Zero-energy Modes in Quantum Field Theories

by

Shovon Biswas

B.Sc., Bangladesh University of Engineering and Technology, 2018
M.Sc., The University of British Columbia, 2019

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

The Faculty of Graduate and Postdoctoral Studies

(Physics)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

August 2025

The following individuals certify that they have read, and recommend to the Faculty of Graduate and Postdoctoral Studies for acceptance, the thesis entitled:

Zero-energy Modes in Quantum Field Theories

submitted by Shovon Biswas

in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Physics.

Examining Committee:

Gordon W. Semenoff, Professor, Physics, University of British Columbia
Supervisor

Joanna Karczmarek, Associate Professor, Physics, University of British Columbia
Co-supervisor

Kristin Schleich, Associate Professor, Physics, University of British Columbia
University Examiner

Kai Behrend, Professor, Mathematics, University of British Columbia
University Examiner

Don N. Page, Professor, Physics, University of Alberta
External Examiner

Additional Supervisory Committee Members:

Mark Van Raamsdonk, Professor, Physics, University of British Columbia
Supervisory Committee Member

Alison Lister, Professor, Physics, University of British Columbia
Supervisory Committee Member

Julio-Parra Martinez, Professor, Physics, IHES
Supervisory Committee Member

Abstract

We discuss three instances where zero-energy or soft modes appear in quantum field theory.

First, we examine massless fermions in a $2+1$ dimensional system with a spatial boundary, specifically graphene in half-space. Two boundary conditions and their interplay with the discrete and continuous symmetries of the system are analyzed. For doubled fermions, we identify a special case that respects CPT symmetry but breaks Lorentz and conformal symmetry, featuring fermion zero mode edge states. These edge states lead to unconventional representations of scale, phase, and translation symmetries, and enforcing symmetry constraints results in edge ferromagnetism.

Second, we investigate the infrared structure of a massless scalar theory coupled to fermions. We demonstrate the existence of a field theory containing massless scalar particles that mirrors the infrared structure of quantum electrodynamics and perturbative quantum gravity but lacks gauge invariance, internal symmetries, or apparent asymptotic symmetry. Unlike soft photons and gravitons, soft scalars do not decouple from dressed states and are generally produced during interactions of hard dressed particles, though their entanglement is minimal.

Lastly, we develop a novel method to calculate changes in an operator's expectation value at asymptotic times, relevant to gravitational wave observations, by exploiting its soft limit. We derive a formula for asymptotic *in-in* observables from the soft limit of five-point amputated response functions. Using this, we re-derive the KMOC formulas for linear impulse and radiated momentum during scattering and provide an unambiguous calculation of radiated angular momentum at leading order. We introduce a causal method of computing classical observables using the Schwinger-Keldysh formalism.

Lay Summary

Quantum field theory provides a powerful framework for describing elementary particles in nature. Among its many features, particles or excitations with zero energy frequently arise. Although they may seem inconsequential, their presence challenges many foundational aspects of the theory. Their presence also gives rise to physical systems with unique properties. This thesis examines these zero-energy modes in three different setups.

Preface

This thesis is based on published research by Shovon Biswas and the collaborators mentioned below.

Chapters 1 of this work is an original synthesis of the relevant literature, laying out the context and motivation of the thesis.

Chapter 2 is based on the work [1] with Gordon W. Semenoff. The idea of the project was proposed by Gordon Semenoff. Both authors contributed equally in the projects.

Chapter 3 is based on the work [2] with Gordon W. Semenoff. The project was proposed by Gordon Semenoff. Both authors contributed equally in the projects.

Chaper 4 is based on the work [3] with Julio Parra-Martinez. The main idea of the project was proposed by the author and then was subsequently refined by Julio Parra-Martinez. Most calculations were performed by the author and checked by Julio Parra-Martinez.

In preparing this thesis, Generative AI tools were employed exclusively for editorial purposes, such as grammatical refinement and restructuring of texts provided by the author. These tools were not utilized in any other way for the actual research presented in this work.

Table of Contents

List of Figures

Acknowledgements

I would like to thank my supervisor Gordon W. Semenoff for his constant guidance throughout my graduate life. Having had the opportunity to observe his way of thinking as a physicist, I attempted to emulate it, only to quickly realize that it is impossible. I am also deeply grateful to Gordon for granting me the academic freedom to explore the magical world of theoretical physics.

My sincere gratitude to my co-supervisor, Joanna Karczmarek, for her invaluable support in my academic journey – and, perhaps even more importantly, for her support and encouragement on topics not related to academia.

I would like to thank Julio Parra-Martinez for his willingness to collaborate with me, despite my lack of experience in scattering amplitudes. Some of my fondest graduate school memories come from our many physics discussions – first in his office during his brief time in Vancouver, and later in Paris at IHES.

I was fortunate to meet some incredible postdocs at UBC, from whom I learned a great deal of physics. My deepest gratitude to James Sully, my first mentor in theoretical physics, whose guidance has been invaluable. I consider myself truly fortunate to have known Panos Betzios, Arjun Kar, Lampros Lamprou, Charles Marteau, and Felipe Rosso. Their influence in my journey extends far beyond theoretical physics.

My deep gratitude to Prof. Arshad Momen for mentoring and inspiring me to pursue a career in theoretical physics, and to my undergrad supervisor Prof. Zunaed Baten for encouraging me to follow the path.

The friendships I built in and through Hennings 418 made graduate school unforgettable. From many nerdy conversations with Abhisek, Chris, Christine, Clara, David, Jonah, Petar, Pompey, Rana, and Wyatt to the endless laughter we shared, these connections have meant a lot to me. A big shout-out to 4M for all the swags from thousands of miles away. Rafid, Pritam, Rahgeer, Tahsin, PC – let us not let our adulthood win over our childhood.

I am truly grateful to John Blachut and Sherry Kennedy for their kindness and support over the years. This journey would have been much more difficult without them.

Now, I would like to take a moment to acknowledge the people to whom I owe everything in my life. My deepest gratitude goes to my parents, Suvash Biswas and Lovely Roy, whose support and sacrifice have been with me through my every decision. I am also blessed with the most wonderful sister, Susmita Biswas. And above all, my heartfelt thanks to my wife Prianjana Deb – for your love, for choosing

to walk this journey with me, and for sharing your life with me. I couldn't have done this without you.

Dedication

To Lovely Roy, who could be many things in her life but chose to be my mother.

Chapter 1

Introduction

Invitation

Quantum field theory (QFT) is a mathematical framework in which the fundamental objects are quantum fields and the particles can be thought of as various excitations of the quantum fields. QFT was developed by Dirac, Pauli, Feynman, Dyson, Schwinger, Tomonaga, Oppenheimer, and many others as a theory to describe the interaction between light and matter. This theory, later known as Quantum Electrodynamics (QED), beautifully unifies the principles of special relativity with quantum mechanics, revealing the intricate relativistic quantum nature of electrons and photons.

Building on the extraordinary success of QED, the concept of elementary particles as excitations of quantum fields was expanded, leading to the formalism of the Standard Model. This framework not only described all elementary particles but also predicted the existence of several new fundamental particles, later confirmed through experiments in particle colliders. Among these the most recent was the Higgs boson, theoretically predicted [4] in 1964 and experimentally discovered [5, 6] in 2012. Countless other checks have been performed at various particle colliders solidifying QFT as the most successful theoretical framework of modern physics.

QFTs have also proven to be extremely successful in describing various condensed matter phenomena such as phase transitions, superconductivity, and non-equilibrium dynamics. The basic idea is to treat particles and quasiparticles, such as electrons and phonons, as excitations of quantum fields. Hence, QFTs are incredibly successful in describing a wide range of phenomena across energy scales, from the behavior of condensed matter systems ($\sim$ eV) to the high-energy interactions in particle colliders ($\sim$ TeV).

In this thesis, we shall focus on a special corner of QFTs – the physics of the zero-energy particles or the zero-energy excitations of quantum fields. At first glance, this subject may appear unassuming, or even overly formal, as zero-energy particles seem inconsequential due to their lack of energy to influence any system. However, we shall try to convince the reader that that is not the case. Rather, these zero energy excitations are enough to challenge some of the theoretical foundations of QFT, have observational consequences, and help obtain classical results necessary for useful for gravitational wave detectors.

Effective field theories

An interesting aspect of our observable universe is the separation between low-energy (large distances) and high-energy (small distances) phenomena. For instance, to analyze the atomic spectra of the hydrogen atom, we can disregard gravitational interactions between the electron and proton, even though gravity is inherently present due to the mutual attraction of all massive particles. The reason for this omission lies in the energy scale — gravitational effects become noticeable at much higher energy levels than those relevant to atomic spectra. In this low-energy regime, gravitational degrees of freedom effectively decouple from the system. In a similar fashion, by identifying the energy scale relevant to a given problem, we can construct a quantum field theory (QFT) tailored to that system. This QFT operates effectively below a specific energy cutoff, denoted $\Lambda_{\text{cut-off}}$, where it provides a fully well-behaved framework for making predictions about the system. However, beyond $\Lambda_{\text{cut-off}}$, the QFT shows theoretical inconsistencies, signaling the emergence of new physics. A QFT description, tailored to a system and valid within a particular energy range using only the relevant degrees of freedom and interactions, is known as an effective field theory (EFT).

Let us now talk about EFTs description of condensed matter systems. A condensed matter system is typically defined on a lattice (lattice constant a), making it fundamentally different from the Lorentz invariant QFTs describing the fundamental particles where space is continuous. However, at distances significantly larger than the lattice spacing, symmetries of continuum field theories, such as translation and rotation, emerge. Consequently, we can formulate EFTs (not necessarily relativistic) valid below $\Lambda_{\text{cut-off}} \sim \frac{1}{a}$ to describe such systems. Examples of many such EFTs can be found in [7]. In this thesis, we will study a fermionic EFT that describes graphene at low energies.

The cutoff dependence of EFTs should be contrasted with QFTs that are valid at *all* energy scales—known as UV-complete QFTs. UV-complete QFTs are notoriously difficult to construct. A well-known example is the search for a UV-complete quantum field theory of gravity, which remains a long-standing open problem in theoretical physics. Fortunately, for specific scenarios, an EFT description of certain gravitational phenomena can be written down. One example is the emission of very low-energy gravitational waves (GWs) by accelerating celestial objects. When the wavelengths of such GWs are much larger than the size of the planets and stars, those objects can be treated as point particles, and such interactions can be described by an EFT of gravitons interacting with point masses. We will explore such systems described by gravitational EFTs.

Zero modes

In this thesis, we will always assume that a QFT description of our system under consideration is available in the sense of EFT. That is, the QFT description is strictly

valid below some high-energy cut-off scale $\Lambda_{\text{cut-off}}$. Once we have written down the EFT, we can concentrate on the energy spectrum.

Zero modes are the particles or quasi-particles with zero energy. Although they can appear in both bosonic and fermionic theories, fermionic zero modes are often more interesting. The existence of zero modes in a system is non-trivial and often requires a boundary or a defect. In a system with a boundary, they usually appear as edge states that are localized close to the boundary. The presence of fermionic zero modes in theory often points to the anomaly of the theory – a situation when the quantum theory fails to respect some symmetries of the classical theory [8].

The zero modes also lead to surprising physical results. For example, they can carry fractional charges [9]. They are particularly stable under local perturbations. In fact, a major pursuit in condensed matter physics is the quest for Majorana zero modes—unique states that act as their own antiparticles and appear at the ends of one-dimensional topological superconducting wires. The stability of these Majorana zero modes against local perturbation makes them ideal candidates for protecting quantum information from noise. In this thesis, we find zero modes in a $2+1$ dimensional fermionic system and discuss their physical significance.

Infrared divergence

Suppose we are interested in the following $2 \to 2$ scattering process where we send in two electrons with momenta p_1, p_2, let them interact and finally measure the amplitude where they end up with final momenta p_1', p_2'.

$$e^- + e^- \to e^- + e^-.$$

We can use QED to describe such processes. In particular, the transition amplitude of the above process is given by the S matrix: $S = \langle p_1' p_2' |\, p_1 p_2 \rangle$ Various physical quantities such as the probability of that event happening or the scattering cross-section depend on the square of the modulus of the S matrix $|S|^2$.

When computing a S matrix, the basic rule of QFT tells us to include processes that correspond to the *virtual exchanges* i.e. the emissions and absorptions of virtual particles. To be more precise, even though we are considering the scattering of electrons, their interactions are mediated by photons – meaning that there are *virtual* photons emitted and reabsorbed by the electrons, and their contributions must be accounted for. The word *virtual* simply means that while the initial and final electrons of mass m_e must obey the relativistic on-shell condition $p_i^2 = m_e^2$, the intermediate photons do not need to be subjected to such on-shell constraints. Infrared divergence is the fact that when a theory has massless particles, such as QED with massless photon, the processes with virtual massless particle exchange give rise to infinite results. These divergences are caused by zero-energy photons, also known as *soft* photons. More precisely, at any finite order in perturbation theory, one encounters integrals of the form $\int_0^{\Lambda_{\text{cut-off}}} d\omega/\omega$ which are divergent at the lower

limit of the integral. Moreover, when all these infinite results are summed over, the full S matrix, which calculates the transition amplitudes between two processes at all-orders in the perturbation constants, vanishes as $S \sim e^{-\infty} = 0$. This is true for any process in QED. Does that mean that QED is ill-defined?

The issue with infrared divergences has been a known challenge since the formulation of QED. There are two main prescriptions to address this issue. The first modifies the problem by introducing an infrared cut-off in such a way that $|S|^2$ is infrared-finite when the cut-off is taken to zero, though the S-matrix itself remains infrared divergence at this limit. The second prescription produces an infrared-finite S-matrix but necessitates technical steps that deviate from the standard mathematical frameworks of QFT. A fully rigorous and mathematically sound infrared-finite formulation of QFT remains an active area of research.

In addition to addressing the conventional challenges of resolving infrared divergences in QFTs, one can explore intriguing questions about the quantum information content of soft photons. These questions further differentiate the two approaches and could lead to observable distinctions. Another avenue of inquiry involves understanding the physical origin of infrared divergence in a theory. Specifically, if a theory exhibits infrared divergence, does this point to a profound underlying symmetry, or is it merely a consequence of the presence of a massless particle in the spectrum? We will discuss some of these issues related to infrared divergence in this thesis.

Classical observables from amplitudes

Einstein's theory of general relativity predicts that accelerating masses produce gravitational waves, akin to how accelerating charges generate electromagnetic waves. In 2015, the LIGO-Virgo collaborations achieved a groundbreaking milestone by directly observing gravitational waves from the merger of two black holes[10]. However, extracting meaningful signals from the data collected by detectors, often buried in noise [11], requires precise theoretical predictions. This necessity has sparked renewed interest in calculating classical gravitational quantities relevant to gravitational wave research. As the next generation of detectors emerges, the demand for increasingly accurate predictions continues to grow.

It is natural to rely on general relativity for such calculations. Since our focus is on a classical result—given that general relativity is devoid of any quantum aspects—the question arises: how does quantum field theory (QFT) come into play in this context? We will address this question shortly. However, let us first remind the reader that quantum field theory (QFT) offers a fine-grained description of nature, enabling us, in principle, to recover any classical result through an appropriate process of coarse graining.

In quantum mechanics, the reduced Planck's constant, $\hbar$, defines the boundary of the quantum domain. To transition from quantum to classical results, the appropriate coarse graining involves taking the limit $\hbar \to 0$, which we will refer to as

the classical limit. Consequently, there has been considerable effort in recent years to compute classical gravitational observables from certain effective field theories (EFTs) of gravity by applying this classical limit.

Given that the gravitational waves of interest are low-energy, with wavelengths significantly larger ($\sim 10^4 - 10^5$ m) than the celestial bodies emitting them ($\sim 10^0 - 10^1$ m), these bodies—such as black holes and neutron stars—can be effectively modeled as point particles. This allows us to employ an effective field theory (EFT) framework, describing point particles interacting through the exchange of gravitons. This description has significant advantages over the general relativity-based approaches. Since general relativity is highly non-linear, it is very hard to use general relativity directly to perform the precision calculation we need. On the other hand, such problems became computationally very demanding in numerical relativity. Using the QFT-based methods, it is possible to do such computations perturbatively in Newton's constant G. This thesis will introduce a new method for performing such calculations.

This thesis

This thesis explores three key aspects of zero-energy modes in quantum field theory. First, we examine a condensed matter system with boundaries, where zero-energy fermion edge states induce ferromagnetic behavior in the material. Next, we delve into a scalar field theory to shed light on certain formal properties of infrared divergences in quantum field theories. Finally, we demonstrate how the zero-energy limit can be utilized to derive a specific class of classical observables from quantum field theories. Through these discussions, we aim to highlight the intricate role of zero-energy modes in both the theoretical and observational domains of QFT. The organization of this thesis is as follows:

In Chapter 2, we first discuss how the low-energy behavior of graphene, a $2+1$ dimensional form of carbon, is governed by a relativistic theory of massless fermions. We then examine the boundary conditions for massless Dirac fermions in 2+1 dimensions, specifically in a half-plane geometry. Among the key results, we highlight a boundary condition that gives rise to fermion zero-mode edge states. These edge states reveal unique representations of scale, phase, and translation symmetries, and by imposing symmetry requirements, we demonstrate how they lead to edge ferromagnetism in the system. Furthermore, we show that these exotic ferromagnetic representations are indeed associated with the ground states of the system when perturbed by certain interaction Hamiltonians, including the non-relativistic Coulomb interaction.

In Chapter 3, after presenting some general discussions on the infrared divergence in QED and the quantum information content of zero-energy photons, we discuss similar issues in a scalar field theory. We present a field theory containing massless scalar particles with an infrared structure closely resembling those found in quantum electrodynamics and perturbative quantum gravity. However, unlike these

theories, it lacks gauge invariance, internal symmetries, or any apparent asymptotic symmetry. We explore the consequences of the reduced symmetry and find that while the absence of gauge symmetries does not affect the infrared structure, it does affect some aspects of the quantum information content.

In Chapter 4, we present a novel approach of calculating classical observables within the quantum field theory (QFT) framework by utilizing the zero-energy limit. This method is implemented through the "in-in" formalism, with all essential pre-requisites discussed in the chapter. Our results demonstrate the effectiveness of the zero-energy limit in deriving classical quantities from QFT. Furthermore, the findings enrich existing literature on using QFT techniques to compute classical quantities, addressing and resolving several technical challenges faced by current methodologies.

In chapter 5, we conclude our work and comment on some future directions.

Chapter 2

Massless Fermions on Semi-infinite Graphene

In this chapter, we will discuss a continuum field theory that describes the low-energy physics of a semi-infinite sheet of graphene. We will find that this system has many interesting properties in the presence of a boundary and allows for zero modes. In the next section, we discuss some prerequisites for understanding the main contents of this chapter. Throughout this chapter, we adopt the mostly positive signature for the flat spacetime metric.

2.1 Background materials

Graphene is a two-dimensional single-layer allotrope of carbon. It is famous for its unusual electronic properties [12]. In particular, graphene can be used as a condensed matter system to simulate various QFT phenomena like anomalies [13–15], as well as models of holographic systems [16]. In this section, we will discuss the emergence of relativistic symmetry in graphene [17].

Let us start with the lattice description of graphene. The bonds that connect the nearest C carbon atoms can be described by sp^2 hybridization. Since the number of valence electrons in carbon is four and three of them are connected to three other carbon atoms, there is one free electron per carbon atom living in the p_z orbital. As a result, graphene has a hexagonal honeycomb lattice as shown in the figure 2.1.

The hexagonal lattice is comprised of two inequivalent sublattices $\mathcal{A}$ and $\mathcal{B}$ depicted by the circle and square in 2.1. Let's take two basis vectors for the $\mathcal{A}$ sublattice:

$$\vec{a}_1 = \frac{1}{2}(3, \sqrt{3})a, \quad \vec{a}_2 = \frac{1}{2}(3, -\sqrt{3})a \tag{2.1}$$

where $a = 1.42\mathring{A}$ is the lattice constant. Each $\mathcal{A}$ atom is connected to three $\mathcal{B}$ atoms via vectors

$$\vec{\delta_1} = \frac{1}{2}(1, \sqrt{3})a, \;\; \vec{\delta}_2 = \frac{1}{2}(1, -\sqrt{3})a, \;\; \vec{\delta}_3 = (-1, 0)a. \tag{2.2}$$

The simplest Hamiltonian for the system is described by the tight-binding model. To the first approximation, we only consider each electron can only hop to the nearest neighbor atom. This gives the following Hamiltonian for graphene in terms of the

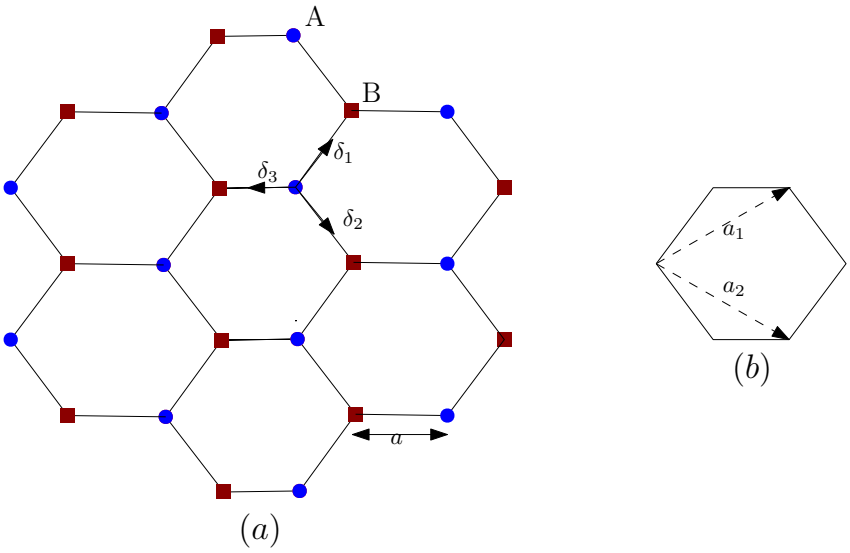


Figure 2.1: a) The hexagonal lattice structure of graphene is comprised of two inequivalent sublattice atoms, denoted here by A and B. b) A choice of basis vector for the A sublattice. Acting on by $m\vec{a}_1 + n\vec{a}_2$ with m, n integers, it is possible to move around between all A points.

fermionic annihilation and creation operators at sites $\mathcal{A}$ and $\mathcal{B}$ as

$$H = \sum_{i=1}^{3} \sum_{\mathbf{n},\sigma} \left[t a^{\dagger}_{\mathbf{n},\sigma} b_{\mathbf{n}+\delta_i,\sigma} + h.c. \right]; \tag{2.3}$$

where $|t| \approx 2.8eV$ is the hopping parameter, $\sigma = \{\uparrow, \downarrow\}$ stands for the spin, and $h.c$ denotes Hermitian conjugation. The $\mathbf{n}$ sum runs over the position of all $\mathcal{A}$ atoms. As an example, $a^{\dagger}_{(3a_1,2a_2)\uparrow}$ creates an electron with spin up at position $(3a_1, 2a_2)$.

To further diagonalize the Hamiltonian, next we write the operators $a_{n,\sigma}, b_{n,\sigma}$ in terms of Bloch waves by defining a column vector of Fourier transformation

$$\psi_{\mathbf{n},\sigma} \equiv \begin{bmatrix} a_{\mathbf{n},\sigma} \\ b_{\mathbf{n},\sigma} \end{bmatrix} = \frac{1}{\sqrt{N}} \sum_{j=1}^{3} \int_B d^2k \begin{bmatrix} \alpha_{\vec{k},\sigma} e^{i\vec{k}\cdot\vec{R}_{\mathbf{n}}} \\ \beta_{\vec{k},\sigma} e^{i\vec{k}\cdot(\vec{R}_{\mathbf{n}}+\delta_j)} \end{bmatrix}, \tag{2.4}$$

where $\vec{R}_{\mathbf{n}} = n_1\vec{a}_1 + n_2\vec{a}_2$ is the position of an arbitrary lattice site $\mathcal{A}$ spanned by the unit vectors, N is the number of unit cells and the $\vec{k}$ integrals are over the Brillouin zone. The Fourier coefficients satisfy

$$\{\alpha^{\dagger}_{\vec{k},\sigma}, \alpha_{\vec{k}',\sigma'}\} = \delta^2(\vec{k} - \vec{k}')\delta_{\sigma\sigma'} = \{\beta^{\dagger}_{\vec{k},\sigma}, \beta_{\vec{k}',\sigma'}\}$$

and all other anti-commutators are zero. The Hamiltonian takes the form

$$H = \sum_{\vec{k}} \begin{bmatrix} \alpha^{\dagger}_{\vec{k},\sigma} & \beta^{\dagger}_{\vec{k},\sigma} \end{bmatrix} \begin{bmatrix} 0 & \Delta(\vec{k}) \\ \Delta^*(\vec{k}) & 0 \end{bmatrix} \begin{bmatrix} \alpha_{\vec{k},\sigma} \\ \beta_{\vec{k},\sigma} \end{bmatrix}, \; \Delta(\vec{k}) = t \sum_{j=1}^{3} e^{i\vec{k}\cdot\vec{\delta}_j}. \tag{2.5}$$

The energy eigenvalues $E_{\vec{k}} = \pm|\Delta(\vec{k})|$ represent two energy bands respectively.

So far our analysis have been exact. Now we will concentrate on the low energy physics of the system. The zero of the energy can be found by solving the following equation

$$\begin{aligned} E_{\vec{k}} &= 0 \\ \text{or, } e^{i\vec{k}\cdot\vec{\delta}_1} + e^{i\vec{k}\cdot\vec{\delta}_2} + e^{i\vec{k}\cdot\vec{\delta}_3} &= 0. \end{aligned}$$

After some rearrangement we get,

$$1 + e^{-i\sqrt{3}k_y a} + e^{-i\frac{3}{2}k_x a - i\frac{\sqrt{3}}{2}k_y a} = 0\,. \tag{2.6}$$

The above equation has the following six solutions

$$(k_x, k_y) = \frac{2\pi}{3a}\left(\pm 1, \pm\frac{1}{\sqrt{3}}\right),\ \frac{2\pi}{3a}\left(0, \pm\frac{2}{\sqrt{3}}\right)\,. \tag{2.7}$$

These six points where the top and bottom bands touch each other i.e the energy is zero are called Dirac points. In fact, not all six points are independent. Only two of them are independent. Let's take them to be

$$(K) = \frac{2\pi}{3a}\left(1, \frac{1}{\sqrt{3}}\right), \quad (K') = \frac{2\pi}{3a}\left(1, -\frac{1}{\sqrt{3}}\right). \tag{2.8}$$

Indeed, if we define two reciprocal lattice vectors[1]

$$\vec{b}_1 = \frac{2\pi}{3a}\left(1, \sqrt{3}\right), \quad \vec{b}_2 = \frac{2\pi}{3a}\left(1, -\sqrt{3}\right), \tag{2.9}$$

then translating the independent points (K) and (K') by these vectors takes us to the remaining Dirac points.

Since we are interested in the low energy excitation of the system, near the (K) and (K') points, we have in the linear order in $\vec{k}$:

$$\begin{aligned} \Delta^{(K)}(K + \vec{k}) &= -ie^{\frac{2\pi i}{3}} v_F(k_x + ik_y) \\ \Delta^{(K')}(K' - \vec{k}) &= -ie^{\frac{2\pi i}{3}} v_F(k_x - ik_y), \end{aligned}$$

where $v_F = \frac{3at}{2}$. Ignoring the overall phase factor $-ie^{\frac{2\pi i}{3}}$, the Hamiltonian near these points takes the form

$$\begin{aligned} H(K + \vec{k}) &= v_F \begin{bmatrix} 0 & k_x - ik_y \\ k_x + ik_y & 0 \end{bmatrix} \\ H(K' + \vec{k}) &= v_F \begin{bmatrix} 0 & k_x + ik_y \\ k_x - ik_y & 0 \end{bmatrix}. \end{aligned} \tag{2.10}$$

[1]These are defined from the real space unit vectors $\vec{a}_i$ by $\vec{a}_i \cdot \vec{b}_i = 2\pi\delta_{ij}$

These two points (K) and (K') are called *valleys*. As it can be seen from the above expressions, the Hamiltonian is different in two valleys. However, from a practical purpose, it is very convenient to use same Hamiltonian but assign the valley index to the two-component spinor instead. The spinors in two valleys are defined as

$$\psi_\sigma^{(K)}(\vec{k}) \equiv \begin{bmatrix} \alpha_\sigma(K+\vec{k}) \\ \beta_\sigma(K+\vec{k}) \end{bmatrix}, \quad \psi_\sigma^{(K')}(\vec{k}) \equiv \begin{bmatrix} \beta_\sigma(K'+\vec{k}) \\ \alpha_\sigma(K'+\vec{k}) \end{bmatrix}. \tag{2.11}$$

Even though we are calling the above column vectors spinors, they have nothing to do with the real spin of the electrons. In fact, the components of the spinor label the wave functions at $\mathcal{A}$ and $\mathcal{B}$ valleys. The Hamiltonian describing the low-energy effective action of the graphene is then

$$H = v_F \sum_{A,\sigma} \int_B d^2k \; \psi_\sigma^{\dagger(A)}(\vec{k}) \begin{bmatrix} 0 & k_x - ik_y \\ k_x + ik_y & 0 \end{bmatrix} \psi_\sigma^{(A)}(\vec{k}), \; A \in \{K, K'\} \tag{2.12}$$

where we integrate over the modes that are in the first Brillouin zone. Taking the inverse Fourier transform we can write the above equation in position space as

$$H = v_F \sum_{A,\sigma} \int d^2x \; \psi_\sigma^{\dagger(A)}(x) \begin{bmatrix} 0 & \partial_x + i\partial_y \\ \partial_x - i\partial_y & 0 \end{bmatrix} \psi_\sigma^{(A)}(x). \tag{2.13}$$

Cutting the graphene

Now we would like to consider a semi-infinite graphene sheet. There are two principal ways of cutting the graphene. The resulting edge configurations are known as zigzag or armchair edges (figure 2.2) The boundary of the zigzag edge is composed of just one type of atom. Mathematically, that means one of the two components of the spinor is zero at the edge. The zigzag edge allows for zero-energy edge states called zero modes. Fermion zero modes give rise to a plethora of remarkable phenomena in condensed matter physics, such as fractional charge [14, 18], topological order [19], and various anomalies. In the context of graphene nano-ribbons, these edge-zero modes have been predicted to show magnetic order when the electron-electron interaction is taken into account [20, 21].

From lattice to continuum

Before discussing the main materials of this chapter, let us briefly discuss what to expect when going from a lattice model to its low-energy field theory description, as we have done above for graphene. To begin with, the symmetry group associated with low-energy EFT is almost always larger than that of lattice theory. In the above example, the lattice model has discreet translational symmetry, but continuum theory has continuous translational as well as spatial rotational symmetry. In fact, the free theory (2.13) has Lorentz as well as conformal symmetry. The latter

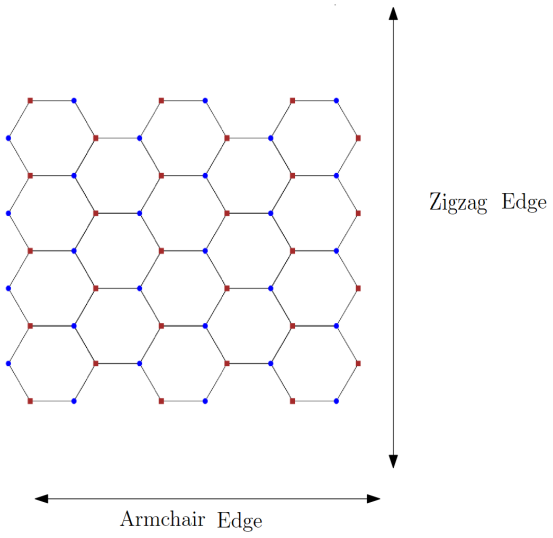


Figure 2.2: Two inequivalent ways of cutting the graphene sheet

is a symmetry group that includes the Lorentz group appended by scale invariance, and another special symmetry transformation called *special conformal transformation.* All these symmetries that are present in the low energy description but not in the lattice model are examples of *emergent* symmetries.

There are intriguing peculiarities that arise when transitioning from a lattice framework to a continuum QFT description. One prominent example is the fracton (see [22, 23] and references therein), where the lattice theories are well-defined, yet the continuum QFT exhibits unusual properties not typically encountered in conventional continuum QFTs. Another interesting question to consider is whether a symmetry observed in a well-defined continuum QFT is also present in the corresponding lattice model. If so, how is this symmetry realized in the lattice system? Recent explorations of these questions have led to very unexpected and non-trivial results [24, 25].

In this chapter, we will see that our low-energy QFT that describes graphene on the semi-infinite plane also exhibits some unique and interesting properties. For example, we will find a situation where we have a theory that breaks Lorentz and conformal symmetry but retains scale invariance. We will find another instance, where the continuum theory predicts two ferromagnetic ground states but the corresponding lattice theory only has one ferromagnetic ground state. Our system can be regarded as yet another example where the transition from lattice to continuum presents some surprises and opens up avenue for further research.

2.2 Introduction

One of the fascinating properties of the Dirac theory of fermions quantized in $2+1$ spacetime dimensions is that the mass term in the Dirac equation can, and in some cases must, break some of the discrete spacetime symmetries. For example, for a single two-component Dirac spinor, the Dirac equation with a mass term, $\left[\not{\partial} + m\right] \psi(x) = 0$, while having the full Poincare symmetry expected of a relativistic wave equation, necessarily violates both spacetime parity (P) and time reversal (T) invariance, although it does have good charge conjugation (C) symmetry, PT and CPT.

Of course the P and T symmetries can be restored by doubling the fermions so that the two species of fermions transform into each other under P or T, but then the signs of the mass term for each of the two species must be different and the mass terms break a global $U(2)$ symmetry (or $O(2)$ symmetry in the case of Majorana fermions) which would rotate the two species into each other. One has the statement that, even for the case of doubled fermions, they can have a mass gap, but that gap must necessarily violate either P and T or an internal $U(2)$ (or $O(2)$) symmetry as $U(2) \to U(1) \times U(1)$ or $O(2) \to Z_2 \times Z_2$[2].

In this chapter we wish to point out another fact, that even in the absence of mass, a single species of Dirac fermion defined on a 2+1-dimensional spacetime with a 1+1-dimensional boundary also violates the discrete spacetime symmetries P and T through its boundary conditions. Like a mass term, it preserves PT. In addition, like for a massive fermion, P and T can be restored by doubling the fermions, but then the boundary condition will violate a $U(2)$ symmetry. Note that we do not include the Majorana option here, since, generically, there isn't one. Unlike a mass term, boundary conditions for a single massless fermion on a half-space generally violate C. Here, and in the rest of this chapter, we will always take the boundary to be as symmetric as possible, specifically, if $x^\mu = (x^0, x^1, x^2)$, are the Cartesian coordinates of open, infinite Minkowski space, we will consider the half-space $x^1 > 0$.

We will show that there is only one essentially unique special case of a boundary condition where a single species of fermion can have C symmetry. That theory, even though it violates P and T, has C and CPT invariance and it also preserves the Poincare and conformal symmetry of the boundary. It has the only boundary condition where a single species fermion can be a boundary conformal field theory. In all other cases, a single fermion species violates C and CPT and, coincidentally, it also violates the Lorentz and conformal symmetries of the boundary.

We will also find that when the fermions are doubled, this special, highly symmetric theory has C, P and T symmetries as well as the full conformal symmetry of the boundary and, amongst all of the possibilities we consider, it is still the unique one which is a boundary conformal field theory. It is the theory described in equations (2.44)-(2.49) of section 2.4 below. We will refer to it as the BCFT. For

[2]We denote the symmetry breaking by the arrow

doubled fermions, we shall call the two species "the valleys", after the terminology used in graphene. This boundary conformal field theory is the low energy theory that models clean graphene near charge neutrality and with an armchair boundary [26, 27].

There is, for doubled fermions, a second boundary condition that preserves C as well as P and T, distinct from the boundary conformal field theory described above. The difference is that its C transformation, like P and T, interchanges the two valleys, whereas in the BCFT it does not. Having C, P and T separately, it has CPT symmetry but it is not Lorentz or conformal invariant. The Lorentz and conformal symmetries are broken by the boundary condition and they cannot be restored by doubling. This theory still has the residual translation symmetries and the scale symmetry of the boundary and it is an example of a scale invariant but not conformally invariant field theory. It is the one described in equations (2.50)-(2.55) of section 2.4 and we shall refer to it as the "zig-zag theory" as it describes the low energy limit of electrons near the charge neutral point in clean graphene with a zig-zag edge. The interesting feature of the zig-zag theory is that it has the unique boundary condition where the Dirac Hamiltonian can have static fermion zero mode edge states. The existence of bound states in a scale invariant theory where there are no dimensionful parameters is already unusual. In addition to that, fermion zero modes have long been associated with exotic realizations of symmetries [14, 18] and this zig-zag theory is no exception. In this case it is the representation of the scaling symmetry which we shall find interesting.

We will show that the only populations of the zero mode edge states with fermions which can be scale invariant are those where complete bands of the states are either completely filled or completely empty. This fact, together with a requirement of $U(1)$ (electric) charge neutrality will lead us to the valley ferromagnetic states. Of course such states are only one or two of the enormous number of candidates amongst the degenerate ground states in a system with an infinite number of fermion zero modes and, in the free fermion theory, there is little besides their special symmetry property that distinguishes them. To improve on this situation, we generalize an old argument, related to the Hund rules of atomic physics, and previously applied to the quantum Hall ferromagnetism that is seen in graphene [28] to prove that a perturbation of the free fermion theory by an instantaneous Coulomb interaction actually favours these ferromagnetic states – in fact they are the exact ground state if the Coulomb interaction is weak enough. This establishes their importance and therefore the importance of scale invariant states.

