



Analytic Structure and Finiteness of Scattering Amplitudes

Citation

Hannesdottir, Holmfridur Sigridar. 2021. Analytic Structure and Finiteness of Scattering Amplitudes. Doctoral dissertation, Harvard University Graduate School of Arts and Sciences.

Permanent link

<https://nrs.harvard.edu/URN-3:HUL.INSTREPOS:37368495>

Terms of Use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at <http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA>

Share Your Story

The Harvard community has made this article openly available.
Please share how this access benefits you. [Submit a story](#).

[Accessibility](#)

HARVARD UNIVERSITY
Graduate School of Arts and Sciences



DISSERTATION ACCEPTANCE CERTIFICATE

The undersigned, appointed by the
Department of Physics
have examined a dissertation entitled

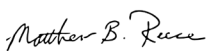
Analytic Structure and Finiteness of Scattering Amplitudes

presented by Holmfridur Sigridar Hannesdottir

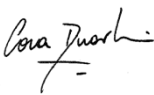
candidate for the degree of Doctor of Philosophy and hereby
certify that it is worthy of acceptance.


Signature _____

Typed name: Professor Matthew Schwartz, Chair


Signature _____

Typed name: Professor Matthew Reece


Signature _____

Typed name: Professor Cora Dvorkin

Date: April 14, 2021

Analytic Structure and Finiteness of Scattering Amplitudes

A DISSERTATION PRESENTED
BY
HOLMFRIDUR SIGRIDAR HANNESDOTTIR
TO
THE DEPARTMENT OF PHYSICS

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY
IN THE SUBJECT OF
PHYSICS

HARVARD UNIVERSITY
CAMBRIDGE, MASSACHUSETTS
APRIL 2021

©2021 – HOLMFRIDUR SIGRIDAR HANNESDOTTIR
ALL RIGHTS RESERVED.

Dissertation advisor:
Professor Matthew D. Schwartz

Author:
Holmfridur Sigrídar Hannesdóttir

Analytic Structure and Finiteness of Scattering Amplitudes

ABSTRACT

Scattering amplitudes are fundamental objects in high energy physics, providing a bridge between theoretical calculations and data from particle colliders. The framework underpinning the study of elementary particles - quantum field theory - has led to remarkably precise scattering amplitude predictions, which, in turn, provide a foundation for fundamental physics discoveries. Despite this success, particles remain poorly described within this framework. Consequently, in theories with massless particles, asymptotic interactions render scattering amplitudes infrared divergent in perturbation theory and ill-defined non-perturbatively. In this dissertation, we take steps towards strengthening the theoretical foundations of quantum field theory, by defining infrared finite cross sections and amplitudes, and by probing the analytic structure of amplitudes through examining their sequential discontinuities. First, we show that infrared finite cross sections are obtained by summing over either initial or final states, and explore how forward scattering diagrams often constitute a crucial contribution to achieve finiteness. Then, using universality and factorization of asymptotic interactions, we demonstrate how to

Dissertation advisor:
Professor Matthew D. Schwartz

Author:
Holmfridur Sigrídar Hannesdóttir

define finite scattering amplitudes in gauge theories. Exploiting freedom in choosing regulators and cutoffs for the asymptotic interactions, these amplitudes can be interpreted alternatively as Wilson coefficients, as remainder functions, or as coherent states. Finally, we extend the traditional cutting rules, that relate discontinuities of amplitudes to cuts of the corresponding Feynman diagrams, to sequential discontinuities and multiple cuts. Our relations provide a new probe of the analytic properties of amplitudes, in addition to a scheme for exploring the finite scattering amplitudes further. Through an enhanced understanding of the analytic structure of finite amplitudes, we hope to unveil a proper description of particles in quantum field theory.

Contents

Title page	i
Copyright	ii
Abstract	iii
Contents	vi
Acknowledgments	vii
Citations to previously published work	x
1 Introduction	1
1.1 Quantum field theory: Probing protons and electrons	2
1.2 Infrared divergences	5
1.3 The S -matrix	7
1.4 Analytic methods for probing amplitudes	9
1.5 Overview of dissertation	11
2 Infrared Finiteness and Forward Scattering	13
2.1 Introduction	14
2.2 KLN theorem revisited	18
2.3 Final or initial state sums only	21
2.4 $Z \rightarrow e^+e^- (+\gamma)$	23
2.5 $e^+e^- \rightarrow Z + X$	25
2.6 Compton scattering	40
2.7 $\gamma\gamma \rightarrow X$	45
2.8 Summary and Conclusions	46
3 A Finite S -matrix	53
4 An S -matrix for Massless Particles	67
4.1 Introduction	68
4.2 The hard S -matrix	78
4.3 Computing the hard S -matrix	99
4.4 QED: Deep Inelastic Scattering	109
4.5 QCD: $e^+e^- \rightarrow \text{jets}$	122

4.6	$\mathcal{N} = 4$ Super Yang-Mills	133
4.7	Summary and Outlook	145
5	Sequential Discontinuities of Feynman Integrals and the Monodromy Group	152
5.1	Introduction	153
5.2	Cutting rules: a review	161
5.3	Discontinuities	171
5.4	Discontinuities as monodromies	184
5.5	Sequential discontinuities	215
5.6	Examples	229
5.7	Conclusions	265
	Appendix A Appendix to Chapter 2	273
A.1	On-shell intermediate propagators	274
A.2	Initial state masses	276
	Appendix B Appendix to Chapter 4	282
B.1	The coproduct from variation matrices	283
B.2	The monodromy and fundamental groups	287
B.3	Single-valued polylogarithms	294
B.4	Permutation symmetry of the triangle integral	297
B.5	Variation matrix of the two-loop box	300
B.6	Cuts of the three-loop triangle	309
B.7	Massless three-point vertices	315
	Bibliography	325

Acknowledgments

My time as a graduate student at Harvard has been wonderful, and I look upon the past five years with gratitude and fondness. I want to thank everyone who helped make this time productive and enjoyable.

First, I want to thank my advisor, Professor Matthew D. Schwartz, for his encouragement and guidance. Through Matt's knowledge and mentorship, my understanding of physics has been deepened, and I have learned to formulate and tackle interesting physics questions. With his profound insight and unique approach to physics, he has taught me how to overcome research obstacles, making our projects both interesting and rewarding. I am immensely grateful to him for giving me the opportunity to develop and grow as a researcher while also making my years in graduate school highly enjoyable.

I would also like to thank the co-authors of the work presented in this dissertation: Jacob L. Bourjaily, Christopher Frye, Andrew J. McLeod, Nisarga Paul, Matthew D. Schwartz, Cristian Vergu and Kai Yan. I am truly privileged to have worked with such knowledgeable and inspiring collaborators. I am grateful to Chris for guiding my first research calculations on infrared divergences in quantum field theory, and I also want to thank him, Kai and Anders Andreassen for their mentorship and guidance, which I benefited greatly from in my second year. During my third and fourth year, I truly enjoyed working with Matt on using factorization to define a finite S -matrix, which transformed my view on effective field

theories and foundations of quantum field theory. In the past couple of years, I have had a blast exploiting analyticity methods to study scattering amplitudes with Jacob, Andrew, Matt and Cristian. I have learned a lot from their insights and knowledge, which I am truly grateful for.

The Harvard community has been warm and welcoming, nurturing both friendships and professional development. During the first few semesters, I was extremely fortunate to live with Mina Himwich, who made me feel at home in Cambridge from the first day. In graduate school, I am privileged to have made wonderful friends, with whom I have had many happy moments, while also having interesting discussions about various scientific fields. I would particularly like to thank Cari Cesarotti, Kelsey DeJesus-Banos, Ana Díaz Rivero, Mina Himwich, Rhine Samajdar, Weilu Shen and Marius Vollberg. During my time in Cambridge, I enjoyed taking runs to unwind and recharge, and I am immensely thankful to Andrea Loizeau and Nishita Parnandi for their friendship and many wonderful Saturday morning runs. Furthermore, I thank members and visitors of the Physics Department, in particular Prateek Agrawal, Nima Arkani-Hamed, Adam Ball, Jacob Barandes, Lisa Cacciabauda, Cora Dvorkin, Howard Georgi, Jenny Hoffman, Qianshu Lu, Sruthi Narayanan, Aditya Parikh, Monica Pate, Ana-Maria Raclariu, Lisa Randall, Matthew Reece, Suzanne Smith, Andrew Strominger, Jesse Thaler and Linda Xu for inspiring conversations and encouragement. I also thank the high energy group at the University of Iceland, particularly Valentina Giangreco M Puletti and Larus Thorlacius, for always welcoming me to the Physics Department.

Lastly, I want to thank my family; my parents Sigridur Olina and Hannes, my brothers

Haraldur Jon and Hrafkell Stefan, my grandparents Holmfridur Kolbrun and Haraldur, my aunts and uncles, my family-in-law, and my friends in Iceland who have all been incredibly encouraging and supporting throughout graduate school. I also thank my partner Olafur Bjarki for his encouragement, and for making this past year full of smiles and wonderful adventures. I will forever be grateful to my family for giving me the opportunity to pursue my dreams, and for filling these years with love, laughter and happiness.

Citations to previously published work

Parts of this dissertation cover research reported in the following articles:

1. Christopher Frye, Holmfridur Hannesdottir, Nisarga Paul, Matthew D. Schwartz and Kai Yan, “Infrared Finiteness and Forward Scattering,” *Phys. Rev. D* **99** no. 5, (2019) 056015, [arXiv: \[1810.10022\]](#) [[hep-ph](#)].
2. Holmfridur Hannesdottir and Matthew D. Schwartz, “A Finite S -matrix,” [arXiv: \[1906.03271\]](#) [[hep-th](#)].
3. Holmfridur Hannesdottir and Matthew D. Schwartz, “ S -matrix for Massless Particles,” *Phys. Rev. D* **101** no. 10, (2020) 105001, [arXiv: \[1911.06821\]](#) [[hep-th](#)].
4. Jacob L. Bourjaily, Holmfridur Hannesdottir, Andrew J. McLeod, Matthew D. Schwartz and Cristian Vergu, “Sequential Discontinuities of Feynman Integrals and the Monodromy Group,” *JHEP* **01**, (2021) 205, [arXiv: \[2007.13747\]](#) [[hep-th](#)].

1

Introduction

1.1 Quantum field theory: Probing protons and electrons

Physics addresses questions about the world in which we live: How did the Milky Way form? Can we make superconductors at room temperature? What kind of particles is our universe made of?

