\gamma_\nu u_1')}{-s}\Big], \tag{8.4.91}$$

where we have suppressed $i\epsilon$ and $s = -(p_1 + p_2)^2$ and $t = -(p_1 - p_1')^2$. Then the

spin-averaged scattering amplitude is

$$\langle|\mathcal{T}|^2\rangle = e^4\Big[\frac{\langle\Phi_{tt}\rangle}{t^2} + \frac{\langle\Phi_{ts}\rangle + \langle\Phi_{st}\rangle}{ts} + \frac{\langle\Phi_{ss}\rangle}{s^2}\Big], \tag{8.4.92}$$

where [5]

$$\begin{aligned}
\langle\Phi_{tt}\rangle &= 2(t^2 + 2st + 2s^2 - 8m^2s + 8m^4),\\
\langle\Phi_{ts}\rangle &= -2(u^2 - 8m^2u + 12m^4),\\
\langle\Phi_{st}\rangle &= -2(u^2 - 8m^2u + 12m^4),\\
\langle\Phi_{ss}\rangle &= 2(s^2 + 2st + 2t^2 - 8m^2t + 8m^4).
\end{aligned} \tag{8.4.93}$$

Now in the CM frame, if we work in the limit $E_{\text{CM}} \gg m$ where m is the mass of the electron (and positron), we can treat the electron as massless so equation (8.4.93) reduces to

$$\begin{aligned}
\langle\Phi_{tt}\rangle &= 2(t^2 + 2st + 2s^2),\\
\langle\Phi_{ts}\rangle &= -2u^2,\\
\langle\Phi_{st}\rangle &= -2u^2,\\
\langle\Phi_{ss}\rangle &= 2(s^2 + 2st + 2t^2).
\end{aligned}$$

In this limit $s + t + u = 0$ so we can solve for t in terms of s and u then insert it into $\langle\Phi_{tt}\rangle$ and similarly, we can solve for s in terms of t and u then insert it into $\langle\Phi_{ss}\rangle$ to get [8]

$$\begin{aligned}
\langle\Phi_{tt}\rangle &= 2(s^2 + u^2)\\
\langle\Phi_{ts}\rangle &= -2u^2,\\
\langle\Phi_{st}\rangle &= -2u^2,\\
\langle\Phi_{ss}\rangle &= 2(t^2 + u^2).
\end{aligned} \tag{8.4.94}$$

Then $\langle|\mathcal{T}|^2\rangle$ in the high energy limit is [8]

$$\begin{aligned}\langle|\mathcal{T}|^2\rangle &= 2e^4\Big[\frac{t^2}{s^2}+\frac{s^2}{t^2}+u^2\Big(\frac{1}{s}+\frac{1}{t}\Big)^2\Big] \\ &= 32\pi^2\alpha^2\Big[\frac{t^2}{s^2}+\frac{s^2}{t^2}+u^2\Big(\frac{1}{s}+\frac{1}{t}\Big)^2\Big]. \end{aligned} \tag{8.4.95}$$

In the CM frame, $|\mathbf{k}_1'| = |\mathbf{k}_1|$ so the differential cross section in the high energy limit is

$$\begin{aligned}\frac{d\sigma}{d\Omega_{\text{CM}}} &= \frac{1}{64\pi^2 s}\frac{|\mathbf{k}_1'|}{|\mathbf{k}_1|}\langle|\mathcal{T}|^2\rangle \\ &= \frac{\alpha^2}{2s}\Big[\frac{t^2}{s^2}+\frac{s^2}{t^2}+u^2\Big(\frac{1}{s}+\frac{1}{t}\Big)^2\Big],\end{aligned}$$

hence [8]

$$\frac{d\sigma}{d\Omega_{\text{CM}}} = \frac{\alpha^2}{2s}\Big[\frac{t^2}{s^2}+\frac{s^2}{t^2}+u^2\Big(\frac{1}{s}+\frac{1}{t}\Big)^2\Big]. \tag{8.4.96}$$

Electron-Muon Scattering

In this section we will explore electron-muon scattering $e^-\mu^- \to e^-\mu^-$. This example is rather unique compared to the other examples we have considered thus far because we are now considering the reaction of two different fermions, namely, an electron and a muon. Also, this process is considered a pedagogical example because since the muon is an unstable particle, it is not realistic to consider the muon as the target in a scattering experiment [1]. This process has only one contributing tree diagram at leading order which is given below [8]:

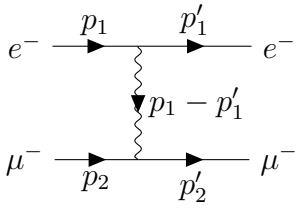


Figure 8.22: Electron-muon scattering $e^-\mu^- \to e^-\mu^-$

For this diagram, we have labeled the external lines according to the particle it belongs to for clarity. The diagram above is a t-channel diagram so we see that

the most probable way an electron and muon might scatter is when the incoming electron with four-momentum p_1 emits a virtual photon with four-momentum $p_1 - p_1'$ to the incoming muon with four-momentum p_2. The electron then scatters with p_1' while the muon absorbs the virtual photon and scatters with four-momentum p_2'. For our calculations, we will use m and M to denote the electron and muon mass, respectively, and the charge of the muon will be Ze where Z is the atomic number [8]. The scattering amplitude for this process is

$$\begin{aligned} i\mathcal{T}_{e^-\mu\to e^-\mu} &= \bar{u}_1'(ie\gamma_\nu)u_1\frac{1}{i}\frac{g^{\mu\nu}}{(p_1-p_1')^2-i\epsilon}\bar{u}_2'(iZe\gamma_\mu)u_2 \\ &= \frac{1}{i}Z(ie)^2\Big[\frac{(\bar{u}_1'\gamma^\mu u_1)(\bar{u}_2'\gamma_\mu u_2)}{-t}\Big] \end{aligned} \tag{8.4.97}$$

and similarly

$$\overline{\mathcal{T}} = Ze^2\Big[\frac{(\bar{u}_1\gamma^\nu u_1')(\bar{u}_2\gamma_\nu u_2')}{-t}\Big], \tag{8.4.98}$$

where we have suppressed $i\epsilon$ and $t = -(p_1 - p_1')^2$. Then the matrix element is

$$\langle|\mathcal{T}|^2\rangle = \frac{Z^2e^4}{t^2}[32M^2m^2 - 4t(M^2-m^2) + 2(s+t)^2 + 2(M^2+m^2-s)^2]. \tag{8.4.99}$$

For this scattering process, we will work in the FT frame. Now the mass of a muon is roughly 200 times that of an electron, $M \approx 200m$, so we will treat the muon as the stationary particle and assume that the recoil is negligible after an electron scatters off the muon which is an accurate approximation [17]. We then have that $|\mathbf{p}_1| = |\mathbf{p}_1'| \equiv |\mathbf{p}|$ which implies that $\mathbf{p}_1 \cdot \mathbf{p}_1' = |\mathbf{p}|^2\cos\theta$ and

$$p_1 = (E, \mathbf{p}), \qquad p_1' = (E, \mathbf{p}), \tag{8.4.100}$$

$$p_2 = (M, \mathbf{0}), \qquad p_2' = (M, \mathbf{0}). \tag{8.4.101}$$

Since the recoil is negligible, the muon hardly moves after the collision so the Mandelstam variable s is essentially the same for both the CM and FT frames and

is approximated as $s = -(p_1 + p_2)^2 \approx -(p_2)^2 = M^2$ [17]. Then our Mandelstam variables become

$$\begin{aligned} s &\approx M^2, \\ t &= -4\,|\mathbf{p}|^2 \sin^2(\theta/2), \\ u &= 2EM + 4\,|\mathbf{p}|^2 \sin^2(\theta/2), \end{aligned} \tag{8.4.102}$$

and the scattering matrix element reduces to

$$\begin{aligned} \langle|\mathcal{T}|^2\rangle &= \Big(\frac{Ze^2M}{|\mathbf{p}|^2 \sin^2(\theta/2)}\Big)^2 (m^2 + |\mathbf{p}|^2 \cos^2(\theta/2)) \\ &= \Big(\frac{4\pi\alpha ZM}{|\mathbf{p}|^2 \sin^2(\theta/2)}\Big)^2 (m^2 + |\mathbf{p}|^2 \cos^2(\theta/2)). \end{aligned} \tag{8.4.103}$$

To get the differential cross section in the FT frame, we take the differential of t to get

$$\begin{aligned} dt &= 2\,|\mathbf{p}|^2 \; d\cos\theta \\ &= 2\,|\mathbf{p}|^2 \; \frac{d\Omega_{\mathrm{FT}}}{2\pi} \\ &= |\mathbf{p}|^2 \; \frac{d\Omega_{\mathrm{FT}}}{\pi}, \end{aligned}$$

then we use our approximation $s \approx M^2$ and the frame independent differential cross section to get

$$\begin{aligned} \frac{d\sigma}{d\Omega_{\mathrm{FT}}} &= \frac{1}{64\pi^2 M^2} \langle|\mathcal{T}|^2\rangle \\ &= \Big(\frac{Z\alpha}{2\,|\mathbf{p}|^2 \sin^2(\theta/2)}\Big)^2 (m^2 + |\mathbf{p}|^2 \cos^2(\theta/2)), \end{aligned} \tag{8.4.104}$$

which is the *Mott formula* [17]. The Mott formula is the approximated differential cross section for electron-proton scattering. In the nonrelativistic regime $|\mathbf{p}| \ll m$,

the Mott formula reduces to

$$\begin{aligned}\frac{d\sigma}{d\Omega_{\mathrm{FT}}} &= \frac{Z^2\alpha^2}{4\left|\mathbf{p}\right|^4\sin^2(\theta/2)}\,m^2 \\ &= \frac{Z^2\alpha^2}{4m^2v^4\sin^2(\theta/2)},\end{aligned} \tag{8.4.105}$$

which is the famous *Rutherford scattering formula* [17]. The Rutherford scattering formula was derived from the Rutherford scattering experiment of alpha particles and gold nuclei. This experiment along with its derived formula gave strong evidence that most of the mass of an atom was concentrated in an interior region called the nucleus. This was instrumental in further developments in the study of atomic structure and ultimately quantum mechanics as well [1]. Interestingly, the Rutherford scattering formula was correctly derived using only classical physics despite the fact that atomic systems are to be treated using quantum mechanics. Indeed when nonrelativistic quantum mechanics was fully developed, the Rutherford scattering formula was derived quantum mechanically and the resulting scattering formula was exactly the same. It is perhaps one of the greatest coincidences of nature that the Rutherford scattering formula agrees with both classical and quantum physics.