Practically any discussion of massless relativistic fermions in 2+1-dimensions should have applications to Dirac materials [29], examples of which are d-wave superconductors, graphene, and topological insulators. Graphene has long been a venue where field theoretical ideas about 2+1 dimensional Dirac fermions could find application [13] and boundary conditions for the Dirac field corresponding to various boundaries have been found [30, 31] and are compatible with what we shall

find in this chapter. Moreover, the occurrence of edge states on zig-zag edges of graphene [20, 32] was discussed years before graphene was readily available in the laboratory and their fascinating properties have inspired literally thousands of research papers. The edge states themselves are readily detected by experimental techniques such as scanning tunnelling microscopy[33, 34], high resolution transmission electron microscopy [35] and atom-by-atom spectroscopy [36]. Most theoretical modelling using a variety of techniques, mean field theory[20, 37, 38], density functional theory[39, 40], numerics [41–43] and perturbation theory [21, 44], indicate that the edge states favour taking up a ferromagnetic configuration, although this has yet to be seen reliably in experiments. Such an edge ferromagnetic state is expected to have important technological applications in spintronics [39, 40]. The theoretical computations normally use the lattice tight binding model of graphene and various types of interactions and mean field theory arguments. Our results are confined to the scale invariant continuum theory that would describe the electrons in a low energy, large wavelength limit of clean graphene near charge neutrality. In that context, our proof that the scale invariant valley or spin ferromagnetic states are actually the true ground states of such a system with a weak Coulomb interaction added is in line with other results in this vein. Affleck [21, 44] et.al. give a strong argument that the ground state of charge neutral graphene with a weak Hubbard interaction is edge ferromagnetic and they also argue, with somewhat less rigour that the same should apply with the Coulomb interaction. Our work uses a similar technique, is in agreement with them and, in addition, it solves the issue of edge-bulk hybridization which they left as an open problem. It also treats the Coulomb interaction directly and it is easily generalized to other interactions. Finally, we emphasize the tantalizing possibility that this edge state ferromagnetism could be confirmed by experiment in the near future.

One interesting property of the valley ferromagnetic states of the system with two species (two valleys and one spin polarization) that we find, and which we would like to understand better, is an induced momentum density for momentum parallel to the boundary (see equation (2.68),

$$\left\langle \mathbf{T}^{02}(x) \right\rangle = \pm \frac{1}{4\pi} \frac{1}{x_1^3}$$

where $\mathbf{T}^{\mu\nu}(x)$ is the stress tensor and the sign on the right-hand-side depends on the polarization of the valley ferromagnet. This equation is relevant to spin polarized graphene. With two degenerate spin states, it can easily cancel between two spin degrees of freedom. It can also easily be shown that this apparent buildup of momentum density is not due to charge transport as the current vanishes,

$$\left\langle J^2(x) \right\rangle = 0$$

The remainder of this chapter is organized as follows. In section 2.3, we will discuss the possible boundary conditions for the Dirac field. This includes a discussion of both the discrete and continuous spacetime symmetries. It is where we show

that there is essentially one unique Lorentz and conformal invariant possibility. In section 2.4, we look at the possibility of restoring the discrete spacetime symmetries by doubling the number of fermions. We find and then we concentrate on two interesting cases, which we will call the BCFT and the zigzag theories. In section 2.5, we explore the zigzag theory in more detail by solving the Dirac equation explicitly and then discussing the possible ground states of the quantum field theory. It is there that we argue that scale invariance is closely tied to valley ferromagnetism. In section 2.6, we prove that the ferromagnetic states are the ground state of the zigzag theory deformed by a weak nonrelativistic Coulomb interaction. In section 2.7, we offer conclusions.

2.3 Single species of massless Dirac Fermion

Consider a single Dirac Fermion which obeys the massless Dirac equation

$$i\gamma^\mu \partial_\mu \psi(x) = 0 \tag{2.14}$$

with the Dirac field obeying equal time anti-commutation relations

$$\left\{\psi(x), \psi^\dagger(x')\right\} \delta(x^0 - x'^0) = \delta^2(x - x') \tag{2.15}$$

and defined on the half-space (x_0, x_1, x_2) with $x_1 > 0$. The spinor field must also obey a boundary condition which we will assume is linear and we will take as

$$\lim_{x^1 \to 0} (1 - \Pi)\, \psi(x) = 0 \tag{2.16}$$

where Π is a 2×2 matrix. In the following we will discuss the possible forms of Π. Note that the anti-commutator (2.15) is valid only in the region where $x^1, x'^1 > 0$.

We will consider linear boundary conditions where the Dirac Hamiltonian

$$H = i\gamma^0 \vec{\gamma} \cdot \vec{\nabla} \tag{2.17}$$

defined on the space with a boundary is a self-adjoint differential operator. A necessary condition is that the normal component of the fermion particle flux vanishes there, that is,

$$-\psi^\dagger(x)\gamma^0\gamma^1\psi(x)\Big|_{x_1=0} = 0 \tag{2.18}$$

This requirement is satisfied by a two-parameter family of local, linear boundary conditions of the sort in equation (2.16) where

$$\Pi = i\cosh\chi\cos\theta\gamma_0 + \cosh\chi\sin\theta\gamma_1 + i\sinh\chi\gamma^2 \tag{2.19}$$

Here the parameter χ can be any real number and $\theta \in [-\pi, \pi]$. A constraint on Π that is solved by the ansatz (2.19) is simply that $(1 - \Pi)$ must be proportional to a

projection operator with rank one, so that, at the boundary, it constrains one of the two degrees of freedom in the Dirac spinor $\psi(x)$. In the next section, we shall see that the boundary condition (2.16) with Π given in (2.19) generally violates both discrete and continuous spacetime symmetries. Then we will focus on a few special cases which can be more symmetric.

2.3.1 Discrete spacetime symmetries

To examine the discrete spacetime symmetries of the Dirac equation, it is useful to observe that the Dirac matrices are 2×2 matrices which have a Majorana representation in which they are real,

$$\gamma_0 = -i\sigma_2 \ , \quad \gamma_1 = \sigma_1 \ , \quad \gamma_2 = \sigma_3 \quad , \quad \gamma_\mu = \gamma_\mu^* \tag{2.20}$$

where

$$\sigma_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \ \sigma_2 = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \ \sigma_3 = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \tag{2.21}$$

are the Pauli matrices. These matrices have the algebraic property

$$\gamma^\mu \gamma^\nu = \epsilon^{\mu\nu\lambda}\gamma_\lambda + \eta^{\mu\nu} \tag{2.22}$$

where $\epsilon^{\mu\nu\lambda}$ is the totally anti-symmetric tensor with $\epsilon^{012} = 1$ and $\eta^{\mu\nu} = \mathrm{diag}(-1, 1, 1)$ is the metric of Minkowski spacetime.

When the Dirac matrices are in the Majorana representation (2.20) the discrete spacetime symmetries, C, P and T of the Dirac equation are given by the transformations

$$C : \psi(x) = \psi^*(x) \qquad \text{when } \gamma_\mu = \gamma_\mu^* \tag{2.23}$$

$$P : \psi(x) = i\gamma^2\psi(x') \qquad x' = (x_0, -x_1, x_2) \tag{2.24}$$

$$T : \psi(x) = \mathcal{T} i\gamma^0\psi(\tilde{x}) \qquad \tilde{x} = (-x_0, x_1, x_2) \tag{2.25}$$

In equation (2.25), $\mathcal{T}$ is the complex conjugation operator. Here we have adjusted the phases of the P and T transformation so that, when acting on fermions, $P^2 = -1$, $T^2 = -1$ and $(CPT)^2 = -1$, although this is not important for the current discussion.

There is no choice of the parameters χ and θ for which the boundary condition (which we recopy here for convenience)

$$\lim_{x^1 \to 0} (1 - \Pi)\, \psi(x) = 0 \tag{2.26}$$

$$\Pi = i \cosh\chi \cos\theta \gamma_0 + \cosh\chi \sin\theta \gamma_1 + i \sinh\chi \gamma^2 \tag{2.27}$$

is compatible with either the P or the T transformation. All possible choices of the linear boundary condition break both P symmetry and T symmetry. However, all of the possible choices preserve the combined PT symmetry.

Since PT is a good symmetry, we could ask about CPT which would require invariance under C. The C transformation is simple in the Majorana representation where the Dirac operator $\not{\partial}$ is real. Invariance under C further requires that the boundary condition is real. Of the family of conditions in equations (2.26) and (2.27), there are only two essentially equivalent boundary conditions that preserve C and therefore CPT. They are gotten from the choice of parameters $\chi = 0$ and $\theta = \pm\frac{\pi}{2}$ in Π above so that

$$\Pi = \pm\gamma_1 \quad , \quad \lim_{x^1 \to 0} \left(1 \mp \gamma^1\right) \psi(x) = 0 \tag{2.28}$$

Coincidentally, we will see shortly that these are also the unique choice of boundary condition that preserves the Lorentz and conformal symmetry of the boundary. It is interesting to find such a strong correlation between CPT and Lorentz invariance in the context of a boundary field theory. The Dirac theory on a half-space with one of the boundary conditions in equation (2.28) is the starting point for construction of a boundary conformal field theory with fermion content.

Another special boundary condition that will be of interest to us is one which allows edge states. These are zero energy eigenstates of the Dirac Hamiltonian (2.17) which are localized at the boundary. In the family of boundary conditions in equations (2.26) and (2.27) there is an essentially unique choice for which this can occur, the choices $\chi = 0$ and $\theta = \pm\pi$,

$$\Pi = \pm i\gamma^0 \quad , \quad \lim_{x^1 \to 0} \left(1 \pm i\gamma^0\right) \psi(x) = 0 \tag{2.29}$$

As we have already discussed, this boundary condition violates all of the C, P, T and CPT symmetries and we will see shortly that it also violates the Lorentz and conformal symmetry of the boundary. On the other hand, the edge states have interesting implications for the physical properties of such a system. A discussion of them and their properties will occupy most of the bulk of this chapter.

2.3.2 Continuous Spacetime symmetries

A continuous space-time symmetry is symmetry under an infinitesimal coordinate transformation $x^\mu \to x^\mu + f^\mu(x)$ where the spinor field transforms as [45]

$$\delta_f \psi(x) = \left[f^\mu(x)\partial_\mu + \frac{1}{4}\epsilon^{\mu\nu\lambda}\partial_\mu f_\nu(x)\gamma_\lambda + \frac{1}{3}\partial_\mu f^\mu(x) \right] \psi(x) \tag{2.30}$$

On infinite $2+1$-dimensional spacetime – without the boundary – the above transformation is a symmetry of the Dirac equation (2.14) and the symplectic structure when $f^\mu(x)$ is a conformal Killing vector of three-dimensional Minkowski space. A conformal Killing vector obeys the conformal Killing equation

$$\partial_\mu f_\nu(x) + \partial_\nu f_\mu(x) - \frac{2}{3}\eta_{\mu\nu}\partial_\lambda f^\lambda(x) = 0 \tag{2.31}$$

The ten solutions of this equation are

$$f^\lambda(x) = \left\{ a^\lambda,\ \omega^\lambda{}_\rho x^\rho,\ cx^\lambda,\ x^\lambda b\cdot x - b^\lambda x^2/2 \right\} \tag{2.32}$$

corresponding to translations $\delta x^\lambda = a^\lambda$, Lorentz transformations $\delta x^\lambda = \omega^\lambda{}_\rho x^\rho$ where $\omega_{\mu\nu} = -\omega_{\nu\mu}$, dilatations $\delta x^\lambda = cx^\lambda$ and conformal transformations $\delta x^\lambda = x^\lambda b_\nu x^\nu - b^\lambda x_\nu x^\nu/2$ with parameters b^λ, respectively.

When the boundary is present, these symmetries are reduced to those which preserve the geometry of the boundary and which also preserve the boundary condition. The symmetries which preserve the geometry of the boundary are just the symmetry transformations of the boundary itself which in our case is the two-dimensional Minkowski spacetime with coordinates $x^\mu = (x^0, 0, x^2)$. The remaining symmetries are then time translations and space translations along the boundary, $a^\mu = (a^0, 0, a^2)$, the Lorentz transformation along the boundary, with the only nonzero components of $\omega^\lambda{}_\rho$ being $\omega_{20} = -\omega_{02}$, the dilatation symmetry with parameter c and $\delta x^\mu = c(x^0, x^1, x^2)$ and the conformal transformations with $b^\mu = (b^0, 0, b^2)$. These transformations are generated by Noether currents related to the stress tensor and Noether charges which generate the transformations, respectively

$$-i\left[P_\mu,\ \psi(x)\right] = \partial_\mu\psi(x)\ ,\quad \mu = 0, 2 \tag{2.33}$$

$$-i\left[M,\ \psi(x)\right] = \left(x_2\partial_0 - x_0\partial_2 + \frac{1}{2}\gamma^1\right)\psi(x) \tag{2.34}$$

$$-i\left[\Delta,\ \psi(x)\right] = (x^\mu\partial_\mu + 1)\ \psi(x) \tag{2.35}$$

$$-i\left[K_\lambda,\ \psi(x)\right] = \left\{ \left(x_\lambda x^\mu - \frac{x^2}{2}\delta^\mu_\lambda\right)\partial_\mu + \frac{1}{2}\epsilon_{\lambda\rho\sigma}x^\rho\gamma^\sigma + x_\lambda \right\}\psi(x),\ \lambda = 0, 2 \tag{2.36}$$

Because of the presence of γ^1 in the transformation, for generic values of χ and θ, the boundary conditions in equations (2.26) and (2.27) are not invariant under the Lorentz transformation generated by M in equation (2.34), nor are they invariant under the conformal transformations generated by K_0 or K_2 in equation (2.36). The transformation from the spin operators in both of those cases would not be compatible with the boundary conditions. The generic boundary condition therefore reduces the symmetry of the theory from the conformal symmetry of the boundary to a subset consisting of time translations, spatial translations along the boundary and dilatations. These symmetries are governed by the sub-algebra.generated by P_0, P_2 and Δ of the full conformal algebra.

There is one special case of the boundary conditions which is compatible with the spin part of the transformation in M, K^0 and K^2. It is the case where $\chi = 0$, $\theta = \pm\frac{\pi}{2}$ and $\Pi = \pm\gamma^1$, where the boundary condition preserves the Lorentz and conformal invariance. As we have discussed earlier in this section, this is also the unique special case where the boundary condition preserves C and CPT symmetries. The single relativistic fermion theory in that special case is a boundary conformal

field theory and as such it is interesting in its own right. It would be very interesting to study it further using conformal techniques. We do not have time or space to do it justice here and we put it off for future work. Instead we will concentrate on another special case.

2.4 Doubled fermions

In the previous section, we learned that P and T can never be good symmetries of a single massless Dirac theory for any acceptable linear boundary condition. In this section, we will double the degrees of freedom so that there are two species of fermions, $\psi^{(1)}(x)$ and $\psi^{(2)}(x)$ satisfying identical massless Dirac equations

$$i\gamma^\mu\partial_\mu\psi^{(1)}(x) = 0 \quad , \quad i\gamma^\mu\partial_\mu\psi^{(2)}(x) = 0 \tag{2.37}$$

and, as is common in the literature on Dirac materials like graphene, we will call the two species the two "valleys".[3] Now, we can do one of two things. We could demand that each valley has the same boundary condition. Then we would simply have two identical copies of the P and T violating fermions that we studied in the previous section. The system would have a "valley symmetry" under the transformation $\psi^{(A)}(x) \to U^A_{\ B}\psi^{(B)}(x)$ where $U^A_{\ B}$ is a 2×2 unitary matrix.

Alternatively, we could allow the valleys to transform into each other under P and T. Cursory examination of the possibilities shows us that the following system is P and T symmetric,

$$i\gamma^\mu\partial_\mu\psi^{(1)}(x) = 0 \quad , \quad i\gamma^\mu\partial_\mu\psi^{(2)}(x) = 0 \tag{2.38}$$

$$\left\{\psi^{(A)}(x), \psi^{(B)\dagger}(x')\right\}\delta(x^0 - x'^0) = \delta^{AB}\delta^2(x - x') \tag{2.39}$$

$$\psi^{(1)}(x)_{x^1\to 0} = (i\cosh\chi\cos\theta\gamma_0 + \cosh\chi\sin\theta\gamma_1 + i\sinh\chi\gamma_2)\psi^{(1)}(x)_{x^1\to 0} \tag{2.40}$$

$$\psi^{(2)}(x)_{x^1\to 0} = (-i\cosh\chi\cos\theta\gamma_0 - \cosh\chi\sin\theta\gamma_1 + i\sinh\chi\gamma_2)\psi^{(2)}(x)_{x^1\to 0} \tag{2.41}$$

where T and P are

$$T : \begin{cases}\psi^{(1)}(x) & \to & \mathcal{T}i\gamma^0\psi^{(2)}(\tilde{x}) \\ \psi^{(2)}(x) & \to & \mathcal{T}i\gamma^0\psi^{(1)}(\tilde{x})\end{cases} \tag{2.42}$$

$$P : \begin{cases}\psi^{(1)}(x) & \to & i\gamma^2\psi^{(2)}(x') \\ \psi^{(2)}(x) & \to & i\gamma^2\psi^{(1)}(x')\end{cases} \tag{2.43}$$

[3] We caution the reader that the valleys that we refer to can differ from the specific species that are commonly referred to as valleys in graphene which are the individual sets of electronic excitations near the two K-points. In our continuum approach we do not have information about the identity of the K points. Instead we will chose a basis for our two species of fermions which diagonalize the boundary conditions in equation (2.40) and (2.41). A more general choice would have a Hermitian matrix with eigenvalues ± 1 acting on the valley indices on the right-hand-side of those equations. Our "valley" fermions are generally linear combinations of the usual graphene valley electrons for the (BCFT) armchair edge, but they coincide with the usual definition for the zig-zag edge.

for any value of the parameters χ and θ. We thus have a large family of boundary conditions which are compatible with P and T symmetry. We note that, on the other hand, for any values of χ and θ, the boundary condition breaks the valley symmetry – reducing $U(2)$ to $U(1) \times U(1)$.

Now that, in equations (2.38)-(2.43), we have restored the P and T symmetries, we can again ask whether any of the family of boundary conditions there is also consistent with imposing a C symmetry. For this, there are only two possibilities. In the first possibility, we assume that C does not interchange the valleys, that is $(1) \to (1)$ and $(2) \to (2)$. Then it can be so only if $\chi = 0$ and $\theta = \pm\pi/2$ and (2.38)-(2.43) become

$$i\gamma^\mu \partial_\mu \psi^{(1)}(x) = 0 \quad , \quad i\gamma^\mu \partial_\mu \psi^{(2)}(x) = 0 \tag{2.44}$$

$$\left\{ \psi^{(A)}(x), \psi^{(B)\dagger}(x') \right\} \delta(x^0 - x'^0) = \delta^{AB} \delta^2(x - x') \tag{2.45}$$

$$[1 - \gamma_1]\, \psi^{(1)}(x)_{x^1 \to 0} = 0 \ , \ [1 + \gamma_1]\, \psi^{(2)}(x)_{x^1 \to 0} = 0 \tag{2.46}$$

where

$$T : \begin{cases} \psi^{(1)}(x) \to \mathcal{T} i\gamma^0 \psi^{(2)}(\tilde{x}) \\ \psi^{(2)}(x) \ \to \ \mathcal{T} i\gamma^0 \psi^{(1)}(\tilde{x}) \end{cases} \tag{2.47}$$

$$P : \begin{cases} \psi^{(1)}(x) \to i\gamma^2 \psi^{(2)}(x') \\ \psi^{(2)}(x) \to i\gamma^2 \psi^{(1)}(x') \end{cases} \tag{2.48}$$

$$C : \begin{cases} \psi^{(1)}(x) \to \psi^{(1)*}(x) \\ \psi^{(2)}(x) \to \psi^{(2)^*}(x) \end{cases} \tag{2.49}$$

The above equations (2.44)-(2.49) now describe a C, P and T invariant boundary conformal field theory which we refer to as BCFT. This theory describes the very low energy limit of clean, charge neutral (spin polarized) graphene with an armchair edge.

The other possibility for obtaining a C invariant theory is to let the C transformation interchange the valleys $(1) \leftrightarrow (2)$. There is also a unique choice in equations (2.38)-(2.43), which obtains C invariance with this interchange, where the parameters are $\chi = 0$ and $\theta = \pm\pi$, and (2.38)-(2.43) become

$$i\gamma^\mu \partial_\mu \psi^{(1)}(x) = 0 \quad , \quad i\gamma^\mu \partial_\mu \psi^{(2)}(x) = 0 \tag{2.50}$$

$$\left\{ \psi^{(A)}(x), \psi^{(B)\dagger}(x') \right\} \delta(x^0 - x'^0) = \delta^{AB} \delta^2(x - x') \tag{2.51}$$

$$[1 + i\gamma_0]\, \psi^{(1)}(x)_{x^1 \to 0} = 0 \ , \ [1 - i\gamma_0]\, \psi^{(2)}(x)_{x^1 \to 0} = 0 \tag{2.52}$$

where

$$T : \begin{cases} \psi^{(1)}(x) \to \mathcal{T} i\gamma^0 \psi^{(2)}(\tilde{x}) \\ \psi^{(2)}(x) \to \mathcal{T} i\gamma^0 \psi^{(1)}(\tilde{x}) \end{cases} \tag{2.53}$$

$$P : \begin{cases} \psi^{(1)}(x) \to i\gamma^2 \psi^{(2)}(x') \\ \psi^{(2)}(x) \to i\gamma^2 \psi^{(1)}(x') \end{cases} \tag{2.54}$$

$$C : \begin{cases} \psi^{(1)}(x) \to \psi^{(2)*}(x) \\ \psi^{(2)}(x) \to \psi^{(1)^*}(x) \end{cases} \tag{2.55}$$

This theory is C, P and T invariant. We note and emphasize here that the doubling of the fermion degrees of freedom cannot restore a continuous space-time symmetry that is broken by boundary conditions. The system described by equations (2.50)-(2.55) boundary conditions still violate the Lorentz and conformal symmetries of the boundary so that its only continuous spacetime symmetries are under time translation, space translation parallel to the boundary and scale transformations. It is still, on the other hand, a doubled version of the unique boundary condition that allows the Dirac equation to have zero energy edge state solutions. A discussion of those solutions and their implications for the realization of the remaining symmetries will occupy the rest of this chapter. This theory describes the very low energy limit of clean, charge neutral (spin polarized) graphene with an zig-zag edge, as well as a host of other possible edges that are related to the zig-zag [30, 31].

2.5 Edge states

Let us consider the zig-zag system described by the equations and boundary conditions in equations (2.50)-(2.55). If we look for solutions of the Dirac equation of the form $\psi^{(A)}(x) = \begin{bmatrix} u^{(A)}(x) \\ v^{(A)}(x) \end{bmatrix}$ the equation and boundary condition become

$$\begin{bmatrix} \partial_y & \partial_x - \partial_t \\ \partial_x + \partial_t & -\partial_y \end{bmatrix} \begin{bmatrix} u^{(A)}(t,x,y) \\ v^{(A)}(t,x,y) \end{bmatrix} = 0$$
$$u^{(1)}(t,0,y) = -iv^{(1)}(t,0,y) \ , \quad u^{(2)}(t,0,y) = iv^{(2)}(t,0,y) \tag{2.56}$$

Note that, in this Majorana representation of the Dirac matrices, the Dirac equation is real. However, the boundary conditions are not real, so the solutions which we shall find are complex.

The equation and boundary conditions in (2.56) have the explicit solutions (with

(x^0, x^1, x^2) labeled (t, x, y))

$$\psi^{(1)}_{k,\ell\pm}(x) = \frac{e^{-i\omega t+iky}}{\sqrt{8\pi\omega^2}} \begin{bmatrix} (ik+\ell+\omega)e^{i\ell x} + (-ik+\ell-\omega)e^{-i\ell x} \\ i(ik+\ell-\omega)e^{i\ell x} + i(-ik+\ell+\omega)e^{-i\ell x} \end{bmatrix} \tag{2.57}$$

$$\psi^{(2)}_{k,\ell\pm}(x) = \frac{e^{-i\omega t+iky}}{\sqrt{8\pi\omega^2}} \begin{bmatrix} (-ik+\ell+\omega)e^{i\ell x} + (ik+\ell-\omega)e^{-i\ell x} \\ i(ik-\ell+\omega)e^{i\ell x} + i(-ik-\ell-\omega)e^{-i\ell x} \end{bmatrix} \tag{2.58}$$

$$\ell \geq 0,\ -\infty < k < \infty,$$
$$\psi^{(A)}_{k,\ell+}(x) \text{ has } \omega = \sqrt{k^2+\ell^2};\ \psi^{(A)}_{k,\ell-}(x) \text{ has } \omega = -\sqrt{k^2+\ell^2} \tag{2.59}$$

$$\psi^{(1)}_{0,k}(x) = \sqrt{\frac{|k|}{2\pi}}\, e^{iky-kx} \begin{bmatrix} 1 \\ i \end{bmatrix} \quad ,\quad k > 0,\ \omega = 0 \tag{2.60}$$

$$\psi^{(2)}_{0,k}(x) = \sqrt{\frac{|k|}{2\pi}}\, e^{iky+kx} \begin{bmatrix} 1 \\ -i \end{bmatrix} \quad ,\quad k < 0,\ \omega = 0 \tag{2.61}$$

The above solutions are Dirac spinors $\psi^{(1)}_{k,\ell\pm}(x)$ and $\psi^{(2)}_{k,\ell\pm}(x)$ which propagate in the bulk of the half-space and we shall refer to them as "bulk states". Their subscripts $\pm$ have two choices of sign, $+$ corresponds to a positive energy solution and $-$ corresponds to a negative energy solution. Then, there are also the spinors $\psi^{(1)}_{0,k}(x)$ and $\psi^{(2)}_{0,k}(x)$ which are localized near the edge of the half-plane at $x^1 = 0$ and we will refer to them as "edge states". They have zero energy.

We have normalized the wave-functions and it is easy to check that they obey the appropriate completeness relation – which indicates that the entire set of bulk and edge states are required by the equal-time anti-commutation relation (2.15). The energy of each of the single particle bulk states is given by its value of ω and, for each valley (1) and (2), there is a band of positive energy bulk states and a band of negative energy bulk states. As well, the infinite number of edge states all of which have zero energy form what is sometimes called a "flat band".

To find these solutions we have used the fact that the system is translation invariant along the $t = x^0$- and $y = x^2$-directions, the latter being parallel to the boundary. The eigenvalues of the energy and momentum operators which generates translations in the t- and y-directions are the variables ω and k which appear in the solutions. It is interesting that the edge states in a given valley, besides having $\omega = 0$, all have momenta with the same sign, either $k > 0$ in valley (1) or $k < 0$ in valley (2).

We are interested in the states of this system near the Dirac vacuum where all of the negative energy states are occupied and all of the positive energy states are empty. This is clearly unambiguous for the bulk states where there is a clear distinction between positive and negative energy states. We begin construction of our candidate ground states for this system by assuming that all of the positive energy bulk states are empty and all of the negative energy bulk states are filled.

There is an equivalent "second quantized" version of this statement – that we choose the state that is the vacuum for both fermions and anti-fermions.

What is left to decide on is the distribution of fermions in the zero energy edge states. There is clearly an enormous vacuum degeneracy since any such distribution of fermions into the zero modes has the same energy as any other, no matter what the number of particles or with which values of momenta or which valleys the particles occupy. In the following, we will concentrate on finding special occupations of the states which have certain symmetry properties, an important one being scale invariance. Later on we will explain why we are particularly interested in scale invariant states.

Let us ask what are the possible scale invariant states of the many-particle system. A scale transformation, $x^\mu \to \Lambda x^\mu$ acts on the positive, negative and zero energy states as the transformation

$$\psi^{(A)}_{k,\ell}(\ x) \ \ \to \ \ \Lambda\psi^{(A)}_{k,\ell\pm}(\Lambda x) \ = \ \Lambda\psi^{(A)}_{\Lambda k,\Lambda\ell\pm}(x) \tag{2.62}$$

$$\psi^{(A)}_{0,k}(x) \ \ \to \ \ \Lambda\psi^{(A)}_{0,k}(\Lambda x) \ = \ \Lambda^{\frac{1}{2}}\psi^{(A)}_{0,\Lambda k}(x) \tag{2.63}$$

An orbit of the scaling transformations acting on a function is the set of all functions that are generated from it by scale transformation with $0 < \Lambda < \infty$. An orbit of a particular bulk state is the infinite set of bulk states with all magnitudes but the same signs of k and ω (remember that $\ell > 0$ in the solutions (2.57)-(2.59)),

$$\text{Orbit of scaling } = \{\Lambda\psi^{(A)}_{\Lambda k,\Lambda\ell\pm}(x), \forall\Lambda \in (0,\infty)\}$$

In each valley the set of all bulk states thus divides into four subsets according to the signs of k and ω, each subset being an orbit of scaling.

On the other hand, as can be seen by inspecting (2.60)-(2.61), the set of all of the edge states in a given valley comprise a single orbit of scaling,

$$\text{Orbit of scaling } = \{\Lambda^{\frac{1}{2}}\psi^{(A)}_{0,\Lambda k}(x), \forall\Lambda \in (0,\infty)\}$$

A scale invariant state of the quantum field theory, that is, of this many-fermion system should have every orbit of the scale transformations either completely filled with fermions or completely empty. For the bulk states in a given valley, we see that the negative energy states consist of two complete orbits, one with $k > 0$ and one with $k < 0$, as do the positive energy states. Thus, completely filling the negative energy bulk states and leaving the positive energy bulk states completely empty is a scale invariant filling of the bulk states.

For occupations of the edge states, on the other hand, the only possibilities are to either completely fill or leave completely empty the zero mode states in a given valley. This leaves us with four possible scale invariant ground states. In all four cases the positive energy bulk states are completely empty and negative energy bulk states are completely filled. Then, either the edge states have both

valleys completely filled, or both valley's completely empty, or the two possibilities where one valley's edge states are completely filled and the other valley's edge states are completely empty. Scale symmetry alone does not tell us which of these four possibilities is the correct one. For that, we need more information.

The Dirac equation and boundary conditions in equation (2.50) has two $U(1)$ symmetries, and two conserved currents, $\bar{\psi}^{(1)}(x)\gamma^\mu\psi^{(1)}(x)$ and $\bar{\psi}^{(2)}(x)\gamma^\mu\psi^{(2)}(x)$, one for each valley. We will call the sum of the Noether charges corresponding to those two symmetries

$$Q = \int d^2x \left\{ \frac{1}{2}\left[\psi^{(1)\dagger}(x), \psi^{(1)}(x)\right] \; + \; \frac{1}{2}\left[\psi^{(2)\dagger}(x), \psi^{(2)}(x)\right] \right\} \tag{2.64}$$

the "electric charge". In this vein, we will also assume that there is a background charge density which is $-1/2$ unit per degree of freedom, so that many fermion states with exactly half of the energy levels occupied are charge neutral. The Dirac commutator prescription for ordering the operators in Q in equation (2.64) is one way to take this neutralizing background charge density into account. As defined in (2.64) it is easy to confirm that $Q \to -Q$ under C and that it is invariant under P and T.

Then, given the symmetry of the spectrum between positive and negative energy bulk states we would conclude that, to achieve overall charge neutrality of the many-fermion system when the bulk negative energy states are filled and positive energy states empty, we would want to fill precisely half of the edge states. As we have argued above, any partial filling of the edge states in a given valley will violate scale invariance. Thus, if we require scale invariance and charge neutrality, there are two candidate ground states, where the edge states of one valley are completely filled and the edge states of the other valley are completely empty. These are "valley ferromagnetic" states. It is interesting that we are led to this valley ferromagnetism by symmetry considerations alone, even in the absence of interactions.

There is a third property of many particle states, which might be bothering the reader at this point, and is worthy of comment. Each single particle state with wavefunction $\psi^{(A)}_{k,\ell\pm}(x)$ or $\psi^{(A)}_{0k}(x)$ carries a linear momentum k along the direction parallel to the boundary. Moreover, as we have already observed, all of the zero mode states in a given valley have the same sign of k. The contribution of a given occupied single particle state to the total momentum is $k\psi^{(A)\dagger}_{k,\ell\pm}(x)\psi^{(A)}_{k,\ell\pm}(x)$ or $k\psi^{(A)\dagger}_{0k}(x)\psi^{(A)}_{0k}(x)$. If we add up the contribution of the occupied bulk states in a single valley it is

$$\int_{-\infty}^{\infty} dk \int_0^\infty d\ell \; k \; \psi^{(A)\dagger}_{k,\ell-}(x)\psi^{(A)}_{k,\ell-}(x) = \lim_{x'\to x} \frac{1}{i}\frac{d}{dx_2} \int_{-\infty}^{\infty} dk \int_0^\infty d\ell \psi^{(A)\dagger}_{k,\ell-}(x')\psi^{(A)}_{k,\ell-}(x)\,. \tag{2.65}$$

We can use a particle-hole-symmetry identity, which we find in the next section in

equation (2.76), to re-write the right-hand-side of the above equations as

$$\lim_{x'\to x}\frac{1}{i}\frac{d}{dx_2}\int_{-\infty}^{\infty}dk\int_0^{\infty}d\ell\;\frac{1}{2}\left\{\psi^{(A)\dagger}_{k,\ell-}(x')\psi^{(A)}_{k,\ell-}(x)+\psi^{(A)\dagger}_{k,\ell+}(x')\psi^{(A)}_{k,\ell+}(x)\right\}. \tag{2.66}$$

Then we can use completeness of the set of all wave-functions to rewrite the right-hand-side of the above equation as

$$\lim_{x'\to x}\frac{1}{i}\frac{d}{dx_2}\;\frac{1}{2}\left\{\delta(x'-x)-\int dk\psi^{(A)\dagger}_{0,k}(x')\psi^{(A)}_{0k}(x)\right\}=-\frac{1}{2}\int dk\;k\;\psi^{(A)\dagger}_{0,k}(x)\psi^{(A)}_{0k}(x)\,, \tag{2.67}$$

where we have taken the $x'\to x$ limit and we have assumed that the delta function (or some parity symmetric regularization of the delta function) obeys $\lim_{x'\to x}\frac{d}{dx_2}\delta(x-x')=0$. We conclude that the bulk states have a deficit of momentum density that can be compensated by the edge states. In some sense, for a given valley, it should be half of the possible momentum that could be stored in the edge states.

Now, if we add the two valleys together, the deficit of bulk momentum cancels, since the edge state contributions would have opposite signs. The Dirac vacuum for the bulk states in the two-valley system contributes zero total momentum density. Then, the momentum that is left is entirely due to the edge states and our valley ferromagnetic configuration of the edge states carries a linear momentum density in the direction parallel to the boundary. It is also easy find the integrated momentum density, which is a component of the stress tensor as

$$\left\langle \mathbf{T}^{0y}(t,x,y)\right\rangle=\int dk\;k\;\psi^{(A)\dagger}_{0k}(x)\psi^{(A)}_{0k}(x)=\pm\frac{1}{4\pi}\frac{1}{x^3} \tag{2.68}$$

where the plus of minus sign depends on which valley edge states are filled and which valley is empty. Remember that the integral over k should be over $(0,\infty)$ for valley (1) and $(-\infty,0)$ for valley (2). This is what leads to the plus or minus sign on the right-hand-side of (2.68).

This induced momentum density should characterize our valley ferromagnetic states. Its source definitely deserves further exploration. We can only comment here that the current density, due to charge transport, by a similar argument to the above is

$$\langle J^y(t,x,y)\rangle=-\int dk\;\;\psi^{(A)\dagger}_{0k}(x)\gamma^0\gamma^2\psi^{(A)}_{0k}(x)=0 \tag{2.69}$$

This tells us that the induced momentum density that we found in equation (2.68) is not associated with charge transport.

There is a consequence of this discussion if we consider the total momentum

$$P^y=\int dxdy\;\mathbf{T}^{0y}(t,x,y)$$

and the generator Δ of scale transformations and the total U(1) charge Q. What we have found is that it is impossible to find a distribution of the fermions into edge states so that our candidate ground state obeys all three of the following

$$\Delta|\psi\rangle = 0 \ , \ P^y|\psi\rangle = 0 \ , \ Q|\psi\rangle = 0$$

It is possible to have two out of three in this system with two valleys of fermions.

Of course if we further double the fermion degrees of freedom, then it is possible to have all three of the above. In graphene, where the electrons have spin, this doubling is done for us, it is just the two spin states which to a good approximation are degenerate there. Once we do this additional doubling of the fermions, we can easily find states where all three of the above hold, the state is scale invariant, charge neutral and has zero momentum density. The two possibilities are

1. Both valleys of one spin polarization are filled and both valleys of the other spin polarization are empty. This is a spin ferromagnetic state.

2. For one spin polarization, valley (1) is filled and valley (2) is empty and for the other spin state, valley (2) is filled and valley (1) is empty. This is a valley, rather than spin ferromagnet.

These are the maximally symmetric states and they are ferromagnetic. It is interesting that the valley ferromagnet competes with the spin ferromagnet. This is a prediction of ours. In condensed matter studies it is usually found that the lowest energy state is the spin ferromagnet. Those studies are usually done at the level of the tight-binding lattice theory. The fact that our continuum theory sees an extra state, the valley ferromagnet, is attributable to the fact that the extra state should be there in the lattice theory too, but perhaps with a small energy, one which scales to zero in the continuum limit. In that case, our prediction would be that, as well as the spin ferromagnet which is widely predicted, there is a competing, metastable valley ferromagnet.

In the next section, we will study the effect of introducing a weak repulsive interaction for the fermions. We will see that our two ferromagnetic states, which in this section we have identified by symmetry arguments alone, survive there and are picked out by the interaction as the true ground states.

2.6 Resolution of the ground state degeneracy

Now we must address the question as to why we are so interested in states which are scale invariant. Of course, at the level of the non-interacting theory, there is little that distinguishes scale invariant or charge neutral states from other states that they are degenerate with. The idea here is that, given the high degree of degeneracy, the introduction of even an infinitesimally weak interaction could split the degeneracy

and favour various states and often these are states which have a higher degree of symmetry. We could, for example, introduce an instantaneous Coulomb interaction,

$$H_I = \int d^2x d^2x' \ \rho(x) \ \frac{e^2}{8\pi\epsilon_0|\vec{x}-\vec{x}'|} \ \rho(x') \, . \tag{2.70}$$

The fact that it is the Coulomb interaction is not overly important to us. Our arguments in the following would be valid for any density-density interaction which is repulsive at all scales, translation invariant and has a positive semi-definite interaction Hamiltonian which depends quadratically on the fermion density as (2.70) does. In the expression (2.70), $\rho(x)$ is the (suitably operator ordered) density operator

$$\rho(x) = \frac{1}{2}\left[\psi^{(1)\dagger}(x), \psi^{(1)}(x)\right] \ + \ \frac{1}{2}\left[\psi^{(2)\dagger}(x), \psi^{(2)}(x)\right] \, . \tag{2.71}$$

Since the boundary conditions already break Lorentz invariance we have no incentive to make the interaction Lorentz invariant. In fact, in a realistic Dirac material, the Coulomb interaction is often the most important one. Moreover, the speed of the Dirac fermions is so much smaller than the speed of light that the electromagnetic interactions are often well approximated by the instantaneous Coulomb interaction contained in (2.70).