Over the course of centuries, scientists have probed ever larger and smaller parts of our universe. While we once thought that Earth was a unique phenomenon situated at its center, we now know that our home planet is at the edge of one out of billions of galaxies, the Milky Way, with each galaxy containing billions of planets orbiting their stars. Similarly, we have probed small scales by decoding the composition of matter; revealing that it is made out of atoms, which, in turn, are made out of electrons, protons and neutrons. In the 20th century, we discovered the constituents of protons and neutrons, called quarks and gluons, in addition to other elementary particles that compose our world, such as the Z boson, the photon and the Higgs boson. Particle physics involves the study of these smallest known constituents of nature.

All elementary particles are described within an elegant mathematical framework called *quantum field theory* (QFT). This framework weaves together two hallmarks of early 20th century physics – Einstein’s special relativity and quantum mechanics – and allows for the creation and destruction of particles using the mass-energy equivalence relation $E = mc^2$. For instance, QFT allows us to calculate the probability for two protons to create a Higgs boson in a particle collider. The QFT framework is guided by *symmetries*, which provide valuable and powerful restrictions since, by *Noether’s theorem* [1], each continuous sym-

metry implies a conservation law. Energy conservation, for example, is a consequence of time-translation invariance, i.e. the fact that the laws of physics are the same today as they were yesterday. These types of conservation laws highly constrain our QFT. Other symmetries are also encountered in QFT, such as the one denoted with $SU(3)$, that governs and constrains interactions between quarks and gluons. Using symmetries as building blocks, we can construct theories incorporating different types of particles.

By specifying the symmetry groups, masses of particles, and some other parameters, we formulate the *Standard Model of Particle Physics*. The Standard Model is therefore a specific QFT with the necessary inputs specified, resulting in a theory that describes all known real-world elementary particles. To observe these tiny, subatomic particles experimentally, we resort to *particle colliders*. For example, the Large Hadron Collider (LHC) at CERN, accelerates protons using magnetic fields in a 27 km long underground tunnel. When two protons have reached speeds close to the speed of light, they collide to create sprays of elementary particles. One of the goals in theoretical particle physics is to calculate, to a high degree of precision, the probabilities of different outcomes at the LHC and other colliders.

When comparing theoretical calculations to data from particle colliders, QFT comprises an essential ingredient in verifying the existence and properties of the elementary particles. The Higgs boson is the most recently verified particle of the Standard Model: its discovery was announced at CERN in 2012 with a significance of more than five standard deviations. In addition to unveiling new particles, QFT has predicted the outcome of many other scattering experiments in colliders to a high precision. In fact, quite remarkably, we have yet to find any striking discrepancy of collider experiment results and the Standard Model. Thus,

the Standard Model is one of the best tested theories in science.

Despite this success, we have ample evidence for new particles, waiting to be discovered. Measurements of rotations of stars in the Milky Way, along with many other experiments, demonstrate the existence of *dark matter*. This form of matter is not comprised of any Standard Model particles, and makes up around 85% of matter in the universe. To probe what constitutes dark matter, physicists look for a discrepancy between collider data and the Standard Model, since such a discrepancy can point to places where the theory can be enhanced. We then hope to formulate an even more fundamental model of nature using QFT, which explains the data by incorporating new particles, and thus sheds light on the particle nature of dark matter.

While QFT is both an elegant and experimentally successful theory, it is still a work in progress. For instance, we can set up a QFT calculation for any intricate process of our choice, but it often results in such a complicated expression that the answer is intractable. Furthermore, a physical understanding is sometimes lacking, for example pertaining to how we can incorporate electromagnetic fields into the description of particles. In this dissertation, I will pinpoint improvable parts of the theory and propose possible developments. With such advancements, we can both improve theoretical scattering calculations and deepen our understanding of particles, which could, in turn, lead to new discoveries in physics.

1.2 Infrared divergences

A proper particle description becomes important when our QFT contains massless particles, such as photons, which can be created without any cost in energy. As a result, a naive application of the theory can result in illogical results of scattering calculations. For instance, the probability for an electron and photon to bounce off one another, which should never exceed 100%, becomes infinite. This well-known problem in QFT, referred to as an *infrared (IR) divergence*, arises since this simple calculation does not account for the electromagnetic field of charged particles. The unphysical, infinite probabilities are therefore a consequence of misinterpreting the mathematical result by not correctly accounting for the effects of massless particles. As an old problem, present since the dawn of QFT, it has a traditional solution: the radiation and absorption of massless particles must be accounted for in scattering calculations to get finite answers.

An example illustrating this solution is as follows: If we detect an electron in a particle detector, and since any real-world detector has some uncertainty in energy, we do not know whether a photon, whose energy was less than the detector resolution, accompanied the electron. In other words, when calculating the probability for measuring an electron hitting our detector, we must also account for the possibility of having had an electron *and* a photon, where the photon had such little energy that it escaped our detector, undiscovered. This construction always works in Quantum Electrodynamics and is referred to as the Bloch-Nordsieck theorem [2]: If we calculate the probability for some scattering process to occur, and add the possibility of having had any number of low-energy, undetected, outgo-

ing photons, the result is free of infrared divergences.

Using the explicit example of $qq \rightarrow qq\mu\mu$, Doria, Frenkel and Taylor [3] showed that the Bloch-Nordsieck theorem is a special feature of Quantum Electrodynamics, and does not hold generally in QFT. In more complicated theories, we must instead resort to the Kinoshita-Lee-Nauenberg [4, 5] theorem: To get a finite cross section for a certain process, one needs to sum over all possible *initial states and final states* consistent with the chosen process. This theorem is also physically motivated; since we cannot know the initial nor final state precisely in any real-world particle detector, we should take all possibilities into account. In this dissertation, we revisit the KLN theorem and show that a more general version can be proven: To get a finite cross section, we do not need to sum over both initial states and final states; an *initial state or final state* sum suffices.

A drawback of the KLN construction, pertaining to comprehending particles in QFT, is that the required sum over *all* initial or final states is not restrictive. In other words, the theorem does not guide us towards an enhanced intuition, as it would if it required only a particular set of states to achieve infrared finiteness. Moreover, an important feature of both of these theorems is that they apply to scattering probabilities, and not probability amplitudes. It is more appealing to have an infrared finite calculation from the outset, without needing to sum over multiple divergent processes to get a finite result. Thus, to gain further insight into particles and our theory, we aim to eliminate infrared divergences from probability amplitudes. To do so, we must examine properties of the S -matrix in QFT.

1.3 The S -matrix

The scattering matrix, or S -matrix, is a fundamental mathematical construct in QFT, underpinning theoretical calculations for particle collider experiments. More precisely, the S -matrix gives the *probability amplitude* for some initial state $|\psi_{\text{in}}\rangle$ at $t \rightarrow -\infty$ to evolve into a final state $|\psi_{\text{out}}\rangle$ at $t \rightarrow +\infty$, where the states $|\psi_{\text{in/out}}\rangle$ specify the initial and final state particles and their momenta. As a consequence, *probabilities* are obtained by squaring the absolute value of the S -matrix, i.e. computing $|S|^2$, and then integrating over all possible momentum configurations of the particles involved in the scattering. In addition to providing accurate predictions for colliders, studies of the S -matrix can lead to new insights into the mathematical foundations of QFT. Symmetries, limiting behavior and analytic properties are analyzed to shortcut calculations and define precision observables for colliders.

Most calculations in QFT are done in perturbation theory, i.e. by expanding in some small parameter. When writing the Hamiltonian of the system as

$$H = H_0 + V, \tag{1.3.1}$$

where H_0 is the *free Hamiltonian*, we expand in the interaction term V , and as a result, we can calculate the S -matrix by summing over *Feynman diagrams*. An example of a Feynman diagram is shown in Figure 1.1. The lines correspond to particles evolving freely with H_0 , while the vertices correspond to interactions between particles. Each diagram represents an integral we would like to calculate; we can read off which integral the Feynman diagram

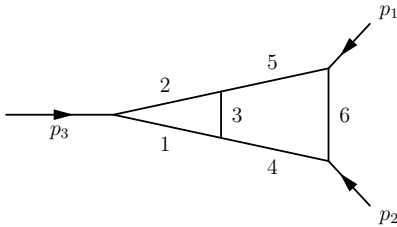


Figure 1.1: An example of a two-loop Feynman diagram.

represents using *Feynman rules*, that dictate which mathematical expressions the lines and vertices correspond to.

An important feature of Feynman diagrams is that the *incoming* and *outgoing* particles are represented by lines, indicating that they evolve without any interactions. In other words, we assume that particles evolve freely at asymptotic times, as $t \rightarrow \pm\infty$. When our theories contain massless particles, however, interactions persist even as $t \rightarrow \pm\infty$, so the free-evolution assumption breaks down. These non-vanishing long-range interactions are to blame for the infrared divergence of the S -matrix in perturbation theory. Thus, despite the importance of the S -matrix both for calculating outcomes of experiments and for the mathematical development of QFT, these divergences prevent the formal existence of the S -matrix when long-range interactions are present.

We propose a solution to the infrared divergence problem in this dissertation by adding asymptotic interactions to H_0 in a universal manner. Then, we obtain finite scattering amplitudes, in contrast to achieving cancellations of infrared divergences for *squared* amplitudes, i.e. probabilities, as proposed by the KLN theorem. This construction results in a finite scattering object which we call the “hard” S -matrix. Furthermore, we show that this object is mathematically equivalent to known and useful quantities in scattering theory, and thus it provides a new starting point for investigating these quantities.

1.4 Analytic methods for probing amplitudes

Armed with a new definition of an infrared finite scattering matrix, we aim to use some clever and modern methods to examine its properties. A useful technique for studying the S -matrix is to examine its *analytic structure*. Instead of performing traditional Feynman-diagrammatic calculations, the goal of this method is to deduce properties of amplitudes by combining known constraints and complex analysis techniques.

A useful constraint on amplitudes is based on combining *Lorentz invariance* and *analyticity*. When the incoming and outgoing particles and their momenta have been specified, the scattering amplitude must depend on frame-invariant parameters only. This dependence follows from special relativity; since the amplitude is independent of the frame in which the calculations are done, it can only depend on frame-invariant quantities. Furthermore, the amplitude is analytic except at certain loci in the complex plane of the invariants, which provides another valuable constraint. Exploiting mathematical techniques for analytic functions can therefore probe physical properties of the S -matrix, and thus shortcut calculations. Modern bootstrapping programs, for example, use these kinds of analyticity constraints to determine amplitudes to high loop orders.