Chapter 9

Conclusion and Afterwards

We conclude this thesis with an overview of what we have covered. We briefly reviewed the key concepts from the three theories that make up quantum field theory, namely, special relativity, classical field theory, and nonrelativistic quantum mechanics. Then we analyzed the topic of symmetries and their relation to conservation laws using Noether's theorem. From this, we built a mathematical framework that explained how quantum fields should be constructed in accordance with Lorentz and CPT symmetry. With this knowledge, we slowly constructed the quantum fields based on the field representations of the Lorentz Lie algebra and proceeded to quantize each field according to their particle statistics. Concurrently, we also provided the physical interpretation on the mathematical and physical results that were derived each time. Then we utilized the path integral formulation to construct the appropriate path integrals for each quantum field and derived the propagators for each equations of motions. Then we derived the LSZ reduction formula along with the scattering and decay formulas using scattering theory that incorporated relativity and quantum mechanics to ensure that the formulas obeyed the principles of both theories in a mathematically and physically consistent way. We explored how classical symmetries can be generalized to quantum symmetries to provide a deeper insight into how classical symmetries arise in quantum fields. We then introduced interacting theories which led to the introduction of the Feynman

diagrams and derived the Feynman rules for each quantum field theory. We then gave both pedagogical and real life examples to demonstrate the power of the Feynman diagrams and the remarkable accuracy of quantum field theory.

The next step forward would be to formally introduce the renormalization through the *renormalization group* to account for the divergent integrals that appear in looped Feynman diagrams. Renormalization would also provide a deeper insight into the many kinds of infinities that appear in quantum field theory such as the ultraviolet divergences. Renormalization would also allow us to gain even more accurate results thus it would be beneficial to explore this route for a deeper understanding of quantum field theory. From here, one could explore the application of the Feynman diagrams and the mathematical tools of quantum field theory to other branches of physics such as condensed matter physics. The renormalization group has been increasingly utilized in condensed matter physics over the past several decades with great success [8]. One could also explore the mathematical structure of the Feynman diagrams in the context of graph theory and combinatorics as well.

If one were to continue quantum field theory after learning about renormalization, the next step would be to explore gauge theories, in particular *Yang-Mills theory*, which would naturally lead to the electroweak theory along with symmetry breaking processes and quantum chromodynamics which is the quantum field theory of the nuclear strong force. The study of gauge theories is an extremely rich field with many physicists and mathematicians conducting research on the full scope of gauge theories from its physical aspects to its mathematical aspects such as differential geometry. Feynman diagrams continue to play an important role in gauge theories and beyond thus the study of Feynman diagrams for more complicated theories also presents another research ground. From gauge theories, the next path is an open ended one because as of the date of this thesis, a *Grand Unified Theory* (GUT), namely, a theory that unites the electroweak force and the nuclear strong force, has not been successfully formulated nor tested. Along with

this, a quantum field theory for gravity has not been successfully formulated as well thus another path to follow is the study of quantum gravity. The study of such topics can lead to nonabelian gauge theories using higher dimensional gauge groups such as SU(5) or one could approach these problems using string theory and/or loop quantum gravity.

Appendix

A

In quantum theory, *Dirac notation*, or *bra-ket notation*, is used ubiquitously due to its simplicity. We will present a brief overview on the use and application of this incredibly useful notation. Excellent sources for further study are [7, 11].

The idea of this notational scheme comes from noticing that one can rewrite the inner product notation as follows

$$\langle u, v\rangle \to \langle u|v\rangle, \tag{A.1}$$

where $\langle u|v\rangle$ is now the new notation for the inner product. In this notation, $|v\rangle$ is a vector in a complex vector space $\mathbb{C}^n$. Note that the following rules can be easily generalized to a real vector space $\mathbb{R}^n$ as well. Adding two kets is simple:

$$|u\rangle + |v\rangle = |w\rangle, \tag{A.2}$$

where $|w\rangle$ is the resulting ket from the addition of the two kets $|u\rangle$ and $|v\rangle$. Multiplying by a complex scalar $c \in \mathbb{C}$ is defined as

$$c|v\rangle = |v\rangle c. \tag{A.3}$$

Note that it is a common convention to express the linearity of the kets in the

following manner:

$$|c_1u + c_2v\rangle = |c_1u\rangle + |c_2v\rangle = c_1|u\rangle + c_2|v\rangle, \tag{A.4}$$

however one must be careful to distinguish between which notation represents a scalar and which notation represents the *labeling* of the ket. The other axioms of a vector space easily generalize.