By its classical engineering dimension, the Coulomb interaction is a marginal deformation of our theory – it contains no dimensionful parameters and it is scale invariant at the classical level. However, it is also well known that, for this theory without the boundary, quantum corrections in the form of renormalization of the massless Dirac theory with this Coulomb interaction added violate scale invariance in such a way that interaction becomes a marginally irrelevant perturbation [46, 47]. That means that the weak coupling limit is stable under the renormalization group flow to the infrared and the fixed point is represented by the scale invariant free field theory that we have been discussing. We expect that the Coulomb interaction introduced to this theory with a boundary behaves in a similar way although the fine details have yet to be explored. Here we will assume that it justifies studying the weak coupling limit in order to resolve the vacuum degeneracy.

Let us consider an ansatz for the ground state which consists of the bulk states having their negative energy levels completely filled and their positive energy levels completely empty and, then, some filling of the edge states. We will study how the Coulomb interaction prefers to fill the edge states. We will denote a state which is such a bulk Dirac vacuum and some definite filling of edge states by the symbol $|\alpha\rangle$. Without interactions, the energy of the fermion system does not depend on how the edge states are filled. There is an enormous degeneracy of ground states in that all states $|\alpha\rangle$ have the same leading order energy.

We shall assume that the interaction in equation (2.70) is sufficiently weak that we can use first order perturbation theory. In the first order of degenerate perturbation theory, we must resolve the ground state degeneracy by diagonalizing the

matrix

$$\delta E_{\alpha\alpha'} \;=\; \langle\alpha|H_I|\alpha'\rangle \tag{2.72}$$

where $|\alpha\rangle, |\alpha'\rangle, \ldots$ are amongst the degenerate ground states of the Hamiltonian in the absence of the interaction.

We will not be able to solve this diagonalization problem in general. However, we will be able to identify the lowest energy states. To do this we examine the operation of the interaction Hamiltonian operator on a state with some definite filling of the edge states. The key to understanding this interaction is to realize that the bulk states in both $|\alpha\rangle$ and $|\alpha'\rangle$ are in the Dirac vacuum with all of the negative energy states filled and positive energy states empty. There are three separate contributions to the matrix element $\langle\alpha|H_I|\alpha'\rangle$ in equation (2.72):

1. The interaction Hamiltonian can create and then re-annihilate a bulk fermion-anti-fermion pair. This process is does not depend on the filling of the edge states and its contribution to the ground state energy is

$$\delta E^{(1)}_{\alpha\alpha'} = \delta E_{\rm bulk}\delta_{\alpha\alpha'} + \ldots \tag{2.73}$$

where

$$\delta E_{\rm bulk} = \int d^2x d^2x' \; \frac{e^2}{8\pi\epsilon_0|x-x'|} \sum_{A=1,2} \int_0^\infty d\ell d\ell' \; \times$$
$$\times \int_{-\infty}^{\infty} dk dk' {\rm Tr}\left[\psi^{(A)}_{k\ell-}(x)\psi^{(A)\dagger}_{k\ell-}(x') \; \psi^{(A)}_{k',\ell'+}(x')\psi^{(A)\dagger}_{k',\ell'+}(x)\right] \tag{2.74}$$

does not depend on the state $|\alpha\rangle$. This is the usual expression for the Coulomb exchange energy which one would compute for the Dirac ground state if the edge states were absent. Generally, this quantity scales like the volume of space and it is ultraviolet divergent. We will assume that it can be defined with a suitable ultraviolet cutoff, the details of which are of no importance to us as this contribution to the matrix that we want to study is already diagonal and it is proportional to the unit matrix, so it is of absolutely no help in distinguishing preferred edge states.

Worthy of note is the fact that the Dirac commutator choice of ordering the operators in the charge density in (2.71) automatically cancels the total charge density of the bulk states when the bulk state is the Dirac vacuum. This is the reason why there is no direct interaction accompanying the exchange interaction in equation (2.75).

2. The interaction Hamiltonian can create and re-annihilate a bulk particle or a bulk anti-particle. This will be accompanied by either annihilation and then creation or creation and then annihilation of edge states, respectively. The

latter processes can happen only when the edge state is suitably occupied or unoccupied. The matrix elements for the two possible processes are

$$\int d^2x d^2x' \ \frac{e^2}{8\pi\epsilon_0|x-x'|}\Bigg\{ \sum_{\substack{k,\ell \\ A,k' \ \text{occupied}}} \mathrm{Tr}\left[\psi^{(A)}_{k\ell-}(x)\psi^{(A)\dagger}_{k\ell-}(x') \ \psi^{(A)}_{0,k'}(x')\psi^{(A)\dagger}_{0,k'}(x)\right]$$

$$+ \sum_{\substack{k,\ell \\ A,k' \ \text{unoccupied}}} \mathrm{Tr}\left[\psi^{(A)}_{k\ell+}(x)\psi^{(A)\dagger}_{k\ell+}(x') \ \psi^{(A)}_{0,k'}(x')\psi^{(A)\dagger}_{0,k'}(x)\right]\Bigg\}, \tag{2.75}$$

where we have taken into account the fact that the particle or hole which are created and then destroyed must have the same k, ℓ quantum numbers and then, that conservation of total x_2-direction momentum requires that the k''s on the edge state wavefunctions are identical. The summation signs stand for integrals over ℓ, k, k' and the bulk wavefunctions in the first term have negative energies whereas those in the second term have positive energies.

It is easy to see from the explicit solutions for the wave-functions that the summations over negative and positive energy bulk states are related by

$$\psi^{(A)}_{k\ell-}(x)\psi^{(A)\dagger}_{k\ell-}(x') = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \ \psi^{(A)}_{k\ell+}(x)\psi^{(A)\dagger}_{k\ell+}(x') \ \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} \tag{2.76}$$

for any k, ℓ and that the matrix $\begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}$ commutes with $\psi^{(A)}_{0,k'}(x')\psi^{(A)\dagger}_{0,k'}(x)$ for any value of k'. This allows us to write equation (2.75) as

$$\int d^2x d^2x' \frac{e^2}{4\pi\epsilon_0|x-x'|} \sum_{\substack{k,\ell \\ A,k'}} \mathrm{Tr}\psi^{(A)}_{k\ell+}(x)\psi^{(A)\dagger}_{k\ell+}(x')\psi^{(A)}_{0,k'}(x')\psi^{(A)\dagger}_{0,k'}(x) \tag{2.77}$$

where the sum over k' is now a sum over all of the edge states in valley A . This sum clearly does not depend on the filling of the edge states. As a result

$$\delta E^{(2)}_{\alpha\alpha'} = \delta E_{\text{bulk}\leftrightarrow\text{edge}}\delta_{\alpha\alpha'} + \ldots$$

In this formula, $\delta E_{\text{bulk}\leftrightarrow\text{edge}}$ does not depend on the states $|\alpha\rangle$ or $|\alpha'\rangle$. This second type of contribution also is insensitive to the structure of the edge states.

3. Finally, there is a term where the interaction Hamiltonian acts only on the edge states. To find its properties, we could restrict the charge densities $\rho(x)$ to operators which act on the edge states only. In second quantization, we could do this by simply dropping all of the bulk state creation and annihilation operators from the charge densities. The restricted charge density operator is

Hermitian and the interaction Hamiltonian that is made from this restricted operator is still a positive semi-definite Hermitian operator. Since it is non-negative, a state in which the expectation value of this restricted interaction Hamiltonian is equal to zero must be an eigenstate with zero eigenvalue and it is thus a lowest energy eigenstate of the edge state restricted interaction Hamiltonian. It is then also a minimal eigenstate of the matrix (2.72) and in linear order of perturbation theory it would be the ground state. This ground state could still have some residual degeneracy as one would expect to happen in a symmetry breaking situation.

Now, it is easy to see that the only possible zero eigenstates of the restricted charge density operator are the states where, for one valley, all of the edge states are full and for the other valley, all of the edge states are empty. These ground states are overall charge neutral, and the background charge implicit in the commutator ordering prescription (2.71) cancels between the valleys. The exchange energy is also zero for those states. This mechanism for finding the lowest energy states is reminiscent of Hund's rule for the population of degenerate, or nearly degenerate atomic orbitals or arguments very similar to the one that we have given here for quantum Hall ferromagnetism [28].

Thus we see that there are two degenerate valley ferromagnetic ground states at this lowest order in perturbation theory. They happen to be identical to the charge neutral and scale invariant states that we constructed in the previous section.

2.7 Discussions

In this chapter, we have done a survey of possible boundary conditions for the 2+1 dimensional Dirac theory defined on a half-space. We have focused on two interesting cases, the BCFT which is a boundary conformal field theory and the zig-zag theory which is not, but where the second is the unique example that supports edge states. We went on to discuss this second example, which we called the zig-zag theory, in some detail.

The boundary condition of the zig-zag theory violates Lorentz invariance and also the special conformal invariance. The problem with these symmetries is the action of the spin operator in their generators. That action is incompatible with the boundary condition. On the other hand, the zig-zag does remain a scale invariant theory and a translation invariant theory, assuming that the boundary of the half-plane is scale and translation invariant, of course.

We found that, with two species of fermions, the only states of the zig-zag theory which are scale invariant and charge neutral are valley ferromagnetic states. In addition, we observed the puzzling fact that these valley ferromagnetic states carry an anomalous linear momentum density. We observed that it is impossible to simultaneously have a state which is scale invariant, charge neutral and vanishing

momentum density.

Of course this does become possible if the fermion species are doubled, which is exactly the case for graphene with spin. To a good approximation the low energy electron dynamics there, the dynamics which is described by the Dirac theory, is spin independent and spin can simply be thought of as an additional degeneracy. On our zig-zag theory on the half-plane, this doubling allows us to find scale invariant, charge neutral states which have vanishing momentum density. Those states are just the edge spin ferromagnetic state which is widely predicted to occur in graphene with a zig-zag edge and another one which is an edge spin neutral valley ferromagnet. The fact that these states are degenerate in the continuum theory implies that they should be almost degenerate at the lattice level where indications are that the spin ferromagnet is favoured. It would be interesting to understand in more detail whether this existence of the valley ferromagnet as a state competing with the spin ferromagnet implies a competing metastable state at the lattice level. This could have very interesting consequences for spintronics applications.

Finally, our proof that our edge ferromagnetic states actually minimize the Coulomb interaction justifies our attention to them. The proof itself has some novel aspects which are interesting could be of use in other applications.

It could be that, as techniques for producing clean Dirac semi-metals in the laboratory evolve, that some of what we say here would be tested by experiments. For a spin polarized case with the zigzag boundary condition, for example, the anomalous momentum density should have some physical consequences which have yet to be worked out.

There has recently been significant discussion of boundary conformal field theories in general [48] and specifically boundary conformal field theories with fermions [49] , higher dimensional versions of the theory that we discuss here. It would be very interesting to see if our some of our ideas generalize there. For example, the existence of edge states of defects or edges of higher than just one dimension is easily seen to occur. For example, it is easy to see that there are boundary conditions for the four dimensional Dirac operator which, like the ones that we have discussed, violate Lorentz and conformal symmetry, but preserve scale invariance and contain edge states very similar to the zigzag theory that we have discussed there. Although no applications to particle physics or cosmology come to mind, for sure three dimensional Weyl semi-metals should have boundaries and this phenomenon would be a possibility there.

Chapter 3

Infrared Divergence of Soft Scalars

In this chapter, we explore the intriguing issue of infrared divergence and its potential link to symmetries within the framework of a scalar model. To set the stage for our discussion, the following section provides the necessary background material, providing the contexts for the main contents of the chapter.

3.1 Background materials

Soft theorem

Infrared divergence occurs in a QFT when there is a massless particle in the spectrum. As a mass term is not allowed by gauge invariance, infrared divergence is present in all gauge theories such as QED, QCD as well as quantum gravity. In such theories, one can write down soft theorems. Let us begin by defining what we mean by a soft particle. A soft particle is simply a *massless* particle with four momentum q^μ with $q := \sqrt{q^2}$ and $\lim q \to 0$[4]. All particles whose momenta are not soft are called *hard* particles. A massless particle is hard when its momentum is not soft. The massive particles are always hard.

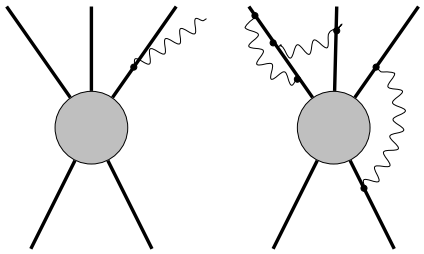


Figure 3.1: In Fig (a) a soft particle is being emitted from an external hard line. Fig. (b) is a typical diagram of virtual exchange of soft particles.

Soft theorem is a statement that a scattering amplitude with $n+1$ particles of which one particle is soft with momentum $q \to 0$ can be written as some function

[4]In this chapter, we are working with $\eta_{\mu\nu} = (-+++)$.

$f(q)$ times the amplitude of n hard particles. Consider a case where we have an infrared finite S matrix

$$S(\{p,k\};\{p',k'\}). \tag{3.1}$$

Here, in the notation $\{p\,,k\}$ we include *incoming* massive hard particles with momenta and helicities in a set $\{p\}$ as well as *incoming* hard massless particles with momenta and helicities $\{k\}$. The primed set $\{p',k'\}$ stands for outgoing hard massive and massless particles with momenta $\{p'\}$ and $\{k'\}$ respectively. We represent such a S-matrix element by the blob. Consider the S-matrix element with an additional outgoing *soft* particle $S(\{p,k\};\{p',k',q\})$ with momentum $\{q\}$, emitted from an *external* line as shown in Fig 3.1(a). Weinberg's soft theorem is the following statement [50]

$$\lim_{q\to 0} S(\{p,k\};\{p',k',q\}) = \left[\mathcal{O}\left(\frac{1}{q}\right) + \mathcal{O}\left(q^0\right) + \mathcal{O}\left(q\right) + ...\right] S(\{p,k\};\{p',k'\}) \tag{3.2}$$

An analogous formula can be derived for $S(p,k,q;p',k')$ when the soft particle is incoming. The S-matrix for an infrared divergent theory follows this general structure. Notably, the leading-order term in (3.2) can be obtained from the Feynman rules governing propagators and the three-point vertex of the theory, along with the on-shell condition of the hard particles. However, deriving the subleading terms is more intricate and depends on the specific details of the theory. In particular, for a gauge theory, the next-to-leading order $\mathcal{O}(q^0)$ terms are fixed by gauge invariance [51–53]. Finally, when a soft particle is emitted from an *internal* line, it does not cause any infrared divergence.

When a soft particle emitted from an external line is reabsorbed by the same or another external line, we get diagrams of the topology shown in Fig 3.1(b) These diagrams calculate the radiative correction due to virtual soft particle exchanges. The S-matrix elements, corresponding to these diagrams, lead to integrals of the form

$$\int_0^\Lambda \frac{d\omega}{\omega^n}, \quad n \geq 1 \tag{3.3}$$

where Λ is an upper energy cut-off. These integrals, as well as the corresponding matrix elements, diverge in the lower limit of the integral leading to infrared divergence in the theory. It was shown by Weinberg [54], these infrared divergent diagrams radiative corrections can be resummed and the resulting S-matrix elements, upon taking into account all the radiative corrections due to infrared soft exchanges is simply

$$S \sim e^{-\infty} = 0.$$

Therefore, in theories like QED or quantum gravity, all scattering processes vanish formally. This is the famous infrared catastrophe in perturbative quantum field

theory. It should be noted that although the full S-matrix is zero, this situation is historically referred to as infrared divergence. That is because at any given order in the perturbation constant, the corresponding Feynman integrals are infrared divergent. We shall stick to the convention and call this situation infrared divergence. Furthermore, a zero S-matrix is also problematic since it implies that there is no unitary operator that takes us from the in/out Fock spaces before and after the scattering.

Dealing with infrared divergence

Since, all observables are finite numbers we need a way to deal with the infrared divergence. The first, known as the *inclusive principle*, is based on the observation that although the elements of the S-matrix exhibit infrared divergence, physical observables – such as scattering cross-sections – are derived from the squared magnitude of the S-matrix. In practice, any measurement has a finite sensitivity, meaning that particles with energy $E < E_{res}$ (where E_{res} represents the minimum energy resolution of our equipment) remain undetectable. Therefore, we must account for the *real* soft particle emissions (that are not reabsorbed like the virtual soft particles) with energies below E_{res}. Incorporating these undetected soft particles results in an infrared-finite $|S|^2$ [55, 56], which suffices for all practical experimental scenarios in particle colliders.

However, it is still unsatisfactory not to have a finite S-matrix but only to have an infrared finite $|S|^2$. It was shown first by Chung [57] and then improved by Faddeev and Kulish [58] that one can indeed achieve an infrared finite S-matrix by dressing the hard particles by a cloud of soft particles. These *soft dressed* are analogous to coherent states in quantum mechanics. The S-matrix elements between these dressed states can be shown to be infrared finite. However, a notable issue with this approach is that these dressed states do not belong to the conventional Fock space of quantum field theory (QFT). A rigorous mathematical framework for understanding these states remains elusive and continues to be an active area of research. For further details, see [59] and the references therein.

Infrared quantum information

As explained above, the inclusive prescription is based on the assumption that certain soft particles remain undetected. So, it is very natural to ask how much quantum information these undetectable soft particles carry away. This can be quantified by computing the von Neumann entropy of the outgoing states by being ignorant about the soft particles in those outgoing states. Formally, this can be done by a mathematical operation known as *tracing out* the undetectable soft particles. Although inclusive principle was established in 1930s, this question has been asked recently in [60]. Interestingly, the authors discovered that these unobservable soft particles carry away a significant amount of quantum information – enough to cause

decoherence in the states of the hard particles in momentum space. Such an effect is absent for the dressing principle suggesting that dressing should be preferred over the inclusive principle [61, 62]. In this chapter, we will go through the details of these claims for a scalar model and the analysis for gauge theories such as QED or gravity is identical.

Physics significance of infrared divergence

We have discussed the infrared divergences and the prescriptions of curing them. Now, let us pose a different question. Is there a deep physical reason why a theory has infrared divergence? Is it as simple as having a massless particle in the spectrum or is there a deep physical reason? In particular, is there some sort of symmetry associated with the issue of infrared divergence? The recent research on this subject can be roughly divided into three categories.

1. A connection between gauge symmetry and infrared divergence has emerged in recent years. We briefly review it here. For more details, the reader is referred to [63]. Consider, electromagnetism where one has a vector field $A_\mu(x)$ and the electromagnetic field tensor $F_{\mu\nu}(x) = \partial_\mu A_\nu(x) - \partial_\nu A_\mu(x)$. Our theory is defined on a manifold $\mathcal{M}$ with boundary denoted by $\partial\mathcal{M}$ and x is the coordinate of a point p of $\mathcal{M}$. Here, $\mathcal{M}$ could be the ordinary Minkowski spacetime and $\partial\mathcal{M}$ would be the conformal null boundaries of the Minkowski space. The theory must be invariant under the gauge transformations

$$A_\mu(x) \to A_\mu(x) + \partial_\mu\lambda(x)$$

for any function $\lambda(x)$. In general, for a well-defined theory we must specify the boundary condition of $A_\mu(x)$, $F_{\mu\nu}(x)$ on $\partial\mathcal{M}$. We can have $\lambda(x)$ that either vanishes on $\partial\mathcal{M}$ or not. When $\lambda(x) = 0$ on $\partial\mathcal{M}$, these gauge transformations trivially preserve the boundary conditions on the fields and are referred to as trivial gauge transformations, as they do not give rise to any conserved charges associated with them. In this sense, the trivial gauge transformations represent redundancies. However, there are gauge transformations, $\lambda(x)$ that do not vanish at the boundary while still preserving the boundary conditions of the fields. These non-trivial or *large* gauge transformations are part of the *asymptotic symmetry group* of large gauge transformations which is an infinite dimensional group. It can be shown that the Noether charges associated with these large gauge transformations constitute an infinite dimensional algebra of conserved charges [64]. For a d dimensional manifold, these large gauge transformations have support in the $d-2$ coordinates x^A. The story does not end here and leads to an interesting subject called *celestial holography*. The above discussion can readily by generalized for other gauge theories like gravity and Yang-Mills theory.

The connection to infrared divergence arises from the interpretation of soft photons and gravitons as Goldstone modes associated with the broken symmetries of certain large gauge transformations. Specifically, the leading and a few sub-leading terms of Weinberg's soft theorem can be derived by computing the commutators of the conserved charges associated with the large gauge transformations with the S-matrix. This intriguing connection suggests that infrared divergence stems from the existence of an infinite number of symmetries within the theory.

2. The above analysis based on the asymptotic symmetry group is particular to the gauge theories. But, what about the non-gauge theories such scalar fields or Fermions? In [65] the authors considered a scalar field theory described by the Lagrangian

$$\mathcal{L} = \frac{1}{2} g_{IJ}(\phi)(\partial_\mu \phi)^I (\partial^\mu \phi)^J - V(\phi) + \lambda_{IJKL}(\phi)(\phi)(\partial_\mu \phi)^I (\partial^\mu \phi)^J (\phi)(\partial_\nu \phi)^K (\partial^\nu \phi)^L + ... \tag{3.4}$$

This theory is very general and devoid of any special symmetry apart from the usual Poincare symmetry. Working on the field space geometry [66, 67], the author showed that it is possible to write down certain soft theorems called *geometric soft theorems* having analogous structure to the scattering amplitude soft theorems. This picture has been extended for scalar theories with fermions, and gauge bosons [68]. The physical interpretation of these geometric soft theorems is that the soft limit is equivalent to an infinitesimal shift of the vacuum expectation value of the corresponding field.

3. Finally, we discuss soft theorems from the point of algebraic quantum field theory. In a QFT, there exists a well-defined algebra of *in* and *out* asymptotic field operators defined at infinite past and infinite future respectively. All the states that are acted on by the *in* and *out* algebra, live on the standard Fock space. One assumes that the *in* and *out* algebras are unitarily equivalent meaning that there is a unitary operator that maps one algebra to the other. The the usual S-matrix is defined to be that unitary operator.

   It is well established that, when there is a massless particle in the spectrum, the leading order massless field at late times does not go back to its original value at early times. This is goes by the name of *memory effect* [69–71]. In particular, the relative separation between a collection of test particles that make up a gravitational wave detector can permanently change due to the passage of a gravitational wave. This phenomenon is known as the gravitational wave memory effect [72]. It can be shown rigorously that in the usual Fock space representation, the norm a state with memory diverges due to low frequency

modes in a way akin to the infrared divergence [73, 74]

$$\int_0^\infty \frac{d\omega}{\omega}\,.$$

The relationship between infrared divergence and the memory effect has further been strengthened in [75] by showing that the memory effect is simply a Fourier transformation of the Weinberg's soft theorem.

Since the states with memory has infinite norm, they do not belong to the standard Fock space[5]. In a recent paper [76], it has been shown that the infrared divergence problem occurs because given an *in* state, we are trying to *preselect* an *out* state with memory that does not belong to an equivalent Fock space. As a result, the S-matrix element that computes the transition amplitude between the two specified states is divergent and ill-defined. This is the physical origin of infrared divergence according to [76].

In the following chapters, we present a scalar field theory model that does not have any special symmetry structure and investigate to what extent this theory mimics the infrared structure of a gauge theory. Our findings are more aligned with the last two groups discussed above.

3.2 Introduction

Infrared divergences [54–56] due to a massless real scalar field in four spacetime dimensions can be, for the most part, structurally identical to those of the massless photons of quantum electrodynamics or the massless gravitons in the perturbative effective field theory of quantum gravity near a flat background. However, there is a big difference between scalar fields and photons or gravitons in that massless scalar fields can in principle occur in quantum field theories which do not exhibit any gauge invariance or even continuous or discrete global symmetry and therefore do not exhibit the asymptotic symmetries which seem intimately tied to the infrared problem of photons and gravitons [63]. In this chapter, we shall study the infrared divergences and their cure in such a scalar field theory and attempt to answer the question as to what is different in that case.

What we will find, in the context of a very specific example, where there are no obvious internal symmetries at all, is one big difference. In this model, one can construct a dressed state of a hard particle where the hard particle is accompanied by a cloud of soft scalars and the soft scalar content of the cloud is fine-tuned in such a way that scattering amplitudes for these dressed particles are free of infrared divergences. Of course such a dressing by soft photons or soft gravitons is already well known, and is already proposed as a solution of the infrared problem of quantum

[5]This is similar to the Faddeev-Kulish dressed states

electrodynamics [57, 58, 77] and perturbative quantum gravity [78]. In those cases, once electrically and gravitationally charged particles are dressed, the dressing can be further fine-tuned in such a way that the soft photons and soft gravitons decouple completely. The S matrix factorizes into a hard sector and a soft sector. Corrections to this factorization are suppressed by powers of an infrared cutoff, E_{res}.

For the soft scalar fields in our model, this is not the case. There is a residual coupling of the scalar-dressed hard particles to soft scalars which is generically as important as the coupling of the dressed hard particles with each other. There is no way to adjust the dressing in order to remove this interaction. This means that even once the hard particles are dressed with soft scalars, an interaction of the hard particles will still produce more soft scalars.

We find that what does decouple is information. Like the case of photons and gravitons, the interaction of undressed hard particles produces infinite numbers of soft scalars. Moreover, the S matrix for scattering processes involving hard particles is infrared divergent. We will confirm these facts in the context of our scalar model. We will also confirm that, like photons and gravitons, the soft scalars which escape detection in a scattering experiment carry away copious amounts of information to the point that their entanglement entropy with the hard particles left behind is itself infrared divergent. The result of this entanglement, like for photons and gravitons, is decoherence of the final state of the hard particles and suppression of interference phenomena [60–62].

When the hard particles are dressed with soft scalars, on the other hand, even if they don't decouple and their interaction still produces some soft scalars, the entanglement between those scalars and the dressed hard particles is negligible, suppressed by powers of the infrared cutoff E_{res}. Even when the soft scalars fly away from a scattering event undetected, the loss of quantum information, which would be manifest in decoherence, is negligible. In this sense it is information that decouples.

The rest of this chapter is an exposition of the results described in the paragraphs above. It is organized as follows. In section 3.3, we will discuss the details of the scalar field theory model that we will use and we will give a derivation of the soft scalar theorem. The soft scalar theorem will be used to study scattering amplitudes in later sections. In section 3.4 we will confirm that, for undressed states, the structure of infrared divergences is practically identical to those of photons or gravitons. We will discuss how the infrared problem is addressed using the Bloch-Nordseick scheme of using inclusive probabilities and we will demonstrate that infrared divergences indeed cancel for the questions that are traditionally asked by particle physicists. Then we will show that some other questions about certain interference phenomena or decoherence are still severely affected by the infrared divergences.

In section 3.5, we will construct dressed states of hard particles. There, we insist that the dressing is added by a formally unitary transformation. This is important since, in the infrared cutoff theory, this is a unitary transform, a simple change

of basis in the Hilbert space. When the fundamental infrared cutoff is removed, it becomes an improper unitary transformation. However, even in that case, if it is formally unitary, it implements a canonical transformation between inequivalent representations of the operator algebra of the quantum field theory. We also discuss the Faddeev-Kulish modification of the S matrix that is to be computed to find the scattering amplitudes for dressed particles. In section 3.6, we demonstrate that the infrared singularities indeed cancel from those amplitudes. In section 3.7, we establish that soft scalars do not decouple from the dressed states. In section 3.8, we discuss the implications. We also argue that, even though soft scalars do not decouple, they do not drive appreciable decoherence.

3.3 Soft Scalar Theorem

Consider a Majorana fermion of mass m coupled to a massless real scalar field ϕ via a Yukawa coupling in four spacetime dimensions. The Lagrangian density is given by

$$\mathcal{L} = -\frac{i}{2}\bar{\psi}\left[\not{\partial} + m - g\phi\right]\psi - \frac{1}{2}\partial_\mu\phi\,\partial^\mu\phi - V(\phi) \tag{3.5}$$

where ψ is the Majorana spinor field, g is the Yukawa coupling constant and ϕ is the scalar field. Since the Yukawa coupling is marginal in four dimensions, g is dimensionless. We will assume that we can always choose counter-terms so that the scalar field tadpole $\sim \phi$, the scalar field mass term $\sim \phi^2$ and the scalar field trivalent coupling $\sim \phi^3$ are all canceled exactly at each order of perturbation theory. To make the Lehmann-Symanzik-Zimmermann (LSZ) reduction formulae simpler, we will also assume that the subtraction scheme can be chosen so that the pole in the scalar field propagator has unit residue, $\sim \frac{-i}{q^2-i\epsilon} + \ldots$. We note that this quantum field theory has no obvious internal symmetries at all. Fermion parity $(-1)^F$ can be regarded as a space time symmetry since it is a rotation by 2π. Our use of the Majorana fermion is not essential as results would be very similar for a complex fermion, which would of course have at least one internal $U(1)$ symmetry. Our use of Majorana fermions is motivated by wanting a model with no internal symmetry at all. We will study the infrared divergences which occur in the analysis of scattering experiments involving the asymptotic particles of this quantum field theory. We will assume that the coupling is weak so that the particle spectrum resembles the tree level one, containing one Majorana fermion with mass m and one massless real scalar field. There has already been some discussion of soft scalar theorems [65], infrared divergences and the possibility of asymptotic symmetries playing a role for scalar fields [79–81]. The latter has been discussed in the context of a dual antisymmetric tensor gauge field representation which, for a Yukawa coupling like we use here, is not related to this theory by a local transformation. This is not the direction that we will pursue here. Instead we will examine the behaviour of the scattering matrix in the quantized theory corresponding to (3.5).

A scattering experiment has an incoming state. We will use the notation $|\{p\}\{k\}\{q\}\rangle$ for such a state, with fermions having momenta and helicity in a set $\{p\}$, hard scalars with momenta $\{k\}$ and soft scalars with momenta $\{q\}$. In the course of the scattering, an incoming state evolves to an outgoing state which is a quantum superposition of the incoming states,

$$|\{p\}\{k\}\{q\}\rangle \implies \sum_{\{p'\}\{k'\}\{q'\}} |\{p'\}\{k'\}\{q'\}\rangle \; S^\dagger(\{p'\}\{k'\}\{q'\};\{p\}\{k\}\{q\}) \tag{3.6}$$

The coefficients in the superposition, $S^\dagger(\{p'\}\{k'\}\{q'\};\{p\}\{k\}\{q\})$, are elements of the S matrix. The dagger in the above formula is there to match conventions,

$$\begin{aligned} &S^\lambda(\{p\}\{k\}\{q\};\{p'\}\{k'\}\{q'\}) \\ &= \langle\{p\}\{k\}\{q\}|\; S^\lambda \;|\{p'\}\{k'\}\{q'\}\rangle \end{aligned} \tag{3.7}$$

We have added a superscript λ to the S matrix to remind ourselves that its definition requires a fundamental infrared cutoff. We will denote this fundamental cutoff by λ.[6] With the cutoff taken into account, the operator S^λ whose matrix elements are discussed above is the Dyson S matrix which is computed using the usual LSZ reduction formulae, time dependent perturbation theory and Feynman diagrams. In the rest of this chapter, what we mean by an infrared finite quantity is that said quantity remains finite as we take $\lambda \to 0$. The matrix elements of S^λ in (3.7) are generally not such a quantity.

Beyond the fundamental infrared cutoff λ, we shall also require a distinction between hard and soft particles. Any particle which has energy above a threshold E_{res} will be called a hard particle. Any particle which has energy less than E_{res} will be called a soft particle. We shall call E_{res} the "detector resolution" . We will require a hierarchy of scales

$$\lambda \ll E_{res}, \Lambda \ll m, \text{ energies of hard particles} \tag{3.8}$$

where Λ is a third infrared cutoff, distinct from λ and E_{res} that we will introduce shortly. Fermions are always hard particles. The massless scalar, on the other hand can either be hard or soft, depending on its energy. The validity of the arguments in the rest of this chapter will need the inequalities in equation (3.8).

Let us consider an amplitude for the scattering of some fermions and hard scalar particles as well as the production of a soft scalar particle. Whenever this occurs, there are some contributions to the amplitude which are singular as the momentum of the soft scalar approaches zero. In direct analogy with the same phenomenon in quantum electrodynamics, which is outlined in beautiful detail in Weinberg's book

[6] An easy, Lorentz invariant way of introducing such a cutoff is to simply allow the scalar field to have a small mass.

[50], the most singular parts come from the emission of the soft scalar from external lines of the amplitude. This singular part, $\sim \frac{1}{q}$ for soft scalar with momentum q, and the next-to leading behaviour due to a scalar emitted from an outgoing fermion line with momentum p', is gotten from the amplitude without the scalar emission by making the replacement

$$\begin{aligned}
\bar{u}_r(p') \quad &\to \quad \lim_{q\to 0} \; \bar{u}_r(p')(-g)\frac{-1}{i\not{p}' + i\not{q} + m - i\epsilon} \\
&= \lim_{q\to 0} \; \bar{u}_r(p')\Bigg(\frac{2gm}{p'^2 + 2p'\cdot q + q^2 + m^2 - i\epsilon} \\
&\qquad\qquad - \frac{g}{-i\not{p}' - i\not{q} + m - i\epsilon}\Bigg) \\
&= \left(\frac{gm}{p'\cdot q - i\epsilon} - \frac{g}{2m}\right)\bar{u}_r(p') + \mathcal{O}(q)
\end{aligned} \tag{3.9}$$

where $\bar{u}_r(p)$ is the momentum space spinor wave-function of the outgoing fermion and we have used the fact that the spinor satisfies the Dirac equation $\bar{u}_r(p')(i\not{p}' + m) = 0$. Here, we note that the contribution to the next-to leading behaviour from the external fermion line has a very simple form for a scalar field. Corrections to this formula go to zero as q goes to zero.

Similarly, when the soft scalar emission is from an incoming fermion line, the singular part and next-to-leading contribution is

$$u_r(p) \;\to\; u_r(p)\;\left(\frac{gm}{-p\cdot q - i\epsilon} - \frac{g}{2m}\right) + \mathcal{O}(q) \tag{3.10}$$

Of course, soft emissions can also take place from internal lines in the Feynman diagrams which contribute to the amplitude. The contribution of such processes is not singular at small q, but it does compete with the next-to-leading terms in (3.9) and (3.10). The effect of one soft emission from an internal line is the same as adding a vertex, together with a vertex counter-term to each of the fermion propagators inside the amputated correlation function that is used to form the S matrix element $S^\lambda(\{p\}\{k\};\{p'\}\{k'\})$. We will denote this process by the symbol $-g\hat{\partial}_m$. We note that the operation $\hat{\partial}_m$ is related to but not exactly the same as taking the derivative of amputated correlation function by the renormalized fermion mass m. The discrepancy is due to the fact that we need a subtraction scheme where some counter-terms are m-dependent in a way that does not preserve the $m - g\phi$ structure that appears in the Lagrangian (3.5). It is easy to see that even though the $m - g\phi$ structure can be maintained at the tree level, it is already violated at the one-loop level.

Putting this together, and a similar one for soft scalar absorption, we have the leading and next-to-leading soft scalar theorem

$$
\begin{aligned}
&S^{\lambda}(\{p\},\{k\};\{p'\}\{k'\},q') \\
&\quad=\left\{\sum_{p_n\in\{p\}\{p'\}}\left(\frac{\eta_n g m}{p_n\cdot q'-i\eta_n\epsilon}-\frac{g}{2m}\right)-g\hat{\partial}_m\right\} \\
&\quad\times S^{\lambda}(\{p\}\{k\};\{p'\}\{k'\})+\mathcal{O}(q')
\end{aligned}
\tag{3.11}
$$

$$
\begin{aligned}
&S^{\lambda}(\{p\},\{k\},q;\{p'\}\{k'\}) \\
&\quad=\left\{\sum_{p_n\in\{p\}\{p'\}}\left(\frac{\eta_n g m}{-p_n\cdot q-i\eta_n\epsilon}-\frac{g}{2m}\right)-g\hat{\partial}_m\right\} \\
&\quad\times S^{\lambda}(\{p\}\{k\};\{p'\}\{k'\})+\mathcal{O}(q)
\end{aligned}
\tag{3.12}
$$

where $\eta_n=+1(-1)$ if p_n is the momentum of an outgoing (incoming) line. Now let us consider the amplitude for a process where an incoming state of hard particles and M additional soft scalars $|\{p\}\{k\},q_1,\ldots,q_M\rangle$ evolves to another state of hard particles but with N additional soft particles $|\{p'\}\{k'\},q'_1,\ldots,q'_N\rangle$. Using the leading parts of the soft scalar theorem in equations (3.11) and (3.12), it follows that the most singular part of the S matrix element is given by[7]

$$
\begin{aligned}
&S^{\lambda}(\{p\}\{k\},q_1,\ldots,q_M;\{p'\}\{k'\},q'_1,...q'_N) \\
&=S^{\lambda}(\{p\}\{k\};\{p'\}\{k'\})\prod_{j=1}^{M}\left(\sum_{p_n\in\{p\}\{p'\}}\frac{\eta_n g m}{-p_n\cdot q_j-i\eta_n\epsilon}\right) \\
&\times\prod_{r=1}^{N}\left(\sum_{p_n\in\{p\}\{p'\}}\frac{\eta_n g m}{p_n\cdot q'_r-i\eta_n\epsilon}\right)+\ldots
\end{aligned}
\tag{3.13}
$$

where the ellipses denote terms less singular at small q or q' than $\frac{1}{q^M}\frac{1}{q'^N}$. The remarkable fact about this soft theorem is that it gives us the most important part of the soft scalar production amplitude for any process if we know the amplitude of the process without the soft scalar production. What is more, it is practically identical to the one for soft photons or gravitons where only the numerators in the singular factors (here it is gm) are slightly different.

[7]We have not attached momentum space wave-functions for the soft scalars. We will take the convention of including them as the appropriate factors in the scalar states.

3.4 Infrared divergence cancellation in the Bloch-Nordseick Scheme

It is easy to see that, in direct parallel with quantum electrodynamics and perturbative quantum gravity, the infrared divergences in the S matrix itself come from loop integrals where both ends of a single scalar propagator in the loop end on external fermion lines. Introduction of the fundamental infrared cutoff λ renders these loop integrals finite and of order $\sim \ln(\lambda)$.