An example of a useful probe of the analytic structure of the S -matrix is its *discontinuity*, defined as the difference between the values of the amplitude on two sides of a branch cut in the complex plane. Its practicality for examining amplitudes stems from a mathematical formula relating it to *cuts*. More precisely, the traditional Cutkosky cutting rules [6] assert that discontinuities of amplitudes are equal to a sum over cuts through the corre-

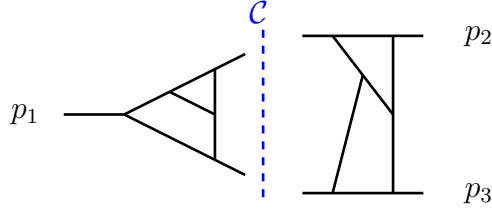


Figure 1.2: Example of a cut \mathcal{C} through a Feynman diagram.

sponding Feynman diagram:

$$\text{Disc}\mathcal{M} = \sum \text{Cut}\mathcal{M}. \quad (1.4.1)$$

Here, \mathcal{M} is the scattering amplitude,¹ Disc is the discontinuity operator, and the Cut operator acts on the Feynman diagram corresponding to \mathcal{M} by putting all cut momenta on shell. An example of a cut \mathcal{C} through a Feynman diagram is shown in Figure 1.2. Since the cuts, denoted with $\sum \text{Cut}\mathcal{M}$, can sometimes be calculated more easily than the full Feynman diagrams corresponding to \mathcal{M} , they provide a simpler computation of its discontinuities, which can consequently be used to deduce analytic properties. Moreover, an even more powerful probe of scattering amplitudes is obtained by examining sequential discontinuities of \mathcal{M} , i.e. discontinuities of discontinuities, and relating them to multiple cuts, as we examine in this dissertation. Our new formulas will hopefully help clear a path towards analyzing both the traditional and the hard S -matrix, and thus bring us closer to a proper particle description.

¹The definition of the amplitude denoted with \mathcal{M} differs from the definition of the S -matrix by a constant, an overall momentum conserving delta function and the absence of forward scattering.

1.5 Overview of dissertation

In this dissertation, I present a path towards studying the analytical properties of an infrared finite S -matrix. The following chapters contain our findings in chronological order.

In Chapter 2, we address infrared divergences by proving a stronger version of the KLN theorem. According to our new theorem, we only need to sum over *either* initial or final states to obtain an infrared finite cross section. A crucial component of the sum, however, is the forward scattering contribution, that represents the process where no scattering occurs; the initial state particles move forward in time without a change in their momenta. Thus, the KLN theorem and its stronger version are a consequence of unitarity: the probability of *anything* happening, either something or nothing, is 1. We illustrate the claims of the stronger version of the KLN theorem with multiple examples, some of which require disconnected Feynman diagrams. In addition, we present other infrared finite cross section calculations, which are physically motivated rather than prompted by the theorem, and thus perhaps more appealing. These findings suggest that further research is needed to determine how infrared finiteness connects to our physical understanding of scattering.

We use the insight gained from studying infrared divergent cross sections to obtain a finite S -matrix in Chapters 3 and 4. Exploiting factorization of soft and collinear divergences in gauge theories, we show how to define an infrared finite S -matrix, S_H . Chapter 3 presents this construction concisely, while Chapter 4 contains details on the derivation of the relevant Feynman rules along with multiple examples. This new scattering object can be interpreted as Wilson coefficients in Soft-Collinear Effective Theory, as remainder

functions in $\mathcal{N} = 4$ Super Yang-Mills Theory, or as coherent states in Quantum Electrodynamics. Probing its properties will therefore provide new insights into familiar scattering objects.

In Chapter 5, we probe the analytic structure of scattering amplitudes by generalizing the cutting rules. While the traditional Cutkosky cutting rules establish the equivalence between discontinuities across branch cuts of a scattering amplitude and the sum of cuts through the corresponding Feynman diagrams, our new formulas relate sequential discontinuities to multiple cuts. We prove our relations in time-ordered perturbation theory (TOPT), and demonstrate them using a number of polylogarithmic examples. As a corollary, we prove the *Steinmann relations* [7], stating that sequential discontinuities in partially-overlapping momentum channels cannot exist, in the case when all external particles are massive. Analogously to the Steinmann relations, our new formulas relating discontinuities and cuts might become useful for constraining scattering amplitudes using bootstrapping methods.

2

Infrared Finiteness and Forward Scattering

2.1 Introduction

The appearance and interpretation of infinities has been an essential ingredient of quantum field theory since its inception. While ultraviolet divergences appearing in perturbation theory are now completely understood through the program of renormalization, infrared divergences remain somewhat mysterious. In contrast to ultraviolet divergences, which drop out when amplitudes are expressed directly in terms of other amplitudes, infrared divergences seem only to cancel at the cross-section level for *sufficiently inclusive* quantities. It is imperative, therefore, to have a precise definition of sufficiently inclusive, i.e. to characterize the minimal set of states that must be included to get a finite cross section.

Part of the reason we find the question of IR finiteness compelling is that its resolution is essential to defining a sensible S matrix. If we define the S matrix in the usual way in quantum field theory, its matrix elements in states of fixed particle number are all infinite at each order in perturbation theory, and zero nonperturbatively. This old problem has not yet limited the applicability of field theory to computing observables at colliders, but is important for studying formal properties of the S -matrix, such as its symmetries. The coherent state approach argues that the problem is that isolated charged particles are not well-defined asymptotic states [8–10], but that electrons dressed with a cloud of photons may be. Although the idea is appealing, it is not clear that the dressed/coherent state approach will work for any theory more complicated than QED with massive electrons. If the approach is to succeed it will likely do so through the same mechanism, with the same set of processes, that the cancellation is achieved at the cross section level. Thus, we focus here

on cross-section level computations where the the path forward is less obscure.

Although working towards a finite S matrix is a noble goal, there are more practical motivations for understanding IR divergence cancellations. One important one is precision collider physics. Over the last several years, there has been renewed interest in understanding factorization, and its violation in various forms. Consequences of factorization violation include the various large logarithms that appear in perturbative calculations, such as non-global logarithms or super-leading logarithms. Non-global logarithms arise when virtual and real-emission contributions end up in different regions of phase space [11–18]. Super-leading logarithms are associated with collinear factorization violation [19–23], which is in turn tied to forward scattering, a focus of this paper. Thus, broadly speaking, an improved understanding of IR divergences is relevant to both formal aspects of quantum field theory and precision collider physics.

One of the earliest important papers on infrared finiteness was by Bloch and Nordsieck in 1937 [2]. They showed that in QED with massive electrons, infrared singularities in loops and real emission graphs have the same functional form with opposite signs. The Bloch-Nordsieck theorem is that an observable that sums over all possible numbers of final state photons with energies $E < \delta$ is “sufficiently inclusive,” i.e. it is infrared finite for any δ . Proofs of the Bloch-Nordsieck theorem were developed sometime later [24–26] and have become textbook material [27, 28]. Essentially, the proof works through the Abelian exponentiation theorem [24, 26]: the soft singularities in QED to all orders in α are given by the exponential of the singularities at 1-loop. With massive electrons, all the singularities in QED are soft in nature and so Abelian exponentiation is all that is needed for the proof.

In theories with massless charged particles, such as QCD, the cancellation of infrared singularities (soft and collinear) is significantly more subtle. The Bloch-Nordsieck theorem fails in QCD: summing inclusively over final state gluon radiation is insufficient to cancel all infrared singularities, even if the initial state consists only of massive quarks [3, 29–31]. Nevertheless, even in QCD infrared divergences can be shown to cancel in certain contexts. For example, in hadronic events in Z boson decays, one can identify “sufficiently inclusive” with “infrared-and-collinear (IRC) safe”: an observable should have the same value if particles with zero energy are added, or if finite energy particles are split into multiple particles going in exactly the same direction. This implies that although the rate for a Z boson to decay to two quarks and nothing else is infinite, the rate for a Z to decay to two “jets”, defined as collections of radiation within an angle θ including all radiation softer than an energy δ , is well-defined (i.e. it is finite) [32]. While this definition of sufficiently inclusive is adequate to remove infrared singularities in $Z \rightarrow \text{hadron}$ events, it is not a sufficient criterion in other contexts. As we will discuss, $Z \rightarrow X$ is special because the $Z \rightarrow Z$ forward scattering amplitude is IR finite to all orders. In most other contexts, IRC safety must be generalized.

A non-minimal definition of “sufficiently inclusive” was proposed by Kinoshita, Lee and Nauenberg in their KLN theorem [4, 5]. The KLN theorem states that summing over all initial and final states with energies in some compact energy window around a reference energy E_0 guarantees finiteness. Stated this way, the theorem is fairly useless as *all* states includes Z ’s, neutrinos, quarks, little red dragons, etc. Fortunately, the KLN theorem derivation involves a sum in a more restricted set: those intermediate particles appearing in any

double-sided-cuts [4, 33, 34] through any given time-ordered perturbation theory diagram gives a finite answer. A double sided cut means summing over all possible initial and final states (we give some examples in Section 2.5.3). There are some caveats to the restriction: a number of different diagrams must be included to maintain gauge invariance and Lorentz invariance, but generally the particles involved can be read off the initial graph.

Despite the importance of the KLN theorem, there are very few explicit computations in the literature showing how the cancellation actually occurs [35–38]. One such example was provided by Lavelle and McMullan who showed that IR divergences cancel in processes with an electron scattering off of a background Coulomb potential [38] (see also [5, 39]). In working out some examples, a number of troublesome features associated with initial state sums emerge. First of all, even though we can define a $n \rightarrow m$ cross section mathematically, it is not clear how to think about it physically. Although one can envision a kind of generalization of IRC safety for initial states, including soft and collinear incoming particles, it is not clear how to identify the physical incoming states in a given experiment. In addition, for the KLN cancellation to occur, not only must disconnected diagrams be included, but also an infinite number of photons can participate at any fixed order in the coupling. How to sum this infinite series, with alternating signs for the divergent and finite pieces, requires careful consideration [38, 40]. We examine some of these issues for the process $e^+e^- + \text{photons} \rightarrow Z + \text{photons}$ in Section 2.5.