Equipping our vector space with an inner product makes our vector space an inner product space. The axioms of the inner product expressed in Dirac notation are

1. Skew-symmetry: $\langle u|v\rangle = \langle v|u\rangle^*$,
2. Positive-semidefinite: $\langle u|u\rangle \geq 0$, $\langle u|u\rangle = 0$ if and only if $u = 0$,
3. Linearity in second argument: $\langle u|c_1v + c_2w\rangle = c_1^*\langle u|v\rangle + c_2^*\langle u|w\rangle$,

for any $|u\rangle, |v\rangle, |w\rangle \in \mathbb{C}^n$ and any $c_1, c_2 \in \mathbb{C}$ [7]. From the axioms of the inner product, one easily prove the *antilinearity* of the first argument as well.

Proof. For any $c_1, c_2 \in \mathbb{C}$, observe that

$$\begin{aligned}\langle c_1u + c_2w|v\rangle &= \langle v|c_1u + c_2w\rangle^* \\ &= c_1^*\langle v|u\rangle^* + c_2^*\langle v|w\rangle^* \\ &= c_1^*\langle u|v\rangle + c_2^*\langle w|v\rangle.\end{aligned}$$

□

The norm of column vector $|v\rangle$ generalizes as

$$||v\rangle|^2 = |\langle v|v\rangle|^2, \tag{A.5}$$

and the Cauchy-Schwarz inequality is

$$|\langle u|v\rangle| \leq ||u\rangle|\,||v\rangle|\,. \tag{A.6}$$

Now we can separate the bra-ket $\langle u|v\rangle$ into the ket $|v\rangle$ and the *bra* $\langle u|$ which are column and row vectors, respectively. In this manner, the inner product $\langle u|v\rangle$ gives a scalar value. The vector space we have been using has been for the ket which is called the *ket space*. The vector space for bras is called the *bra space* spanned by the basis vectors $\{\langle i|\}_{i=1}^{n}$ and it is the vector space that is dual to the ket space. The previous rules of addition of multiplication of kets easily generalize to bras as well. There exists a bijection between the ket space and bra space such that

$$\begin{aligned} |v\rangle &\overset{\text{DC}}{\longleftrightarrow} \langle v| \\ |1\rangle, |2\rangle, \ldots &\overset{\text{DC}}{\longleftrightarrow} \langle 1|, \langle 2|, \ldots \\ c_1|u\rangle + c_2|v\rangle &\overset{\text{DC}}{\longleftrightarrow} c_1^*\langle u| + c_2^*\langle v|, \end{aligned} \tag{A.7}$$

where DC stands for *dual correspondence* [11]. It is important to note that the dual of $c|v\rangle$ is $c^*\langle v|$.

In elementary linear algebra, all vectors can be written as a linear combination of some basis vectors. To express this in Dirac notation, let $\{|i\rangle\}_{i=1}^{n}$ be a set of basis vectors then for an arbitrary ket $|v\rangle$ we have

$$|v\rangle = \sum_{i=1}^{n} v_i|i\rangle, \tag{A.8}$$

and we see that

$$|v\rangle + |u\rangle = \sum_{i=1}^{n} (v_i + u_i)|i\rangle, \tag{A.9}$$

where v_i and u_i are the components of kets $|v\rangle$ and $|u\rangle$, respectively, in that basis. In quantum mechanics, one usually works with an *orthonormal* basis and if a given set of basis vectors is not orthonormal, one can apply the *Gram-Schmidt* process to turn an orthogonal set of basis vectors into an orthonormal set of basis vectors. To see how this is done in Dirac notation, let $\{|i'\rangle\}_{i'=1}^{n}$ be a set of basis vectors.

Then we first choose an arbitrary basis vector, say $|1'\rangle$, to normalize and we get

$$|1\rangle = \frac{|1'\rangle}{||1'\rangle|}. \tag{A.10}$$

Then we choose a second vector, say $|2'\rangle$, and we first make this basis vector orthogonal to $|1\rangle$

$$|\widetilde{2}\rangle = |2'\rangle - |1\rangle\langle 1|2'\rangle, \tag{A.11}$$

and we see that

$$\begin{aligned}\langle 1|\widetilde{2}\rangle &= \langle 1|2'\rangle - \langle 1|1\rangle\langle 1|2'\rangle \\ &= \langle 1|2'\rangle - \langle 1|2'\rangle \\ &= 0.\end{aligned}$$

We then normalize $|\widetilde{2}\rangle$ to get our second orthonormal basis vector

$$|2\rangle = \frac{|\widetilde{2}\rangle}{\left||\widetilde{2}\rangle\right|}. \tag{A.12}$$

Repeating this process for n basis vectors gives us

$$\begin{aligned}|\widetilde{n}\rangle &= |n'\rangle - \sum_{i=1}^{n-1} |i\rangle\langle i|n'\rangle \\ |n\rangle &= \frac{|\widetilde{n}\rangle}{||\widetilde{n}\rangle|}.\end{aligned} \tag{A.13}$$