We imagine that, in these loop integrals, the infrared cutoff λ could be replaced by a more convenient one which we shall call Λ. This is a third infrared cutoff, distinct from E_{res} and λ. with the assumption that it is much larger than the fundamental cutoff, $\Lambda \gg \lambda$, but it would still work as an infrared cutoff in that $\Lambda \ll m$ and that Λ is much smaller than the momentum scales of any of the hard particles. It is in the same interval of the hierarchy (3.8) as the detector resolution E_{res}. Then with the new cutoff the logarithmically divergent integral goes as $\sim \ln(\Lambda)$ and the original integral with cutoff λ goes like $\sim \ln(\lambda) = \ln(\Lambda) + \ln(\lambda/\Lambda)$, the first logarithm being produced by the integration over loop momenta from Λ to infinity and the second being produced by the integration over loop momenta between λ and Λ. Moreover, it is well known how to separate and sum up the latter, $\ln(\lambda/\Lambda)$ contributions. We refer the reader to Weinberg's book [50] for the details in the case of photons and we note that, for massless scalars, the argument is practically identical. The result is a relationship between the S matrix defined with the two different infrared cutoffs

$$
\begin{aligned}
&S^{\lambda}(\{p\}\{k\};\{p'\}\{k'\}) \\
&\quad = S^{\Lambda}(\{p\}\{k\};\{p'\}\{k'\}) e^{-i\bar{\Phi}(\{p\})-i\bar{\Phi}(\{p'\})} \\
&\quad \times \exp\left\{ -\frac{1}{2}\int_{\lambda}^{\Lambda} \frac{d^3q}{(2\pi)^3 2|\vec{q}|} \sum_{p_n\in\{p\}\{p'\}} \frac{gm\eta_n}{p_n\cdot q} \sum_{p_m\in\{p\}\{p'\}} \frac{gm\eta_m}{p_m\cdot q} \right\} \qquad (3.14) \\
&\quad = S^{\Lambda}(\{p\}\{k\};\{p'\}\{k'\}) \left(\frac{\lambda}{\Lambda}\right)^{\frac{1}{2}\bar{A}(\{p\},\{p'\})} e^{-i\bar{\Phi}(\{p\})-i\bar{\Phi}(\{p'\})}\,; \qquad (3.15)
\end{aligned}
$$

$$
\text{where } \bar{A}(\{p\},\{p'\}) = -\frac{1}{8\pi^2} \sum_{p_n p_m\in\{p\}\{p'\}} \frac{g^2m^2\eta_n\eta_m}{\gamma_{nm}} \ln\left(\frac{1+\xi_{nm}}{1-\xi_{nm}}\right),
$$

$$
\bar{\Phi}(\{p\}) = \frac{1}{8\pi} \sum_{\substack{p_m p_n\in\{p\} \\ m\neq n}} \frac{g^2m^2}{\gamma_{nm}} \ln\frac{\Lambda}{\lambda}\,, \qquad (3.16)
$$

$$
\xi_{nm} = \sqrt{1-\frac{m^4}{(p_n\cdot p_m)^2}}\ , \quad \gamma_{nm} = (p_n\cdot p_m)\sqrt{1-\frac{m^4}{(p_n\cdot p_m)^2}}\ . \qquad (3.17)
$$

To be clear, the beautiful formula (3.14) does not remove the fundamental in-

frared cutoff. It simply gives us a relationship between S matrices computed with different infrared cutoffs. Furthermore, it is strictly valid only when $\lambda \ll \Lambda \ll m$. We note that the exponent, $\bar{A}(\{p\},\{p'\})$, of the ratio of cutoffs contains data about the incoming and outgoing fermions only. It is independent of the hard incoming or outgoing scalar particles. The phases, $\bar{\Phi}(\{p\}), \bar{\Phi}(\{p'\})$ are separated into two functions, one of incoming and one of outgoing fermion momenta. Notice that the two functions have the same sign.

For reasons which will become clear shortly, it is useful to express the evolution from the initial to final state in equation (3.6) in the language of density matrices where a more general incoming state composed entirely of hard particles would be

$$\sum_{\substack{\{p\}\{k\}\\ \{\tilde{p}\}\{\tilde{k}\}}} |\{p\}\{k\}\rangle \rho_{\{p\}\{k\}\{\tilde{p}\}\{\tilde{k}\}} \langle\{\tilde{p}\}\{\tilde{k}\}|$$

with $\rho_{\{p\}\{k\}\{\tilde{p}\}\{\tilde{k}\}}$ an incoming density matrix. In the course of a scattering experiment, this incoming density matrix evolves to an outgoing one where the evolution is governed by the S matrix,

$$\begin{aligned}
&\sum_{\substack{\{p\}\{k\}\\ \{\tilde{p}\}\{\tilde{k}\}}} |\{p\}\{k\}\rangle \rho_{\{p\}\{k\};\{\tilde{p}\}\{\tilde{k}\}} \langle\{\tilde{p}\}\{\tilde{k}\}| \\
&\Longrightarrow \sum_{\substack{\{p\}\{k\}\\ \{\tilde{p}\}\{\tilde{k}\}}} \rho_{\{p\}\{k\}\{\tilde{p}\}\{\tilde{k}\}} \sum_{\substack{\{p'\}\{k'\}\{q'\}\\ \{\tilde{p}'\}\{\tilde{k}'\}\{\tilde{q}'\}}} |\{p'\}\{k'\}\{q'\}\rangle\langle\{\tilde{p}'\}\{\tilde{k}'\}\{\tilde{q}'\}| \\
&\qquad \times S^{\lambda\dagger}(\{p'\}\{k'\}\{q'\};\{p\}\{k\})\ \ S^{\lambda}(\{\tilde{p}\}\{\tilde{k}\};\{\tilde{p}'\}\{\tilde{k}'\}\{\tilde{q}'\})\,. \qquad (3.18)
\end{aligned}$$

We have assumed that there are no soft particles in the incoming states. However, soft particles are produced when the hard particles interact and they must appear in the final state density matrix. We remember that the S matrix elements in the expression above are infrared divergent and they are defined with a fundamental infrared cutoff, λ. We will implement the Bloch-Nordsieck mechanism where we make the assumption that, due to the limitations of detector resolution, the soft particles which are produced by the scattering are unobservable. They fly away from the scattering experiment undetected. What is left behind are the hard particles. All of the experimentally accessible properties of the quantum state of the hard particles which remain are embedded in the reduced density matrix that is gotten from the final state density matrix in (3.18) by taking a trace over all of the soft scalar states. The reduced density matrix of the final state is thus

$$\rho_{\text{final}} = \sum_{\substack{\{p\}\{k\}\\ \{\tilde{p}\}\{\tilde{k}\}}} \rho_{\{p\}\{k\};\{\tilde{p}\}\{\tilde{k}\}} \sum_{\substack{\{p'\}\{k'\}\\ \{\tilde{p}'\}\{\tilde{k}'\}}} |\{p'\}\{k'\}\rangle\langle\{\tilde{p}'\}\{\tilde{k}'\}|$$
$$\times \sum_{\{q\}} S^{\lambda\dagger}(\{p'\}\{k'\}\{q\};\{p\}\{k\}) S^{\lambda}(\{\tilde{p}\}\{\tilde{k}\};\{\tilde{p}'\}\{\tilde{k}'\}\{q\}) \tag{3.19}$$

where $\sum_{\{q\}}$ denotes integration and summation over all possible soft scalar states. We can use the soft scalar theorem (3.13) to simplify equation (3.19). To take the trace, we identify pairs of ingoing and outgoing q's and we integrate each identified pair over all values of $\vec{q}$ with $\lambda < |\vec{q}| < E_{res}$. Using the soft scalar theorem yields the expression[8]

$$\sum_{\{q\}} S^{\lambda\dagger}(\{p'\}\{k'\}\{q\};\{p\}\{k\}) \;\; S^{\lambda}(\{\tilde{p}\}\{\tilde{k}\};\{\tilde{p}'\}\{\tilde{k}'\}\{q\})$$
$$= \sum_{N=0}^{\infty} \frac{1}{N!} \int_{\lambda}^{E_{res}} \frac{d^3\vec{q_1}}{(2\pi)^3 2|\vec{q_1}|} \cdots \frac{d^3\vec{q_N}}{(2\pi)^3 2|\vec{q_N}|}$$
$$\times \prod_{r=1}^{N} \left(\sum_{p_n \in \{p\}\{p'\}} \frac{gm\eta_n}{p_n \cdot q} \sum_{p_n \in \{p\}\{p'\}} \frac{gm\eta_m}{-p_m \cdot q} \right)$$
$$\times S^{\lambda\dagger}(\{p'\}\{k'\};\{p\}\{k\}) \;\; S^{\lambda}(\{\tilde{p}\}\{\tilde{k}\};\{\tilde{p}'\}\{\tilde{k}'\}) \,.$$

Notice that, in each factor in the product of integrals in the equation above, the three dimensional integration over $\vec{q}$ is over a narrow shell with $\lambda < |\vec{q}| < E_{res}$ and the result would be small, $\sim E_{res}^2$, if it were not for the singular terms due to soft scalar emission. It is those singular terms which allow the integrals to be appreciable, in fact logarithmically infrared divergent $\sim \ln E_{res}/\lambda$. Corrections to the above formula due to the non-singular next-to-leading contributions to the soft scalar theorem would be relatively suppressed by positive powers of E_{res}. The summation in the equation above exponentiates and we find the expression for the reduced final state density matrix

$$\rho_{\text{final}} = \sum_{\substack{\{p\}\{k\}\\ \{\tilde{p}\}\{\tilde{k}\}}} \rho_{\{p\}\{k\};\{\tilde{p}\}\{\tilde{k}\}} \sum_{\substack{\{p'\}\{k'\}\\ \{\tilde{p}'\}\{\tilde{k}'\}}} |\{p'\}\{k'\}\rangle\langle\{\tilde{p}'\}\{\tilde{k}'\}|$$
$$\times S^{\lambda\dagger}(\{p'\}\{k'\};\{p\}\{k\}) S^{\lambda}(\{\tilde{p}\}\{\tilde{k}\};\{\tilde{p}'\}\{\tilde{k}'\})$$
$$\times \exp\left\{ \int_{\lambda}^{E_{res}} \frac{d^3\vec{q}}{(2\pi)^3 2|\vec{q}|} \sum_{p_n \text{in} \{p\}\{p'\}} \frac{gm\eta_n}{p_n \cdot q} \sum_{p_m \in \{\tilde{p}\}\{\tilde{p}'\}} \frac{gm\eta_m}{-p_m \cdot q} \right\}. \tag{3.20}$$

[8] Since, once the scalar momenta q are on-shell, $p_n \cdot q > 0$ and we can drop the ie's from the denominators.

Now, we must examine the infrared cutoff dependence of the S matrix elements on the right-hand-side of equation (3.20). For this, we must use the λ-dependence of the S matrix elements that is summarized in equation (3.14). We then get

$$\begin{aligned}
\rho_{\text{final}} = & \sum_{\substack{\{p\}\{k\} \\ \{\tilde{p}\}\{\tilde{k}\}}} \rho_{\{p\}\{k\}\{\tilde{p}\}\{\tilde{k}\}} \sum_{\substack{\{p'\}\{k'\} \\ \{\tilde{p}'\}\{\tilde{k}'\}}} |\{p'\}\{k'\}\rangle\langle\{\tilde{p}'\}\{\tilde{k}'\}| \\
& \times S^{\Lambda\dagger}(\{p'\}\{k'\};\{p\}\{k\}) S^{\Lambda}(\{\tilde{p}\}\{\tilde{k}\};\{\tilde{p}'\}\{\tilde{k}'\}) e^{i\bar{\Phi}(\{p\})+i\bar{\Phi}(\{p'\})-i\bar{\Phi}(\{\tilde{p}\})-i\bar{\Phi}(\{\tilde{p}'\})} \\
& \times \exp\left\{ \int_{\lambda}^{E_{res}} \frac{d^3\vec{q}}{(2\pi)^3 2|\vec{q}|} \sum_{p_n \in \{p\}\{p'\}} \frac{gm\eta_n}{p_n \cdot q} \sum_{p_m \in \{\tilde{p}\}\{\tilde{p}'\}} \frac{gm\eta_m}{-p_m \cdot q} \right\} \\
& \times \exp\left\{ -\int_{\lambda}^{\Lambda} \frac{d^3 q}{(2\pi)^3 2|\vec{q}|} \left[\frac{1}{2} \left[\sum_{p_n \in \{p\}\{p'\}} \frac{gm\eta_n}{p_n \cdot q} \right]^2 + \frac{1}{2} \left[\sum_{p_n \in \{\tilde{p}\}\{\tilde{p}'\}} \frac{gm\eta_n}{p_n \cdot q} \right]^2 \right] \right\}. \qquad (3.21)
\end{aligned}$$

Upon combining the exponentials in the last two lines, we can write the above expression as

$$\begin{aligned}
\rho_{\text{final}} = & \sum_{\substack{\{p\}\{k\} \\ \{\tilde{p}\}\{\tilde{k}\}}} \rho_{\{p\}\{k\}\{\tilde{p}\}\{\tilde{k}\}} \sum_{\substack{\{p'\}\{k'\} \\ \{\tilde{p}'\}\{\tilde{k}'\}}} |\{p'\}\{k'\}\rangle\langle\{\tilde{p}'\}\{\tilde{k}'\}| \\
& \times S^{\Lambda\dagger}(\{p'\}\{k'\};\{p\}\{k\}) S^{\Lambda}(\{\tilde{p}\}\{\tilde{k}\};\{\tilde{p}'\}\{\tilde{k}'\}) e^{i\bar{\Phi}(\{p\})+i\bar{\Phi}(\{p'\})-i\bar{\Phi}(\{\tilde{p}\})-i\bar{\Phi}(\{\tilde{p}'\})} \\
& \times \exp\left\{ \int_{\Lambda}^{E_{res}} \frac{d^3\vec{q}}{(2\pi)^3 2|\vec{q}|} \sum_{p_n \in \{p\}\{p'\}} \frac{gm\eta_n}{p_n \cdot q} \sum_{p_m \in \{\tilde{p}\}\{\tilde{p}'\}} \frac{gm\eta_m}{-p_m \cdot q} \right\} \\
& \times \exp\left\{ -\frac{1}{2} \int_{\lambda}^{\Lambda} \frac{d^3 q}{(2\pi)^3 2|\vec{q}|} \left[\sum_{p_n \in \{p\}\{p'\}} \frac{gm\eta_n}{p_n \cdot q} - \sum_{p_m \in \{\tilde{p}\}\{\tilde{p}'\}} \frac{gm\eta_m}{p_m \cdot q} \right]^2 \right\}. \qquad (3.22)
\end{aligned}$$

Now we want to examine the right-hand-side of equation (3.22) as the fundamental infrared cutoff $\lambda \to 0$. The second line contains phases which are separately infrared divergent and we reserve comment on them for later. The third line is λ-independent and infrared finite. The fourth (last) line has a negative semi-definite

exponent which can be written as

$$\exp\Bigg\{-\frac{1}{2}\int_\lambda^\Lambda \frac{d^3q}{(2\pi)^3 2|\vec{q}|}\Bigg[\sum_{p_n\in\{p\}\{p'\}}\frac{gm\eta_n}{p_n\cdot q} - \sum_{p_m\in\{\tilde{p}\}\{\tilde{p}'\}}\frac{gm\eta_m}{p_m\cdot q}\Bigg]^2\Bigg\}$$
$$= \exp\Bigg\{-\frac{1}{32\pi^3}\ln\frac{\Lambda}{\lambda}\cdot\int d\hat{q}\Bigg[\sum_{p_n\in\{p\}\{p'\}}\frac{gm\eta_n}{p_n\cdot v} - \sum_{p_m\in\{\tilde{p}\}\{\tilde{p}'\}}\frac{gm\eta_m}{p_m\cdot v}\Bigg]^2\Bigg\}\ ;\ v^\mu=(1,\hat{q})\ ,\quad \hat{q}\equiv\vec{q}/|\vec{q}|\,.$$

This exponent is either negative and logarithmically divergent or it vanishes. It can vanish only if the integrand in the integration over unit vectors vanishes, that is, if

$$\sum_{p_n\in\{p\}\{\tilde{p}'\}}\frac{gm}{p_n\cdot v} = \sum_{p_m\in\{p'\}\{\tilde{p}\}}\frac{gm}{p_m\cdot v}\,. \tag{3.23}$$

It is only in this case where the $|\{p'\}\{k'\}\rangle\langle\{\tilde{p}'\}\{\tilde{q}'\}|$ matrix element of the reduced outgoing density matrix can be nonzero when the matrix element $|\{p\}\{k\}\rangle\langle\{\tilde{p}\}\{\tilde{q}\}|$ of the incoming density matrix was nonzero.

Remember that the sums in (3.23) are over terms containing hard fermion momenta only. Hard scalar momenta do not enter in these expressions. The above equation must be so for all values of the null four-vector $v^\mu=(1,\hat{q})$. If we Taylor expand the above in powers off $\frac{\vec{p}_n}{\sqrt{\vec{p}_n^2+m^2}}<1$ and equate each order we see that

$$\sum_{p_n\in\{p\}\{\tilde{p}'\}}\frac{(\hat{q}\cdot\vec{p}_n)^\ell}{(\sqrt{\vec{p}_n^2+m^2})^{\ell+1}} = \sum_{p_n\in\{\tilde{p}\}\{p'\}}\frac{(\hat{q}\cdot\vec{p}_n)^\ell}{(\sqrt{\vec{p}_n^2+m^2})^{\ell+1}},\qquad \forall\ell,\ \forall\hat{q}\,.$$

which implies that all multipole moments of the set of fermion momentum vectors $\{p\}\cup\{\tilde{p}'\}$ are equal to all moments of the set $\{p'\}\cup\{\tilde{p}\}$ which can only be so if the two sets of vectors are identical

$$\{p\}\cup\{\tilde{p}'\} = \{p'\}\cup\{\tilde{p}\}\,. \tag{3.24}$$

Only those elements of the density matrix for which this criterion is satisfied survive the limit $\lambda\to 0$.

Now, notice that, when the constraint (3.24) is obeyed, the phases in the third line of equation (3.22) also cancel exactly. The result is, for evolutions from the initial to the final state for which (3.24) holds, the final density matrix is free of infrared divergences. Let us examine some of the consequences of this result.

Since, by unitarity, both the incoming and reduced outgoing density matrices must have unit trace, they both must have nonzero diagonal matrix elements. Moreover, a question about the evolution of a diagonal element of the density matrix element to another diagonal element of the density matrix is unaffected by the constraint (3.24). This is due to the fact that, in such a case, $\{p\} = \{p'\}$ and $\{\tilde{p}\} = \{\tilde{p}'\}$ and (3.24) is automatically satisfied. This is the question that is usually asked in particle physics: what is the probability that the state $|\{p\}\{k\}\rangle\langle\{p\}\{k\}|$ will evolve to the state $|\{p'\}\{k'\}\rangle\langle\{p'\}\{k'\}|$? This quantity is finite in the limit $\lambda \to 0$ and it is unconstrained by the condition (3.24). We have nothing new to say about it.

On the other hand, if we ask the questions which probe the off-diagonal elements of the density matrix, the $\lambda \to 0$ limit can have drastic consequences. If we ask what is the probability of the process

$$\begin{aligned}&\frac{1}{\sqrt{2}}\left(|\{p_1\}\{k_1\}\rangle + |\{p_2\}\{k_2\}\rangle\right) \times \\ &\frac{1}{\sqrt{2}}\left(\langle\{p_1\}\{k_1\}| + \langle\{p_2\}\{k_2\}|\right) \implies |\{p'\}\{k'\}\rangle\langle\{p'\}\{k'\}|\,,\end{aligned}$$

we find that, unless $\{p_1\} = \{p_2\}$, it is

$$\begin{aligned}&\frac{1}{2}\text{ Prob. of } |\{p_1\}\{k_1\}\rangle\langle\{p_1\}\{k_1\}| \implies |\{p'\}\{k'\}\rangle\langle\{p'\}\{k'\}| \\ &\qquad\qquad + \\ &\frac{1}{2}\text{ Prob. of } |\{p_2\}\{k_2\}\rangle\langle\{p_2\}\{k_2\}| \implies |\{p'\}\{k'\}\rangle\langle\{p'\}\{k'\}|\,.\end{aligned}$$

There is no interference between the incoming states.

The probability for evolving to a superposition, on the other hand

$$\begin{aligned}|\{p\}\{k\}\rangle\langle\{p\}\{k\}| \implies &\frac{1}{\sqrt{2}}\left(|\{p_1'\}\{k_1'\}\rangle + |\{p_2'\}\{k_2'\}\rangle\right) \\ &\times \frac{1}{\sqrt{2}}\left(\langle\{p_1'\}\{k_1'\}| + \langle\{p_2'\}\{k_2'\}|\right)\end{aligned}$$

unless $\{p_1'\} = \{p_2'\}$, is simply the sum

$$\begin{aligned}&\frac{1}{2}\text{ Prob. of } |\{p\}\{k\}\rangle\langle\{p\}\{k\}| \implies |\{p_1'\}\{k_1'\}\rangle\langle\{p_1'\}\{k_1'\}| \\ &\qquad\qquad + \\ &\frac{1}{2}\text{ Prob. of } |\{p\}\{k\}\rangle\langle\{p\}\{k\}| \implies |\{p_2'\}\{k_2'\}\rangle\langle\{p_2'\}\{k_2'\}|\,.\end{aligned}$$

The state experiences complete decoherence.

In order to get the results outlined above, we reduced the final state density matrix by tracing it over the Fock space states of the soft scalar fields. Of course,

since we are tracing over this entire subspace of the total Hilbert space, tracing over any redefinition of the Fock basis for soft scalars by a unitary transformation must give the same answer. To find a result that is any different than what we have obtained, one would have to use a redefinition of the basis that is not implemented by a proper unitary transformation. Of course, it is easy to find such improper unitary transformations in a Fock space. Here, the relevant one is the basis constructed around certain coherent states which become improper coherent states when the fundamental infrared cutoff is removed. Indeed, this redefinition of the basis for soft scalar states is easily implemented and, when one subsequently traces in such a basis, the reduced final state density matrix differs in ways that are physically consequential. In that basis the evolution does not exhibit the severe decoherence or suppression of interference that we found for the Fock state basis.

However, then we would be in a situation where the S matrix evolves an initial soft scalar Fock vacuum to soft scalar coherent states which live in a different Hilbert space and the S matrix itself is therefore not a proper unitary operator. The only way to preserve unitarity of S in this context is to also use the coherent states as incoming states, so that S evolves coherent states to coherent states in such a way that it is unitary. This is the gist of what is done in the dressed state formalism which we will discuss in the context of soft scalar fields in the next section. We expect that the dressed states, in the limit where the cutoff is removed, will have the same problems with violations of Lorentz invariance as photon and graviton dressed states [82–84].

3.5 Soft Scalar Dressing

In quantum electrodynamics, a dressed state [57, 77] is a modification of the quantum state of a charged hard particle which attaches a coherent state of soft photons to it. As well, it must be accompanied by a singular redefinition of the phases of the S matrix [58]. A similar idea can be used to obtain dressed states of gravitationally charged particles in quantum gravity when that theory is written as an effective field theory for perturbations of flat spacetime [78, 85]. In both electrodynamics and gravity, the dressing of states can be done in such a way that the S matrix that describes the scattering of hard dressed particles is infrared finite. In both cases, it has been argued that the soft photons or gravitons simply decouple from the dressed states [58, 86] in that their interactions are suppressed by factors of E_{res}/(hard particle scales) which can be very small. In the following, we will argue that the same dressing procedure can be implemented for hard particles which interact with soft scalar fields. In our simple model (3.5) it is a close parallel to the construction for photons or gravitons. We will reserve the discussoin of decoupling or non-decoupling for a later section.

Consider an incoming Fock space state of hard particles $|\{p\}\{k\}\rangle$. We will consider our quantum field theory with a fundamental infrared cutoff λ. Following

Chung and Faddeev and Kulish [57, 58], we define the dressed state, which we denote by $|\{p\}\{k\}\rangle\rangle$, as

$$|\{p\}\{k\}\rangle\rangle \equiv W(\{p\})\,|\{p\}\{k\}\rangle \tag{3.25}$$

where $W(\{p\})$ is the unitary operation implemented on an incoming state with fermion quantum numbers $\{p\}$ as

$$W(\{p\}) = \exp\left(R(\{p\})\right) \tag{3.26}$$

$$R(\{p\}) = \int_{\lambda}^{E_{res}} \frac{d^3k}{\sqrt{(2\pi)^3 2|\vec{k}|}} \left\{ \left[\sum_{p_n \in \{p\}} f(k,p_n) \right] a^\dagger(k) - \left[\sum_{p_n \in \{p\}} f^*(k,p_n) \right] a(k) \right\} \tag{3.27}$$

$$f(k,p) = \frac{gm}{k\cdot p}\,. \tag{3.28}$$

Note that $R(\{p\})$ is anti-Hermitian and $W(\{p\})$ is unitary. In this expression, λ is a fundamental infrared cutoff, and E_{res} is a second cutoff analogous to the detector resolution of the previous section, and we use the same symbol for it here [9].

Our normalization of the creation and annihilation operators is such that free field is

$$\phi_{\rm in}(x) = \int \frac{d^3k}{\sqrt{(2\pi)^3 2|\vec{k}|}} \left(e^{ikx} a(k) + e^{-ikx} a^\dagger(k) \right) \tag{3.29}$$

and

$$\left[a(k), a^\dagger(k') \right] = \delta^3(\vec{k} - \vec{k}')\,.$$

We have omitted the wave-function, $\frac{1}{\sqrt{(2\pi)^3 2|\vec{k}|}}$, for the scalar fields from our expressions for the S matrix. This means that we must compensate by taking the normalization integral for the states to be the Lorentz invariant measure $\int \frac{d^3k}{(2\pi)^3 2|\vec{k}|}$ when we finally sum over scalar field states.

In addition to the dressing of states described in equations (3.26)-(3.28), the S matrix must be modified. As an operator on Fock space, the modified S matrix is similar to the Dyson S matrix that is computed in Feynman-Dyson-Wick perturbation theory and which we used in the previous sections, the only difference is that it should be multiplied by some phases which take into account the infinite range of interactions. If we consider the transition between a dressed state $|\{p\}\{k\}\rangle\rangle$ and a

[9] Note that f defined in (3.28) is real. Thus f^* in (3.27) appears to over-complicate the expression. However, we prefer to write it this way to allow for any (possibly complex) subleading term in the definition of the dressed states, should we require it.

dressed state $|\{p'\}\{k'\}\rangle\rangle$, modified S matrix, which we shall denote by the symbol $\mathbf{S}$ is defined by

$$\mathbf{S}(\{p\}\{k\};\{p'\}\{k'\}) \;=\; \langle\langle\{p'\}\{k'\}|\;\mathbf{S}\;|\{p\}\{k\}\rangle\rangle \tag{3.30}$$

$$\equiv\; e^{i\Phi(\{p\})}\;\langle\langle\{p'\}\{k'\}|\;S^{\lambda}\;|\{p\}\{k\}\rangle\rangle\;e^{i\Phi(\{p'\})}\,; \tag{3.31}$$

$$\Phi(\{p\}) = -\frac{1}{8\pi}\sum_{\substack{p_n,p_m\in\{p\}\\ m\neq n}}\frac{g^2m^2}{\sqrt{(p_n\cdot p_m)^2-m^4}}\ln\frac{E_{res}}{\lambda}\,. \tag{3.32}$$

As we shall see, the infrared diverging phases serve to cancel the infrared divergent parts of the phases due to infrared divergent loop integrals encountered in the computation of S^{λ}.

3.6 Infrared Finiteness of the Dressed S Matrix

In this section, we shall show that matrix elements of the $\mathbf{S}$ matrix between dressed states are free of infrared divergences. The proof is very similar to the analogous one for quantum electrodynamics and for perturbative quantum gravity [57, 78]. Consider the Dyson S matrix operator S^{λ} computed in renormalized perturbation theory and the dressed states defined in equations (3.25)-(3.28). The matrix element of the $\mathbf{S}$ matrix between dressed states is given in equation (3.31) which we recopy here for the reader's convenience:

$$\begin{aligned}&\mathbf{S}(\{p\}\{k\};\{p'\}\{k'\})\\ &= e^{i\Phi(\{p\})}\;\langle\langle\{p'\}\{k'\}|\;S^{\lambda}\;|\{p\}\{k\}\rangle\rangle\;e^{i\Phi(\{p'\})}\,.\end{aligned} \tag{3.33}$$

The superscript λ on S^{λ} indicates that the matrix elements on the right-hand-side of the above equation are computed while using λ as a fundamental infrared cutoff. The dressed states $|\{p\}\{k\}\rangle\rangle$ and the phases $\Phi(\{p\})$ are also defined with this cutoff. Cancellation of this singular dependence on λ and finiteness as λ is put to zero on the right-hand-side of equation (3.33) is the "infrared finiteness" that we are seeking in this section.

We can use the Baker-Campbell-Hausdorff formula[10] to rewrite the exponential operators in the dressed states defined in equations (3.25)-(3.28) as

[10]For operators A and B with the properties $[A,[A,B]] = 0$ and $[B,[A,B]] = 0$, the Baker-Campbell-Hausdorff formula is

$$e^Ae^B = e^{\frac{1}{2}[A,B]}e^{A+B} = e^{[A,B]}e^Be^A$$

$$
\begin{aligned}
&|\{p'\}\{k'\}\rangle\rangle \\
&= \exp\Bigg\{\int_\lambda^{E_{res}} \frac{d^3\ell}{(2\pi)^3 2|\vec{\ell}|}\Bigg[-\frac{1}{2}\Bigg|\sum_{p'_n \text{in}\{p'\}} f(\ell,p'_n)\Bigg|^2 \\
&+ \sum_{p'_n \text{in}\{p'\}} f(\ell,p'_n)a^\dagger(\ell)\Bigg]\Bigg\}|\{p'\}\{k'\}\rangle\,, \qquad (3.34)
\end{aligned}
$$

$$
\begin{aligned}
&\langle\langle\{p\}\{k\}| \\
&= \langle\{p\}\{k\}|\exp\Bigg\{\int_\lambda^{E_{res}} \frac{d^3\ell}{(2\pi)^3 2|\vec{\ell}|} \\
&\times\Bigg[-\frac{1}{2}\Bigg|\sum_{p_n\in\{p\}} f(\ell,p_n)\Bigg|^2 - \sum_{p_n\in\{p\}} f^*(\ell,p_n)a(\ell)\Bigg]\Bigg\}\,. \qquad (3.35)
\end{aligned}
$$

Then, using these states, equation (3.33) becomes

$$
\begin{aligned}
&\mathbf{S}(\{p\}\{k\};\{p'\}\{k'\}) = \\
&\exp\Bigg\{i\Phi(\{p\}) + i\Phi(\{p'\}) + \int_\lambda^{E_{res}} \frac{d^3\ell}{(2\pi)^3 2|\vec{\ell}|}\Bigg[-\frac{1}{2}\Bigg|\sum_{p_n\text{in}\{p\}} f(\ell,p_n)\Bigg|^2 - \frac{1}{2}\Bigg|\sum_{p'_n\text{in}\{p'\}} f(\ell,p'_n)\Bigg|^2\ \Bigg]\Bigg\} \\
&\times\langle\{p\}\{k\}|\exp\Bigg\{-\int_\lambda^{E_{res}} \frac{d^3\ell}{\sqrt{(2\pi)^3 2|\vec{\ell}|}}\sum_{p_n\in\{p\}} f^*(\ell,p_n)a(\ell)\Bigg\}S^\lambda \\
&\times\exp\Bigg\{\int_\lambda^{E_{res}} \frac{d^3\ell}{\sqrt{(2\pi)^3 2|\vec{\ell}|}}\sum_{p'_m\in\{p'\}} f(\ell,p'_m)a^\dagger(\ell)\Bigg\}|\{p'\}\{k'\}\rangle\,. \qquad (3.36)
\end{aligned}
$$

A compact form for the LSZ formula for the scalar field part of the Dyson S matrix is obtained using a generating functional

$$
S^\lambda = :e^{\int dz\phi_{\text{in}}(z)(-\partial^2)\frac{\delta}{\delta J(z)}}: \ \ldots\mathcal{T}\ldots e^{i\int J\phi}\ldots\ldots\Big|_{J=0}
$$

where $\phi_{\text{in}}(x)$ is the asymptotic free field as in equation (3.29) and $:\ldots:$ denotes the normal ordering. With this formula, we can find the action of the dressing operator on S^λ as

$$
\begin{aligned}
&\langle\{p\}\{k\}|e^{-\int_\lambda^{Eres}\frac{d^3\ell}{\sqrt{(2\pi)^3 2|\vec{\ell}|}}\sum_{p_n\in\{p\}} f^*(\ell,p_n)a(\ell)} S^\lambda e^{\int_\lambda^{Eres}\frac{d^3\ell}{\sqrt{(2\pi)^3 2|\vec{\ell}|}}\sum_{p'_m\in\{p'\}} f(\ell,p'_m)a^\dagger(\ell)}|\{p'\}\{k'\}\rangle \\
&= \exp\Bigg\{-\int_\lambda^{E_{res}} \frac{d^3\ell}{(2\pi)^3 2|\vec{\ell}|}\sum_{p_n\in\{p\}} f^*(\ell,p_n)\sum_{p'_m\in\{p'\}} f(\ell,p'_m)\Bigg\}\exp\left(\int dz f(z)(-\partial^2)\frac{\delta}{\delta J(z)}\right) \\
&\times\langle\{p\}\{k\}|\ :e^{\int dz\phi_{\text{in}}(z)(-\partial^2)\frac{\delta}{\delta J(z)}}:\ \ldots\mathcal{T}\ldots e^{i\int J\phi}\ldots\ldots\ |\{p'\}\{k'\}\rangle\ \Big|_{J=0}
\end{aligned}
$$

In the above formula, we have used the equations

$$e^{-\int_\lambda^{E_{res}} \frac{d^3\ell}{\sqrt{(2\pi)^3 2|\vec{\ell}|}} \sum_{p_n\in\{p\}} f^*(\ell,p_n)a(\ell)} a^\dagger(k)$$
$$= \left(a^\dagger(k) - \sum_{p_n\in\{p\}} f^*(p_n,k)\right) e^{-\int_\lambda^{E_{res}} \frac{d^3\ell}{\sqrt{(2\pi)^3 2|\vec{\ell}|}} \sum_{p_n\in\{p\}} f^*(\ell,p_n)a(\ell)}, \tag{3.37}$$

$$a(k)e^{\int_\lambda^{E_{res}} \frac{d^3\ell}{\sqrt{(2\pi)^3 2|\vec{\ell}|}} \sum_{p'_m\in\{p'\}} f(\ell,p'_m)a^\dagger(\ell)}$$
$$= e^{\int_\lambda^{E_{res}} \frac{d^3\ell}{\sqrt{(2\pi)^3 2|\vec{\ell}|}} \sum_{p'_m\in\{p'\}} f(\ell,p'_m)a^\dagger(\ell)} \left(a(k) + \sum_{p'_n\in\{p'\}} f(k,p'_n)\right), \tag{3.38}$$

and the result is the classical scalar field,

$$f(z) = \int_\lambda^{E_{res}} \frac{d^3\ell}{(2\pi)^3 2|\vec{\ell}|} \left(\sum_{p'_n\in\{p'\}} f(\ell,p'_n)e^{i\ell z} - \sum_{p_m\in\{p\}} f^*(\ell,p_m)e^{-i\ell z}\right)$$

occurring in the exponential functional derivative operator in the second line,

$$\exp\left(\int dz f(z)(-\partial^2)\frac{\delta}{\delta J(z)}\right).$$

This operation inserts ingoing and outgoing soft scalars with wave-functions $\sim f, -f^*$ into the S matrix element that it operates on. We can use the soft scalar theorem (3.13) to re-write these as the S matrix element without the soft scalars and with an exponential factor,

$$\exp\Bigg\{\int_\lambda^{E_{res}} \frac{d^3\ell}{(2\pi)^3 2|\vec{\ell}|} \left[\sum_{p_n\in\{p\}} \frac{gm}{p_n\cdot\ell} - \sum_{p'_n\in\{p'\}} \frac{gm}{p'_n\cdot\ell}\right] \times$$
$$\times \left[\sum_{p_m\in\{p\}} f(\ell,p_m) - \sum_{p'_m\in\{p'\}} f^*(-\ell,p'_m)\right]\Bigg\}$$

which then leads us to

$$\mathbf{S}(\{p\}\{k\};\{p'\}\{k'\}) = S^\lambda(\{p\}\{k\};\{p'\}\{k'\})\; e^{i\Phi(\{p\})+i\Phi(\{p'\})}$$
$$\times \exp\Bigg\{\int_\lambda^{E_{res}} \frac{d^3\ell}{(2\pi)^3 2|\vec{\ell}|} \Bigg[-\frac{1}{2}\left|\sum_{p'_n\in\{p\}} f(\ell,p_n)\right|^2$$
$$-\frac{1}{2}\left|\sum_{p'_n\in\{p'\}} f(\ell,p'_n)\right|^2 - \sum_{p_n\in\{p\}} f^*(\ell,p_n)\sum_{p'_m\in\{p'\}} f(\ell,p'_m)$$
$$+\left[\sum_{p_n\in\{p\}} \frac{gm}{p_n\cdot\ell} - \sum_{p'_n\in\{p'\}} \frac{gm}{p'_n\cdot\ell}\right]\left[\sum_{p_m\in\{p\}} f(\ell,p_m) - \sum_{p'_m\in\{p'\}} f^*(-\ell,p'_m)\right]\Bigg]\Bigg\} \tag{3.39}$$

Then, we can use equations (3.14) to write the above equation as

$$\begin{aligned}
&\mathbf{S}^{\lambda}(\{p\}\{k\};\{p'\}\{k'\}) = S^{\Lambda}(\{p\}\{k\};\{p'\}\{k'\})\; e^{i\Phi(\{p\})-i\bar{\Phi}(\{p\})+i\Phi(\{p'\})-i\bar{\Phi}(\{p'\})} \\
&\times \exp\Bigg\{\int_{\lambda}^{E_{res}} \frac{d^3\ell}{(2\pi)^3 2|\vec{\ell}|}\Bigg[-\frac{1}{2}\left|\sum_{p'_n\in\{p\}} f(\ell,p_n)\right|^2 - \frac{1}{2}\left|\sum_{p'_n\in\{p'\}} f(\ell,p'_n)\right|^2 \\
&- \sum_{p_n\in\{p\}} f^*(\ell,p_n) \sum_{p'_m\in\{p'\}} f(\ell,p'_m) \\
&+ \left[\sum_{p_n\in\{p\}} \frac{gm}{p_n\cdot\ell} - \sum_{p'_n\in\{p'\}} \frac{gm}{p'_n\cdot\ell}\right]\left[\sum_{p_m\in\{p\}} f(\ell,p_m) - \sum_{p'_m\in\{p'\}} f^*(-\ell,p'_m)\right] \\
&+ \int_{\lambda}^{\Lambda} \frac{d^3\ell}{(2\pi)^3 2|\vec{\ell}|}\Bigg[-\frac{1}{2}\left(\sum_{p_n\in\{p\}} \frac{gm}{p_n\cdot\ell}\right)^2 - \frac{1}{2}\left(\sum_{p'_n\in\{p'\}} \frac{gm}{p'_n\cdot\ell}\right)^2 \\
&+ \sum_{p_n\in\{p\}} \frac{gm}{p_n\cdot\ell} \sum_{p'_m\in\{p'\}} \frac{gm}{p'_m\cdot\ell}\Bigg]\Bigg\}
\end{aligned} \tag{3.40}$$

The first line in equation (3.40) contains the phases that come from the Faddeev-Kulish prescription plus from the internal loop contributions to the infrared cutoff S^{λ}. The logarithmic infrared divergences cancel in the combinations in which the phases appear there, leaving behind finite parts which we will display shortly. The second and third lines in equation (3.40) comes from the normalizations and overlaps of the coherent states. The fourth line in equation (3.40) contains the result of using the soft scalar theorem to take into account the soft scalars coming from the coherent states. The last line in equation (3.40) is the contribution of internal loops encountered in the computation of S^{λ}. It is easy to see that the λ-dependence of the sum of all of these terms cancels when we put $f(p,\ell) = \frac{gm}{p\cdot\ell}$. The latter exponential factors simplify as

$$\begin{aligned}
&\exp\Bigg\{\int_{E_{res}}^{\Lambda} \frac{d^3\ell}{(2\pi)^3 2|\vec{\ell}|}\Bigg[-\frac{1}{2}\left(\sum_{p_n\in\{p\}} \frac{gm}{p_n\cdot\ell}\right)^2 \\
&- \frac{1}{2}\left(\sum_{p'_n\in\{p'\}} \frac{gm}{p'_n\cdot\ell}\right)^2 + \sum_{p_n\in\{p\}} \frac{gm}{p_n\cdot\ell} \sum_{p'_m\in\{p'\}} \frac{gm}{p'_m\cdot\ell}\Bigg]\Bigg\} \\
&= \left(\frac{E_{\text{res}}}{\Lambda}\right)^{A(\{p\}\{p'\})}
\end{aligned}$$

where

$$A(\{p\}\{p'\}) = \frac{1}{8\pi^2}\sum_{nm}\frac{g^2m^2\eta_m\eta_n}{\sqrt{(p_m\cdot p_n)^2-m^4}}\left[i\left(\frac{1+\eta_m\eta_n}{2}\right)\right.$$
$$\left.+\ln\frac{1+\sqrt{1-\frac{m^4}{(p_m\cdot p_n)^2}}}{1-\sqrt{1-\frac{m^4}{(p_m\cdot p_n)^2}}}\right].$$

The final result for the element of the S matrix in dressed states is simply

$$\mathbf{S}(\{p\}\{k\};\{p'\}\{k'\})$$
$$=\left(\frac{E_{\rm res}}{\Lambda}\right)^{A(\{p\}\{p'\})}S^{\Lambda}(\{p\}\{k\};\{p'\}\{k'\})\,. \tag{3.41}$$

The left-hand-side of the above formula (3.41) is the matrix element of the modified S matrix, $\mathbf{S}$, computed with dressed states. It is the amplitude for the transition from dressed state $|\{p\}\{k\}\rangle\rangle$ to dressed state $|\{p'\}\{k'\}\rangle\rangle$. The right-hand-side simply contains the Dyson S matrix, S^{λ}, its matrix elements $S^{\Lambda}(\{p\}\{k\};\{p'\}\{k'\})$ computed with undressed Fock space states and with an infrared cutoff Λ. It is multiplied the factor $\left(\frac{E_{\rm res}}{\Lambda}\right)^{A(\{p\}\{p'\})}$ made from the ratio of cutoffs raised to a complex, momentum-dependent exponent. The right-hand-side does not depend on Λ in that the Λ-dependence of S^{Λ} is compensated by the Λ dependence of the factor. However, the right-hand-side does depend on the cutoff, E_{res} which is now a parameter of the theory.