Although one can demonstrate the cancellation of IR divergences when summing over initial and final states following the KLN theorem, a careful examination of the proof of theorem provides two revelations: 1) The processes that contribute to assure the cancellation in-

clude exactly forward scattering and 2) infrared divergences cancel when summing over final states alone for fixed initial state *or* summing over initial states for a fixed final state. This second point is a relief: one can avoid the troublesome aspects of initial state sums. The first point is less of a relief: it requires us to revise our intuition for what states are physically distinguishable. For example, to resolve infrared divergences in $\gamma\gamma \rightarrow e^+e^-$ one must also include $\gamma\gamma$ in the final state. In the end, it seems there are multiple ways to achieve finiteness for this process at next-to-leading order: a final state sum, an initial state sum, or a partial final and initial state sum. We discuss this example in depth in Section 2.5 and some related QED processes are considered in Sections 2.6 and 2.7. A summary of the various results of this paper and some additional thoughts are presented in Section 2.8. Appendix A.1 shows how to compute diagrams with on-shell intermediate states that occur from cuts with disconnected pieces. Appendix A.2 gives some details of an initial-state jet mass calculation from Section 2.5.

2.2 KLN theorem revisited

We begin by reviewing the KLN theorem and showing that the initial state sum is not necessary. The KLN theorem is attributed to two papers [4], the first by Kinoshita and the second by Lee and Nauenberg [5]. The Kinoshita paper follows after a paper by Kinoshita and Sirlin [41] that considered muon decay $\mu^- \rightarrow e^- \nu_\mu \bar{\nu}_e$ in the 4-Fermi theory. They observed that while the exclusive cross section for this process is infrared divergent in the limit of a massless electron, the inclusive cross section is finite when the virtual contribu-

tion in $\mu^- \rightarrow e^- \nu_\mu \bar{\nu}_e$ is combined with the $\mu^- \rightarrow e^- \nu_\mu \bar{\nu}_e \gamma$ cross section. The Kinoshita paper, which builds on work by Nakanishi [34], proves the finiteness of $\mu^- \rightarrow e^- \nu_\mu \bar{\nu}_e$ with $m_e = 0$ to all orders in perturbative QED. In this way, it generalizes Bloch-Nordsieck to include collinear divergences (mass singularities) as well as soft divergences. Kinoshita also discusses the sum over initial and final states as necessary to cancel mass singularities associated with the muon being massless. Lee and Nauenberg (LN) generalize Kinoshita's result, providing a simple proof for any quantum mechanical system that all infrared divergences (soft and collinear) cancel when initial and all degenerate final states are summed over. Since the LN approach is simple and includes Kinoshita's result, we will focus on it here.

The theorem proved by Lee and Nauenberg is that the transition amplitude squared is IR finite when summed over initial and final states,

$$\sum_{a \in D(E), b \in D(E)} |\langle b | U(\infty, -\infty) | a \rangle|^2 < \infty. \quad (2.2.1)$$

Here $a \in D(E)$ means that the energy of the state a is in the range specified by $D(E)$, e.g. $|E_a - E| < \delta$ for some $\delta > 0$. For Eq. (2.2.1) to be true, we must define the sum over states $|b\rangle$ to include the state where $|b\rangle = |a\rangle$, i.e. the forward scattering contribution. Including $|a\rangle$ in the sum is critical – without it the proof does not hold.

The operator $U(t_2, t_1)$ is the unitary operator that evolves the system from time t_1 to t_2 . In the interaction picture we write the Hamiltonian as $H(t) = H_0 + V(t)$ with H_0 the free

Hamiltonian of which the Fock-states $|a\rangle$ are eigenstates and

$$U(t_2, t_1) = \mathcal{T} \left\{ \exp \left[-i \int_{t_1}^{t_2} dt' V_I(t') \right] \right\}, \quad (2.2.2)$$

with $V_I(t) = e^{iH_0(t-t_1)} V(t_1) e^{-iH_0(t-t_1)}$ the interaction picture potential and t_0 an arbitrary reference time.

To prove Eq. (2.2.1), LN observe that

$$\begin{aligned} |\langle b|U(\infty, -\infty)|a\rangle|^2 &= \sum_{i,j} \left[\langle b|U(\infty, 0)|j\rangle \langle j|U(0, -\infty)|a\rangle \right] \left[\langle b|U(\infty, 0)|i\rangle \langle i|U(0, -\infty)|a\rangle \right]^* \\ &= R_{bij}^+ R_{aij}^-, \end{aligned} \quad (2.2.3)$$

with

$$R_{bij}^+ = \langle i|U(0, \infty)|b\rangle \langle b|U(\infty, 0)|j\rangle, \quad (2.2.4)$$

$$R_{aij}^- = \langle j|U(0, -\infty)|a\rangle \langle a|U(-\infty, 0)|i\rangle. \quad (2.2.5)$$

So, it is enough to show that

$$R_{ij}^+ = \sum_{b \in D(E)} R_{bij}^+ < \infty, \quad R_{ij}^- = \sum_{a \in D(E)} R_{aij}^- < \infty, \quad (2.2.6)$$

for Eq. (2.2.1) to hold.

LN prove Eq. (2.2.6) inductively on the number of singular intermediate states. The singularities come from time-ordered-perturbation theory propagators of the form $\frac{1}{E_i - E_j \pm i\varepsilon}$

when $E_i = E_j$. By unitarity, we get a finite answer by summing over all states b :

$$R_{ij,\text{all}}^+ = \sum_b \langle i|U(0, \infty)|b\rangle \langle b|U(\infty, 0)|j\rangle = \delta_{ij} < \infty \quad (2.2.7)$$

Thus to show Eq. (2.2.6), we only need to consider states b with energy outside of $D(E)$.

But any contribution with $E_b \neq E_i$ or $E_b \neq E_j$ must have at least one non-singular propagator. In this way, LN reduce the number of singular propagators using unitarity and are thereby able to use mathematical induction to complete their proof.

2.3 Final or initial state sums only

The key step in the LN proof is the employment of unitarity, in Eq. (2.2.7). Note that the sum over all states b includes the intermediate state where $|b\rangle = |i\rangle$, namely forward scattering. Once we accept that forward scattering must be included in the sum, we can prove a stronger result than the KLN theorem. Say we have an initial state $|a\rangle$, at $t = -\infty$ with energy E . Then the rate to produce any final states $\langle b|$ in an energy window $D(E)$ around E is finite:

$$R_{ab}^E = \sum_{j, E_j \in D(E)} \langle b|U(-\infty, \infty)|j\rangle \langle j|U(\infty, -\infty)|a\rangle < \infty \quad (2.3.1)$$

To prove this, we only need unitarity and energy conservation. Note that for the LN theorem, the matrix element $\langle b|U(\infty, 0)|a\rangle$ appeared. This matrix element can be non-vanishing when $E_a = E_b$. Thus the restriction to an energy window was non-trivial. Since $\langle b|U(\infty, -\infty)|a\rangle \propto \delta(E_b - E_a)$, energy must be conserved and the restriction on $E_j \in D(E)$

is the same as summing over all states. Removing the restriction, we then find

$$R_{ab}^E = \sum_j \langle b|U(-\infty, \infty)|j\rangle \langle j|U(\infty, -\infty)|a\rangle = \delta_{ab}, \quad (2.3.2)$$

which is finite. Note that we are not trying to make this trivial proof seem more complicated than it is – it really does just require completeness and unitarity. This is in contrast to the LN proof, which is less simple because of the required induction step due to energy non-conservation at finite time.

In words, we have shown that

- For a given fixed initial state $|a\rangle$, the cross section for $|a\rangle$ to go to anything is IR finite.

An analogous proof shows that

- For a given fixed final state $|b\rangle$, the cross section for anything to go to $|b\rangle$ is IR finite.

Note that in both cases, the states summed over include when the initial state and final state are the same, i.e. forward scattering. Importantly, however, we do not need to sum over final *and* initial states for IR finiteness.

Obviously, we do not want to sum over all possible states all the time: the probability for *anything*, including both something and nothing, to happen is 1. To get a physical prediction, we must remove a set of states from the sum whose production cross section is finite on its own. The question is then, what is the minimal sufficiently inclusive set of final states required for a finite cross section? For perturbative unitarity to hold, the virtual

emission graphs give

$$\sigma_R = \left| \text{---} \begin{array}{c} \nearrow \\ \searrow \end{array} + \text{---} \begin{array}{c} \nearrow \\ \searrow \end{array} \right|^2 = \sigma_0^d \Gamma_d \frac{e^2}{\pi^2} \left\{ \frac{1}{4\varepsilon^2} + \frac{3}{8\varepsilon} - \frac{7\pi^2}{48} + \frac{19}{16} \right\} \delta(1-z). \quad (2.4.2)$$

The IR singularities cancel between these two, giving the textbook result $\sigma_V + \sigma_R =$

$$\sigma_0 \frac{3e^2}{16\pi^2} \delta(1-z).$$

Note that for this process, the cross section is finite without including the $Z \rightarrow Z$ forward scattering contribution $Z \rightarrow Z$. Indeed, the forward-scattering amplitude for $Z \rightarrow Z$ is IR-finite to all orders in perturbation theory. This follows from the Kinoshita-Poggio-Quinn theorem [4, 28, 42–44]. It is also easy to see from general features of infrared divergences [45–50]: there are no massless external states, so there are no collinear divergences and the external lines are not charged, so there are no soft divergences.

Note that unitarity alone does not guarantee that these diagrams together are infrared finite. Strictly speaking, unitarity holds when summing over cuts of a fixed topology only if the on-shell states in the cut correspond to the those in the loop. In a covariant gauge, the photon propagator does not represent the sum over physical states. Thus only when a gauge invariant combination of all the relevant topologies is summed will unitarity hold. In this case, graphs with external-leg self-energy contributions are amputated while cuts though them (the individual graphs-squared in Eq. (2.4.2)) are included. This is the correct procedure as dictated by the LSZ reduction theorem, and the final result is gauge-invariant and subtraction-scheme-independent as it must be.