With the orthonormality of the basis vectors, we see that the inner product of two basis vectors is

$$\langle i|j\rangle = \delta_{ij} = \begin{cases} 1, & \text{if } i = j \\ 0, & \text{if } i \neq j \end{cases} \tag{A.14}$$

where δ_{ij} is the Kronecker delta function. We can then find the components of an

arbitrary ket $|v\rangle$ in the following manner:

$$\begin{aligned}
\langle j|\Big(|v\rangle &= \sum_{i=1}^{n} v_i|i\rangle\Big) \\
\langle j|v\rangle &= \sum_{i=1}^{n} v_i\langle i|j\rangle \\
\langle j|v\rangle &= \sum_{i=1}^{n} v_i\,\delta_{ij} \\
\langle j|v\rangle &= v_j, \qquad \text{(A.15)}
\end{aligned}$$

thus

$$|v\rangle = \sum_{i=1}^{n} |i\rangle\langle i|v\rangle. \qquad \text{(A.16)}$$

If we observe from equation (A.16)

$$|v\rangle = \Big(\sum_{i=1}^{n} |i\rangle\langle i|\Big)|v\rangle, \qquad \text{(A.17)}$$

we see that

$$\sum_{i=1}^{n} |i\rangle\langle i| = \hat{1} \equiv \hat{I}, \qquad \text{(A.18)}$$

where $\hat{1} \equiv \hat{I}$ is called the *identity operator* and equation (A.18) is called the *completeness relation*. The completeness relation is one of the most important operations in quantum mechanics because we can insert the identity operator at any place in a long chain of kets and/or bras using any set of orthonormal basis. This will almost always massively simplify a problem and gives physical insight into the nature a quantum state when expanded. From equation (A.18), consider the term

$$\hat{\mathbb{P}}_i = |i\rangle\langle i|. \qquad \text{(A.19)}$$

Expression (A.19) is clearly a Hermitian operator which is called the *projection operator* since for any ket $|v\rangle$

$$\hat{\mathbb{P}}_i|v\rangle = |i\rangle\langle i|v\rangle = |i\rangle v_i, \qquad \text{(A.20)}$$

thus the projection operator projects the i^{th} component of a ket $|v\rangle$. It is easy to see that

$$\hat{\mathbb{P}}_i\hat{\mathbb{P}}_i = |i\rangle\langle i|i\rangle\langle i| = |i\rangle\langle i| = \hat{\mathbb{P}}_i, \tag{A.21}$$

and the completeness relation can be written as

$$\sum_i \hat{\mathbb{P}}_i = \hat{1} \equiv \hat{I}, \tag{A.22}$$

for some basis $|i\rangle$.

From the postulates of quantum mechanics, observables are Hermitian operators that act on the elements of the Hilbert space. Since we have thus far only covered linear algebra, Hermitian operators are simply Hermitian matrices which are denoted as $\hat{A}$ for any arbitrary observable A. For an arbitrary operator $\hat{X}$ and some arbitrary ket $|v\rangle$, we have

$$\hat{X}|v\rangle \overset{\text{DC}}{\longleftrightarrow} \langle v|\hat{X}^\dagger, \tag{A.23}$$

where $\hat{X}^\dagger$ is the *Hermitian adjoint*, or *Hermitian conjugate*, of $\hat{X}$ and the dagger $\dagger$ is the Hermition adjoint, or Hermitian conjugate, operation [11]. We see that an operator is defined as Hermitian if

$$\hat{X} = \hat{X}^\dagger. \tag{A.24}$$

Then we see that

$$\big(|v\rangle\big)^\dagger = \langle v|, \tag{A.25}$$

$$\big(\langle v|\big)^\dagger = |v\rangle. \tag{A.26}$$

It is common practice to denote the operator and its *eigenket*, which is the Dirac notation version of the eigenvector, by a labeling scheme that clearly shows that the ket is an eigenket of the operator. One common convention is to use uppercase

letters to denote an operator and label its respective eigenket and eigenvalue with the equivalent lowercase letter. It is common to use Greek and/or Latin letters. For example, for arbitrary operators $\hat{A}$ and $\hat{\Omega}$, we have

$$\hat{A}|a\rangle = a|a\rangle, \tag{A.27}$$

$$\hat{\Omega}|\omega\rangle = \omega|\omega\rangle, \tag{A.28}$$

where a and ω are the eigenvalues. There are some exceptions to this, such as the labeling of creation and annihilation operators along with their eigenkets, but it will usually be known from the context whether or not a ket is an eigenket of some operator. It is known from elementary linear algebra that the eigenkets $\{|a_i\rangle\}_{i=1}^n \equiv \{|i\rangle\}_{i=1}^n$ of any $n \times n$ Hermitian matrix $\hat{A}$ form an orthogonal basis of $\mathbb{C}^n$. We can normalize the eigenkets, if they are not already normalized, to form an orthonormal basis then for any arbitrary operator $\hat{X}$