3.7 Non-decoupling of Soft Scalar Emission

Having constructed the infrared finite S matrix for dressed states, we are now interested in computing the amplitude of emission of additional soft scalars beyond those in the dressing. The vanishing of such amplitudes and the factorization of the S matrix into hard and soft sectors for the scattering amplitudes of dressed states is an already well-known feature of quantum electrodynamics and perturbative quantum gravity [58, 61, 85, 86]. There, the factorization has to do with the fact that one could correct the dressing factors to take into account the next-to-leading contributions to the soft theorem. Then amplitude for the interaction of suitably dressed states to emit or absorb an additional soft photon or soft graviton is indeed suppressed by powers of the detector resolution cutoff E_{res}. This decoupling also has sound physical reasoning. Photons or gravitons with wavelengths the size of the solar system should have nothing to do with electrodynamic or gravitational physics at a subatomic scale.

As we shall see, the case of the massless scalar field is a little different. We will find that soft scalars do not decouple. A scattering event for hard dressed particles

can produce soft scalars with an amplitude of the same order as other radiative corrections to the amplitude and at the same order as the interactions between the hard particles themselves. We will outline the argument for this in the following. Later, in the next section, we will examine the entanglement of the dressed hard particles and the soft particles that are produced in the scattering of hard particles.

Let us consider an outgoing dressed state which contains an additional soft scalar

$$|\{p'\}\{k'\}q'\rangle\rangle = e^{R(\{p'\})}a^\dagger(q')|\{p'\}\{k'\}\rangle \tag{3.42}$$

We would like to compute the dressed S matrix element $\mathbf{S}(\{p\}\{k\};\{p'\}\{k'\}q')$ which is defined as the quantity

$$\begin{aligned}&\mathbf{S}(\{p\}\{k\};\{p'\}\{k'\}q') = e^{i\Phi(\{p\})}\times\\&\langle\{p\}\{k\}|W^\dagger(\{p\})S^\lambda W(\{p'\})a^\dagger(q')|\{p'\}\{k'\}\rangle e^{i\Phi(\{p'\})}\end{aligned} \tag{3.43}$$

We can move the scalar creation operator past the operator $W(\{p'\})$ by using the identity (3.37) to get

$$\begin{aligned}&\mathbf{S}(\{p\}\{k\};\{p'\}\{k'\}q')\\&= e^{i\Phi(\{p\})+i\Phi(\{p'\})}\langle\{p\}\{k\}|W^\dagger(\{p\})S^\lambda\\&\times\left(a^\dagger(q') - \sum_{p'_n\in\{p'\}} f(p'_n,q')\right)W(\{p'\}))|\{p'\}\{k'\}\rangle\end{aligned} \tag{3.44}$$

Then, the $a^\dagger(q')$ can either act on $W^\dagger(\{p\})$, for which we use equation (3.38) or it could be absorbed into the S matrix, in which case we can use the soft scalar theorem (3.11), which we will keep to next-to-leading order. The result is a cancellation between the singular factors and the f's, leaving only the contributions from the next-to-leading soft scalar theorem,

$$\begin{aligned}\mathbf{S}(\{p\}\{k\};\{p'\}\{k'\}q') = &\left(\frac{E_{\rm res}}{\Lambda}\right)^{A(\{p\}\{p'\})}\\&\times\left\{-\sum_{p_n\in\{p\}\{p'\}}\left(\frac{g}{2m}\right) - g\hat{\partial}_m\right\}S^\Lambda(\{p\}\{k\};\{p'\}\{k'\})\end{aligned} \tag{3.45}$$

where quantities on the right-had-side are defined in the discussions around equation (3.11)-(3.12) and (3.41).

This equation is one of our central results. The right-hand-side, unlike what occurs for photons or gravitons, cannot be written in a form that depends only on the initial and final states in a way that its effect can be absorbed into $W(\{p'\})$ or $W^\dagger(\{p\})$. Even absorbing the first term with $-\frac{g}{2m}$ for each external fermion would modify W in such a way that it is no longer unitary. It is easy to confirm by a simple

tree level computation that $-g\hat{\partial}_m$ operating on an amputated correlation function simply cannot in general be written as something that depends only on $\{p\}\{k\}$ plus something that depends only on $\{p'\}\{k'\}$ and its action therefore cannot be absorbed into the W's either. Our conclusion is that the soft scalars couple to the dressed S matrix at order g in the coupling constant, which makes them just as coupled as the other particles.

3.8 Discussions

We have shown that, in the context of our admittedly rather specialized model, the infrared problem due to massless scalar fields is practically identical to that for photons in quantum electrodynamics and for gravitons in perturbative quantum gravity. This is in spite of the fact that there are no apparent internal symmetries, either continuous or discrete, whatsoever. This means that there are no conserved Noether currents beyond the energy-momentum tensor, no Ward-Takahashi identities and no apparent asymptotic symmetries. We have found what we conjecture is different in the soft scalar theory. The difference lies in the next-to-leading soft scalar theorem. Unlike the case of photons or gravitons where Ward-Takahashi identities help to write those terms as referring only to the initial and final states [85], their scalar analog cannot be written that way. Then, unlike for photons and gravitons, the next-to-leading behaviour cannot be absorbed by modifying the hard particle dressing. Soft scalars are still produced when dressed hard particles interact.

Now that we have demonstrated that soft scalars do not decouple from the interactions of dressed states, we can revisit the question as to whether they carry any significant amount of information. We still expect the scenario where the interactions of hard particles in a scattering event also produces a cloud of soft particles and those soft particles are undetectable. Our experimental resources only have access to the dressed hard particles. To proceed, we could simply ask the same question that we did for scattering of undressed hard particles. We consider an incoming density matrix state

$$\sum_{\substack{\{p\}\{k\}\\ \{\tilde{p}\}\{\tilde{k}\}}} |\{p\}\{k\}\rangle\rangle \rho_{\{p\}\{k\};\{\tilde{p}\}\{\tilde{k}\}} \langle\langle\{\tilde{p}\}\{\tilde{k}\}| \tag{3.46}$$

which should evolve to an outgoing density matrix

$$\begin{aligned}\sum_{\substack{\{p\}\{k\}\\ \{\tilde{p}\}\{\tilde{k}\}}} \sum_{\substack{\{p'\}\{k'\}\{q'\}\\ \{\tilde{p}'\}\{\tilde{k}'\}\{\tilde{q}'\}}} & |\{p'\}\{k'\}\{q'\}\rangle\rangle\langle\langle\{\tilde{p}'\}\{\tilde{k}'\}\{\tilde{q}'\}| \\ & \times \mathbf{S}^\dagger(\{p'\}\{k'\}\{q\};\{p\}\{k\})\ \rho_{\{p\}\{k\};\{\tilde{p}\}\{\tilde{k}\}} \\ & \times \mathbf{S}(\{\tilde{p}\}\{\tilde{k}\};\{\tilde{p}'\}\{\tilde{k}'\}\{q'\})\end{aligned} \tag{3.47}$$

which we now trace over the outgoing soft scalars to obtain a reduced density matrix which describes all of the accessible physics of the hard particles in the outgoing state,

$$\begin{aligned}\rho_{\text{final}} = & \sum_{\substack{\{p\}\{k\}\\ \{\tilde{p}\}\{\tilde{k}\}}} \sum_{\substack{\{p'\}\{k'\}\\ \{\tilde{p}'\}\{\tilde{k}'\}}} |\{p'\}\{k'\}\rangle\rangle\langle\langle\{\tilde{p}'\}\{\tilde{k}'\}| \\ & \times \sum_{\{q\}} \mathbf{S}^\dagger(\{p'\}\{k'\}\{q\};\{p\}\{k\})\ \rho_{\{p\}\{k\};\{\tilde{p}\}\{\tilde{k}\}} \\ & \mathbf{S}(\{\tilde{p}\}\{\tilde{k}\};\{\tilde{p}'\}\{\tilde{k}'\}\{q\})\,. \end{aligned} \tag{3.48}$$

Then, we observe that the trace over the soft scalars involves

$$\begin{aligned}&\sum_{\{q\}} \mathbf{S}^\dagger(\{p'\}\{k'\}\{q\};\{p\}\{k\})\ \rho_{\{p\}\{k\};\{\tilde{p}\}\{\tilde{k}\}}\ \mathbf{S}(\{\tilde{p}\}\{\tilde{k}\};\{\tilde{p}'\}\{\tilde{k}'\}\{q\}) \\ &= \mathbf{S}^\dagger(\{p'\}\{k'\};\{p\}\{k\})\ \rho_{\{p\}\{k\};\{\tilde{p}\}\{\tilde{k}\}}\ \mathbf{S}(\{\tilde{p}\}\{\tilde{k}\};\{\tilde{p}'\}\{\tilde{k}'\}) \\ &+ \int_\lambda^{E_{res}} \frac{d^3q'}{(2\pi)^3 2|\vec{q}'|} \mathbf{S}^\dagger(\{p'\}\{k'\}q';\{p\}\{k\})\ \rho_{\{p\}\{k\};\{\tilde{p}\}\{\tilde{k}\}}\ \mathbf{S}(\{\tilde{p}\}\{\tilde{k}\};\{\tilde{p}'\}\{\tilde{k}'\}q') \\ &+ \dots \end{aligned} \tag{3.49}$$

Since the combination

$$\begin{aligned}&\mathbf{S}^\dagger(\{p'\}\{k'\}q';\{p\}\{k\})\rho_{\{p\}\{k\};\{\tilde{p}\}\{\tilde{k}\}}\mathbf{S}(\{\tilde{p}\}\{\tilde{k}\};\{\tilde{p}'\}\{\tilde{k}'\}q') \\ &\sim\ (q')^0\end{aligned}$$

is not singular as $q \to 0$, it goes like a constant, the integration in the last line of the above equation produces a factor of E_{res}^2 and the higher order terms represented by the ellipses there have multiple volume integrals which produce higher orders of E_{res}. All of these are suppressed. In the approximation where we neglect contributions with positive powers of E_{res}, we neglect such terms. Then equation (3.48) becomes

$$\begin{aligned}\rho_{\text{final}} = & \sum_{\substack{\{p\}\{k\}\\ \{\tilde{p}\}\{\tilde{k}\}}} \sum_{\substack{\{p'\}\{k'\}\\ \{\tilde{p}'\}\{\tilde{k}'\}}} |\{p'\}\{k'\}\rangle\rangle\langle\langle\{\tilde{p}'\}\{\tilde{k}'\}| \\ & \times \mathbf{S}^\dagger(\{p'\}\{k'\};\{p\}\{k\})\ \rho_{\{p\}\{k\};\{\tilde{p}\}\{\tilde{k}\}} \\ & \times \mathbf{S}(\{\tilde{p}\}\{\tilde{k}\};\{\tilde{p}'\}\{\tilde{k}'\}) + \mathcal{O}(E_{res}^2) \end{aligned} \tag{3.50}$$

with no reference to soft particles at all. This is the sense in which the information contained in the soft particles decouples. We could produce soft scalars in a scattering experiment. However, if we do not have the detector resolution to see them directly, we have no way of knowing that they are there. It will also be interesting

to understand the implications of our result in the context of Hawking, Perry, and Strominger's proposal [87, 88] of the resolution of the black hole information paradox which has been heavily criticized based on the decoupling of soft gravitons in the dressed formalism [86].

We should note that recent work on the memory effect [74] suggests that, as well as the well-known cases of massless QED and Yang-Mills theory, Feddeev-Kulish like dressings may not work for the full nonlinear diffeomorphism invariant quantum gravity. Perhaps studying this issue in the simpler context of trivalently coupled massless scalar fields could shed some light on these complex issues.

Chapter 4

Classical Observables from Causal response Functions

In this chapter, we will discuss how the soft limit can be used to calculate classical observables from a quantum field theory. The calculation of classical results from quantum amplitudes has recently received a lot of attention in recent years, particularly due to its application in gravitational wave astronomy. In section 4.1, we review some background materials that would be helpful to understand a new technique to calculate classical observables from causal response functions that is presented in the subsequent sections.

4.1 Background materials

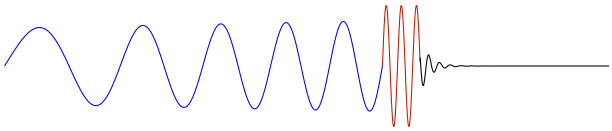
The first experimental detection of gravitational waveforms from the merger of two black holes by LIGO-Virgo collaboration [10] in 2015 marks a new era of gravitational wave (GW) astronomy. The output signal of a gravitational wave detector $s(t)$ is in the form [89]

$$s(t) = h(t, \boldsymbol{\lambda}) + n(t) \tag{4.1}$$

where $h(t, \boldsymbol{\lambda})$ represents a deterministic gravitational wave (GW) signal, which depends on various parameters of the source of the wave (such as the mass, spin, size etc) collectively denoted by $\boldsymbol{\lambda}$, and $n(t)$ is noise. In general, the actual GW signal $h(t, \boldsymbol{\lambda})$ is buried deep in the noise and the first step of GW data analysis is to determine whether there is any true GW signal is present in the observed data. To do this, it is extremely essential to have a bank of templates, where each template is a theoretical prediction of the waveform $h(t, \boldsymbol{\lambda}_i)$ with a set of parameters $\boldsymbol{\lambda}_i$ for a potential source. The number of templates in the bank is chosen to ensure a sufficiently dense coverage of the parameter space [89]. Having a bank of templates in hand, one can then use various statistical methods to extract true GW information from the observed data. To analyze the GW detector data, we must, therefore, have a template or prediction for the waveform of the system.

Finding templates using GR

A typical gravitational waveform during a binary black hole or neutron star coalescence is shown in Fig 4.1. To find the waveforms, we need to solve Einstein's



(Inspiral) (Merger) (Ringdown)

Figure 4.1: A typical shape of the gravitational waveform of a binary coalescence is shown. The inspiral phase describes the stage where two compact objects, such as black holes or neutron stars, orbit each other while steadily losing energy through the emission of gravitational waves. As this energy loss causes the objects to spiral closer together, they eventually enter the merger phase, where they combine. Following the merger, the newly formed object dissipates the excess energy, a process known as the ringdown phase.

equation

$$R_{\mu\nu} - \frac{1}{2} g_{\mu\nu} = \frac{8\pi G}{c^4} T_{\mu\nu} \tag{4.2}$$

where $g_{\mu\nu}$ is the metric,$R_{\mu\nu}$ is the Ricci tensor, and $T_{\mu\nu}$ is the source. The gravitational waves are found by writing the metric as a perturbation around the flat metric as

$$g_{\mu\nu} = \eta_{\mu\nu} + h_{\mu\nu} \tag{4.3}$$

However, solving the Einstein's equations directly to find the waveform for a binary coalescence is an extremely hard problem and we must resort to some approximate schemes. Common methods using classical general relativity (GR) to find classical gravitational observables such as the waveform are the following.

1 *Post-Newtonian (PN) method:* In this method, one expands a gravitational observable $\mathbb{O}$ such as the waveform in powers of $\frac{v^2}{c^2}$ where v is the relative velocity of the binaries.

$$\mathbb{O}^{PN} = \mathcal{O}\left(\frac{v^2}{c^2}\right) + \mathcal{O}\left(\frac{v^4}{c^4}\right) + ...$$

Additionally, this method assumes weak gravity.

2. *Post-Minkowskian (PM) method:* In this method, the expansion parameter is the gravitational constant G, i.e.

$$\mathbb{O}^{PM} = \mathcal{O}\left(G\right) + \mathcal{O}\left(G^2\right) + ...$$

PM method also assumes weak gravity.

3. *Self-force (SF) method:* Here one considers the expansion in terms of the mass ratio m_L/m_H where m_L and m_H are the lighter and heavier objects of the binary. In this method, one assumes that the lighter object moves in the Schwartzchild metric sourced by the heavier object. This is one of the key methods that is able to deal with strong gravity.

4. *EOB method:* The effective of body (EOB) approach [90] is an analytic method that can be regarded as the relativistic generalization of mapping a two-body problem to an one-body problem. This method is particularly successful in describing various phases of the binary coalescence.

5. *Numerical relativity (NR):* This is the method of solving Einstein's equation numerically. This method is the most accurate method of computing gravitational observables but it is computationally very expensive.

In describing the binary coalescence, the post-Newtonian (PN) method proves highly effective during the early inspiral phase, as the velocity v remains relatively low. However, as v increases, the PN approximation loses accuracy, and PM methods become more reliable. Despite this, both PN and PM assume the weak-field limit, which breaks down as the two objects approach merger. At this stage, numerical relativity (NR) is required, but NR simulations are computationally demanding and require significant computational resources to cover the full range of source parameters $\boldsymbol{\lambda}$. Furthermore, in the ringdown phase black hole perturbation theory is needed to be used. In practice, constructing a complete waveform template necessitates combining multiple methods, leveraging their overlapping regions and integrating hybrid and numerical approaches to achieve high precision.

The need for precision

All of the above analytical methods, including other hybrid analytic methods [90] based on general relativity, are quite computationally involved. On the other hand, the future GW detectors such as Einstein telescope [91], Cosmic explorer [92], LISA [93] as well as the future runs of LIGO-Virgo-KARGA are expected to make measurements within the next decade with at least two orders of magnitude in precision compared to the current measurements. They will be capable of detecting binary black hole systems with masses up to $\sim 1000 M_\odot$ where $M_\odot$ is the solar mass. These massive systems, currently beyond the reach of existing detectors, produce lower-frequency gravitational waves during the inspiral phase. At the theoretical front, this means that we need to significantly improve our ability to calculate gravitational observables at all phases of the merger. For example, perturbative calculations in G are expected to need to be improved by at least two orders compared to what we currently have, which translates to the need for $\mathcal{O}(G^6) - \mathcal{O}(G^7)$ results for the

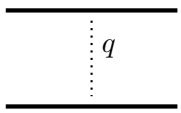
inspiral phase [11]. At the same time, we need to develop novel methods to describe various stages of coalescence where the conventional analytic methods fall short.

Gravitational wave physics from QFT amplitudes

In recent years, remarkable progress has been made in computing *classical* gravitational observables using quantum field theory (QFT) methods. Before reviewing some of these advancements, it is important to emphasize that our focus is on deriving *classical* observables from the effective field theory of gravity coupled with matter fields. This puts us in a much more comfortable position, given that a fully formulated quantum field theory of gravity remains elusive as discussed in Chapter 1. We will explore QFT-based techniques that enable us to efficiently take the classical limit when $\hbar \to 0$, where quantum effects disappear, leaving behind purely classical results. Compared to traditional general relativistic methods, these alternative approaches provide a more suitable framework for the precise computation of gravitational observables.

The idea of calculating classical quantities from scattering amplitudes is a familiar one. For example, the potential $V(\boldsymbol{b})$ between two massive objects interacting via a massless exchange (such as photon or graviton) with momentum $q^\mu = (q^0, \boldsymbol{q})$ can be found from the Fourier transform of the 4-point tree amplitude

$$V(\boldsymbol{b}) \sim \int d^3\boldsymbol{q} e^{-i\boldsymbol{q}\cdot\boldsymbol{b}} \times \quad \propto \frac{1}{\boldsymbol{b}} \tag{4.4}$$



where we assumed that the massless propagator is of the form q^{-2}. It should be of no surprise that the same calculation for two massive particles with mass m_1, m_2 interacting via graviton will give the Newton's formula of gravitational potential

$$V(\boldsymbol{b}) = -\frac{Gm_1m_2}{\boldsymbol{b}}.$$

In the above formula, G comes from two powers of the gravitational coupling constant $\sqrt{16\pi G}$ as required by the tree diagram. However, this formula receives correction in GR and take the form [94]

$$V(\boldsymbol{b}) = -\frac{Gm_1m_2}{\boldsymbol{b}}\left(1 + \alpha\frac{G(m_1+m_2)}{c^2\boldsymbol{b}}\right) + \mathcal{O}(G^3) \tag{4.5}$$

where α is a constant not important for our discussion. We again emphasize that (4.5) is derived using general relativity and is a purely classical result.

Can we derive the $\mathcal{O}(G^2)$ and possibly higher order terms in (4.5) from a classical limit of an EFT of gravity? In fact, can we develop a EFT approach the aligns

[11] An analogous calculation within an effective field theory (EFT) of gravity would involve evaluating five- and six-loop Feynman integrals with graviton interactions.

with the PM results discussed earlier? In such an EFT, we do computations perturbatively in G so it seems plausible. However, at first glance, there seems to be a problem, as $\mathcal{O}(G^2)$ and higher terms would originate from loop diagrams. This seems to contradict the common notion that "loops are quantum." Nevertheless, a careful analysis reveals a one-loop calculation in EFT gives [95]

$$V(\boldsymbol{b}) = -\frac{Gm_1m_2}{\boldsymbol{b}}\left(1+\alpha\frac{G(m_1+m_2)}{c^2\boldsymbol{b}}+\beta\frac{G\hbar}{c^3\boldsymbol{b}^2}\right)+\mathcal{O}(G^3) \tag{4.6}$$

where β is a constant. Indeed (4.6) is equal to (4.5) in the $\hbar \to 0$ limit.

The main takeaway here is that loop corrections contain classical pieces when both massive and massless particles are involved[12]. Coming back to our the binary interaction problem, we need a method within perturbative quantum field theory (QFT) that systematically separates classical and quantum contributions. Indeed, tackling the classical relativistic two-body problem from the classical limit of quantum field theory amplitudes has proven very effective [98–114] and complemented classical field theory methods [90, 115, 116]. In a seminal paper [117], Kosower, Maybee and O'Connell (KMOC) developed a formalism for computing classical observables from quantum scattering amplitudes in the $\hbar \to 0$ limit. The main advantage of the scattering-amplitude-based method is that it can borrow various powerful tools like the double-copy [118, 119], unitarity methods [120, 121], and modern techniques for computing Feynman integrals [122–126], originally developed in the context of collider physics to compute classical gravitational quantities.

In the KMOC framework, one aims to compute various classical observables such as linear and angular impulses of a particle, or the radiative waveform itself, perturbatively in the gravitational coupling constant. Such classical observables are defined from classical limits of changes in the expectation value of an operator $\mathbb{O}$ in the quantum theory

$$\Delta\mathbb{O} = \langle \text{in}|\, \mathbb{O}^{\text{out}}\, |\text{in}\rangle - \langle \text{in}|\, \mathbb{O}^{\text{in}}\, |\text{in}\rangle \ , \tag{4.7}$$

where objects with *in* and *out* labels are defined at the asymptotic past and future infinity respectively, and $|\text{in}\rangle$ are two-particle states with classically-peaked wavepackets. Such observables are *in-in* quantities by definition, but in the KMOC framework these are expressed in terms of *in-out* scattering amplitudes, using the relation

$$\mathbb{O}^{\text{out}} - \mathbb{O}^{\text{in}} = S^\dagger \mathbb{O}^{\text{in}} S - \mathbb{O}^{\text{in}} = i[\mathbb{O}^{\text{in}}, T] + T^\dagger[\mathbb{O}^{\text{in}}, T]\,, \tag{4.8}$$

where $S = 1 + iT$ is the familiar unitary S-matrix. We review the KMOC formalism in more detail later in the chapter.

Let us now understand why a clean separation of quantum and classical pieces is possible for the two-body problem. We have two massive objects interacting via graviton. There are three length scales in the problem (we are working in a unit $c = 1$):

[12] This is a relatively old result, see [95–97]

1. The Compton length $l_c = \frac{\hbar}{M}, \quad M \sim (m_1 + m_2)$
2. The Schwarzschild radius $l_s = GM$
3. The impact parameter[13] $b = \sqrt{-b^2}$

We impose a hierarchy of scale

$$l_c \ll l_s \ll b\,. \tag{4.9}$$

In this limit, the two particles remain sufficiently distant, ensuring that the exchanged gravitons are soft and quantum effects are effectively suppressed. This also ensures that the wavepackets associated with the massive particles do not grow sufficiently to affect the results [117]. Defining the conjugate momentum $|q|$ such that

$$|q| \sim \frac{\hbar}{|b|}$$

we have,

$$\frac{q}{M} \ll \frac{GMq}{\hbar} \ll 1 \tag{4.10}$$

In the limit $q = \hbar\bar{q}$ where $\bar{q}$ is the wavenumber, factors of $\frac{\hbar\bar{q}}{m}$ keeps track of the quantum pieces of the amplitude. The factor $\frac{GMq}{\hbar} = GM\bar{q}$ does not change the $\hbar$ scaling and keeps track of the perturbative corrections. In the computation of an observable, the leading order classical results come from the tree amplitude which is $1/\bar{q}^2$. Then the one-loop, two-loop and higher order classical results are encoded in $1/\bar{q},\ \bar{q}\ln|\bar{q}|, ...$ terms of the amplitudes. The analytic terms $\bar{q}^n$ with $n > 0$ includes the quantum results [14]. Moreover, one expands the matter propagators according to the method of regions [127] and adopts the same $\hbar$ scaling for the loop momenta $\ell \sim q$ for the massless particle (graviton). Then by power counting its a simple exercise to show that the UV divergent diagrams, such as the self- energy diagrams, vertex corrections are all quantum since they go like $(\hbar\bar{q})^n,\ n > 0$ and therefore do not contribute in the classical limit $\hbar \to 0$[103, 117]. Notice that it is sufficient to keep track of q, so we set $\hbar = 1$ form now on. This analysis applies to any theory where such a meaningful separation of scale exists such as scalar QED where the scales are Compton wavelength of the massive particles, classical radius of the electron, and the impact parameter. All the other conclusions in the above discussion remain unchanged.

In this chapter, we develop a new method to compute classical observable form QFT. The later sections of this chapter explore the motivation, outcomes, and advantages of this method over the KMOC approach.

[13] We are working in the mostly negative signature.

[14] Their Fourier transform in the b space gives delta function and its derivatives ensuring that these are contact terms

4.2 Introduction

In the KMOC formalism, the UV divergent amplitudes do not contribute to the classical results. However, delicate manipulations are required to manifest the cancellation of infrared singular contributions in the classical limit between different terms in Eq. (4.8) [103, 117]. Furthermore, despite classical observables being causal quantities by definition, when expressed in terms of the in-out scattering amplitudes, causality is hidden in the intermediate steps of the computation.

An alternative way of computing these observables is to directly use the *in-in* or *Schwinger-Keldysh* formalism [128, 129], which is commonly utilized in worldline-based approaches to the two-body problem [130–138]. In field theory, the relation between asymptotic observables and generalized scattering amplitudes has been discussed recently in [139], where it was shown that a class of in-in *response functions* [140, 141], computes KMOC-like expectation values such as the classical gravitational waveform. In this chapter, we further explore this connection and show that other observables can also be computed from appropriate soft limits of response functions. We explain how the KMOC formula for other observables, such as the linear impulse and radiated momentum, $\Delta\mathbb{P}$, can be derived from a causal response function.

A particularly interesting classical gravitational observable is the angular momentum loss for the binary system, $\Delta\mathbb{J}$. A position space calculation by Damour [142] showed that the angular momentum loss in General Relativity begins at $\mathcal{O}\left(G^2\right)$, whereas the energy loss starts at $\mathcal{O}\left(G^3\right)$ [102]. This result was reproduced in momentum space using Weinberg's soft theorem [143, 144]. However, the naive application of the KMOC framework produces an ambiguous result, as it requires the inclusion of a unitary cut involving an on-shell three-point amplitude, which is zero for real kinematics, times the singular soft-limit of a five-point amplitude. In this chapter, we will show that the soft limit of the response function unveils new integration regions thereby enabling us to present an unambiguous calculation of the radiated angular momentum, within the KMOC framework.

Response functions are causal observables. Does manifest causality simplify the calculations of classical observables from amplitudes? In this chapter, we find an affirmative answer to this question by computing classical observables using causal Feynman rules, i.e., amplitudes with a retarded $i\epsilon$ prescription, which appear naturally within the in-in formalism in the so-called Keldysh basis [116, 141].

We show that the causal representation manifests various properties of certain classical observables, such as the linear impulse, the scattering waveform and its variance. In particular, we find that the retarded $i\epsilon$ prescription ensures the manifest cancellation of $\mathcal{O}(1/\hbar^n)$ terms within classical observables, as well as the vanishing of the classical variance of the scattering waveform.

The organization of the reminder of this chapter is as follows. In Section 4.3, we review causal response functions and discuss their computation using the Schwinger-

Keldysh closed time contour. We also review the relation between amputated response functions and the KMOC waveform from Ref. [139]. In Section 4.4, we present a derivation of the KMOC formulas for the linear impulse and radiated momentum from response functions, and explain the subtle calculation of the angular momentum loss. In Section 4.5, we revisit the calculation of impulse, waveform, and variance up to one loop in the causal basis. Finally in Section 4.6, we summarize our results.

For simplicity, we present example calculations in scalar QED, described by the Lagrangian

$$\mathcal{L} = -\frac{1}{4}F_{\mu\nu}F^{\mu\nu} + \sum_{i=1,2}\left[(D_\mu\Phi_i)\,(D^\mu\Phi_i)^\dagger - m_i^2\Phi_i\Phi_i^\dagger\right] \tag{4.11}$$

where $D_\mu = \partial_\mu + ie\,Q_iA_\mu$ is the covariant derivative, $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu$ is the photon field strength, Q_i and m_i are the charges and masses of the fields Φ_i, and e is the gauge coupling. The generalization to the more interesting case of gravity is straightforward.

4.3 Review: in-in formalism and asymptotic observables

In this section, we review causal response functions which are familiar objects in nonlinear response theory. More details can be found in refs. [140, 141, 145–147]. We present our discussion in terms of a generic field $\varphi(x)$, as a stand-in for Φ_i and A_μ. Then, we review the KMOC formula and its application for the scattering waveform from Refs. [139, 148], as well as the connection of the latter with a five-point causal response functions.

4.3.1 Local response functions and in-in formalism

Let us denote the expectation value of a local operator $\mathcal{O}(x_i)$ at some time $t = x_i^0$ by $\langle\mathcal{O}(x_i)\rangle = \langle 0|\,\mathcal{O}(x_i)\,|0\rangle$ where $|0\rangle$ is the ground state of the theory. A perturbation $\mathcal{H}_j = \int d^dx\varphi(x)j(x)$ with source $j(x)$ is added to the system and the expectation value of the operator $\mathcal{O}(x)$ at $t = x^0$ as a result of the perturbation is given by [141]

$$\langle\mathcal{O}_j(x)\rangle = \langle 0|\,\bar{\mathcal{T}}e^{i\int d^Dx\varphi(x)j(x)}\mathcal{O}(x)\mathcal{T}e^{-i\int d^Dx\varphi(x)j(x)}\,|0\rangle \tag{4.12}$$

where the time integral is from x_i^0 to x^0, $\mathcal{T}$ ($\bar{\mathcal{T}}$) is the time ordering (anti-time-ordering) symbol, and the operator $\mathcal{O}(x)$ is in the interaction picture. We shall describe the path-integral computation of the above quantity momentarily. The difference between $\langle\mathcal{O}_j(x)\rangle$ and $\langle\mathcal{O}(x)\rangle$ can be thought of as a measure of the response of the system to the external perturbation sourced by $j(x)$ [146]. The $n+1$-point

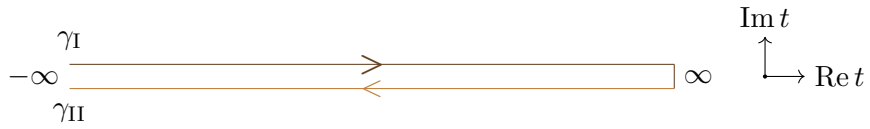


Figure 4.2: Closed-timed contour for computing in-in expectation values.

response function is defined as

$$\begin{aligned} R_{n+1}[\mathcal{O}(x);\varphi(x_1),\varphi(x_2),\cdots,\varphi(x_n)] &= \frac{i^n\delta^n}{\delta j(x_1)\cdots\delta j(x_n)}\langle\mathcal{O}_j(x)\rangle \\ &= \sum_P \theta(x^0-x^0_{P(1)})\theta(x^0_{P(2)}-x^0_{P(2)})\cdots\theta(x^0_{P(n-1)}-x^0_{P(n)}) \\ &\times\langle 0|\left[\left[\cdots\left[\left[\mathcal{O}(x),\varphi(x_{P(1)})\right],\varphi(x_{P(2)})\right]\cdots\right],\varphi(x_{P(n)})\right]|0\rangle\ , \end{aligned} \tag{4.13}$$

where the sum is over all permutations of $(1,2,\cdots,n)$. The definition in Eq. (4.13) is manifestly causal, and symmetric in $x_1,\cdots,x_n$. In particular, for the retarded 2-point function of the field one has

$$R_2[\varphi(x_1);\varphi(x_2)] = \theta(x_1^0-x_2^0)\,\langle 0|\left[\varphi(x_1),\varphi(x_2)\right]|0\rangle \tag{4.14}$$

which is the usual retarded propagator. Now, let us take $x_i^0\to-\infty$, $x^0\to\infty$ and discuss the computation of the path integral in Eq. (4.12). One can use a single closed time contour as shown in Fig. 4.2 with $\Gamma=\gamma^{\rm I}\cup\gamma^{\rm II}$ that goes from $\gamma^{\rm I}:\ -\infty\to\infty$ and back to $\gamma^{\rm II}:\ \infty\to-\infty$ along with two field variables $\varphi_{\rm I}$, $\varphi_{\rm II}$ to write the unitary time evolution operator and its conjugate as path-integrals. These two fields are identified at $x^0=\infty$ i.e

$$\varphi^{\rm I}-\varphi^{\rm II}=0,\quad x^0=\infty. \tag{4.15}$$

The full partition function in terms of the doubled fields is given by [141]

$$Z=\int D\varphi^{\rm I}D\varphi^{\rm II}\ e^{i\int\left(S[\varphi^{\rm I}]-S[\varphi^{\rm II}]\right)}\,. \tag{4.16}$$

Eq. (4.16) allows for the computation of contour-ordered correlation functions by inserting fields with contour labels $\rm I,II$ on the path-integral in Eq. (4.16). For example, an ordinary n-point time-ordered (or anti time-ordered) correlation function can be computed by inserting n type I (or type II) fields in Eq. (4.16). It is useful to adopt notation of [139] to denote the contour ordering of various fields: for a string of operators $\mathcal{O}_1^{\rm I,II},\mathcal{O}_2^{\rm I,II},\mathcal{O}_3^{\rm I,II},\cdots$, the contour-ordering symbol $\mathcal{C}$ sorts all type II fields to the left of type I fields and then time orders or anti-time orders the I and II fields respectively. For example,

$$\mathcal{C}\left\{\mathcal{O}_1^{\rm II}\mathcal{O}_2^{\rm I}\hat{\mathcal{O}}_3^{\rm I}\mathcal{O}_4^{\rm I}\mathcal{O}_5^{\rm II}\right\}=\bar{\mathcal{T}}\left\{\mathcal{O}_1\mathcal{O}_5\right\}\mathcal{T}\left\{\mathcal{O}_2\mathcal{O}_3\mathcal{O}_4\right\}.$$

The boundary condition in Eq. (4.15) implies that the correlation functions including all difference fields vanishes identically

$$\langle 0|\,\mathcal{C}\big\{(\mathcal{O}_1^{\rm I}-\mathcal{O}_1^{\rm II})(\mathcal{O}_2^{\rm I}-\mathcal{O}_2^{\rm II})(\mathcal{O}_3^{\rm I}-\mathcal{O}_3^{\rm II})\cdots\big\}\,|0\rangle=0\,, \tag{4.17}$$

which is known as the *largest time equation*. Finally, it can be shown that the response function in Eq. (4.13) can be written as [140]

$$\begin{aligned}
&R_{n+1}[\mathcal{O}(x);\varphi(x_1),\varphi(x_2),\cdots,\varphi(x_n)]\\
&\quad=\langle 0|\,\mathcal{C}\left\{\frac{1}{2}\left(\mathcal{O}^{\rm I}(x)+\mathcal{O}^{\rm II}(x)\right)\left(\varphi^{\rm I}\left(x_1\right)-\varphi^{\rm II}(x_1)\right)\cdots\left(\varphi^{\rm I}(x_{n-1})-\varphi^{\rm II}(x_n)\right)\right\}|0\rangle\\
&\quad=\langle 0|\,\mathcal{C}\left\{\mathcal{O}^{\rm I}(x)(\varphi^{\rm I}\left(x_1\right)-\varphi^{\rm II}(x_1)\right)\cdots\left(\varphi^{\rm I}(x_{n-1})-\varphi^{\rm II}(x_n)\right)\big\}\,|0\rangle\ ,
\end{aligned} \tag{4.18}$$

using the largest time equation in Eq. (4.17), and with the normalization condition of theta functions $\sum_P\theta(x^0_{P(1)}-x^0_{P(2)})\cdots\theta(x^0_{P(n-1)}-x^0_{P(n)})=1$.