2.5 $e^+e^- \rightarrow Z + X$

Next let us consider the crossed process, $e^+e^- \rightarrow Z + X$. The virtual graphs are the same as for $Z \rightarrow e^+e^-$:

$$\begin{aligned}
 \tilde{\sigma}_{00} = & \text{Diagram 1} + \text{Diagram 2} \\
 & = \sigma_0^d \Gamma_d \frac{e^2}{\pi^2} \left\{ -\frac{1}{4\epsilon^2} - \frac{3}{8\epsilon} + \frac{7\pi^2}{48} - 1 \right\} \delta(1-z),
 \end{aligned} \tag{2.5.1}$$

where $\Gamma_d = \left(\frac{4\pi e^{-\gamma_E} \mu^2}{Q^2} \right)^{\frac{4-d}{2}}$, $\sigma_0^d = \sigma_0 \frac{d-2}{2} \mu^{4-d}$, and $\sigma_0 = \frac{4\pi g^2}{Q^2}$ with $\sigma_0 \delta(1-z)$ the tree-level cross section at center-of-mass energy Q . The real emission graphs give

$$\begin{aligned}
 \tilde{\sigma}_{01} = & \left| \text{Diagram 3} + \text{Diagram 4} \right|^2 \\
 & = \sigma_0^d \frac{e^2}{\pi^2} \Gamma_d \left\{ \delta(1-z) \left(\frac{1}{4\epsilon^2} - \frac{\pi^2}{16} \right) + \frac{1+z^2}{4} \left(-\frac{1}{\epsilon} \left[\frac{1}{1-z} \right]_+ + 2 \left[\frac{\ln(1-z)}{1-z} \right]_+ \right) \right\}.
 \end{aligned} \tag{2.5.2}$$

The sum of these graphs does not vanish: $\tilde{\sigma}_{00} + \tilde{\sigma}_{01} = \infty$. Here, our notation $\tilde{\sigma}_{nm}$ refers to the generalized cross section with n incoming photons and m outgoing photons (the generalized cross section is the same as the regular cross section for $2 \rightarrow n$ scattering where the incoming particles are massless, see Eq. (2.5.4) below).

What is different about radiation off incoming and outgoing electrons that changes the singularity structure? Note that for $Z \rightarrow e^+e^-\gamma$, both the soft and collinear singularities

have support only at $z = 1$, as can be seen in Eq. (2.4.2). For $e^+e^- \rightarrow Z\gamma$, if the photon is soft then $z = 1$, since a soft photon induces no recoil so the kinematics is the same as for $e^+e^- \rightarrow Z$. However, for a hard collinear photon, additional energy is needed in the final state above that in the Z boson, so we must have $z < 1$. Thus the $\frac{1}{\varepsilon} \frac{1}{1-z}$ pole in Eq. (2.5.2) is of collinear origin, and different from the $\frac{1}{\varepsilon} \delta(1-z)$ structure of the loop so cannot cancel it. That collinear photons are the origin of the difference is consistent with the Bloch-Nordsieck theorem: if the electron were massive, then there would be no collinear singularities and the cross section would be IR finite with either incoming or outgoing electrons.

2.5.1 Generalized cross section

In order to cancel the singularities coming from the loop, we can instead sum over initial states. To sum over initial states, we need a generalization of cross section that can apply to $n \rightarrow m$ scattering processes. First of all, we want to allow for forward scattering, so instead of writing $S = 1 + (2\pi)^d \delta^d(P_i^\mu - P_f^\mu) i\mathcal{M}$, we write

$$S = (2\pi)^d \delta^d(P_i^\mu - P_f^\mu) i\widetilde{\mathcal{M}}, \quad (2.5.3)$$

so that $\widetilde{\mathcal{M}}$ includes the forward scattering contribution.¹ Here P_i^μ is the sum of all the incoming particles' momenta and P_f^μ is the sum of all the outgoing particles' momenta. Then,

¹One might hope that IR finiteness could be achieved using \mathcal{M} in the conventional way, rather than $\widetilde{\mathcal{M}}$. Unfortunately, arguments based on cluster decomposition and analyticity that allow us to discard the $\mathbb{1}$ in S , and more generally the disconnected components, do not apply with massless particles, when the S matrix is IR divergent. A brief discussion can be found in [51, pp. 191-192].

rather than computing a cross section, we integrate over both initial and final state phase space. Because the result is Lorentz invariant, it is convenient to work in the center-of-mass frame. So we define

$$\tilde{\sigma} \equiv \frac{2^{2d-4} \pi^{2d-2}}{Q^{d-2}} \sum_{\text{spins}} \int d\Pi_i d\Pi_f |\widetilde{\mathcal{M}}|^2 (2\pi)^d \delta^d(P_i^\mu - P_f^\mu) \delta^{d-1}(\vec{P}_i + \vec{P}_f) \delta(P_i^0 - Q) \delta^{d-2}(\Omega_{d-1}^{(1)}), \quad (2.5.4)$$

where $\Omega_{d-1}^{(1)}$ corresponds to the angle of particle 1 and

$$d\Pi_i = \prod_{\text{initial states } j} \frac{d^{d-1}p_j}{(2\pi)^{d-1}} \frac{1}{2E_{p_j}}, \quad d\Pi_f = \prod_{\text{final states } j} \frac{d^{d-1}p_j}{(2\pi)^{d-1}} \frac{1}{2E_{p_j}}. \quad (2.5.5)$$

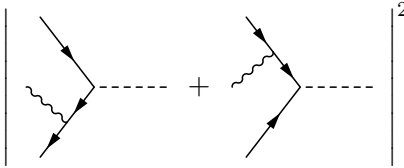
We always sum over initial and final state spins. For a fixed initial or final state, one can always divide by the number of spins to turn the sum into an average. We do not include this averaging factor so that $\tilde{\sigma}$ corresponds more precisely to what is proven to be infrared finite in Section 2.3.

The normalization is set so that this definition reduces to the usual definition of a cross section for $2 \rightarrow n$ processes where the incoming particles are massless. For example, Eqs. (2.5.1) and (2.5.2) still hold.

Note that we always sum over spins, for simplicity. One can consider more exclusive cross sections without the spin sum, but since all spins are summed in virtual contributions, we will often need to perform a spin sum to get a finite answer.

2.5.2 Initial and final state sum

Integrating inclusively over the initial state photon phase space at fixed center-of-mass energy Q gives

$$\tilde{\sigma}_{10} = \left| \begin{array}{c} \text{Diagram 1} \\ \text{Diagram 2} \end{array} \right|^2 = \sigma_0^d \frac{e^2}{\pi^2} \Gamma_d \left\{ \frac{1}{4\varepsilon^2} + \frac{3}{8\varepsilon} - \frac{7\pi^2}{48} + \frac{19}{16} \right\} \delta(1-z). \quad (2.5.6)$$


This is identical to Eq. (2.4.2) and the infrared divergences (soft and collinear) of these absorption graphs exactly cancel those from the loop in Eq. (2.5.1): $\tilde{\sigma}_{00} + \tilde{\sigma}_{10} < \infty$. This is not surprising as we are doing the identical integrals as for $Z \rightarrow e^+e^- (+\gamma)$.

Although the IR divergences of the loop are cancelled by absorption graphs in this way, the emission graphs in Eq. (2.5.2) cannot simply be ignored. There is no reason not to include final state radiation in the physical cross section. But since we have already used the loop to cancel the absorption singularities, what is left to cancel them? Since we have now accepted processes with additional photons in the initial state, we should also allow for all such processes. For example, we can have a diagram with an incoming and outgoing photon interfered with a disconnected graph (the importance of disconnected diagrams has been

observed in many contexts [5, 38, 52]). These diagrams give:

$$\begin{aligned}
\tilde{\sigma}_{11} = & \left(\begin{array}{c} \text{Diagram 1} + \text{Diagram 2} + \text{Diagram 3} \\ \text{Diagram 4} + \text{Diagram 5} + \text{Diagram 6} \end{array} \right) \text{---} \text{Diagram 7} + \text{c.c.} \\
= & \sigma_0^d \frac{e^2}{\pi^2} \Gamma_d \left\{ \delta(1-z) \left(-\frac{1}{2\varepsilon^2} + \frac{\pi^2}{8} \right) - \frac{1-z}{2} \right. \\
& \left. + \frac{3z^2 - 2z + 1}{2} \left(\frac{1}{\varepsilon} \left[\frac{1}{1-z} \right]_+ - \ln z \left[\frac{1}{1-z} \right]_+ - 2 \left[\frac{\ln(1-z)}{1-z} \right]_+ \right) \right\} \quad (2.5.7)
\end{aligned}$$

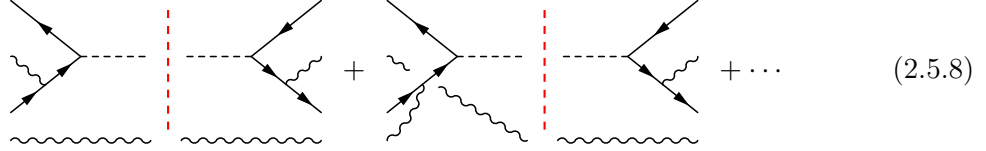
Evaluating these diagrams requires some care. Consider the first diagram for example.

When interfered with the diagram on the right, the outgoing photon momentum is forced to be the same as the incoming photon momentum. This puts one of the intermediate electron propagators on-shell. Normally, on-shell propagators are amputated, but in this case, the on-shell propagator is internal. Such singular propagators were handled by Lee and Nauenberg by including subleading terms in ϵ using the $i\epsilon$ prescription in time-ordered perturbation theory [5]. We find however, that their prescription does not work in our case. An alternative method was suggested by Lavelle and McMullan [38]. A similar situation also occurs when trying to factorize on-shell top production from decay [53]. Our approach is most similar to that of [53].

To deal with the on-shell intermediate state, we must recall that a propagator $\frac{i}{p^2 + i\varepsilon}$ is technically a distribution, defined only after integration. Similarly, the $\delta(p^2)$ putting the cut

electron on-shell is also a distribution. The product of these distributions must be treated as a distribution, proportional to $\delta'(p_0 - \omega_p)$, as we show with an explicit computation of the first diagram above in Appendix A.1. The sum of all the diagrams gives the result in Eq. (2.5.7). It is intriguing that one cannot interpret these cut diagrams as the product of an amplitude and a conjugate amplitude: the δ' distribution is only meaningful under the integral of the cut.