$$\begin{aligned} \hat{X} &= \hat{1}\hat{X} \\ &= \sum_i |i\rangle\langle i|\hat{X}\hat{1} \\ &= \sum_i \sum_j |i\rangle\langle i|\hat{X}|j\rangle\langle j|, \end{aligned} \tag{A.29}$$

where we have insert the identity operator $\hat{1}$ using the *same* set of basis vectors except with a different labeling. Then the term

$$\underbrace{\langle i|}_{\text{row}} \hat{X} \underbrace{|j\rangle}_{\text{column}}, \tag{A.30}$$

is the matrix element of $\hat{X}$ in row i column j. Then we have that any Hermitian

operator $\hat{X}$ (including matrices) can be represented in matrix form

$$\hat{X} \doteq \begin{pmatrix} \langle 1|\hat{X}|1\rangle & \langle 1|\hat{X}|2\rangle & \dots \\ \langle 2|\hat{X}|1\rangle & \langle 2|\hat{X}|2\rangle & \dots \\ \vdots & \vdots & \ddots \end{pmatrix}, \tag{A.31}$$

where the symbol $\doteq$ means "represented by."

In quantum mechanics, we work with infinite-dimensional Hilbert spaces which is a type of function space equipped with an inner product defined as

$$\langle f, g\rangle = \int dx\, f^*(x)g(x), \tag{A.32}$$

or in Dirac notation

$$\langle f|g\rangle = \int dx\, f^*(x)g(x), \tag{A.33}$$

for any $f(x), g(x) \in \mathbb{C}$ on the interval $(-\infty, +\infty)$ where we denote an integral with no labeled bounds as being integrated on $(-\infty, +\infty)$ [7]. With Hilbert spaces, a ket $|v\rangle$ is still a vector in a complex vector space $\mathbb{C}^n$ but the bra is different. A bra $\langle u|$ is a linear map $\langle u| : V \to \mathbb{C}$, or in Dirac notation, $\langle u| : |v\rangle \to \langle u|v\rangle$, where V is a vector space. Before we explain this further, note that the completeness relation easily generalizes in one dimension as

$$\int dx\, |x\rangle\langle x| = \hat{1}, \tag{A.34}$$

where the basis kets are labeled with the *continuous* label x. Then for an arbitrary ket $|f\rangle$ and bra $\langle x|$, we have

$$\begin{aligned} \langle x|f\rangle &= \langle x|\hat{1}|f\rangle \\ &= \int dx'\, \langle x|x'\rangle\langle x'|f\rangle, \end{aligned} \tag{A.35}$$

which is the projection of $|f\rangle$ onto the basis ket $|x\rangle$ so

$$\langle x|f\rangle = f(x). \tag{A.36}$$

Then we see that equation (A.33) is derived as

$$\begin{aligned}\langle f|g\rangle &= \langle f|\hat{\mathbb{1}}|g\rangle \\ &= \int dx\, \langle f|x\rangle\langle x|g\rangle \\ &= \int dx\, f^*(x)g(x).\end{aligned} \tag{A.37}$$

In equation (A.35), we now know that $\langle x|f\rangle = f(x)$ and $\langle x'|f\rangle = f(x')$ so we must have that

$$\langle x|x'\rangle = \frac{1}{2\pi}\int dp\, e^{i(x-x')p} = \delta(x-x'), \tag{A.38}$$

where $\delta(x-x')$ is the Dirac delta function and in n dimensions, we have

$$\langle \mathbf{x}|\mathbf{x}'\rangle = \frac{1}{(2\pi)^n}\int d^n p\, e^{i(\mathbf{x}-\mathbf{x}')\cdot\mathbf{p}} = \delta^n(\mathbf{x}-\mathbf{x}'). \tag{A.39}$$

Equation (A.39) tells us how to normalize basis kets labeled with a continuous index.

With the basics of Dirac notation established, we will derive the canonical commutation relations in three dimensions

$$\begin{aligned}[\hat{X}_i, \hat{X}_j] &= 0, \\ [\hat{P}_i, \hat{P}_j] &= 0, \\ [\hat{X}_i, \hat{P}_j] &= i\hbar\delta_{ij}.\end{aligned} \tag{A.40}$$

This example will demonstrate the full scope on how to use Dirac notation along with its simplicity. Before we can even derive the commutation relation (A.40), we must derive many prerequisite results. Using the postulates of quantum mechanics,

let $\hat{\mathbf{X}}$ and $\hat{\mathbf{P}}$ be the Hermitian position and momentum operators, respectively, such that

$$\hat{\mathbf{X}}|\mathbf{x}\rangle = \mathbf{x}|\mathbf{x}\rangle, \tag{A.41}$$

$$\hat{\mathbf{P}}|\mathbf{p}\rangle = \mathbf{p}|\mathbf{p}\rangle, \tag{A.42}$$

where

$$\begin{aligned} |\mathbf{x}\rangle &= |x_1, x_2, x_3\rangle \equiv |x, y, z\rangle, \\ |\mathbf{p}\rangle &= |p_1, p_2, p_3\rangle \equiv |p_x, p_y, p_z\rangle, \end{aligned} \tag{A.43}$$