We conclude this section by describing the momentum space in-in Feynman rules to compute correlation functions in Eq. (4.18) using the in-in path integral [139, 149]:

- Assign a label I to the response field $\mathcal{O}$. The source fields φ should be assigned all possible I/II labels. External points must be connected with the same type of internal points.
- Include an additional minus sign for a type II vertex.
- For an internal line connecting two type II fields, insert a factor of $i/(p^2-m^2+i\epsilon)$, whereas for an internal line connecting two type I fields, insert a factor of $-i/(p^2-m^2-i\epsilon)$.
- Draw a positive energy cut for an internal line connecting a type I vertex to a type II vertex. Include a factor of $\hat{\delta}_+(p^2-m^2)=\theta(p^0)\hat{\delta}(p^2-m^2)$ for the propagator.

4.3.2 Waveforms from KMOC and response functions

A systematic approach to computing classical observables from amplitudes was developed in [117, 148] in the context of $2\to 2$ scattering. Suppose we are interested in computing the change in expectation value of some *classical* observable $\mathbb{O}$ due to the interaction between two massive objects mediated by some massless particle. The change in the expectation value of $\mathbb{O}$ is defined as

$$\Delta\mathbb{O}=\langle{\rm in}|\,\mathbb{O}^{\rm out}\,|{\rm in}\rangle-\langle{\rm in}|\,\mathbb{O}^{\rm in}\,|{\rm in}\rangle=\langle{\rm in}|\,S^\dagger\mathbb{O}^{\rm in}S\,|{\rm in}\rangle-\langle{\rm in}|\,\mathbb{O}^{\rm in}\,|{\rm in}\rangle\ , \tag{4.19}$$

where, $\mathbb{O}^{\rm out},\mathbb{O}^{\rm in}$ are $\mathbb{O}(t)$ at evaluated at $t=\infty, t=\ -\infty$ respectively, $S=1+iT$ is the unitary S matrix. The two-particle states, $|{\rm in}\rangle$, are carefully chosen so that they

correspond to two classical particles separated by impact parameter $b^\mu = b_1^\mu - b_2^\mu$ in the following way

$$|\text{in}\rangle = \int \Big[\prod_{i,1,2} \frac{d^4p_i}{(2\pi)^4} \hat{\delta}_+(p_i^2 - m_i^2) e^{-ip_i b_i} \Psi_i(p_i) \Big] |12\rangle \ , \tag{4.20}$$

$$\langle \text{in}| = \int \Big[\prod_{i,1,2} \frac{d^4p_i'}{(2\pi)^4} \hat{\delta}_+(p_i'^2 - m_i^2) e^{ip_i' b_i} \Psi_i^*(p_i) \Big] \langle 1'2'| \ , \tag{4.21}$$

where $|ij\rangle = a_{p_i}^\dagger a_{p_j}^\dagger |0\rangle$ is a two-particle state, $\hat{\delta}_+(p_i^2 - m_i^2) = (2\pi)\theta(p_i^0)\delta(p_i^2 - m_i^2)$ the Lorentz invariant phase space measure, and $\Psi_i(p_i)$ are wavepackets sharply peaked around the classical momenta $p_i^\mu = m_i u_i^\mu$ so that for any function $f(p_i)$,

$$f(p_i)\big|_{\text{cl.}} = \int \Big[\prod_{i,1,2} \frac{d^4p_i'}{(2\pi)^4} \hat{\delta}_+(p_i'^2 - m_i^2) \Big] |\Psi_i(p_i)|^2 f(p_i) = f(m_i u_i) \, , \tag{4.22}$$

where m_i, u_i are the masses and four velocities of the particles, and $u_i^2 = 1$. Then, the change in an observable during the two-particle-scattering process is given by

$$\Delta\mathbb{O} = \langle 1'2'| \, \mathbb{O}^{\text{out}} - \mathbb{O}^{\text{in}} \, |12\rangle \, \big|_{\text{cl.}} \, , \tag{4.23}$$

where $\cdot\big|_{\text{cl.}}$ denotes the integral against the classical wavefunctions as in Eq. (4.22), with an appropriate impact parameter phase. KMOC then instruct us to write $\mathbb{O}^{\text{out}} = S^\dagger \mathbb{O}^{\text{in}} S$, and expand the S-matrix as $S = 1 + iT$, yielding

$$\begin{aligned} \Delta\mathbb{O} &= i \langle 1'2'| \, [\mathbb{O}^{\text{in}}, T] \, |12\rangle + \langle 1'2'| \, T^\dagger [\mathbb{O}^{\text{in}}, T \, |12\rangle] \\ &= i \langle 1'2'| \, [\mathbb{O}^{\text{in}}, T] \, |12\rangle + \sum_X \langle 1'2'| \, T^\dagger \, |X\rangle \, \langle X| \, [\mathbb{O}^{\text{in}}, T] \, |12\rangle \ , \end{aligned} \tag{4.24}$$

upon the use of the unitarity of the S-matrix

$$SS^\dagger = 1 \quad \to \quad i(T - T^\dagger) = T^\dagger T \, . \tag{4.25}$$

and inserting a complete set of states $\sum_X |X\rangle \langle X|$.

Perhaps the simplest example of such an observable is the classical spectral waveform of the massless particle. This is defined as the change in expectation value of the annihilation operator $a^h(k)$ of the photon with helicity h as

$$\mathcal{W}_k^h = \langle 1'2'| \, a_k^{h \, \text{out}} - a_k^{h \, \text{in}} \, |12\rangle \, \big|_{\text{cl.}} \, . \tag{4.26}$$

The absence of photons in the initial state, $a_k^{h \, \text{in}} |12\rangle = 0$, simplifies the calculation of the various terms

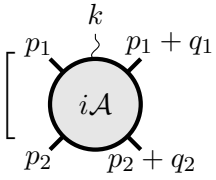
$$\langle 1'2'| \, [a_k^{h \, \text{in}}, T] \, |12\rangle = \langle 1'2'| \, a_k^{h \, \text{in}} T \, |12\rangle = \langle 1'2'k| \, T \, |12\rangle = \mathcal{A}(12 \to 1'2'k) \, . \tag{4.27}$$

The momentum conserving delta functions are implicit in the definition of the amplitude $\mathcal{A}$. similarly,

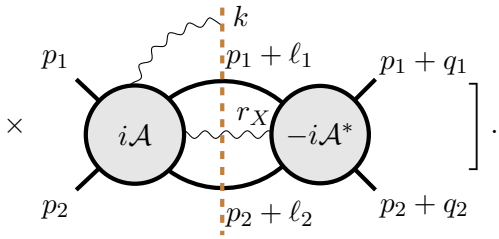
$$\sum_X \langle 1'2'|\, T^\dagger\, |X\rangle\, \langle X|\, a_k^{h\,\mathrm{in}} T\, |12\rangle = \sum_X \mathcal{A}(1,2\to Xk)\mathcal{A}^*(X\to 1'2')\,. \tag{4.28}$$

Thus, by this procedure, we obtain a formula for the spectral waveform in terms of the scattering amplitudes which was first derived in Ref. [150]

$$\mathcal{W}_k^h = \int \prod_{i=1,2} \hat{d}^4 q_i\; \hat{\delta}_+\left(2p_1\cdot q_i + q_i^2\right) e^{-ib_i\cdot q_i}\hat{\delta}^4(k+q_1+q_2)\Big[$$



$$+\sum_X \int \prod_{i=1,2} \hat{d}^4\ell_i\; \hat{\delta}_+\left(2p_1\cdot \ell_i + \ell_i^2\right)\; \hat{\delta}(\ell_1+\ell_2+r_X)\times$$



Let us now review how this formula is derived from a local causal response function, as explained in Ref. [139]. Consider the LSZ amputations

$$\begin{aligned}
&\mathrm{LSZ}_x^{\mathrm{out}}\Phi_i(x) = +i\int d^4x e^{+ip_i x}\left(\partial^2+m^2\right)\Phi_i(x) = a_{p_i}^{\mathrm{out}} - a_{p_i}^{\mathrm{in}} \quad \text{(outgoing)}\,,\\
&\mathrm{LSZ}_x^{\mathrm{in}}\;\Phi_i^\dagger(x) = -i\int d^4x e^{-ip_i x}\left(\partial^2+m^2\right)\Phi_i^\dagger(x) = a_{p_i}^{\dagger\mathrm{out}} - a_{p_i}^{\dagger\mathrm{in}} \quad \text{(incoming)}\,.
\end{aligned} \tag{4.30}$$

Similarly for the photon with helicity h

$$\begin{aligned}
&\mathrm{LSZ}_x^{\mathrm{out}} A^h(x) = +i\int d^4k e^{+ikx}\left(\partial^2+m^2\right)\varepsilon_\mu^{h*}(k)A^\mu(x) = a_k^{h\;\mathrm{out}} - a_k^{h\;\mathrm{in}} \quad \text{(outgoing)}\,,\\
&\mathrm{LSZ}_x^{\mathrm{in}}\; A^h(x) = -i\int d^4x e^{-ikx}\left(\partial^2+m^2\right)\varepsilon_\mu^{h}(k)A^\mu(x) = a_k^{h\dagger\;\mathrm{out}} - a_k^{h\dagger\;\mathrm{in}} \;\text{(incoming)}\,.
\end{aligned} \tag{4.31}$$

Then, by amputating the five-point response function we obtain the key relation

$$\begin{aligned}
&\mathrm{LSZ}_x^{\mathrm{out}}\prod_{i=1,2}\mathrm{LSZ}_{x_i'}^{\mathrm{out}}\,\mathrm{LSZ}_{x_i}^{\mathrm{in}} R_5[A^h(x);\Phi_1^\dagger(x_1),\Phi_2^\dagger(x_2),\Phi_1(x_1'),\Phi_2(x_2')]\\
&= \langle 0|\,\mathcal{C}\left\{(a_k^{\mathrm{I}\,h\,\mathrm{out}} - a_k^{\mathrm{I}\,h\,\mathrm{in}})(a_{1'}^{\mathrm{II}\;\mathrm{in}} - a_{1'}^{\mathrm{I}\;\mathrm{in}})(a_{2'}^{\mathrm{II}\;\mathrm{in}} - a_{2'}^{\mathrm{I}\;\mathrm{in}})(a_1^{\mathrm{II}\;\dagger\mathrm{in}} - a_1^{\mathrm{I}\;\dagger\mathrm{in}})(a_2^{\mathrm{II}\;\dagger\mathrm{in}} - a_1^{\mathrm{I}\;\dagger\mathrm{in}})\right\}|0\rangle\\
&= \langle 0|\, a_{1'}^{\mathrm{in}} a_{2'}^{\mathrm{in}}(a_k^{h\;\mathrm{out}} - a_k^{h\;\mathrm{in}})a_1^{\mathrm{in}\dagger} a_2^{\mathrm{in}\dagger}\,|0\rangle = \langle 1'2'|\, a_k^{h\;\mathrm{out}}\,|12\rangle = \mathcal{W}_k^h\,.
\end{aligned} \tag{4.32}$$

In the second line, we have adopted the contour ordering prescription in Eq. (4.18). As a result of the boundary condition in Eq. (4.15), we obtain, for instance, from the amputated heavy fields [139]

$$\begin{aligned}\mathrm{LSZ}^{\mathrm{out}}_{x_i'}\left[\Phi^{\mathrm{I}}_i(x_i') - \Phi^{\mathrm{II}}_i(x_i')\right] &= a^{\mathrm{II\,in}}_{p_i'} - a^{\mathrm{I\,in}}_{p_i'}\,,\\ \mathrm{LSZ}^{\mathrm{in}}_{x_i}\left[\Phi^{\mathrm{I}\dagger}_i(x_i) - \Phi^{\mathrm{II}\dagger}_i(x_i)\right] &= a^{\mathrm{II\,in}\,\dagger}_{p_i} - a^{\mathrm{I\,in}\,\dagger}_{p_i}\,,\end{aligned}$$

and similarly for the photon. Using the Feynman rules to compute the response functions outlined in Section 4.3.1, it is easy to see that indeed, upon amputation, the five-point function reproduces precisely the KMOC formula for the spectral waveform in Eq. (4.30). The first term Eq. (4.30) corresponds to diagrams with only I fields and the cuts in the second term correspond to diagrams which connect I and II fields.

4.4 KMOC formulas from causal response functions

In this section, we explain in detail the relationship between in-in expectation values of more general (composite) operators and causal response functions. This was briefly touched upon in [139], but here we further clarify the precise relationship and explain various subtleties. In particular, we show that the change in the expectation value of an observable $\mathbb{O}$ is given by a particular soft limit of a five-point response function.

We shall consider observables, $\mathbb{O}$ which are integrals of a local density given by an operator $\mathcal{O}(x)$

$$\mathbb{O}(t) = \int d^3\boldsymbol{x}\; \mathcal{O}(t,\boldsymbol{x})\,. \tag{4.33}$$

Indeed, such is the case for the linear impulse and radiated momentum, as well as the radiated angular momentum loss, which are suitable integrals of the stress-energy tensor, as we review below.

Just as for the waveform, following [139], we can write the expectation value the local density $\mathcal{O}$ in terms of an amputated response function

$$\begin{aligned}&\langle 1'2'|\,\mathcal{O}(x)\,|12\rangle = \langle 0|\,a^{\mathrm{in}}_{1'}a^{\mathrm{in}}_{2'}\mathcal{O}(x)a^{\mathrm{in}\dagger}_{1}a^{\mathrm{in}\dagger}_{2}\,|0\rangle\\ &= \prod_{i=1,2}\mathrm{LSZ}^{\mathrm{out}}_{x_i'}\,\mathrm{LSZ}^{\mathrm{in}}_{x_i} R_5[\mathcal{O}(x);\Phi^\dagger_1(x_1),\Phi^\dagger_2(x_2),\Phi_1(x_1'),\Phi_2(x_2')]\,.\end{aligned} \tag{4.34}$$

Let us now define the Fourier transform of this expectation value

$$\langle\mathcal{O}(\omega,\boldsymbol{k})\rangle = \int d^4x e^{i\omega t - i\boldsymbol{k}\boldsymbol{x}}\,\langle 1'2'|\,\mathcal{O}(t,\boldsymbol{x})\,|12\rangle\;, \tag{4.35}$$

with $\omega > 0$, which at zero three-momentum, $\boldsymbol{k}$, yields

$$\lim_{\boldsymbol{k}\to 0}\langle \mathcal{O}(\omega,\boldsymbol{k})\rangle = \int dt e^{i\omega t}\,\langle 1'2'|\,\mathbb{O}(t)\,|12\rangle\ . \tag{4.36}$$

Finally, by taking the zero-frequency limit we obtain the change of the observable

$$\begin{aligned} -\lim_{\omega\to 0}\lim_{\boldsymbol{k}\to 0} i\omega\ \langle \mathcal{O}(\omega,\boldsymbol{k})\rangle &= \lim_{\omega\to 0}\int dt\ e^{i\omega t}\,\langle 1'2'|\,\partial_t\mathbb{O}(t)\,|12\rangle \\ &= \langle 1'2'|\,\mathbb{O}^{\rm out}-\mathbb{O}^{\rm in}\,|12\rangle\,. \end{aligned} \tag{4.37}$$

where the first equality follows from integration by parts, which is carried out before taking the $\omega \to 0$ limit.

Therefore, we find that the change of the observable between the *in* and *out* states is simply given by the soft limit of an amputated five-point local response function

$$\begin{aligned} &\langle 1'2'|\,\mathbb{O}^{\rm out}-\mathbb{O}^{\rm in}\,|12\rangle = -\lim_{\omega\to 0}\lim_{\boldsymbol{k}\to 0} i\omega\int d^4x\ e^{i\omega t-i\boldsymbol{k}\boldsymbol{x}}\times \\ &\times\prod_{i=1,2}\mathrm{LSZ}^{\rm out}_{x_i'}\,\mathrm{LSZ}^{\rm in}_{x_i} R_5[\mathcal{O}(x);\Phi_1^\dagger(x_1),\Phi_2^\dagger(x_2),\Phi_1(x_1'),\Phi_2(x_2')] \end{aligned} \tag{4.38}$$

Eq. (4.38) is our key formula to compute the change in classical observables from causal response functions. Next, we will explain how this soft limit does indeed reproduce the KMOC formula for the linear impulse and reveals a subtlety in the computation of the angular impulse.

4.4.1 Linear impulse

Let us first demonstrate how the KMOC formula for the impulse of a particle undergoing scattering follows from Eq. (4.38) using the in-in Feynman rules to compute the response functions discussed in section 4.3.1. In the KMOC framework the impulse is defined as the difference in expectation value of momentum of particle 1 defined at $x^0 \to -\infty$ from its value at $x^0 \to \infty$, that is,

$$\Delta\mathbb{P}_1^\mu = \langle 1'2'|\,\mathbb{P}_1^{\mu\,\rm out}-\mathbb{P}_1^{\mu\,\rm in}\,|12\rangle\,\big|_{\rm cl.}\,, \tag{4.39}$$

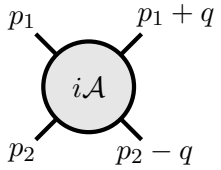
with

$$\mathbb{P}_1^{\mu\,\rm in} = \int \hat{d}^4p_1\ p_1^\mu\hat{\delta}(p_1^2-m_1^2)a_{p_1}^{\rm in\dagger}a_{p_1}^{\rm in}\,, \tag{4.40}$$

$$\mathbb{P}_1^{\mu\,\rm out} = \int \hat{d}^4p_1\ p_1^\mu\hat{\delta}(p_1^2-m_1^2)a_{p_1}^{\rm out\dagger}a_{p_1}^{\rm out}\,. \tag{4.41}$$

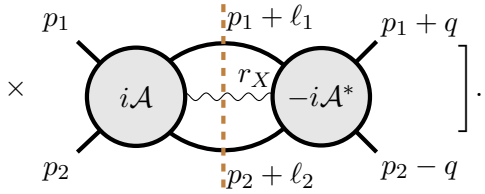
The KMOC impulse formula in terms of the in-out scattering amplitudes is [117]

$$\Delta\mathbb{P}_1^\mu = \int \hat{d}^4q\, \hat{\delta}\left(2p_1\cdot q + q^2\right)\hat{\delta}\left(2p_2\cdot q - q^2\right)\; e^{-ib\cdot q}\Bigg[\; q^\mu$$



(4.42)

$$+\sum_X \int \prod_{i=1,2} \hat{d}^4\ell_i\, \hat{\delta}_+\left(2p_1\cdot\ell_i + \ell_i^2\right)\; \hat{\delta}(\ell_1+\ell_2+r_X)\; \ell_1^\mu \times$$



where the momentum transfer q is

$$q^\mu = p_1'^\mu - p_1^\mu = p_2^\mu - p_2'^\mu. \tag{4.43}$$

and is taken to be small in the classical limit, $q \ll p_i$. We refer the reader to [117] for a detailed derivation.

Next, we will derive this formula from amputated response functions. The relevant local density for $\mathbb{P}_1$ is the momentum density

$$\mathbb{P}_1^\mu(t) = \int d^3\boldsymbol{x}\mathcal{P}_1(t,\boldsymbol{x}) = \int d^3\boldsymbol{x} T_1^{0\mu}(t,\boldsymbol{x})\,, \tag{4.44}$$

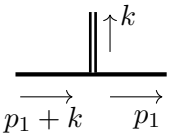
given in terms of the stress-energy tensor of particle 1

$$T_1^{\mu\nu} = 2\partial^\mu\Phi_1\partial^\nu\Phi_1^\dagger - \eta^{\mu\nu}\left(\partial_\rho\Phi_1\partial^\rho\Phi_1^\dagger - m_1^2\Phi_1\Phi_i^\dagger\right)\,. \tag{4.45}$$

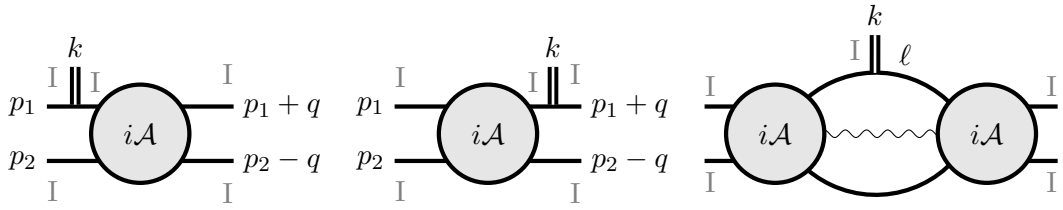
Thus, to derive the linear impulse we consider the amputated response function

$$\langle\mathcal{P}_1^\mu(\omega,\boldsymbol{k})\rangle = \int d^4x e^{i\omega t - i\boldsymbol{k}\boldsymbol{x}}\,\langle 1'2'|\, T_1^{0\mu}(t,\boldsymbol{x})\,|12\rangle\,. \tag{4.46}$$

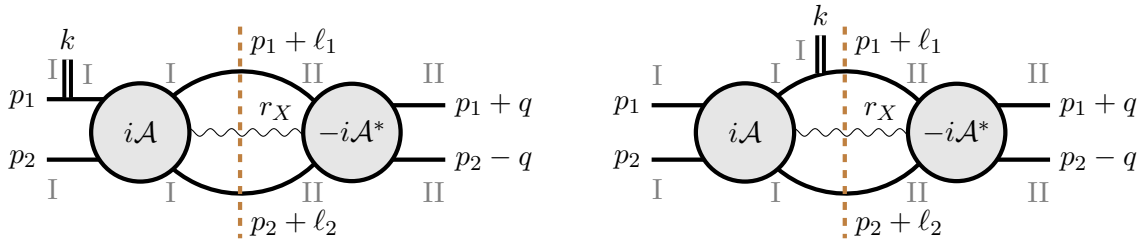
The Feynman rule corresponding to the heavy-scalar stress-tensor insertion can be easily obtained



$$= (p_1+k)^{(0}p_1^{\mu)} = 2p_1^0p_1^\mu + \mathcal{O}(k)\,, \tag{4.47}$$



(a) Operator insertions on uncut diagrams



(b) Cut diagrams with operator insertions

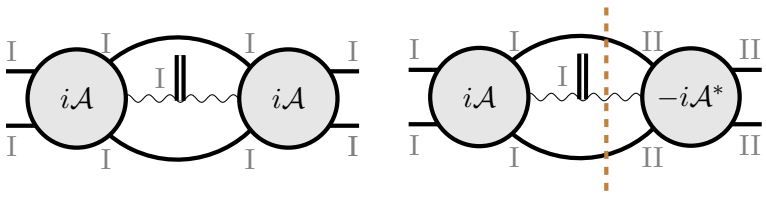
Figure 4.3: Two different types of diagrams with the momentum operator insertions are shown. We have taken $p_1,\ p_2$ to be incoming and $p_1+q,\ p_2-q,\ k$ to be outgoing.

where we used an unnormalized symmetrization symbol $A^{(\mu}B^{\nu)} = A^\mu B^\nu + A^\nu B^\mu$ and on the second equality we show the terms which are relevant in the soft limit $k \to 0$. Importantly, the energy of the scalar legs is conserved only up to $\omega = k^0$.

In writing the diagrams for this response function we assign a type I vertex for the operator insertion with energy ω flowing *out* of the diagrams and all possible assignments of I/II vertices for the remaining vertices. The in-in Feynman rules lead to various classes diagrams with all type I fields as well as cut diagrams involving type I and type II fields as shown in Fig. 4.3. Consider the types of diagrams in Fig. 4.3a where $\mathcal{O}$ is inserted on the external legs. The first diagram corresponds to multiplying four-point the amplitude $i\mathcal{A}$ without the soft operator insertion by an additional propagator. Taking the soft limit we find

$$\lim_{\omega\to 0}(-2i\omega p_1^0 p_1^\mu)\frac{i}{(\omega-p_1^0)^2-\boldsymbol{p}_1^2-m_1^2+i\epsilon} = \lim_{\omega\to 0}\frac{2\omega p_1^0 p_1^\mu}{-2p^0\omega+\omega^2+i\epsilon} = -p_1^\mu\,, \quad (4.48)$$

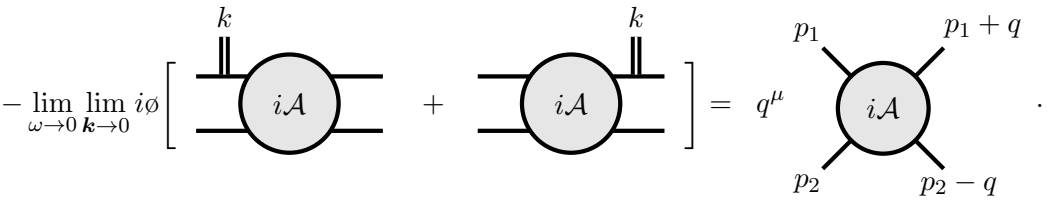
where in the second line we have used the on-shell condition $p_1^2 = m_1^2$. Similarly, the second diagram with the momentum operator inserted on the outgoing external



(a) (b)

Figure 4.4: Two types of diagrams contributing to the causal response function for the photon stress tensor.

line produces a factor of $p_1^\mu + q^\mu$. These two classes of diagrams combined yield



(4.49)

which reproduces the first term in the KMOC formula in Eq. (4.43). The uncut diagram on the right of Fig. 4.3a, which features the operator inserted on an internal line with momentum ℓ does not contribute in the soft limit due to the identity

$$\lim_{\omega\to 0} \frac{2i\omega\ell^0\ell^\mu}{\left((\omega+\ell^0)^2-\boldsymbol{\ell}^2-m_1^2+i\epsilon\right)\left(\ell^2-m_1^2+i\epsilon\right)} = 0 \tag{4.50}$$

Finally, let us consider the cut diagrams obtained by connecting a type I vertex with a type II vertex in Fig. 4.3b. The stability condition, which requires the vanishing of the three-particle cuts, and the positivity of energy flow along the cuts guarantee that only the unitarity cuts survive as depicted in Fig. 4.3. Using the identity in Eq. 4.49 on the left-hand side of the cut one easily finds that these two classes of diagrams combined give the cut term in the second line of Eq. (4.43). All together, we have reproduced the KMOC formula for the linear impulse from the soft limit of a causal response function of the stress tensor.

4.4.2 Radiated linear momentum

It is easy to generalize the above discussion to give a formula for the radiated momentum, which is carried away by the photons:

$$\Delta\mathbb{P}_\gamma^\mu = \langle 1'2'|\,\mathbb{P}_\gamma^{\mu\,\text{out}} - \mathbb{P}_\gamma^{\mu\,\text{in}}\,|12\rangle\,\big|_{\text{cl.}} = \langle 1'2'|\,\mathbb{P}_\gamma^{\mu\,\text{out}}\,|12\rangle\,\big|_{\text{cl.}}\,, \tag{4.51}$$

where we used the absence of photons in the initial state to set $\mathbb{P}_\gamma^{\mu\,\text{in}}|12\rangle = 0$ and

$$\mathbb{P}_\gamma^{\mu\,\text{out}} = \sum_h \int \hat{d}^4 p\, p^\mu \hat{\delta}(p^2) a_p^{h\,\text{out}\dagger} a_p^{h\,\text{out}}\,. \tag{4.52}$$

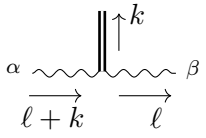
This can be written in terms of the local density

$$\mathbb{P}_\gamma^\mu(t) = \int d^3\boldsymbol{x} \mathcal{P}_\gamma^\mu(t,\boldsymbol{x}) = \int d^3\boldsymbol{x} T_\gamma^{0\mu}(t,\boldsymbol{x})\,, \tag{4.53}$$

where the stress energy tensor of the photon is

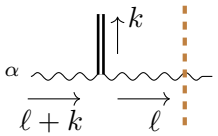
$$T_\gamma^{\mu\nu} = F^{\mu\rho}F^\nu_\rho - \frac{1}{4}\eta^{\nu\nu}F^{\rho\sigma}F_{\rho\sigma}\,. \tag{4.54}$$

The Feynman rule corresponding to the insertion of this operator is



$$\begin{aligned} &= (\ell+k)^\alpha \ell^{(\mu}\eta^{0)\beta} + \ell^\beta \eta^{\alpha(\mu}(\ell+k)^{0)} \\ &\quad - \eta^{\alpha(\mu}\eta^{0)\beta}\ell\cdot(\ell+k) - \eta^{\alpha\beta}\ell^{(\mu}(\ell+k)^{0)}\,. \end{aligned} \tag{4.55}$$

This simplifies considerably when one of the legs is put on-shell and contracted with a transverse polarization



$$= -2\ell^0\ell^\mu\varepsilon^\alpha + \ell^\alpha\varepsilon^{(\mu}\ell^{0)} + \mathcal{O}(k)\,, \tag{4.56}$$

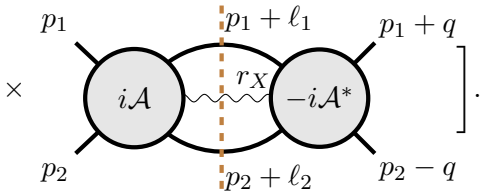
where we only display terms that remain in the soft limit $k \to 0$.

The absence of photons in the initial state implies that the only classes of diagrams that can contribute to the response function of the photon stress tensor are those in Fig. 4.4. The diagrams with only I or II fields in Fig. 4.4a vanish in the soft limit by a manipulation analogous to Eq. (4.50). On the other hand, the diagram with both I or II fields yields the cut diagrams of the form of Fig. 4.4a where on the support of the on-shell delta function for the photon, $\delta(\ell^2)$, we can use the completeness relation $\eta^{\beta\delta} = \sum_h \epsilon_h^\beta \epsilon_h^{*\,\delta}$, which holds up to terms that vanish due to the Ward-Takahashi identity. In the soft limit these diagrams produce terms on the form

$$\lim_{\omega\to 0}(-i\omega)\mathcal{A}_\alpha \frac{-i(-2\ell^0\ell^\mu\varepsilon^\alpha + \ell^\alpha\varepsilon^{(\mu}\ell^{0)})}{(\omega+\ell^0)^2 - \boldsymbol{\ell}^2 + i\epsilon} = \lim_{\omega\to 0}\mathcal{A}_\alpha\varepsilon^\alpha \frac{2\omega\ell^0\ell^\mu}{2\omega\ell^0 + \omega^2 + i\epsilon} = \ell^\mu\mathcal{A}\,, \tag{4.57}$$

where $\mathcal{A} = \varepsilon^\alpha \mathcal{A}_\alpha$ denotes the amplitude on the left of the cut, and we used its on-shell Ward identity, $\ell^\alpha \mathcal{A}_\alpha = 0$. Thus, from the soft limit of the causal response function we obtain the KMOC formula for the radiated momentum

$$\Delta \mathbb{P}^\mu_\gamma = \int \hat{d}^4 q \; \hat{\delta}\left(2p_1 \cdot q + q^2\right) \hat{\delta}\left(2p_2 \cdot q - q^2\right) \; e^{-ib\cdot q} \Bigg[\tag{4.58}$$
$$\sum_X \int \prod_{i=1,2} \hat{d}^4 \ell_i \; \hat{\delta}_+\left(2p_1 \cdot \ell_i + \ell_i^2\right) \; \hat{\delta}(\ell_1 + \ell_2 + r_X) \; r_X^\mu \times$$



where we recall that $r_X = \sum_a \ell_a$ is the sum over the momenta of all photons in the cut.

Finally, we note that in deriving the above results, we assumed $\omega \ll \ell$, and expanded the Feynman rules and propagators using the method of regions [127]. For the linear impulse and momentum loss there are no other regions in the expansion that contribute, but we will see next that this is not true when considering the angular momentum loss.

4.4.3 Radiated angular momentum

In this section, we compute angular momentum loss in the form of photons during scattering, from a response function. A key difference between this calculation and that of the linear impulse or energy loss is that the soft limit $\omega \to 0$ of the response function is subtle.

We take the canonical definition of the angular momentum operator for a massive or massless particle as an integral of the corresponding density

$$\mathbb{J}^{ij}_\gamma(t) = \int d^3\boldsymbol{x} \, \mathcal{J}^{ij}(t, \boldsymbol{x}) = \int d^3\boldsymbol{x} \, \boldsymbol{x}^{[i} T^{0j]}_\gamma(x) = \int d^3\boldsymbol{x} \left(\boldsymbol{x}^i T^{0j}_\gamma(x) - \boldsymbol{x}^j T^{0i}_\gamma(x)\right), \tag{4.59}$$

where $T^{\mu\nu}$ is the appropriate energy-momentum tensor, and i, j are spatial indices. Thus we conclude that the angular impulse is given as the soft limit of a derivative of the same amputated causal response function

$$\Delta \mathbb{J}^{ij}_\gamma = -\lim_{\omega\to 0} \lim_{\boldsymbol{k}\to 0} i\omega \, \langle \mathcal{J}^{ij}_\gamma(\omega, \boldsymbol{k}) \rangle \big|_{\text{cl.}} = -\lim_{\omega\to 0} \lim_{\boldsymbol{k}\to 0} \omega \, \partial^{[i}_{\boldsymbol{k}} \langle \mathcal{P}^{j]}_\gamma(\omega, \boldsymbol{k}) \rangle \big|_{\text{cl.}} . \tag{4.60}$$

Naively, for this computation, we can simply consider the diagrams in Fig. 4.4 and compute the appropriate derivatives and soft limit as prescribed in Eq. (4.60).

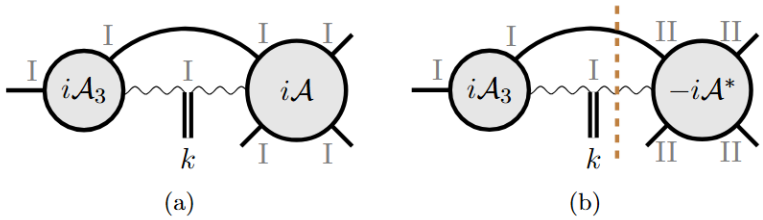


Figure 4.5: Diagrams relevant for the computation of the angular momentum loss.