Once we have allowed for disconnected diagrams, nothing prevents contributions with disconnected photons both in $\widetilde{\mathcal{M}}$ and $\widetilde{\mathcal{M}}^*$, such as



$$(2.5.8)$$

We have to be careful in evaluating such graphs. If we contract the disconnected photons with each other, as shown in the first graph, then an extra $\delta^4(0)$ results. This extra infinity is expected by cluster decomposition as the S matrix must factorize into disconnected non-interfering pieces for separated processes [51]. We are not interested in those contractions here, and indeed they are not required by the KLN theorem, as they do not come from double-cut diagrams (see Section 2.5.3 below). Thus, when we draw diagrams like this we refer to only the connected interference component, like the second diagram in Eq. (2.5.8), where the disconnected photon in $\widetilde{\mathcal{M}}$ contracts with the absorbed photon in $\widetilde{\mathcal{M}}^*$ or vice-versa.² Note that the connected component must be gauge-invariant on its own as contribu-

²The connected interference component means that the uncut full double-cut diagram, wrapped on a cylinder (see Section 2.5.3), is connected. Thus a contribution to $\widetilde{\sigma}$ can be connected even if the contribution $\widetilde{\mathcal{M}}$ or $\widetilde{\mathcal{M}}^*$ is disconnected.

tions with different numbers of δ functions cannot cancel.

Focusing on the connected interference terms, in the center of mass frame, the outgoing Z and γ have energies $E_Z = \frac{Q^2+m^2}{2Q}$ and $E_\gamma = \frac{Q^2-m^2}{2Q}$ respectively. Since this photon contracts with the absorbed photon, both photons have the same momentum and so the e^+e^- pair has twice the 3-momentum of the Z and energy $E_{ee} \geq 2E_\gamma$. For energy to be conserved in the $e^+e^-\gamma \rightarrow Z$ subdiagram we must then have $E_Z = 3E_\gamma$, which only has solution for $Q > \sqrt{2}m_Z$ or equivalently $z > \frac{1}{2}$. This kinematic regime is the one we are interested in anyway as the singularity in the original $e^+e^- \rightarrow Z$ loop diagram occurred at $z = 1$ and so we want to focus on singularities in the $z \approx 1$ regime. The result is:

$$\begin{aligned}
\tilde{\sigma}_{21} &= \left| \begin{array}{c} \text{Diagram 1} \\ \text{Diagram 2} \end{array} \right|_{\text{connected}}^2 \\
&= \sigma_0^d \frac{e^2}{\pi^2} \Gamma_d \Theta(2z-1) \left\{ \delta(1-z) \left(\frac{1}{4\varepsilon^2} - \frac{\pi^2}{16} \right) \right. \\
&\quad \left. + \frac{5z^2 - 4z + 1}{4} \left(-\frac{1}{\varepsilon} \left[\frac{1}{1-z} \right]_+ + \ln(2z-1) \left[\frac{1}{1-z} \right]_+ + 2 \left[\frac{\ln(1-z)}{1-z} \right]_+ \right) \right\} \quad (2.5.9)
\end{aligned}$$

Note the θ -function enforcing the kinematical limit. In the kinematic regime we are interested in, $z > \frac{1}{2}$, the IR divergences in the sum of Eqs. (2.5.2), (2.5.7) and (2.5.9) exactly cancel: $\tilde{\sigma}_{01} + \tilde{\sigma}_{11} + \tilde{\sigma}_{21} < \infty$.

This is, however, not the end of the story. Once we agree that connected interference diagrams involving disconnected photons are allowed, we must also allow for such photons to be added to any of the diagrams we have already included. Since disconnected photons

do not change the order in the coupling, one can have an arbitrary number of them.³ We find for a process with m incoming photons and n outgoing photons, for $z > \frac{1}{2}$ and $n > 0$ that

$$\begin{aligned} \tilde{\sigma}_{mn} = \sigma_0^d \frac{e^2}{\pi^2} \Gamma_d (\delta_{m-1,n} - 2\delta_{m,n} + \delta_{m+1,n}) & \left\{ \delta(1-z) \left(\frac{1}{4\varepsilon^2} + \frac{\ln n}{\varepsilon} - \frac{\pi^2}{16} + \frac{\ln^2 n}{2} \right) \right. \\ & + \frac{2nz[n - m(1-z)] + (1-z)^2}{4n^2} \left[\left(-\frac{1}{\varepsilon} + \ln \left(\frac{n - m(1-z)}{n^3} \right) \right) \left[\frac{1}{1-z} \right]_+ \right. \\ & \left. \left. + 2 \left[\frac{\ln(1-z)}{1-z} \right]_+ \right] + \frac{1-z}{4n^2} \delta_{mn} \right\}. \quad (2.5.10) \end{aligned}$$

Note that the number of incoming and outgoing photons can differ by at most 1 at this order in the coupling. At higher order, there will be additional terms in $\tilde{\sigma}_{mn}$ farther from the diagonal.

It is easy to check from this formula that the IR divergences cancel for any fixed n , i.e. $\tilde{\sigma}_{n-1,n} + \tilde{\sigma}_{n,n} + \tilde{\sigma}_{n+1,n}$ is finite. Moreover, we find that if we sum over m first, then the sum over n is convergent. Indeed, at large n , the asymptotic behavior is

$$\tilde{\sigma}_{n-1,n} + \tilde{\sigma}_{n,n} + \tilde{\sigma}_{n+1,n} = \sigma_0^d \frac{e^2}{\pi^2} \Gamma_d \left\{ -\frac{(1-z)^3}{6z^2n^4} + \mathcal{O}\left(\frac{1}{n^6}\right) \right\}, \quad (2.5.11)$$

which is summable. Unfortunately, the series is not absolutely convergent and thus there is an ambiguity on the answer depending on the order in which the terms are summed [38, 40]. The ambiguity can be easily seen by considering reversing the order of the sum. Hold-

³One way to understand these multi-photon processes from the KLN theorem is that they original originate from diagrams where the photon wraps around the double-cut cylinder more than once [38].

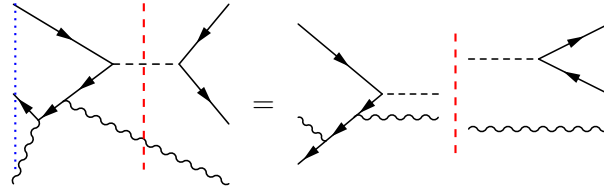
ing the number of initial-state photons fixed, we find that the sum over m of $\tilde{\sigma}_{m,m-1} + \tilde{\sigma}_{m,m} + \tilde{\sigma}_{m,m+1}$ is finite and scales like $\frac{1}{m^2}$ at large m . However, the sum of the $m = 0$ terms, $\tilde{\sigma}_{00} + \tilde{\sigma}_{01}$, is IR divergent: this is the original problematic sum $e^+e^- \rightarrow Z(+\gamma)$. So summing n first, then m we get infinity, while summing m first, then n we get a finite answer.

Even if one could come up with a consistent justification for how to sum the infinite series of $m \rightarrow n$ photon contributions, the physical interpretation is still unsettling. Even the $\tilde{\sigma}_{10}$ contribution, $e^+e^-\gamma \rightarrow Z$ is disturbing. The way we have done the calculation involved integrating over the entire kinematically accessible phase space for the incoming photon, including the region of hard, large-angle (non-collinear) photons. It is hard to justify why such photons should be involved in any experimental measurement of $e^+e^- \rightarrow Z$. Instead, we might try to restrict the integral to some infrared-and-collinear safe region. For example, we can consider scattering only incoming “jets” with invariant mass less than some cutoff m . Details using a hemisphere-jet mass definition are provided in Appendix A.2. We find that as with the total cross section, the jet mass cross section is also IR finite for any fixed number n of outgoing photons. Although the infinite sum retains the same ordering ambiguity as for the full cross section, it is closer to something that could conceivably be measured. Indeed, one can think of the initial-state jet mass calculation as a matching calculation if we set the jet masses equal to the physical electron mass. While this line of inquiry might ultimately be fruitful, it is not clear that at higher order in perturbation theory, or in more complicated theories like QCD, the IR divergences will still cancel without including forward scattering.

2.5.3 KLN interpretation

We saw that summing over $e^+e^- + m\gamma \rightarrow Z + n\gamma$ cross sections was infrared finite when summed over m and n . This is exactly the kind of cancellation the KLN theorem predicts: including all degenerate initial and final states guarantees finiteness. Although we found the relevant set of graphs by guessing all the relevant physical processes that might contribute at the same order in perturbation theory the KLN theorem actually tells us which subsets of diagrams should cancel: those coming from the double cuts of the same Feynman graph.⁴

The KLN theorem says that if we take a particular graph and identify the initial and final states, then all possible cuts of that graph should add up to a finite result. For example, the first $\tilde{\sigma}_{11}$ graph in Eq. (2.5.7) can be represented as:


(2.5.12)

where the red dashed line is the usual final-state cut and the blue dotted line represents an initial state cut. The diagram should be viewed as on a cylinder, with the right-hand side identified with the left-hand side. Then, for example, the square of the first real emission

⁴As noted before, although the proof works diagram-by-diagram, it requires unitarity which only holds if the propagator-numerators are the same as the sum over physical on-shell spin states. This is true for gauge theories in physical gauges (like axial gauge) but not true in covariant gauges (like Feynman gauge). One can work in Feynman gauge as long as all of the diagrams required to ensure gauge invariance are included.

graph in Eq. (2.5.2) can be drawn as

$$\text{Diagram with blue and red dashed lines} = \text{Diagram with red dashed line} \quad (2.5.13)$$

A different double-cut diagram can produce the third diagram in Eq. (2.5.7) or the disconnected diagram in Eq. (2.5.9):

$$\text{Diagram 1} = \text{Diagram 2}, \quad \text{Diagram 3} = \left| \text{Diagram 4} \right|^2 \quad (2.5.14)$$

Now, as we have observed, the double-cut sum in the KLN theorem also includes contributions where both cuts are in the same place, giving forward scattering contributions. For example, the double cut diagram in Eqs. (2.5.12) and (2.5.13) also generates forward-scattering cuts,

$$\text{Diagram with two vertical dashed lines} = \text{Diagram with square loop} \quad (2.5.15)$$

The double-cut sum also includes contributions where the initial and final state cuts are

swapped from Eq. (2.5.13)

$$\sigma_{\gamma Z ee} = \text{[Diagram 1]} = \text{[Diagram 2]} \quad (2.5.16)$$

The process corresponding to these cuts are $\gamma Z \rightarrow \gamma Z$ and $\gamma Z \rightarrow e^+e^-$ respectively, neither of which seems very much like the original $e^+e^- \rightarrow Z$ process whose singularities we were trying to cancel.