and

$$\hat{X}_i|x_i\rangle = x_i|x_i\rangle, \tag{A.44}$$

$$\hat{P}_i|p_i\rangle = p_i|p_i\rangle, \tag{A.45}$$

for $i = 1, 2, 3$. The fact that $|\mathbf{x}\rangle$ and $|\mathbf{p}\rangle$ are simultaneous eigenkets implies that the position and momentum operators commute with amongst each other in different directions so the first two commutations relations of (A.40) are trivially established. Now from Wigner's theorem, momentum is the generator of translations so let $\hat{T}(\mathbf{x})$ be the translation operator then an infinitesimal translation by $d\mathbf{x}$ is

$$\hat{T}(d\mathbf{x}) = 1 - \frac{i\hat{\mathbf{P}}}{\hbar} \cdot d\mathbf{x}, \tag{A.46}$$

then

$$\hat{T}(d\mathbf{x})|\mathbf{x}\rangle = |\mathbf{x} + d\mathbf{x}\rangle, \tag{A.47}$$

for some position eigenket $|\mathbf{x}\rangle$ [11]. Let $|\psi\rangle$ be an arbitrary state ket then in one

dimension

$$\begin{aligned}\hat{T}(\Delta x)|\psi\rangle &= \Big(1 - \frac{i\hat{P}}{\hbar}\Delta x\Big)|\psi\rangle \\ &= \int dx \,\Big(1 - \frac{i\hat{P}}{\hbar}\Delta x\Big)|x\rangle\langle x\|\psi\rangle \\ &= \int dx \,|x + \Delta x\rangle\langle x|\psi\rangle \\ &= \int dx \,|x\rangle\langle x - \Delta x|\psi\rangle \\ &= \int dx \,|x\rangle\Big(\langle x|\psi\rangle - \langle \Delta x|\psi\rangle\Big) \\ &= \int dx \,|x\rangle\Big(\langle x|\psi\rangle - \Delta x\frac{\partial}{\partial x}\langle x|\psi\rangle\Big), \end{aligned} \tag{A.48}$$

where Δx is infinitesimal and we have changed variables in the fourth equality but since the choice is arbitrary, we have relabeled the variable to its original labeling [11]. Comparing both sides of equation (A.48) yields

$$\hat{P}|\psi\rangle = -i\hbar\frac{\partial}{\partial x}|\psi\rangle, \tag{A.49}$$

and operating with $\langle x|$ on the left yields

$$\langle x|\hat{P}|\psi\rangle = -i\hbar\frac{\partial}{\partial x}\langle x|\psi\rangle. \tag{A.50}$$

Since $|\psi\rangle$ was arbitrary, we have that the momentum operator in the position basis is

$$\hat{P} \xrightarrow{x-basis} -i\hbar\frac{\partial}{\partial x}, \tag{A.51}$$

and in three dimensions

$$\hat{\mathbf{P}} \xrightarrow{x-basis} -i\hbar\nabla. \tag{A.52}$$

With expressions (A.51) and (A.52), we can now derive the scalar product of

position and momentum, namely $\langle x|p\rangle$. Observe that in one dimension

$$\begin{aligned}\langle x|\hat{P}|p\rangle &= p\langle x|p\rangle \\ -i\hbar\frac{\partial}{\partial x}\langle x|p\rangle &= p\langle x|p\rangle \\ \langle x|p\rangle &= Ne^{ixp/\hbar},\end{aligned} \tag{A.53}$$

where N is some normalization constant. To find N, note that

$$\begin{aligned}\delta(x-x') &= \langle x|x'\rangle \\ &= \int dp\,\langle x|p\rangle\langle p|x'\rangle \\ &= |N|^2\int dp\,e^{i(x-x')p/\hbar} \\ &= |N|^2\,(2\pi\hbar)\,\delta(x-x').\end{aligned} \tag{A.54}$$

Equating both sides of the equation and choosing the constant to be real and positive by convention, we get

$$N = \frac{1}{\sqrt{2\pi\hbar}}, \tag{A.55}$$

thus

$$\langle x|p\rangle = \frac{1}{\sqrt{2\pi\hbar}}\,e^{ixp/\hbar} \tag{A.56}$$

and in three dimensions

$$\langle x|p\rangle = \frac{1}{(2\pi\hbar)^{3/2}}\,e^{i\mathbf{x}\cdot\mathbf{p}/\hbar}. \tag{A.57}$$

Finally, we can derive the third commutation relation, namely, $[\hat{X}_i, \hat{P}_j] = i\hbar\delta_{ij}$. In one dimension, observe that

$$[\hat{X},\hat{P}] = \int dx\,dx'\,|x\rangle\langle x|[\hat{X},\hat{P}]|x'\rangle\langle x'|. \tag{A.58}$$