But the set of diagrams that contribute will be slightly different. In the calculation of the radiated momentum $\Delta\mathbb{P}_\gamma$ we tacitly discarded the diagrams in Fig. **??**, because in the soft limit, they yield an on-shell three-point amplitude, which vanishes for real kinematics. However, to reach this conclusion we expanded the integrand in the limit $\omega \ll \ell$, which implicitly assumes that there are no relevant regions in which the loop momentum is or order $\ell \sim \omega$. This assumption is valid for the class of diagrams without cut propagators in Fig. 4.5(a), which indeed vanish; but it fails when considering the diagrams in Fig. 4.5(b). The region $\ell \sim \omega$ of the loop integration in these diagrams in fact completely captures the leading contribution to the angular momentum loss. The mechanism for this is that the integrals develop a pole in ω in the soft limit, thanks to the cut photon propagator

$$\lim_{\boldsymbol{k}\to 0} \frac{\hat{\delta}_+(\ell^2)\hat{\delta}\left((p_1-k-\ell)^2-m_1^2\right)}{(\ell+k)^2+i\epsilon} = \frac{\hat{\delta}_+(\ell^2)\hat{\delta}\left(2p_1\cdot\ell-2p_1^0\omega\right)}{2\ell^0\omega+\omega^2+i\epsilon}\,.$$

In the region $\omega \sim \ell$, one must not drop the $p_1^0\omega$ term in the delta function, and power-counting the measure as $d^4\ell \sim \omega^4$ we find that the result is indeed of order $1/\omega$. The shift in the on-shell delta function for the massive particle can be interpreted as it being off-shell by an infinitesimal amount of $\mathcal{O}(\omega)$. We will see momentaritly that this off-shell-ness makes the evaluation of related integrals completely unambiguous. Furthermore, since we are interested in small $\varnothing \sim \ell$ we can write $(q_2-\ell)^2 \sim q_2^2$, so the amplitude on the other side of the cut effectively becomes a contact term, as it cannot be resolved by waves with small frequency of $\mathcal{O}(\omega)$. In hindsight, this justifies the calculations of the angular momentum loss in refs. [143, 144] using Weinberg's soft theorem. Let us now compute this quantity at leading order in electromagnetism. The leading diagrams we can draw are the cut diagrams shown in Fig. 4.6. These can be straightforwardly computed using the Feynman rule in Eq. (4.55). Although the latter looks rather complicated, the calculation simplifies greatly by using power counting to read off the contributing terms after applying the derivatives. For more details, we refer the reader to Appendix B where we present the calculation in detail for a scalar toy model. In electromagnetism, summing up

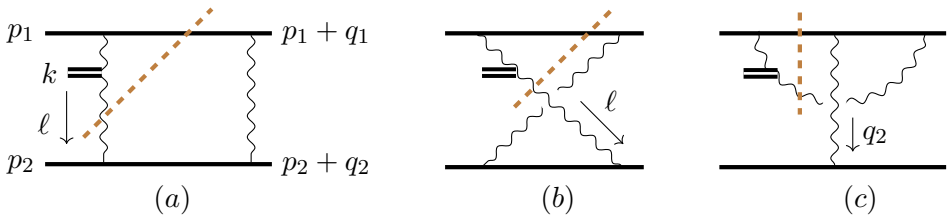


Figure 4.6: Three diagrams contributing to the angular impulse are shown. We take the p_1, p_2 incoming and $p_1+q_1, p_2+q_2,\ k$ outgoing.

the diagrams in Fig. 4.6 we get

$$\mathcal{J}^{ij}_{ph,1} = \frac{ie^4}{4}\lim_{\omega\to 0} e^4 Q_1 Q_2\, p_1\cdot p_2 \int \hat{d}^4 q\, e^{ib\cdot q}\,\hat{\delta}\left(p_1\cdot q\right)\hat{\delta}\left(p_2\cdot q\right)\times \\ \times\, p_1^{[i}\left[Q_1 Q_2\left(4p_1\cdot p_2 \mathcal{K}^{j]}_{12} + 2\mathcal{I}_{12} q^{j]}\right) - Q_1^2\left(4m_1^2\mathcal{K}^{j]}_{11} + 2\mathcal{I}_{11}q^{j]}\right)\right], \tag{4.61}$$

where we have defined the following integrals

$$\mathcal{K}^{\mu}_{ab} = \int \hat{d}^4\ell \frac{\hat{\delta}_+(\ell^2)\hat{\delta}'(p_a\ell - p_1^0\omega)\; q\cdot\ell}{(p_b\ell + i\epsilon)^2}\ell^{\mu}\,, \tag{4.62}$$

$$\mathcal{I}_{ab} = \int \hat{d}^4\ell \frac{\hat{\delta}_+(\ell^2)\hat{\delta}(p_a\ell - p_a^0\omega)}{(p_b\ell+i\epsilon)} + \int \hat{d}^4\ell \frac{\hat{\delta}_+(\ell^2)\hat{\delta}'(p_a\ell - p_a^0\omega)(p_a\cdot\ell)}{(p_b\ell+i\epsilon)}\,. \tag{4.63}$$

These can be computed by using integration by parts identities (IBP) [151, 152] to reduce them to a basis of master integrals

$$I_a = \int \hat{d}^4\ell\hat{\delta}_+(\ell^2)\hat{\delta}(p_a\cdot\ell - p_a^0\omega)\,, \tag{4.64}$$

$$I_{ab} = \int \hat{d}^4\ell \frac{\hat{\delta}_+(\ell^2)\hat{\delta}(p_a\cdot\ell - p_a^0\omega)}{p_b\cdot\ell + i\epsilon}\,, \tag{4.65}$$

which can be unambiguously evaluated. For instance, the first master integral yields

$$I_1 = \int \hat{d}^4\ell\hat{\delta}_+(\ell^2)\hat{\delta}(p_1\cdot\ell - p_1^0\omega) = \frac{1}{4\pi^2 m_1}\int_0^\infty \frac{d^3\boldsymbol{\ell}}{2|\boldsymbol{\ell}|}\delta(|\boldsymbol{\ell}|-\omega) = \frac{\omega}{2\pi m_1}\,. \tag{4.66}$$

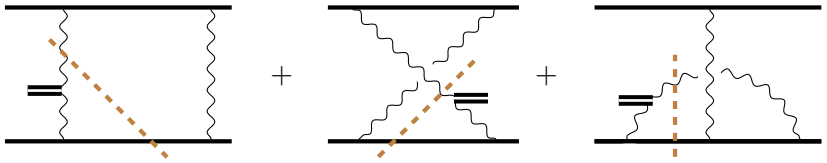
More details about the integration are given in Appendix A. Here we quote the results:

$$\mathcal{I}_{12} = \frac{1}{\pi m_1 m_2}\frac{\cosh^{-1}\gamma}{\sqrt{\gamma^2-1}}\,, \qquad \mathcal{I}_{11} = \frac{1}{\pi m_1^2}\,, \tag{4.67}$$

$$\mathcal{K}^{\mu}_{12} = \frac{q^\mu}{2\pi m_1^2 m_2^2}\left[\frac{1}{\gamma^2-1} - \frac{\gamma\cosh^{-1}\gamma}{(\gamma^2-1)^{3/2}}\right], \qquad \mathcal{K}^{\mu}_{11} = -\frac{q^\mu}{6\pi m_1^4}\,. \tag{4.68}$$

where $\gamma = p_1 \cdot p_2/m_1 m_2$ is the relative Lorentz factor of the particles.

There are three additional diagrams shown below that can be evaluated by simply changing labels $1 \leftrightarrow 2$:

$\mathcal{J}^{ij}_{\gamma,2} =$  $+$ $+$

$$= \mathcal{J}^{ij}_{\gamma,1}\,(1 \leftrightarrow 2)\ . \tag{4.69}$$

Furthermore, simple power counting in ω shows that the triangle diagrams do not contribute to angular impulse. Finally, using the expression for the leading order electromagnetic linear impulse of the heavy particle 1 [117]

$$\begin{aligned}\Delta\mathbb{P}^i_1 = -\Delta\mathbb{P}^i_2 &= ie^2 Q_1 Q_2 p_1\cdot p_2 \int \frac{\hat{d}^4 q}{q^2}\hat{\delta}\,(p_1\cdot q)\,\hat{\delta}\,(p_2\cdot q)\,q^i e^{-iq\cdot b}\\ &= -\frac{e^2 Q_1 Q_2}{2\pi}\frac{\gamma}{\sqrt{\gamma^2-1}}\frac{b^i}{b^2}\,,\end{aligned} \tag{4.70}$$

we find the leading order electromagnetic angular momentum loss

$$\begin{aligned}\Delta\mathbb{J}^{ij}_\gamma &= \mathcal{J}^{ij}_{\gamma,1} + \mathcal{J}^{ij}_{\gamma,2}\\ &= \frac{e^2}{2\pi}\left[-\frac{2Q_1^2}{3m_1^2} + \frac{Q_1 Q_2}{m_1 m_2}\left(\frac{\gamma}{\gamma^2-1} - \frac{\cosh^{-1}\gamma}{(\gamma^2-1)^{3/2}}\right)\right] p_1^{[i}\,\Delta\mathbb{P}_1^{j]} + (1\leftrightarrow 2).\end{aligned} \tag{4.71}$$

which agrees with [153].

4.5 Classical observables in the causal basis

In this section, we will present examples of classical observables computed using manifestly causal diagrams, and point out the various simplifications that this method provides.

4.5.1 Amputated response functions in causal basis

Since we are interested in causal observables, we shall now choose a basis in the in-in path integral in which the causal structure of these quantities becomes manifest. This basis, called the Keldysh basis or r/a basis, is defined via a rotation of the I/II basis in the path integral in Eq. (4.16) as

$$\varphi^r = \frac{1}{2}(\varphi^{\rm I} + \varphi^{\rm II})\,, \qquad \varphi^a = \varphi^{\rm I} - \varphi^{\rm II}\,, \tag{4.72}$$

and for a generic operator, $\mathcal{O}$, we also define

$$\mathcal{O}^r = \frac{1}{2}(\mathcal{O}^{\rm I} + \mathcal{O}^{\rm II}), \qquad \mathcal{O}^a = \mathcal{O}^{\rm I} - \mathcal{O}^{\rm II}. \tag{4.73}$$

The response function in Eq. (4.18) takes a simple form in this basis

$$R_{n+1}[\mathcal{O}(x); \varphi(x_1), \cdots, \varphi(x_n)] = \langle \mathcal{C}\left\{\mathcal{O}^r(x)\varphi^a(x_1)\cdots\varphi^a(x_n)\right\}\rangle. \tag{4.74}$$

More generally, one can insert n_r number of r-type fields and $n - n_r$ number of a-type fields to construct general n-point causal response function

$$\begin{aligned} &R_n[\mathcal{O}^r(x_1), \cdots, \mathcal{O}^r(x_{n_r}); \varphi^a(x_{n_r+1}), \cdots, \varphi^a(x_n)] \\ &= \langle \mathcal{C}\left\{\mathcal{O}^r(x_1)\cdots\mathcal{O}^r(x_{n_r})\varphi^a(x_{n_r+1})\cdots\varphi^a(x_n)\right\}\rangle\,. \end{aligned} \tag{4.75}$$

The boundary condition in Eq. (4.15) becomes $\varphi^a(t = +\infty) = 0$ in the causal basis. Thanks to this, the largest time equation in Eq. (4.17) simply becomes the vanishing of correlation functions of all a fields.

The r/a basis serves as a natural basis for taking the semi-classical limit. One can think of the difference in the fields along two contours to be much smaller compared to their sum i.e. $\varphi^a = (\varphi^{\rm I} - \varphi^{\rm II}) \ll \frac{1}{2}\left(\varphi^{\rm I} + \varphi^{\rm II}\right) = \varphi^r$ in this limit. Thus, the action can be expanded in small φ^a

$$S[\varphi^{\rm I}] - S[\varphi^{\rm II}] = \frac{\delta S[\varphi]}{\delta\varphi}\Big|_{\varphi=\varphi_r}\varphi_a + \frac{\delta^3 S[\varphi]}{\delta\varphi^3}\Big|_{\varphi=\varphi_r}\varphi_a^3 + \cdots\,. \tag{4.76}$$

At the leading order in φ^a, we have

$$\int D\varphi_r D\varphi_a\, e^{i\frac{\delta S[\varphi]}{\delta\varphi}\big|_{\varphi=\varphi_r}\varphi_a}\,. \tag{4.77}$$

The path integral over φ^a can be performed exactly and it imposes the equation of motion for φ^r. For this reason, we will refer to all vertices linear in a fields as *classical vertices*. Since the exponent in Eq. (4.16) has to be odd in φ^a, the next-order term is cubic in φ^a, etc, which we will call *quantum vertices*.

The free-field propagators in the causal basis are [154]

$$\begin{pmatrix} G_{rr}(p) & G_{ra}(p) \\ G_{ar}(p) & G_{aa}(p) \end{pmatrix} = \begin{pmatrix} \frac{1}{2}\hat{\delta}(p^2 - m_i^2) & \frac{i}{p^2 - m_i^2 + i\epsilon p^0} \\ \frac{i}{p^2 - m_i^2 - i\epsilon p^0} & 0 \end{pmatrix}. \tag{4.78}$$

The vanishing of the G_{aa} propagator follows non-perturbatively from the largest time equation. Note the appearance of causal, that is, advanced and retarded, propagators

$$G_R(p) = G_A(-p) = \quad \xrightarrow{\quad\longrightarrow p\quad} \quad = \frac{i}{p^2 - m_i^2 + i\epsilon p^0}\,, \tag{4.79}$$

which make the causal properties of the amplitude manifest, as well as the "cut" (or Hadamard) propagator

$$G_{rr}(p) = \quad = \frac{1}{2}\hat{\delta}(p^2 - m_i^2)\,, \tag{4.80}$$

which we note does not include a positive-energy condition.

We are interested in momentum-space n-point amputated causal amplitudes with n_r number of r-type external fields, which can be computed using the following diagrammatic rules: On each diagram,

- Put r, a labels to the r or a-type external fields. The internal lines should be dressed with all possible assignments of r/a labels.
- Draw arrows such that all the causal flows are from a to r. In doing so, G_{rr} propagators can be treated as sources for the arrows.
- Align the directions of momenta of the heavy lines with the directions of the particle flow unless otherwise specified. Amputate the external lines.

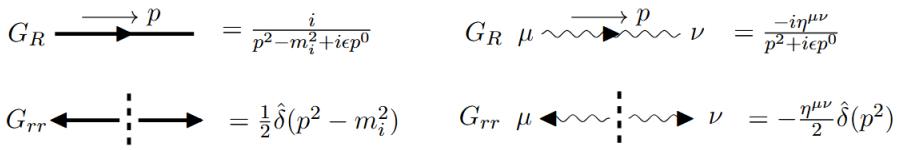
For the case of electrodynamics, the corresponding Feynman rules are given in Fig. 4.7.

The calculation of the observables in the r/a basis is simplified by a very useful counting rule derived in [141] which we quote without proof: for a n-point L-loop diagram with n_r number of r-type external legs, let us denote the number of rr propagators by P_{rr}, and the number of *additional* a-type fields at the internal vertices by n_a (which is zero for the classical vertex, and always even for the quantum ones). These are related by [141]

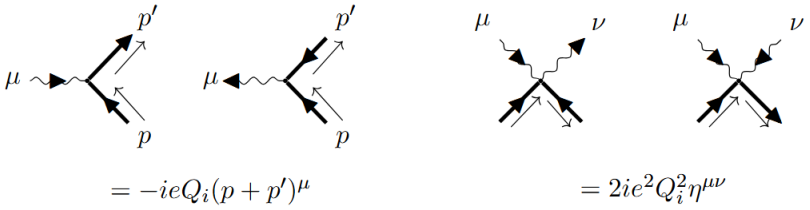
$$L + n_r - 1 = P_{rr} + n_a\,. \tag{4.81}$$

Crucially, the above counting only depends on the number of a and r-type fields, but not on their species (i.e whether they come from massive or massless fields in our context). It was argued in [141] using Eq. (4.81) that a n-point response function does not receive any contribution from the quantum vertices up to one loop. Indeed, inserting a cubic all-a vertex in a 1-loop diagram results in a closed loop of retarded propagators which vanishes due to causality. Thus, for the causal amplitudes with $n_r = 1$, Eq. (4.81) implies that the tree amplitudes are connected and the one-loop amplitudes have a single cut. More generally, it implies that diagrams with $n_a = 0$, that is, those which only feature the classical vertex (i.e., the vertex linear in φ^a) in Eq. (4.77), have at least one cut propagator per loop.

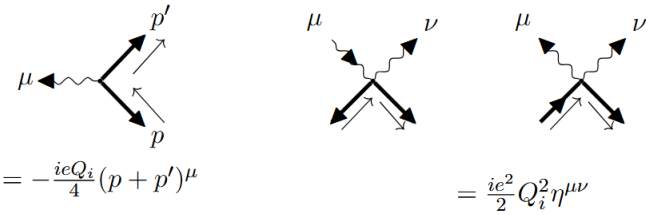
Finally, let us comment on the $i\epsilon$ prescription of the causal massive propagators. A crucial feature of the computations in the classical limit is the soft expansion of



(a) Bare propagators



(b) Classical vertices



(c) Quantum vertices

Figure 4.7: Feynman rules for scalar QED are shown above. The arrow on the retarded propagators G_R corresponds to the direction of the causal flow. The G_{rr} type propagators are represented by cuts. The a and r-type fields in the vertices are represented by outgoing and incoming arrows respectively. The directions of momenta going in or out of the vertices are aligned with the directions of particle flows shown by the thin arrows. The quantum vertices get an additional factor of $\frac{1}{4}$ owing to two additional a fields.

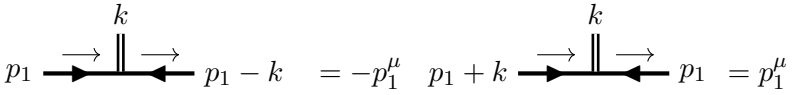
the matter propagators [103, 155] in the soft limit $\ell \sim q \sim \hbar \ll p_i$, with ℓ being a loop momentum. For a massive retarded propagator we get

$$\frac{1}{(p_i+\ell-q)^2-m_i^2+i\epsilon(p_i^0+\ell^0-q^0)} = \frac{1}{2p_i\cdot\ell+i\epsilon\, p_i^0}+\cdots\,. \tag{4.82}$$

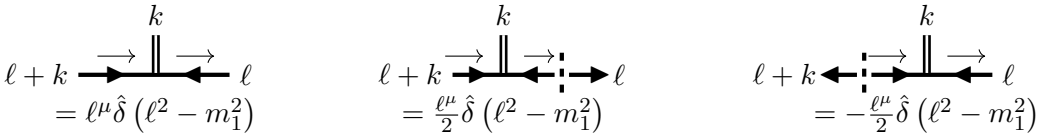
The key observation is that in the soft limit, the sign of $(p_i^0+\ell^0+q^0)$ is determined by the sign of the massive particle energy, p_i^0, since $\ell^0,\ q^0 \ll p_i^0$. Thus, the $i\epsilon$ prescription for the massive propagator is fixed by the causal flow of the heavy particle in the classical limit.

4.5.2 Linear impulse

In this section, we compute the impulse of particle 1 in the r/a basis. Let us begin by stating the Feynman rules for inserting the momentum operator into the internal and external lines in Fig. 4.8.



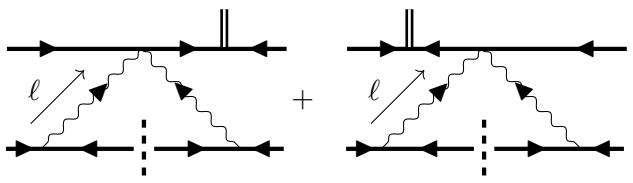
(a) Operator insertion on the external lines



(b) Operator insertion on an internal line

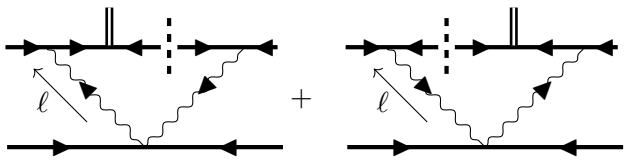
Figure 4.8: Five possible cases are shown when the momentum vertex is inserted on the external and internal lines. The directions of momenta of the heavy lines are denoted by the thin arrows. We take $k^\mu=(\omega,\mathbf{0})$ outgoing.

The derivation of these rules is similar to the ones described in sec 4.4. Notice, that the operator, represented by a double line, is r-type since causal arrows flow towards it. According to Eq. (4.81), we have single-cut diagrams at one-loop one external r leg. The simplest are the triangle diagrams



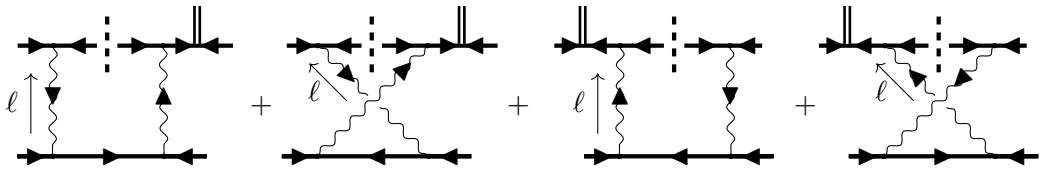
$$= 4ie^4Q_1^2Q_2^2p_1^2\int \hat{d}^4\ell\ \frac{\hat{\delta}(2p_2\cdot\ell)q^\mu}{\ell^2(\ell-q)^2}\,. \tag{4.83}$$

We have dropped the $i\epsilon$ for the photon propagators since they can not go on-shell at one-loop. For the same reason, we do not need to consider the diagrams where the photon propagators are cut. The inverted triangle can be found by replacing $1 \leftrightarrow 2$. However, for the inverted diagram, we have two additional diagrams that sum to zero:



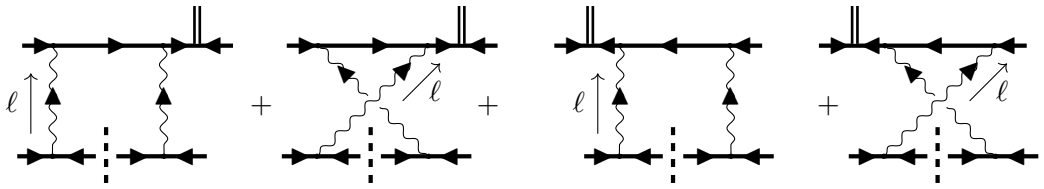
$$= 0, \tag{4.84}$$

thanks to the last two Feynman rules in Fig. 4.8b. Next, we consider the box and crossed-box diagrams. When the operator is inserted on the external lines, we have



$$= -i\mathcal{N} \int \hat{d}^4\ell \, \frac{\hat{\delta}(2p_1 \cdot \ell)}{\ell^2(\ell - q)^2} \left[\frac{(p_1 + q)^\mu 2q \cdot \ell}{(2p_2 \cdot \ell - i\epsilon)^2} - \frac{p_1^\mu 2q \cdot \ell}{(2p_2 \cdot \ell + i\epsilon)^2} \right], \tag{4.85}$$

with $\mathcal{N} = 8e^4 Q_1^2 Q_2^2 (p_1 \cdot p_2)^2$. Notice that leading-order terms from the expansion of the propagators cancel automatically due to the causal $i\epsilon$ prescription. We have also dropped all the scaleless integrals from the expansion for simplicity. Similarly,

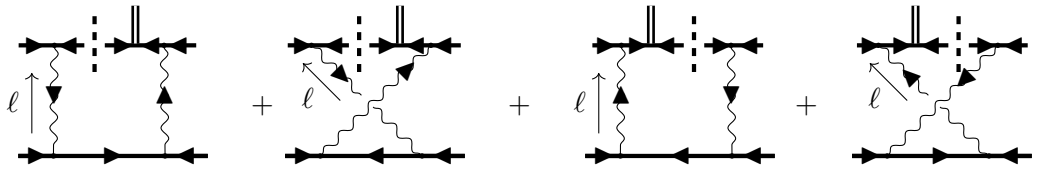


$$= -\mathcal{N} \int \hat{d}^4\ell \, \frac{\hat{\delta}(2p_2 \cdot \ell)}{\ell^2(\ell - q)^2} \left[\hat{\delta}(2p_1 \cdot \ell)(2p_1^\mu + q^\mu) + i \left(\frac{(p_1 + q)^\mu 2q \cdot \ell}{(2p_1 \cdot \ell - i\epsilon)^2} - \frac{p_1^\mu 2q \cdot \ell}{(2p_1 \cdot \ell + i\epsilon)^2} \right) \right], \tag{4.86}$$

where we have used the distributional identity

$$\frac{1}{x - i\epsilon} - \frac{1}{x + i\epsilon} = i\hat{\delta}(x)\,. \tag{4.87}$$

Next, we have diagrams from inserting the momentum operator on the internal lines:

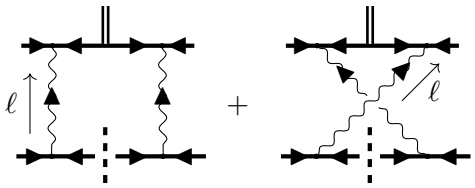


$$= \mathcal{N} \int \hat{d}^4\ell \, \frac{\hat{\delta}(2p_1 \cdot \ell)}{\ell^2 (\ell - q)^2} \hat{\delta}'(2p_2 \cdot \ell)(2q \cdot \ell)(p_1^\mu + \ell^\mu) \,, \tag{4.88}$$

where we have used

$$\frac{i}{(x - i\epsilon)^2} - \frac{i}{(x + i\epsilon)^2} = i\hat{\delta}'(x) \,. \tag{4.89}$$

Notice that the leading-order terms cancel similar to Eq. (4.85). Finally, we have



(4.90)

$$= \mathcal{N} \int \hat{d}^4\ell \, \frac{\hat{\delta}(2p_2 \cdot \ell)}{\ell^2 (\ell - q)^2} \left[\hat{\delta}(2p_1 \cdot \ell)(2p_1 + q)^\mu + \hat{\delta}'(2p_1 \cdot \ell)(p_1 + q - \ell)^\mu (2q \cdot \ell) \right] .$$

Adding all the contributions together, we see that the super-classical terms cancel manifestly. The sub-leading terms give the next to leading-order impulse

$$\begin{aligned} \Delta \mathbb{P}^\mu_{1,\mathrm{NLO}} = \frac{ie^4}{2} Q_1^2 Q_2^2 \int \frac{\hat{d}^4\ell}{\ell^2 (\ell - q)^2} \hat{d}^4 q \; \hat{\delta}\left(u_1 \cdot q\right) \hat{\delta}\left(u_2 \cdot q\right) e^{-ib \cdot q} \times \\ \times \left[q^\mu \left(\frac{\hat{\delta}\left(u_1 \cdot \ell\right)}{m_1} + \frac{\hat{\delta}\left(u_2 \cdot \ell\right)}{m_2} - \gamma^2 \left(\frac{\hat{\delta}\left(u_2 \cdot \ell\right)}{m_1} \frac{q \cdot \ell}{\left(u_1 \cdot \ell + i\epsilon\right)^2} + \frac{\hat{\delta}\left(u_1 \cdot \ell\right)}{m_2} \frac{q \cdot \ell}{\left(u_2 \cdot \ell - i\epsilon\right)^2} \right) \right) \\ + i\gamma^2 \ell^\mu \; q \cdot \ell \left(\frac{\hat{\delta}'\left(u_1 \cdot \ell\right)\left(u_2 \cdot \ell\right)}{m_1} - \frac{\hat{\delta}'\left(u_2 \cdot \ell\right)\left(u_1 \cdot \ell\right)}{m_2} \right) \Bigg], \end{aligned} \tag{4.91}$$

which agrees with [117] modulo scaleless integrals.

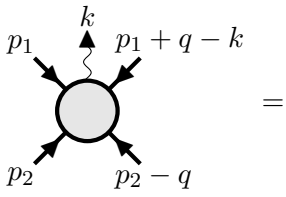
4.5.3 Waveform and its variance

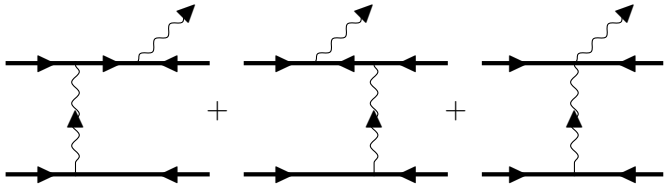
In this section, we will discuss the leading order waveform and its variance in the causal basis. As discussed in 4.3.2, the waveform is given by the amputated five-point response function in Eq. (4.32). The leading-order waveform in the III basis

involves $\mathcal{O}\left(q^{-3}\right)$ super-classical terms upon the expansion of the massive propagators, the cancellation of which formally requires including a cut with a three-particle amplitude: where we have used the on-shell condition $2p_1 \cdot q = 2p_1 \cdot k + \mathcal{O}\left(q^2\right)$, and Eq. (4.87). The analogous calculation in the r/a basis is given by the five-point amputated response function of one r-type photon field and four $a-$ type heavy fields

$$\mathcal{W}_k^h = \mathrm{LSZ}_x^{\mathrm{out}} \prod_{i=1,2} \mathrm{LSZ}_{x_i'}^{\mathrm{out}} \, \mathrm{LSZ}_{x_i}^{\mathrm{in}} R_5[A^{r\,h}(x); \Phi_1^{a\dagger}(x_1), \Phi_2^{a\dagger(x_2)}, \Phi_1^a(x_1'), \Phi_2^a(x_2')]. \tag{4.92}$$

At the leading order, we have the following diagrams

$\mathcal{W}_k^h(q,k) =$ 

$=$ 
$+ \ (1 \leftrightarrow 2)\,.$ (4.93)

The $\mathcal{O}\left(q^{-3}\right)$ terms in Eq. (4.93) are

$$-8ie^3 Q_1^2 Q_2 \frac{(p_1 \cdot p_2)(p_1 \cdot \varepsilon^h)}{q^2} \left[\frac{1}{2p_1 \cdot q + i\epsilon} + \frac{1}{-2p_1 \cdot k - i\epsilon}\right] = 0. \tag{4.94}$$

where we have used the on-shell condition $2p_1 \cdot q = 2p_1 \cdot k + \mathcal{O}\left(q^2\right)$. Similar to the computation of the linear impulse in the causal basis, the $\mathcal{O}\left(q^{-3}\right)$ terms, which are singular as $\hbar \to 0$, manifestly cancel, thanks to the causal $i\epsilon$ prescription. We have also verified that the subleading terms of these diagrams in the soft expansion of the waveform yield precisely the classical answer.

So far we have considered causal response functions with one r type field. However, correlation functions with multiple r-type insertions which measure correlations between different measurements after a system is perturbed, are also familiar objects in non-linear response theory. The simplest of them in our context is the six-point amputated response function with two on-shell r-type photon fields and four a-type massive fields. Explicitly, we want to compute

$$\begin{aligned}
&\langle 1'2'|\, a_{k_1}^{h_1\,\mathrm{out}} a_{k_2}^{h_2\,\mathrm{out}} \,|12\rangle \\
&= \prod_{i=1}^{2} \mathrm{LSZ}_{x_i'}^{\mathrm{in}} \prod_{j=1}^{2} \mathrm{LSZ}_{x_j}^{\mathrm{out}} \prod_{k=1}^{2} \mathrm{LSZ}_{z_k}^{\mathrm{out}} R_6 \left[A^{rh_1}(z_1), A^{rh_2}(z_2); \Phi_1^{a\dagger}(x_1), \Phi_2^{a\dagger}(x_2), \Phi_1^a(x_1'), \Phi_2^a(x_2')\right].
\end{aligned} \tag{4.95}$$

The expectation value in Eq. (4.95) is used to compute the variance of the waveform in [150] which can be defined as

$$\mathcal{V}(k_1,k_2) = \langle 1'2'|\, a_{k_1}^{h_1\,\text{out}} a_{k_2}^{h_2\,\text{out}} \,|12\rangle\,\big|_{\text{cl.}} - \langle 1'2'|\, a_{k_1}^{h_1\,\text{out}} \,|12\rangle\,\big|_{\text{cl.}}\, \langle 1'2'|\, a_{k_2}^{h_2\,\text{out}} \,|12\rangle\,\big|_{\text{cl.}}. \tag{4.96}$$

appropriately integrated against classical wavepackets for the heavy particles. We refer the reader to [150] for details. Here, we concentrate on the properties of the six-point amputated response function in Eq. (4.95) in the classical limit. It was shown in [150] that at the tree-level the expectation value in Eq. (4.96) is quantum. More specifically, working in a gauge

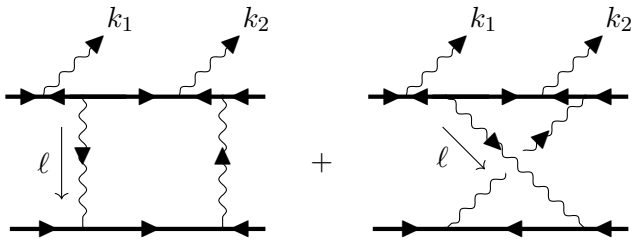
$$p_1 \cdot \epsilon^{h_i}(k_i) = 0, \quad i = 1,2. \tag{4.97}$$

it was shown that the $\mathcal{O}\left(q^{-4}\right)$, $\mathcal{O}\left(q^{-3}\right)$ diagrams in the six-point tree amplitudes cancel, and the leading order six-point connected tree amplitude is, in fact, $\mathcal{O}\left(q^{-2}\right)$. The particular gauge choice was crucial to simplify the analysis in [150].

In this section, we demonstrate the advantage of the causal r/a basis for the calculation of the variance. We will present our result in full generality without adopting any particular gauge. The counting rule in Eq. (4.81) for $n_r = 2$ and $n_a = 0$ gives $P_{rr} = 1$ at the tree level. Therefore, at leading order one can only draw various single-cut diagrams that isolate a three-point on-shell amplitude or diagrams with horizontal cuts as shown below. When the external photons are on-shell, the diagrams with three-particle cuts vanish due to the stability condition. As the t-channel is spacelike in the classical limit, the diagrams with horizontal cuts do not contribute to the classical results. Thus, the connected part in Eq. (4.13) at the leading order is entirely quantum. Notice that the configurations in which external photons are emitted from the quartic vertex in QED (or a cubic graviton vertex in gravity) are prohibited by the Feynman rules in the r/a basis. This is because such emissions must originate from quantum vertices, which cannot contribute at tree level as shown by the counting rule in Eq. (4.81). In the usual in-out formalism, one must include such diagrams. This analysis is rather involved in the I/II basis since we need to involve various cut diagrams. For instance, at leading order in the classical expansion, the following diagrams, which include a Compton scattering sub-process, must cancel among themselves

At one-loop we can either have diagrams with $P_{rr} = 2,\ n_a = 0$ or $P_{rr} = 0,\ n_a = 2$. The combination of these diagrams leaves behind factorized six-point diagrams in the classical limit. We discuss one such example and leave the rest as an exercise for the reader. Consider the box topology with two photon emissions. The cut diagram at the leading order is where the outgoing heavy momenta are $p_1 + q - k_1 - k_2$ and $p_2 - q$. The ellipses include various cut diagrams such as cross-boxes and the photon emission from various internal and external lines. In the classical limit, this becomes a convolution of two leading-order waveforms, thereby demonstrating the classical factorization of such an observable after Fourier transform to impact

parameter space and the time domain. Finally, we have two diagrams of the same topology involving quantum vertices



The leading order terms in the classical limit cancel due to the causal $i\epsilon$ prescription. To see this, notice the two diagrams only differ by the bottom matter propagator

$$\frac{1}{(p_2+\ell)^2-m_2^2+i\epsilon}+\frac{1}{(\ell+q-p_2)^2-m_2^2-i\epsilon}=\mathcal{O}(q^0,\ell^0)\,. \tag{4.98}$$

Therefore, the connected diagrams with quantum vertices are subleading compared to the diagram involving only classical vertices.

4.6 Discussions

In this chapter, we have explained how to compute the change in the expectation value of a generic asymptotic observable $\mathbb{O}$ during a scattering process from an amputated causal response function of its local density $\mathcal{O}(x)$, computed using the in-in formalism. For instance, the soft limit of a five-point response function of the stress-energy tensor of a particle i with four amputated legs computes the linear impulse $\Delta\mathbb{P}_i$ of particle i during scattering with an initial two-particle state. We demonstrate that this procedure is exactly equivalent to the KMOC formalism, thus generalizing the findings of Ref. [139].

Although our formalism is general, we have applied it to the computation of classical observables. We find that the soft limit acts as a useful regulator which carefully parameterizes the approach to the asymptotic past and future. This becomes particularly important for certain observables, such as the angular momentum loss, which receive contribution from the long-range interactions present in theories with massless particles. As an example, we computed the angular momentum loss in scalar QED, for which the soft frequency unveils a new region in the loop integration that yields unambiguously the leading contribution to this quantity. It would be interesting to use our method to reproduce the angular momentum loss in gravity at $\mathcal{O}\left(G^3\right)$ from [143], and produce new results beyond this order. Furthermore, we expect that similar considerations will be important to understand other phenomena related to long-range forces, such as violation of peeling [70, 156] from the perspective of scattering amplitudes.

We also have initiated a study of classical observables from amplitudes with retarded propagators within the Schwinger-Keldysh formalism. This approach ensures that causality is manifest at every step of the calculation, setting it apart from the conventional method of computing classical observables from in-out scattering amplitudes with Feynman propagators. We find that doing the calculations in this way, terms which are singular as $\hbar \to 0$ cancel manifestly for the linear impulse at one-loop, and for waveform and its variance at tree-level. In particular, for the variance, the classical vertices naturally give rise to factorized diagrams as expected for a classical quantity. It would be interesting to extend our analysis beyond one loop. In particular, starting at two loops, both classical and quantum vertices contribute. We expect the contributions from the diagrams with quantum vertices to be $q \sim \hbar$ suppressed compared to those with classical vertices at all loops. We leave the explicit checks of this statement for future work.

Chapter 5

Conclusion

In this thesis, we discussed three instances of zero modes or soft modes which are the zero-energy excitations in quantum field theory. We discussed a condensed matter system, infrared divergence in a scalar field theory, and finally we discussed how soft modes encode the classical physics within quantum field theories.

In chapter 2, we discussed a continuum boundary quantum field theory that describes a semi-infinite graphene sheet. We discussed the linear boundary conditions that can be imposed on the field and analyzed the compatibility of the boundary conditions with the discrete symmetries of the system. Notably, under a specific boundary condition involving zigzag states, the system exhibits scale invariance while allowing for zero-energy states localized at the edge, leading to significant ground-state degeneracy. By leveraging this scale invariance, we identified two distinct ground states that suggest a ferromagnetic edge. One of these states is well established in the lattice model, while our analysis predicts the existence of an additional state, which we refer to as the *valley ferromagnetic* state.

In chapter 3, we first discuss various issues with infrared divergence in QFTs and its possible connections with symmetries. Then, we presented a scalar model that mimics the infrared structure of gauge theories such as QED and quantum gravity. This example illustrates that having an infrared divergence in a theory does not point to any special symmetry structure of the theory. However, we identified where the scalar theory differs from the gauge theories. In contrast with the gauge theories, the soft scalars remain coupled to dressed states and are inevitably generated during interactions involving hard dressed particles. However, we found that the entanglement between the hard particles and the resulting soft particles is negligibly small.

In chapter 4, we developed a new technique to compute classical observables from amputated response functions. Our new method supplements the existing literature of computing classical observables from the classical limit of quantum field theories. We showed that such computations can be performed by exploiting the soft-limit of a generally off-shell local operator inserted on an amputated causal response function. Using this method, we compute the classical angular momentum carried away by the radiated photon in electromagnetism during a binary interaction of point charges. We also employed a manifestly causal basis to perform such classical calculations in the quantum field theory setup. We established that our method overcomes various shortcomings of the existing methods.

In this thesis, we have examined the physics of soft modes in three distinct con-

texts that initially seem unrelated. To conclude, we comment on some potential connections between the discussions in Chapters 3 and 4, specifically concerning infrared divergences and the computation of classical observables from quantum field theory. The dressed states proposed by Faddeev and Kulish [58] ensure a well-defined, finite S-matrix. However, this requires introducing infrared-divergent phase factors and adopting a modified definition of the S-matrix. To the best of the author's knowledge, no existing framework naturally incorporates these divergent phase integrals into the definition of hard dressing. Interestingly, these integrals arise naturally when the classical wave packets in (4.20) exchange one soft particle in a $2 \to 2$ tree-level amplitude involving four of these dressed particles. A more in-depth investigation may lead to a dressing principle for hard particles that extends the Faddeev-Kulish approach, providing a resolution to infrared divergences in both magnitude and phase. We also anticipate the correspondence to be bidirectional—suggesting that Faddeev-Kulish-like dressed states should naturally serve as candidates for computing classical observables. This is because they incorporate a cloud of soft particles and soft particles are known to determine the classical physics within scattering amplitudes[15]. A closer exploration of these aspects is left for future investigation.

[15]We direct the reader to equation (4) of [157] where the state closely resembles a Faddeev-Kulish dressed state.

Bibliography

[1] Shovon Biswas and Gordon W. Semenoff. Massless fermions on a half-space: the curious case of 2+1-dimensions. *JHEP*, 10:045, 2022.

[2] Shovon Biswas and Gordon W. Semenoff. Soft scalars do not decouple. *Phys. Rev. D*, 106(10):105023, 2022.

[3] Shovon Biswas and Julio Parra-Martinez. Classical Observables from Causal Response Functions. 11 2024.

[4] Peter W. Higgs. Broken Symmetries and the Masses of Gauge Bosons. *Phys. Rev. Lett.*, 13:508–509, 1964.