If the KLN theorem requires us to sum over this large number of contributions, why do a subset of them cancel among themselves? In some cases, we can find a clear answer. For example, we found that the virtual contribution to $e^+e^- \rightarrow Z$ was IR finite when summed with $e^+e^-\gamma \rightarrow Z$; these are contributions with a fixed final state, namely the Z . Since $Z \rightarrow Z$ is IR finite on its own, the sum of contributions of anything $\rightarrow Z$ will be finite whether or not we include forward scattering.

Similarly, we found $e^+e^- \rightarrow Z\gamma$ canceled against $\gamma e^+e^- \rightarrow Z\gamma$ and $\gamma\gamma e^+e^- \rightarrow Z\gamma$. These contributions all have final states with $Z\gamma$. Thus we could explain the cancellation among these terms alone if the forward scattering contribution $Z\gamma \rightarrow Z\gamma$ were infrared

finite. Evaluating the loop, we find:

$$\begin{aligned}
\tilde{\sigma}_{\gamma Z} &= \text{[Diagram: A square loop with a vertical dashed red line on the left and wavy lines on the bottom and right. Arrows indicate a clockwise flow.] } + \text{crossings} + \text{c.c.} \\
&= \sigma_0^d \frac{e^2}{\pi^2} \Gamma_d \times \begin{cases} \frac{5z^2 - 4z + 1}{4(1-z)} \left[-\frac{1}{\varepsilon} + 2 \ln(1-z) \right] + \frac{3z^2 - 2z + 1}{2(1-z)} \ln z + \frac{1-z}{2} & z < \frac{1}{2} \\ \frac{3z^2 - 2z + 1}{2} \ln z \left[\frac{1}{1-z} \right]_+ \\ \quad + \frac{-5z^2 + 4z - 1}{4} \ln(2z-1) \left[\frac{1}{1-z} \right]_+ + \frac{1-z}{2}, & z > \frac{1}{2} \end{cases}
\end{aligned} \tag{2.5.17}$$

In this expression, the $\mathcal{O}(\varepsilon)$ piece is evaluated at $z = 1$ since it only contributes when multiplying the $\frac{1}{\varepsilon}\delta(1-z)$ term from the expansion of the prefactor.

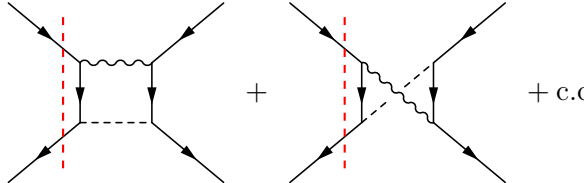
We see that the forward scattering contribution is infrared divergent, but only for $z < \frac{1}{2}$ ($Q^2 > 2m_Z^2$). For $z \approx 1$ ($Q \approx m_Z$), which is the limit in which we can examine the IR divergence associated with $e^+e^- \rightarrow Z$, the $\gamma Z \rightarrow \gamma Z$ forward scattering contribution is IR finite. This explains why the the sum over all other $X \rightarrow \gamma Z$ diagrams, $\sum \tilde{\sigma}_{n1}$ will be IR finite, as we have seen.

At high energy, $Q^2 > 2m_Z^2$, the forward scattering process is IR divergent. Note however that this is the identical to threshold above which the $2 \rightarrow 1$ diagrams $\tilde{\sigma}_{21}$ vanish. Since the singularities of $\tilde{\sigma}_{\gamma Z}$ and $\tilde{\sigma}_{21}$ are identical, the cross section to produce γZ is IR finite smoothly through the threshold. Despite the IR finiteness, the physical interpretation of the cancellation at high energy is a little strange: to produce a γZ from e^+e^- initial states, we

must also include initial states with γZ in them. On the other hand, we were not originally interested in γZ final states, but e^+e^- initial states, so a more relevant question is what states must we include along with e^+e^- to make a finite cross section?

2.5.4 Final states only

Unitarity implies that the sum over final states only, including forward scattering, is IR finite for any initial state. In this case, $e^+e^- \rightarrow X$ summed over all states X coming from cuts of the $e^+e^- \rightarrow Z$ diagrams must be finite on their own. We already computed $e^+e^- \rightarrow Z$ and $e^+e^- \rightarrow Z\gamma$, and they did not cancel by themselves. We cannot add additional photons in the initial state, but we must also consider e^+e^- forward scattering with γZ intermediate states. Adding the 2 box diagrams gives:

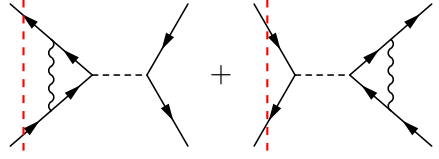


$$= \sigma_0^d \frac{e^2}{\pi^2} \Gamma_d \left\{ \delta(1-z) \left(-\frac{1}{4\varepsilon^2} + \frac{\pi^2}{16} \right) + \frac{1+z^2}{4} \left(\frac{1}{\varepsilon} \left[\frac{1}{1-z} \right]_+ - 2 \left[\frac{\ln(1-z)}{1-z} \right]_+ \right) \right\} \quad (2.5.18)$$

These contributions exactly cancel the real emission graphs in Eq. (2.5.2).

We saw that the IR divergences in the $e^+e^- \rightarrow Z$ 1-loop amplitude were canceled by real absorption graphs. To cancel these divergences without absorption graphs we need a different set of forward scattering diagrams, namely those containing the troublesome loop.

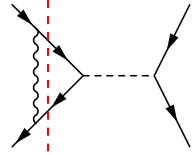
We find in this case



$$+ \text{c.c.} = \sigma_0^d \frac{e^2}{\pi^2} \Gamma_d$$

$$\times \left\{ \delta(1-z) \left(\frac{1}{4\epsilon^2} + \frac{3}{8\epsilon} + 1 - \frac{7\pi^2}{48} \right) + \frac{1-z}{(1-z)^2 + \hat{\Gamma}_z^2} \left(-\frac{1}{4\epsilon} - \frac{3}{8} \right) \right\}, \quad (2.5.19)$$

where $\hat{\Gamma}_z \equiv \frac{\Gamma_z m_z}{Q^2}$. There is also an additional cut to this diagram, representing a non-forward scattering $e^+e^- \rightarrow e^+e^-$ contribution:



$$+ \text{c.c.} = \sigma_0^d \frac{e^2}{\pi^2} \Gamma_d \left\{ \frac{1-z}{(1-z)^2 + \hat{\Gamma}_z^2} \left(\frac{1}{4\epsilon} + \frac{3}{8} \right) \right\} \quad (2.5.20)$$

The sum of these graphs cancel the cut graphs in Eq. (2.5.1).

Although the cancellation confirms the general theorem from Sec. 2.3, it is still a bit surprising: to cancel the infrared singularities in $e^+e^- \rightarrow Z + \text{photons}$ we must include states without a Z in them, namely $e^+e^- \rightarrow e^+e^-$. One way to explain this observation is that a Z boson can mix with an e^+e^- pair, so the two states are not distinguishable. Actually, this case has extra complications over other QED processes because the Z is massive and unstable (as it must be if it can be produced by massless particles). If the Z were massless, like the photon, the disconnected diagrams with $Z \rightarrow e^+e^-$ would not be allowed. We thus turn next to pure QED processes, Compton scattering and light-by-light scattering, to further explore the physics of infrared finiteness and forward scattering.

2.6 Compton scattering

Compton scattering is a simple example where where the KLN theorem, summing over degenerate initial and final states, but not including forward scattering, fails. At leading order, Compton scattering has an s - and t -channel contribution:

$$\tilde{\sigma}_{\text{C}} = \left| \begin{array}{c} \text{Diagram 1: } s\text{-channel Compton scattering} \\ \text{Diagram 2: } t\text{-channel Compton scattering} \end{array} \right|^2 \quad (2.6.1)$$

The s -channel graph makes a non-singular contribution at finite Q , but the t -channel graph has a pole. Regulating the divergence in d dimensions and working in Feynman gauge we find that the t -channel contribution is

$$\tilde{\sigma}_{\text{C}}^t = \frac{e^4}{\pi Q^2} \Gamma_d \left\{ -\frac{1}{2\varepsilon} + 1 \right\}, \quad (2.6.2)$$

with $\Gamma_d = \left(\frac{4\pi e^{-\gamma_E} \mu^4}{Q^2} \right)^{\frac{4-d}{2}}$.

What could cancel this singularity? We cannot dress the initial or final state electron with additional photons, as the contribution would be higher order in e . Because of unitarity, and the proof of cancellation, we can find the answer by simply drawing all possible

cuts:

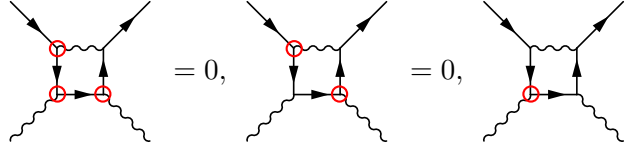
(2.6.3)

Actually, rather drawing the cuts as lines (or shaded lines), we find it clearest to enumerate the possible cuts as all possible circlings of the vertices, following 't Hooft and Veltman [54, 55]. For example,

(2.6.4)

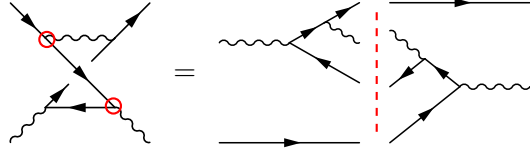
Lines going from uncircled vertices to uncircled vertices get a $+i\varepsilon$, lines going from circled to circled get a $-i\varepsilon$ and lines going from uncircled to circled are cut, so they get δ functions. Although not explained explicitly in [54, 55], an incoming line connecting to a circled vertex or an outgoing line connecting to an uncircled vertex, as in diagram D_1 gives a disconnected line. In this way, we see that there are 16 possible cuts. Most of these vanish. Indeed, if a connected set of circled vertices attaches only to incoming lines or a connected set of uncircled vertices attaches only to outgoing lines, the graph vanishes by energy con-

servation. Thus,



$$= 0, \quad = 0, \quad = 0, \quad \text{etc.} \quad (2.6.5)$$

Enumerating the cuts through circled vertices is particularly helpful for disconnected diagrams, like the crossed-box graphs in Eq. (2.5.18), where drawing lines through the graph is ambiguous. For example,



$$= \quad (2.6.6)$$

This is the connected component of the interference between the disconnected and connected graph (see discussion in Sec. 2.5.2). It vanishes for Compton scattering since the photon is massless even though the equivalent topology for $\gamma Z \rightarrow e^+e^-Z$ in Eq. (2.5.9), does not vanish.