Then

$$\begin{aligned}\langle x|[\hat{X},\hat{P}]|x'\rangle &= \langle x|(\hat{X}\hat{P}-\hat{P}\hat{X})|x'\rangle \\ &= \langle x|\hat{X}\hat{P}|x'\rangle - \langle x|\hat{P}\hat{X}|x'\rangle \\ &= x\langle x|\hat{P}|x'\rangle + i\hbar\frac{\partial}{\partial x}\Big(\langle x|\hat{X}|x'\rangle\Big) \\ &= x\langle x|\hat{P}|x'\rangle + i\hbar\frac{\partial}{\partial x}\Big(x\langle x|x'\rangle\Big) \\ &= -ix\hbar\frac{\partial}{\partial x}\Big(\langle x|x'\rangle\Big) + ix\hbar\frac{\partial}{\partial x}\Big(\langle x|x'\rangle\Big) + i\hbar\langle x|x'\rangle \\ &= i\hbar\langle x|x'\rangle, \end{aligned} \tag{A.59}$$

thus

$$\begin{aligned}[\hat{X},\hat{P}] &= i\hbar\int dx\,dx'\,|x\rangle\langle x|x'\rangle\langle x'| \\ &= i\hbar\hat{I}. \end{aligned} \tag{A.60}$$

Therefore, we have the canonical commutation relations

$$\begin{aligned}[\hat{X}_i,\hat{X}_j] &= 0, \\ [\hat{P}_i,\hat{P}_j] &= 0, \\ [\hat{X}_i,\hat{P}_j] &= i\hbar\delta_{ij}. \end{aligned} \tag{A.61}$$

List of References

[1] T. Stephen Thornton and Andrew Rex. *Modern Physics for Scientists and Engineers*. 4th ed. Cengage Learning, 2013. ISBN: 978-1-133-10372-1.

[2] David J. Griffiths. *Introduction to Electrodynamics*. 4th ed. Cambridge University Press, 2017. ISBN: 978-0-321-85656-2. DOI: 10.1017/9781108333511.

[3] Albert Einstein. "Zur Elektrodynamik bewegter Körper. (German) [On the electrodynamics of moving bodies]". In: *Annalen der Physik* 322.10 (1905), pp. 891–921. DOI: http://dx.doi.org/10.1002/andp.19053221004.

[4] R. John Taylor. *Classical Mechanics*. University Science Books, 2005. ISBN: 189138922X.

[5] M. Srednicki. *Quantum Field Theory*. Cambridge University Press, Jan. 2007. ISBN: 978-0-521-86449-7.

[6] Tai-Kai Ng. *Introduction to Classical and Quantum Field Theory*. Wiley-VCH, 2009. ISBN: 978-3527407262.

[7] Ramamurti Shankar. *Principles of Quantum Mechanics*. 2nd ed. Plenum, 1994. ISBN: 0-306-44790-8.

[8] Michael E. Peskin and Daniel V. Schroeder. *An Introduction to Quantum Field Theory*. Reading, USA: Addison-Wesley, 1995. ISBN: 978-0-201-50397-5.

[9] Claude Cohen-Tannoudji, Bernard. Diu, and Franck. Laloe. *Quantum Mechanics Vol.1:Basic Concepts, Tools, and Applications*. 2nd ed. Wiley-VCH, 2019. ISBN: 978-3-527-34553-3.

[10] Erik Aurell Giulio Chiribella and Karol Zyczkowski. “A Wigner Theorem for Quantum Evolutions”. In: (Jan. 2021).

[11] J. J. Sakurai and Jim Napolitano. *Modern Quantum Mechanics*. 2nd ed. Cambridge University Press, 2017. ISBN: 978-1-108-42241-3. DOI: 10.1017/9781108499996.

[12] Brian C. Hall. *Lie Groups, Lie Algebras, and Representations: An Elementary Introduction*. Graduate Texts in Mathematics. Springer, 2016. ISBN: 978-3-319-37433-8.

[13] John W. Milnor and James D. Stasheff. *Characteristic Classes*. Annals of Mathematics Studies. Princeton University Press, Aug. 1974. ISBN: 978-0691081229.

[14] Mikio Nakahara. *Geometry, Topology and Physics*. 2nd ed. Graduate Student Series in Physics. CRC Press, 2003. ISBN: ISBN-13: 978-0750306065.

[15] Michele Maggiore. *A Modern Introduction to Quantum Field Theory*. Oxford Master Series in Statistical, Computational, and Theoretical Physics. Oxford University Press, 2005. ISBN: 978-0198520740.

[16] Xavier Bekaert and Nicolas Boulanger. “The Unitary Representations of the Poincare group In Any Spacetime Dimension”. In: (Dec. 2006).

[17] David Griffiths. *Introduction to Elementary Particles*. 2nd ed. New York, USA: Wiley-VCH, 2008. ISBN: 978-3-527-40601-2.

Biography

Junhyup Sung, who also goes by the name Eric, was born in Seoul, Republic of Korea. He moved to the United States of America at a young age and graduated from Georgetown Day High School in Washington D.C. in 2015. He subsequently enrolled at Tulane University where after experimenting with many different fields, he eventually found his lifelong passion in mathematics and received his Bachelor of Science in mathematics in May 2020. In September 2020, he entered the mathematics graduate program at Tulane University under the $4+1$ program and will be graduating with a M.S. in Applied Mathematics in May 2021.