[5] Georges Aad et al. Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. *Phys. Lett. B*, 716:1–29, 2012.

[6] Serguei Chatrchyan et al. Observation of a New Boson at a Mass of 125 GeV with the CMS Experiment at the LHC. *Phys. Lett. B*, 716:30–61, 2012.

[7] Eduardo Fradkin. *Field Theories of Condensed Matter Physics*. Cambridge University Press, 2 2013.

[8] Edward Witten. Fermion Path Integrals And Topological Phases. *Rev. Mod. Phys.*, 88(3):035001, 2016.

[9] R. Jackiw and C. Rebbi. Solitons with Fermion Number 1/2. *Phys. Rev. D*, 13:3398–3409, 1976.

[10] B. P. Abbott et al. Observation of Gravitational Waves from a Binary Black Hole Merger. *Phys. Rev. Lett.*, 116(6):061102, 2016.

[11] Niayesh Afshordi et al. Waveform Modelling for the Laser Interferometer Space Antenna. 11 2023.

[12] A. H. Castro Neto, F. Guinea, N. M. R. Peres, K. S. Novoselov, and A. K. Geim. The electronic properties of graphene. *Rev. Mod. Phys.*, 81:109–162, 2009.

[13] Gordon W. Semenoff. Condensed-matter simulation of a three-dimensional anomaly. *Physical Review Letters*, 53(26):2449–2452, dec 1984.

[14] A.J. Niemi and G.W. Semenoff. Fermion number fractionization in quantum field theory. *Physics Reports*, 135(3):99–193, apr 1986.

[15] F. D. M. Haldane. Model for a Quantum Hall Effect without Landau Levels: Condensed-Matter Realization of the 'Parity Anomaly'. *Phys. Rev. Lett.*, 61:2015–2018, 1988.

[16] Gordon W. Semenoff. Engineering holographic graphene. *AIP Conf. Proc.*, 1483(1):305–329, 2012.

[17] Jean-Noel Fuchs and Mark Oliver Goerbig. Introduction to the physical properties of graphene. *Lecture notes*, 10(11), 2008.

[18] R. Jackiw and C. Rebbi. Solitons with fermion number $\frac{1}{2}$. *Physical Review D*, 13(12):3398–3409, jun 1976.

[19] Masatoshi Sato, Yoshiro Takahashi, and Satoshi Fujimoto. Non-abelian topological order in s-wave superfluids of ultracold fermionic atoms. *Physical review letters*, 103(2):020401, 2009.

[20] Mitsutaka Fujita, Katsunori Wakabayashi, Kyoko Nakada, and Koichi Kusakabe. Peculiar localized state at zigzag graphite edge. *Journal of the Physical Society of Japan*, 65(7):1920–1923, jul 1996.

[21] Hamed Karimi and Ian Affleck. Towards a rigorous proof of magnetism on the edges of graphene nanoribbons. *Physical Review B*, 86(11):115446, sep 2012.

[22] Rahul M. Nandkishore and Michael Hermele. Fractons. *Ann. Rev. Condensed Matter Phys.*, 10:295–313, 2019.

[23] Michael Pretko, Xie Chen, and Yizhi You. Fracton Phases of Matter. *Int. J. Mod. Phys. A*, 35(06):2030003, 2020.

[24] Nathan Seiberg and Shu-Heng Shao. Majorana chain and Ising model – (non-invertible) translations, anomalies, and emanant symmetries. *SciPost Phys.*, 16:064, 2024.

[25] Arkya Chatterjee, Salvatore D. Pace, and Shu-Heng Shao. Quantized Axial Charge of Staggered Fermions and the Chiral Anomaly. *Phys. Rev. Lett.*, 134(2):021601, 2025.

[26] Sudipta Dutta and Swapan K. Pati. Novel properties of graphene nanoribbons: a review. *Journal of Materials Chemistry*, 20(38):8207, 2010.

[27] Haomin Wang, Hui Shan Wang, Chuanxu Ma, Lingxiu Chen, Chengxin Jiang, Chen Chen, Xiaoming Xie, An-Ping Li, and Xinran Wang. Graphene nanoribbons for quantum electronics. *Nature Reviews Physics*, 3(12):791–802, sep 2021.

[28] Gordon W. Semenoff and Fei Zhou. Magnetic catalysis and quantum Hall ferromagnetism in weakly coupled graphene. *Journal of High Energy Physics*, 2011(7), jul 2011.

[29] T.O. Wehling, A.M. Black-Schaffer, and A.V. Balatsky. Dirac materials. *Advances in Physics*, 63(1):1–76, jan 2014.

[30] A. R. Akhmerov and C. W. J. Beenakker. Boundary conditions for Dirac fermions on a terminated honeycomb lattice. *Physical Review B*, 77(8):085423, feb 2008.

[31] J. A. M. van Ostaay, A. R. Akhmerov, C. W. J. Beenakker, and M. Wimmer. Dirac boundary condition at the reconstructed zigzag edge of graphene. *Physical Review B*, 84(19):195434, nov 2011.

[32] Kyoko Nakada, Mitsutaka Fujita, Gene Dresselhaus, and Mildred S. Dresselhaus. Edge state in graphene ribbons: Nanometer size effect and edge shape dependence. *Physical Review B*, 54(24):17954–17961, dec 1996.

[33] Y. Niimi, T. Matsui, H. Kambara, K. Tagami, M. Tsukada, and Hiroshi Fukuyama. Scanning tunneling microscopy and spectroscopy of the electronic local density of states of graphite surfaces near monoatomic step edges. *Physical Review B*, 73(8):085421, feb 2006.

[34] Yousuke Kobayashi, Ken ichi Fukui, Toshiaki Enoki, and Koichi Kusakabe. Edge state on hydrogen-terminated graphite edges investigated by scanning tunneling microscopy. *Physical Review B*, 73(12):125415, mar 2006.

[35] Zheng Liu, Kazu Suenaga, Peter J. F. Harris, and Sumio Iijima. Open and closed edges of graphene layers. *Physical Review Letters*, 102(1):015501, jan 2009.

[36] Kazu Suenaga and Masanori Koshino. Atom-by-atom spectroscopy at graphene edge. *Nature*, 468(7327):1088–1090, dec 2010.

[37] J. Jung, T. Pereg-Barnea, and A. H. MacDonald. Theory of interedge superexchange in zigzag edge magnetism. *Physical Review Letters*, 102(22):227205, jun 2009.

[38] J. Jung and A. H. MacDonald. Carrier density and magnetism in graphene zigzag nanoribbons. *Physical Review B*, 79(23):235433, jun 2009.

[39] Young-Woo Son, Marvin L. Cohen, and Steven G. Louie. Energy gaps in graphene nanoribbons. *Physical Review Letters*, 97(21):216803, nov 2006.

[40] Young-Woo Son, Marvin L. Cohen, and Steven G. Louie. Half-metallic graphene nanoribbons. *Nature*, 444(7117):347–349, nov 2006.

[41] Toshiya Hikihara, Xiao Hu, Hsiu-Hau Lin, and Chung-Yu Mou. Ground-state properties of nanographite systems with zigzag edges. *Physical Review B*, 68(3):035432, jul 2003.

[42] Sudipta Dutta, S. Lakshmi, and Swapan K. Pati. Electron-electron interactions on the edge states of graphene: A many-body configuration interaction study. *Physical Review B*, 77(7):073412, feb 2008.

[43] Hélène Feldner, Zi Yang Meng, Andreas Honecker, Daniel Cabra, Stefan Wessel, and Fakher F. Assaad. Magnetism of finite graphene samples: Mean-field theory compared with exact diagonalization and quantum monte carlo simulations. *Physical Review B*, 81(11):115416, mar 2010.

[44] Zheng Shi and Ian Affleck. Effect of long-range interaction on graphene edge magnetism. *Physical Review B*, 95(19):195420, may 2017.

[45] Gordon Walter Semenoff. *Quantum Field Theory.* Springer, 2023.

[46] Daniel E. Sheehy and Jörg Schmalian. Quantum critical scaling in graphene. *Physical Review Letters*, 99(22):226803, nov 2007.

[47] D. T. Son. Quantum critical point in graphene approached in the limit of infinitely strong coulomb interaction. *Physical Review B*, 75(23):235423, jun 2007.

[48] N Andrei, A Bissi, M Buican, J Cardy, P Dorey, N Drukker, J Erdmenger, D Friedan, D Fursaev, A Konechny, C Kristjansen, I Makabe, Y Nakayama, A O'Bannon, R Parini, B Robinson, S Ryu, C Schmidt-Colinet, V Schomerus, C Schweigert, and G M T Watts. Boundary and defect CFT: open problems and applications. *Journal of Physics A: Mathematical and Theoretical*, 53(45):453002, oct 2020.

[49] Yoshiki Sato. Free energy and defect c-theorem in free fermion. *Journal of High Energy Physics*, 2021(5), may 2021.

[50] Steven Weinberg. *The Quantum Theory of Fields, Volume 1.* Cambridge University Press, 2005.

[51] F. E. Low. Scattering of light of very low frequency by systems of spin 1/2. *Phys. Rev.*, 96:1428–1432, 1954.

[52] F. E. Low. Bremsstrahlung of very low-energy quanta in elementary particle collisions. *Phys. Rev.*, 110:974–977, 1958.

[53] T. H. Burnett and Norman M. Kroll. Extension of the low soft photon theorem. *Phys. Rev. Lett.*, 20:86, 1968.

[54] Steven Weinberg. Infrared photons and gravitons. *Physical Review*, 140(2B):B516–B524, oct 1965.

[55] F. Bloch and A. Nordsieck. Note on the radiation field of the electron. *Physical Review*, 52(2):54–59, jul 1937.

[56] D.R Yennie, S.C Frautschi, and H Suura. The infrared divergence phenomena and high-energy processes. *Annals of Physics*, 13(3):379–452, jun 1961.

[57] Victor Chung. Infrared divergence in quantum electrodynamics. *Physical Review*, 140(4B):B1110–B1122, nov 1965.

[58] P. P. Kulish and L. D. Faddeev. Asymptotic conditions and infrared divergences in quantum electrodynamics. *Theoretical and Mathematical Physics*, 4(2):745–757, 1971.

[59] Holmfridur Hannesdottir and Matthew D. Schwartz. S -Matrix for massless particles. *Phys. Rev. D*, 101(10):105001, 2020.

[60] Daniel Carney, Laurent Chaurette, Dominik Neuenfeld, and Gordon Walter Semenoff. Infrared quantum information. *Physical Review Letters*, 119(18):180502, oct 2017.

[61] Daniel Carney, Laurent Chaurette, Dominik Neuenfeld, and Gordon Walter Semenoff. Dressed infrared quantum information. *Physical Review D*, 97(2):025007, jan 2018.

[62] Daniel Carney, Laurent Chaurette, Dominik Neuenfeld, and Gordon Semenoff. On the need for soft dressing. *Journal of High Energy Physics*, 2018(9), sep 2018.

[63] Andrew Strominger. *Lectures on the Infrared Structure of Gravity and Gauge Theory*. Princeton University Press.

[64] Noah Miller. From Noether's Theorem to Bremsstrahlung: a pedagogical introduction to large gauge transformations and classical soft theorems. 12 2021.

[65] Clifford Cheung, Andreas Helset, and Julio Parra-Martinez. Geometric soft theorems. *Journal of High Energy Physics*, 2022(4), apr 2022.

[66] Rodrigo Alonso, Elizabeth E. Jenkins, and Aneesh V. Manohar. A Geometric Formulation of Higgs Effective Field Theory: Measuring the Curvature of Scalar Field Space. *Phys. Lett. B*, 754:335–342, 2016.

[67] Andreas Helset, Elizabeth E. Jenkins, and Aneesh V. Manohar. Geometry in scattering amplitudes. *Phys. Rev. D*, 106(11):116018, 2022.

[68] Maria Derda, Andreas Helset, and Julio Parra-Martinez. Soft scalars in effective field theory. *JHEP*, 06:133, 2024.

[69] Y. B. Zel'dovich and A. G. Polnarev. Radiation of gravitational waves by a cluster of superdense stars. *Sov. Astron.*, 18:17, 1974.

[70] Demetrios Christodoulou. The global initial value problem in general relativity. *The Ninth Marcel Grossmann Meeting*, pages 44–54, 2002.

[71] Lydia Bieri and David Garfinkle. An electromagnetic analogue of gravitational wave memory. *Class. Quant. Grav.*, 30:195009, 2013.

[72] Alexander Tolish and Robert M. Wald. Cosmological memory effect. *Phys. Rev. D*, 94(4):044009, 2016.

[73] Abhay Ashtekar, Miguel Campiglia, and Alok Laddha. Null infinity, the BMS group and infrared issues. *Gen. Rel. Grav.*, 50(11):140–163, 2018.

[74] Kartik Prabhu, Gautam Satishchandran, and Robert M Wald. Infrared finite scattering theory in quantum field theory and quantum gravity. *arXiv: 2203.14334*, 2022.

[75] Andrew Strominger and Alexander Zhiboedov. Gravitational Memory, BMS Supertranslations and Soft Theorems. *JHEP*, 01:086, 2016.

[76] Kartik Prabhu and Gautam Satishchandran. Infrared finite scattering theory: Amplitudes and soft theorems. *Phys. Rev. D*, 110(8):085022, 2024.

[77] T. W. B. Kibble. Coherent soft-photon states and infrared divergences. i. classical currents. *Journal of Mathematical Physics*, 9(2):315–324, feb 1968.

[78] John Ware, Ryo Saotome, and Ratindranath Akhoury. Construction of an asymptotic s matrix for perturbative quantum gravity. *Journal of High Energy Physics*, 2013(10), oct 2013.

[79] Marc Henneaux and Cédric Troessaert. Asymptotic structure of a massless scalar field and its dual two-form field at spatial infinity. *Journal of High Energy Physics*, 2019(5), may 2019.

[80] Miguel Campiglia, Leonardo Coito, and Sebastian Mizera. Can scalars have asymptotic symmetries? *Physical Review D*, 97(4):046002, feb 2018.

[81] Miguel Campiglia and Leonardo Coito. Asymptotic charges from soft scalars in even dimensions. *Physical Review D*, 97(6):066009, mar 2018.

[82] J Fröhlich, G Morchio, and F Strocchi. Charged sectors and scattering states in quantum electrodynamics. *Annals of Physics*, 119(2):241–284, jun 1979.

[83] J. Fröhlich, G. Morchio, and F. Strocchi. Infrared problem and spontaneous breaking of the lorentz group in QED. *Physics Letters B*, 89(1):61–64, dec 1979.

[84] A. P. Balachandran and S. Vaidya. Spontaneous Lorentz violation in gauge theories. *The European Physical Journal Plus*, 128(10), oct 2013.

[85] Sangmin Choi and Ratindranath Akhoury. Subleading soft dressings of asymptotic states in QED and perturbative quantum gravity. *Journal of High Energy Physics*, 2019(9), sep 2019.

[86] Mehrdad Mirbabayi and Massimo Porrati. Dressed hard states and black hole soft hair. *Physical Review Letters*, 117(21):211301, nov 2016.

[87] Stephen W. Hawking, Malcolm J. Perry, and Andrew Strominger. Soft hair on black holes. *Physical Review Letters*, 116(23):231301, jun 2016.

[88] Andrew Strominger. Black hole information revisited. In *Jacob Bekenstein*, pages 109–117. WORLD SCIENTIFIC, sep 2019.

[89] Lorenzo Speri, Nikolaos Karnesis, Arianna I. Renzini, and Jonathan R. Gair. A roadmap of gravitational wave data analysis. *Nature Astron.*, 6(12):1356–1363, 2022.

[90] A. Buonanno and T. Damour. Effective one-body approach to general relativistic two-body dynamics. *Phys. Rev. D*, 59:084006, 1999.

[91] M. Punturo et al. The Einstein Telescope: A third-generation gravitational wave observatory. *Class. Quant. Grav.*, 27:194002, 2010.

[92] David Reitze et al. Cosmic Explorer: The U.S. Contribution to Gravitational-Wave Astronomy beyond LIGO. *Bull. Am. Astron. Soc.*, 51(7):035, 2019.

[93] K. Danzmann and A. Rudiger. LISA technology - Concept, status, prospects. *Class. Quant. Grav.*, 20:S1–S9, 2003.

[94] Steven Weinberg. *Gravitation and Cosmology: Principles and Applications of the General Theory of Relativity.* John Wiley and Sons, New York, 1972.

[95] John F. Donoghue. General relativity as an effective field theory: The leading quantum corrections. *Phys. Rev. D*, 50:3874–3888, 1994.

[96] John F. Donoghue. Leading quantum correction to the Newtonian potential. *Phys. Rev. Lett.*, 72:2996–2999, 1994.

[97] Barry R. Holstein and John F. Donoghue. Classical physics and quantum loops. *Phys. Rev. Lett.*, 93:201602, 2004.

[98] Zvi Bern, Clifford Cheung, Radu Roiban, Chia-Hsien Shen, Mikhail P. Solon, and Mao Zeng. Scattering Amplitudes and the Conservative Hamiltonian for Binary Systems at Third Post-Minkowskian Order. *Phys. Rev. Lett.*, 122(20):201603, 2019.

[99] Zvi Bern, Clifford Cheung, Radu Roiban, Chia-Hsien Shen, Mikhail P. Solon, and Mao Zeng. Black Hole Binary Dynamics from the Double Copy and Effective Theory. *JHEP*, 10:206, 2019.

[100] Zvi Bern, Julio Parra-Martinez, Radu Roiban, Michael S. Ruf, Chia-Hsien Shen, Mikhail P. Solon, and Mao Zeng. Scattering Amplitudes and Conservative Binary Dynamics at $\mathcal{O}(G^4)$. *Phys. Rev. Lett.*, 126(17):171601, 2021.

[101] Zvi Bern, Julio Parra-Martinez, Radu Roiban, Michael S. Ruf, Chia-Hsien Shen, Mikhail P. Solon, and Mao Zeng. Scattering Amplitudes, the Tail Effect, and Conservative Binary Dynamics at $\mathcal{O}(G^4)$. *Phys. Rev. Lett.*, 128(16):161103, 2022.

[102] Enrico Herrmann, Julio Parra-Martinez, Michael S. Ruf, and Mao Zeng. Gravitational Bremsstrahlung from Reverse Unitarity. *Phys. Rev. Lett.*, 126(20):201602, 2021.

[103] Enrico Herrmann, Julio Parra-Martinez, Michael S. Ruf, and Mao Zeng. Radiative classical gravitational observables at $\mathcal{O}(\mathrm{G}^3)$ from scattering amplitudes. *JHEP*, 10:148, 2021.

[104] Paolo Di Vecchia, Carlo Heissenberg, Rodolfo Russo, and Gabriele Veneziano. Radiation Reaction from Soft Theorems. *Phys. Lett. B*, 818:136379, 2021.

[105] Paolo Di Vecchia, Carlo Heissenberg, Rodolfo Russo, and Gabriele Veneziano. The eikonal approach to gravitational scattering and radiation at $\mathcal{O}(\mathrm{G}^3)$. *JHEP*, 07:169, 2021.

[106] Paolo Di Vecchia, Carlo Heissenberg, Rodolfo Russo, and Gabriele Veneziano. The eikonal operator at arbitrary velocities I: the soft-radiation limit. *JHEP*, 07:039, 2022.

[107] Paolo Di Vecchia, Carlo Heissenberg, Rodolfo Russo, and Gabriele Veneziano. Classical gravitational observables from the Eikonal operator. *Phys. Lett. B*, 843:138049, 2023.

[108] Andreas Brandhuber, Graham R. Brown, Gang Chen, Stefano De Angelis, Joshua Gowdy, and Gabriele Travaglini. One-loop gravitational bremsstrahlung and waveforms from a heavy-mass effective field theory. *JHEP*, 06:048, 2023.

[109] Aidan Herderschee, Radu Roiban, and Fei Teng. The sub-leading scattering waveform from amplitudes. *JHEP*, 06:004, 2023.

[110] Alessandro Georgoudis, Carlo Heissenberg, and Rodolfo Russo. An eikonal-inspired approach to the gravitational scattering waveform. *JHEP*, 03:089, 2024.

[111] Alessandro Georgoudis, Carlo Heissenberg, and Rodolfo Russo. Post-Newtonian multipoles from the next-to-leading post-Minkowskian gravitational waveform. *Phys. Rev. D*, 109(10):106020, 2024.

[112] Giacomo Brunello and Stefano De Angelis. An improved framework for computing waveforms. *JHEP*, 07:062, 2024.

[113] Donato Bini, Thibault Damour, Stefano De Angelis, Andrea Geralico, Aidan Herderschee, Radu Roiban, and Fei Teng. Gravitational waveforms: A tale of two formalisms. *Phys. Rev. D*, 109(12):125008, 2024.

[114] Mikhail M. Ivanov, Yue-Zhou Li, Julio Parra-Martinez, and Zihan Zhou. Gravitational Raman Scattering in Effective Field Theory: A Scalar Tidal Matching at O(G3). *Phys. Rev. Lett.*, 132(13):131401, 2024.

[115] Rafael A. Porto. The effective field theorist's approach to gravitational dynamics. *Phys. Rept.*, 633:1–104, 2016.

[116] Luc Blanchet. Gravitational Radiation from Post-Newtonian Sources and Inspiralling Compact Binaries. *Living Rev. Rel.*, 17:2, 2014.

[117] David A. Kosower, Ben Maybee, and Donal O'Connell. Amplitudes, Observables, and Classical Scattering. *JHEP*, 02:137, 2019.

[118] Z. Bern, J. J. M. Carrasco, and Henrik Johansson. New Relations for Gauge-Theory Amplitudes. *Phys. Rev. D*, 78:085011, 2008.

[119] Zvi Bern, John Joseph Carrasco, Marco Chiodaroli, Henrik Johansson, and Radu Roiban. The duality between color and kinematics and its applications. *J. Phys. A*, 57(33):333002, 2024.

[120] Zvi Bern, Lance J. Dixon, David C. Dunbar, and David A. Kosower. One loop n point gauge theory amplitudes, unitarity and collinear limits. *Nucl. Phys. B*, 425:217–260, 1994.

[121] Ruth Britto, Freddy Cachazo, and Bo Feng. Generalized unitarity and one-loop amplitudes in N=4 super-Yang-Mills. *Nucl. Phys. B*, 725:275–305, 2005.

[122] A. V. Kotikov. Differential equations method: New technique for massive Feynman diagrams calculation. *Phys. Lett. B*, 254:158–164, 1991.

[123] Zvi Bern, Lance J. Dixon, and David A. Kosower. Dimensionally regulated pentagon integrals. *Nucl. Phys. B*, 412:751–816, 1994.

[124] Ettore Remiddi. Differential equations for Feynman graph amplitudes. *Nuovo Cim. A*, 110:1435–1452, 1997.

[125] T. Gehrmann and E. Remiddi. Differential equations for two-loop four-point functions. *Nucl. Phys. B*, 580:485–518, 2000.

[126] Johannes M. Henn. Multiloop integrals in dimensional regularization made simple. *Phys. Rev. Lett.*, 110:251601, 2013.

[127] Vladimir A. Smirnov. Applied asymptotic expansions in momenta and masses. *Springer Tracts Mod. Phys.*, 177:1–262, 2002.

[128] Julian S. Schwinger. Brownian motion of a quantum oscillator. *J. Math. Phys.*, 2:407–432, 1961.

[129] L. V. Keldysh. Diagram technique for nonequilibrium processes. *Zh. Eksp. Teor. Fiz.*, 47:1515–1527, 1964.

[130] Gregor Kälin and Rafael A. Porto. Post-Minkowskian Effective Field Theory for Conservative Binary Dynamics. *JHEP*, 11:106, 2020.

[131] Gustav Mogull, Jan Plefka, and Jan Steinhoff. Classical black hole scattering from a worldline quantum field theory. *JHEP*, 02:048, 2021.

[132] Gregor Kälin, Zhengwen Liu, and Rafael A. Porto. Conservative Dynamics of Binary Systems to Third Post-Minkowskian Order from the Effective Field Theory Approach. *Phys. Rev. Lett.*, 125(26):261103, 2020.

[133] Gregor Kälin, Jakob Neef, and Rafael A. Porto. Radiation-reaction in the Effective Field Theory approach to Post-Minkowskian dynamics. *JHEP*, 01:140, 2023.

[134] Christoph Dlapa, Gregor Kälin, Zhengwen Liu, and Rafael A. Porto. Conservative Dynamics of Binary Systems at Fourth Post-Minkowskian Order in the Large-Eccentricity Expansion. *Phys. Rev. Lett.*, 128(16):161104, 2022.

[135] Gustav Uhre Jakobsen, Gustav Mogull, Jan Plefka, and Benjamin Sauer. All things retarded: radiation-reaction in worldline quantum field theory. *JHEP*, 10:128, 2022.

[136] Christoph Dlapa, Gregor Kälin, Zhengwen Liu, Jakob Neef, and Rafael A. Porto. Radiation Reaction and Gravitational Waves at Fourth Post-Minkowskian Order. *Phys. Rev. Lett.*, 130(10):101401, 2023.

[137] Poul H. Damgaard, Elias Roos Hansen, Ludovic Planté, and Pierre Vanhove. The relation between KMOC and worldline formalisms for classical gravity. *JHEP*, 09:059, 2023.

[138] Mathias Driesse, Gustav Uhre Jakobsen, Gustav Mogull, Jan Plefka, Benjamin Sauer, and Johann Usovitsch. Conservative Black Hole Scattering at Fifth Post-Minkowskian and First Self-Force Order. *Phys. Rev. Lett.*, 132(24):241402, 2024.

[139] Simon Caron-Huot, Mathieu Giroux, Holmfridur S. Hannesdottir, and Sebastian Mizera. What can be measured asymptotically? *JHEP*, 01:139, 2024.

[140] Kuang-chao Chou, Zhao-bin Su, Bai-lin Hao, and Lu Yu. Equilibrium and Nonequilibrium Formalisms Made Unified. *Phys. Rept.*, 118:1–131, 1985.

[141] Simon Caron-Huot. Loops and trees. *JHEP*, 05:080, 2011.

[142] Thibault Damour. Radiative contribution to classical gravitational scattering at the third order in G. *Phys. Rev. D*, 102(12):124008, 2020.

[143] Aneesh V. Manohar, Alexander K. Ridgway, and Chia-Hsien Shen. Radiated Angular Momentum and Dissipative Effects in Classical Scattering. *Phys. Rev. Lett.*, 129(12):121601, 2022.

[144] Paolo Di Vecchia, Carlo Heissenberg, and Rodolfo Russo. Angular momentum of zero-frequency gravitons. *JHEP*, 08:172, 2022.

[145] Michael Crossley, Paolo Glorioso, and Hong Liu. Effective field theory of dissipative fluids. *JHEP*, 09:095, 2017.

[146] Jean-Paul Blaizot and Edmond Iancu. The Quark gluon plasma: Collective dynamics and hard thermal loops. *Phys. Rept.*, 359:355–528, 2002.

[147] Jacopo Ghiglieri, Aleksi Kurkela, Michael Strickland, and Aleksi Vuorinen. Perturbative Thermal QCD: Formalism and Applications. *Phys. Rept.*, 880:1–73, 2020.

[148] Andrea Cristofoli, Riccardo Gonzo, David A. Kosower, and Donal O'Connell. Waveforms from amplitudes. *Phys. Rev. D*, 106(5):056007, 2022.

[149] Simon Caron-Huot, Murat Kologlu, Petr Kravchuk, David Meltzer, and David Simmons-Duffin. Detectors in weakly-coupled field theories. *JHEP*, 04:014, 2023.

[150] Andrea Cristofoli, Riccardo Gonzo, Nathan Moynihan, Donal O'Connell, Alasdair Ross, Matteo Sergola, and Chris D. White. The uncertainty principle and classical amplitudes. *JHEP*, 06:181, 2024.

[151] K. G. Chetyrkin and F. V. Tkachov. Integration by parts: The algorithm to calculate β-functions in 4 loops. *Nucl. Phys. B*, 192:159–204, 1981.

[152] F. V. Tkachov. A Theorem on Analytical Calculability of Four Loop Renormalization Group Functions. *Phys. Lett. B*, 100:65–68, 1981.

[153] M. V. S. Saketh, Justin Vines, Jan Steinhoff, and Alessandra Buonanno. Conservative and radiative dynamics in classical relativistic scattering and bound systems. *Phys. Rev. Res.*, 4(1):013127, 2022.

[154] A. H. Mueller and D. T. Son. On the Equivalence between the Boltzmann equation and classical field theory at large occupation numbers. *Phys. Lett. B*, 582:279–287, 2004.

[155] Zvi Bern, Juan Pablo Gatica, Enrico Herrmann, Andres Luna, and Mao Zeng. Scalar QED as a toy model for higher-order effects in classical gravitational scattering. *JHEP*, 08:131, 2022.

[156] Thibault Damour. ANALYTICAL CALCULATIONS OF GRAVITATIONAL RADIATION. In *4th Marcel Grossmann Meeting on the Recent Developments of General Relativity*, 6 1985.

[157] Asaad Elkhidir, Donal O'Connell, and Radu Roiban. Supertranslations from Scattering Amplitudes. 8 2024.

[158] Roman N. Lee. LiteRed 1.4: a powerful tool for reduction of multiloop integrals. *J. Phys. Conf. Ser.*, 523:012059, 2014.

Appendix A

Integrals

The integrals appearing in section 4.4.3 and Appendix B all belong to the integral family

$$I_{a_1,a_2,a_3,a_4} = \int \hat{d}^D \ell \frac{1}{(p_a \cdot \ell - p_a^0 \omega \pm i\epsilon)^{a_1} (p_b \cdot \ell \pm i\epsilon)^{a_2} (\ell^2) ((q+\ell)^2)^{a_4}} \,. \tag{A.1}$$

It is important to note that IBP relations are insensitive to the $i\epsilon$ prescriptions of the integrals. In particular, the delta function and its derivative appearing in various integrals that we are interested in can be reduced to the above form by using the distributional identity

$$\frac{2\pi i}{(-1)^n n!} \delta^{(n)}(x) = \frac{1}{(x - i\epsilon)^{n+1}} - \frac{1}{(x + i\epsilon)^{n+1}}.$$

Using `LiteRed` [158], we find two master integrals

$$I_{1,0,1,0} \equiv I_a, \quad I_{1,1,1,0} \equiv I_{ab} \,, \quad a, b \text{in} \{1, 2\}. \tag{A.2}$$

To evaluate I_{ab}, we work in the rest frame of a specified by $p_a^\mu = m_a(1, \mathbf{0})$, $p_b^\mu = \gamma m_b (1, \boldsymbol{v})$

$$\begin{aligned} I_{ab} &= \int \hat{d}^4 \ell \frac{\hat{\delta}_+(\ell^2) \hat{\delta}(p_a \cdot \ell - p_a^0 \omega)}{p_2 \cdot \ell \pm i\epsilon} = \int \frac{\hat{d}^3 \ell}{2|\boldsymbol{\ell}|} \frac{\hat{\delta}\left(p_a^0(|\boldsymbol{\ell}| - \omega) - \boldsymbol{p}_1 \cdot \boldsymbol{\ell}\right)}{\left(p_b^0 |\boldsymbol{\ell}| - \boldsymbol{p}_b \cdot \boldsymbol{\ell}\right)} \\ &= \frac{1}{8\pi^2 \gamma m_a m_b} \int_0^\infty d|\boldsymbol{\ell}| \delta(|\boldsymbol{\ell}| - \omega) \int \frac{d\Omega}{1 - v \cos\theta} \\ &= \frac{1}{2\pi m_a m_b \gamma v} \cosh^{-1} \gamma = \frac{1}{2\pi m_a m_b} \frac{\cosh^{-1} \gamma}{\sqrt{\gamma^2 - 1}} \,. \end{aligned} \tag{A.3}$$

I_{aa} is given by the $\gamma \to 1$ limit of Eq. (A.3)

$$I_{aa} = \frac{1}{2\pi m_a^2} \,. \tag{A.4}$$

Now we discuss the integrals appearing in the calculation of the angular impulse. We start with the evaluation of $\mathcal{I}_{12}$, which upon IBP reduction yields

$$\mathcal{I}_{12} = 2 I_{12} = \frac{1}{\pi m_1 m_2} \frac{\cosh^{-1} \gamma}{\sqrt{\gamma^2 - 1}} \,. \tag{A.5}$$

Next, we have

$$\mathcal{K}^{\mu}_{ab} = \int \hat{d}^4\ell \frac{\hat{\delta}_+(\ell^2)\hat{\delta}'(p_a\cdot\ell - p_a^0\omega)\; q\cdot\ell}{(p_b\cdot\ell \pm i\epsilon)^2}\ell^{\mu} = Ap_a^{\mu} + Bp_b^{\mu} + Cq^{\mu}\,. \tag{A.6}$$

The coefficients A, B, C can be expressed in terms of scalar integrals $p_a\cdot\mathcal{K}_{ab}$, $p_b\cdot\mathcal{K}_{ab}$, $q\cdot\mathcal{K}_{ab}$. The scalar integrals can be IBP reduced in terms of the master integrals which gives

$$p_a\cdot\mathcal{K}_{ab} = p_b\cdot\mathcal{K}_{ab} = 0. \tag{A.7}$$

Taking into account $p_a\cdot q = 0 = p_b\cdot q$, we obtain $A = B = 0$ and

$$C = \frac{1}{4q^2}\left(I_{2,2,1,-2} - 2q^2 I_{2,2,1,-1} + q^4 I_{2,2,1,0}\right) = \frac{1}{2\pi m_a^2 m_b^2}\left[\frac{1}{\gamma^2-1} - \frac{\gamma\cosh^{-1}\gamma}{(\gamma^2-1)^{3/2}}\right] \tag{A.8}$$

Finally, $\mathcal{K}^{\mu}_{aa}$ is found by taking the $\gamma\to 1$ limit of K^{μ}_{ab}.

Appendix B

Radiated scalar angular momentum

In this section, we compute angular momentum loss at one-loop in a simple model of real massive scalars Φ_i interacting with a real massless scalar field ϕ described by the Lagrangian

$$\mathcal{L} = \frac{1}{2}\partial_\mu\phi\partial^\mu\phi + \sum_{i=1,2}\left[\frac{1}{2}\partial_\mu\Phi_i\partial^\mu\Phi_i - m_i^2\Phi^2 + \frac{g_i}{2}\Phi_i^2\phi\right]. \tag{B.1}$$

We take the canonical definition of the stress-energy tensor for the massless scalar ϕ. The Feynman rule corresponding to the insertion of the stress-energy tensor of the massless scalar is the same as Eq. (4.47). The relevant diagrams are shown in Fig. B.1 where we have represented the massless scalar propagator by a dotted line. The first diagram gives

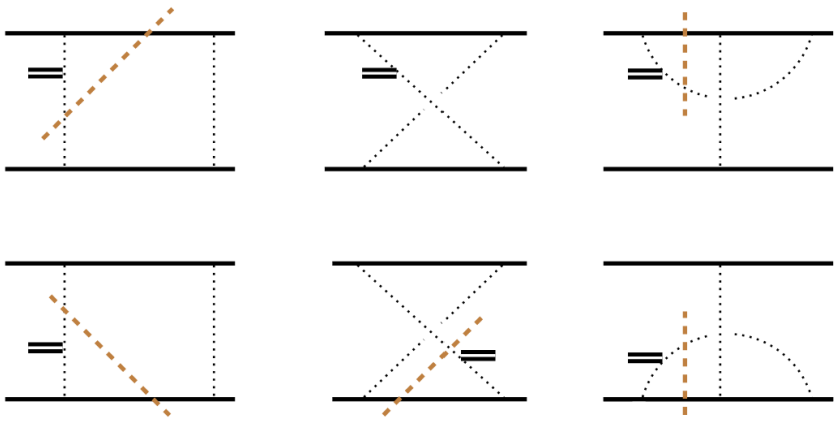


Figure B.1: The diagrams contributing to the scalar angular momentum loss are shown. We label the momenta as in fig 4.6.

$$\mathcal{J}_a = ig_1^2 g_2^2 \lim_{\omega\to 0}\lim_{\boldsymbol{k}\to 0} \partial_{\boldsymbol{k}}^{[i}\Bigg[\int \hat{d}^4q_1\hat{d}^4q_2\hat{\delta}\left(2p_1\cdot q_1+q_1^2\right)\hat{\delta}\left(2p_2\cdot q_2+q_2^2\right)e^{-ib_1\cdot q_1-ib_2\cdot q_2}$$
$$\times\int \hat{d}^4\ell\frac{2\omega\ell^0\hat{\delta}_+(\ell^2)\hat{\delta}(2p_1\cdot\ell+2k\cdot(p_1-\ell)+k^2)\,\ell^{j]}}{(2\ell\cdot k+k^2+i\epsilon)\,(2p_2\cdot\ell-i\epsilon)\,((q_2-\ell)^2-i\epsilon)}\times\hat{\delta}^4(k+q_1+q_2)\Bigg]. \tag{B.2}$$

In evaluating Eq. (B.2), we need to be careful since the $\boldsymbol{k}$ derivative hits the delta function and one must integrate by parts before imposing the momentum conservation. Taking the soft limits and adding all the diagrams in Fig. B.1, we obtain the angular impulse of the massless scalar field

$$\Delta\mathbb{J}_\phi^{ij} = -\lim_{\omega\to 0}\frac{ig_1g_2}{16}\left[\int\frac{\hat{d}^4q}{q^2}e^{-ib_1\cdot q}\,\hat{\delta}\left(p_1\cdot q\right)\hat{\delta}\left(p_2\cdot q\right)p_1^{[i}\left[g_1g_2\,\mathcal{K}_{12}^{j]}-g_1^2\mathcal{K}_{11}^{j]}\right]\right.$$
$$\left.=\frac{1}{8\pi}\left[\frac{g_1^2}{3m_1^4}+\frac{g_1g_2}{m_1^2m_2^2}\left(\frac{1}{\gamma^2-1}-\frac{\gamma\cosh^{-1}\gamma}{(\gamma^2-1)^{3/2}}\right)\right]p_1^{[i}\,\Delta p_1^{j]}+(1\leftrightarrow 2)\,, \right. \tag{B.3}$$

where we have used the on-shell conditions $2p_i\cdot q_i+q_i^2=0$. We have also defined the leading order linear impulse of the heavy particle 1:

$$\Delta p_1^i = -\frac{ig_1g_2}{4}\int\frac{\hat{d}^4q}{q^2}\hat{\delta}\left(p_2\cdot q\right)\hat{\delta}\left(p_1\cdot q\right)q^ie^{-iq\cdot b}. \tag{B.4}$$

This agrees with the small-deflection limit of the result in [144].