Now let us look at some of the graphs. Cut D_1 is

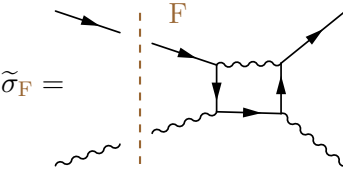


$$(2.6.7)$$

This diagram is the product of a disconnected graph for $e^- \rightarrow e^-$ and $\gamma \rightarrow e^+e^-$ and a connected graph for $e^-\gamma \rightarrow e^-e^+e^-$. The disconnected part only has support if the e^+e^- pair

is collinear to the incoming photon. In this phase space region, which is a set of measure zero over all of phase space, the connected part is non-singular and thus the product vanishes when integrated over phase space. (The disconnected diagrams in Eq. (2.5.7) had an $e^+e^- \rightarrow Z$ component which did not vanish because the Z boson is massive.) It is not hard to see that for Compton scattering at this order, all the cuts giving disconnected graphs are exactly 0.

Thus the only remaining contribution to cancel the divergence in the $\tilde{\sigma}_{\text{C}}$ is forward scattering. The forward scattering contribution, labeled F in Eq. (2.6.3) is the interference of the forward scattering non-interacting diagram and the box

$$\tilde{\sigma}_{\text{F}} = \text{F} = \frac{e^4}{\pi Q^2} \Gamma_d \left\{ \frac{i}{4\pi\varepsilon^2} + \frac{-2i + \pi}{4\pi\varepsilon} - \frac{1}{2} + \frac{i}{4\pi} - \frac{7i\pi}{48} \right\}. \quad (2.6.8)$$


In this contribution, the tree-level part only has support for forward scattering and the loop however is singular at $t = 0$. Their product is integrable in d dimensions, leading to the above result.

Adding to this the cut graph F^* , which is the complex conjugate to F gives

$$\sigma_{\text{F}} + \sigma_{\text{F}^*} = \frac{e^4}{\pi Q^2} \Gamma_d \left\{ \frac{1}{2\varepsilon} - 1 \right\}, \quad (2.6.9)$$

which exactly cancels the tree-level cross section, as expected.

On the one hand, this result should not come as a surprise. It is guaranteed by unitarity. However, note that the Compton diagrams had a singularity in the Bjorken $x = 1$ region,

where the entire momentum of the incoming electron is transferred to the outgoing photon. In contrast, the forward scattering contribution is at $x = 0$, where the momentum of the electron stays with the electron. Thus in the two cancelling contributions the hard electron is going off in entirely different directions. It seems like the question of whether a hard particle is an electron or photon should be physical. We find that instead, only the cross section for an hard electron *or* hard photon is finite.

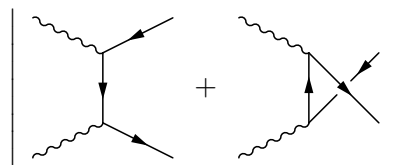
One way to understand why electrons and photons are effectively indistinguishable at high energy is that when the electron is massless, there is no energetic penalty to produce additional e^+e^- pairs from the vacuum. Thus, the state with a photon and electron can mix with one where a soft positron is created in the hard electron's direction, neutralizing its charge to produce a photon, and a soft electron is created in the hard photon's direction. Thus a hard electron going left and a hard photon going right can mix with the state of a hard electron going right and hard photon going left. Such mixing are exactly the degeneracies that must be summed over in the KLN theorem to get a finite result.

One could object to the reasoning here because the electron is in fact massive. Indeed, there are no massless particles in nature with nonzero electric charge, and such particles may not even be consistent (although gluons are, of course, massless particles charged under a different force). The point, however, is not to envision some fictitious theory with massless electrons. Rather, we want to understand when and how large contributions to the cross section are counterbalanced by superficially distinguishable processes. If the electron had a small mass, the collinear divergence would be regulated. We would then find the rate for producing a hard electron scales like $e^4 \ln \frac{Q}{m_e}$ at large $\frac{Q}{m_e}$ and the rate for producing

a hard forward photon scales like $1 - e^4 \ln \frac{Q}{m_e}$. Thus at high energy, when the logarithms become large, the two contributions should be added to restore perturbativity. In terms of the physical picture, when the center of mass energy of the collision becomes high enough, the energetic penalty to produce e^+e^- pairs takes a negligible amount of the total energy. In this way, a hard electron and hard photon become indistinguishable and their cross sections must be combined, according to the same logic as when $m_e = 0$.

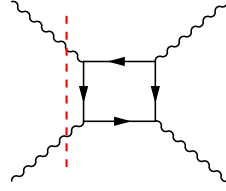
2.7 $\gamma\gamma \rightarrow X$

Perhaps the most powerful example of the failure of the initial-and-final-state sum picture is for light-by-light scattering – it is hard to argue that a photon is not a well-defined asymptotic state and that additional initial state particles must be added in $\gamma\gamma \rightarrow \gamma\gamma$. Let us consider then the process $\gamma\gamma \rightarrow e^+e^-$. The total cross section for this process is IR divergent due to the forward scattering region:

$$\left| \begin{array}{c} \text{Diagram 1} \\ \text{Diagram 2} \end{array} \right|^2 = \frac{e^4}{\pi Q^2} \Gamma_d \left\{ -\frac{1}{\varepsilon} + 1 \right\} \quad (2.7.1)$$


with $\Gamma_d = \left(\frac{4\pi e^{-\gamma_E} \mu^4}{Q^2} \right)^{\frac{4-d}{2}}$. As with Compton scattering, since the divergence is at tree-level, there can be no loop or bremsstrahlung contributions to cancel this singularity. Instead, the

singularity is canceled by the forward scattering amplitude



$$+ \text{crossings} + \text{c.c.} = \frac{e^4}{\pi Q^2} \Gamma_d \left\{ \frac{1}{\varepsilon} - 1 \right\} . \quad (2.7.2)$$

The sum of these is exactly zero, as expected by unitarity. Note that for this process, as for Compton scattering, all the diagrams with $1 \rightarrow 2$ disconnected pieces vanish exactly since all the particles are massless.

We have found something shocking: the total rate for photons to annihilate into charged particles is undefined. Similarly, the total rate for photons to annihilate into photons is undefined. Only the total cross section including photon final states and e^+e^- final states is IR finite.

2.8 Summary and Conclusions

This paper explores the question of which cross sections must be summed along with the cross section for a given process to produce an infrared finite result. Some of the main results of this paper are that

1. One never needs to sum over initial *and* final states to achieve IR finiteness, in contrast to expectations from the KLN theorem.
2. IR finiteness often requires the inclusion of forward scattering and the interference between disconnected and connected Feynman diagrams.

3. In QED with massless electrons, $e^+e^- \rightarrow Z$ can be made IR finite at the first nontrivial order by including
 - (a) outgoing photons and $e^+e^- \rightarrow e^+e^-$,
 - (b) incoming photons and $Z \rightarrow Z$, or
 - (c) an infinite number of processes dressing $e^+e^- \rightarrow Z$ with additional incoming or outgoing photons. Summing all the contributions, the cross section with fixed initial-state jet masses is convergent.
4. In $e^-\gamma \rightarrow e^-\gamma$, the tree-level IR divergence from the region with the outgoing γ collinear to the incoming e^- is canceled by the region with the outgoing γ collinear to the incoming γ .
5. The IR divergence in $\gamma\gamma \rightarrow \gamma\gamma$ scattering is cancelled by $\gamma\gamma \rightarrow e^+e^-$.

The first point is perhaps the most important observation in this paper. Although the KLN theorem instructs us to sum over degenerate initial and final states to produce an infrared finite cross section, in fact only the sum over initial *or* final states is necessary. Moreover, finiteness is only guaranteed if the forward scattering contribution is included.

In some cases, the forward scattering contribution is infrared finite on its own. An important example is $Z \rightarrow Z$. Its finiteness allows the rate for $Z \rightarrow e^+e^- + \text{photons}$ to be finite, or $Z \rightarrow \text{hadrons}$ to be finite in QCD. That the infrared singularities (in particular the collinear singularities associated with massless electrons or quarks) cancel is in a sense an accident of the simplicity of the $Z \rightarrow Z$ amplitude. For most other processes, forward

scattering is singular and must be included for infrared finiteness. For example, $Z\gamma \rightarrow Z\gamma$ at 1-loop is IR finite only in a small kinematic window but otherwise divergent.

Although one may sum only over final states, it may be important to consider initial state sums in some contexts. The example we studied in detail here was $e^+e^- \rightarrow Z$. For this process, one can add photons to the initial state and the cross section will be finite, as the process is then the exact crossing of $Z \rightarrow e^+e^-$. However, for $e^+e^- \rightarrow Z$, one cannot prevent the electrons from radiating photons into the final state. These $e^+e^- \rightarrow Z\gamma$ processes are infrared divergent, with their infrared divergences canceled in turn by additional diagrams with disconnected photons. We understood this cancellation as a initial-state sum cancellation because the $Z\gamma$ forward scattering amplitude is infrared finite when the center-of-mass energy is close to m_Z . Moreover, we found an infinite number of diagrams contributing at next-to-leading order in perturbation theory. Summing these diagrams, all the infrared divergences cancel. The infinite sum over the finite parts of all the diagrams appears to be convergent, although it is not clear how to interpret the result as hard wide-angle initial state photons are included. In Appendix A.2, we refined the calculation to initial-state jet masses, to eliminate the hard wide-angle photons, and still found convergence. It will certainly be interesting to consider connecting these infinite sums to experimental observables, as this is an example where an initial and final state sum gives a non-trivial result. Two reasons this may be challenging are that 1) the infinite sum over the finite parts is convergent but not absolutely convergent, so the result depends on how the terms are ordered and 2) it is not clear if the cancellation will hold at higher orders, or in more complicated theories like QCD.

An alternative to summing over initial and final states is to sum over just final states, but to include also the $e^+e^- \rightarrow Z \rightarrow e^+e^-$ forward scattering contribution. This contribution is IR divergent and cancels the IR divergence of the $e^+e^- \rightarrow Z$ virtual corrections as well as the $e^+e^- \rightarrow Z\gamma$ bremsstrahlung graphs. Thus, the forward scattering in this case achieves the same cancellation as the multiple initial and final state photons did, but avoids having to include disconnected diagrams and perform an infinite sum. On the other hand, when forward scattering is included, the total cross section is exactly zero at this order (as required by unitarity).

Additional insight came from examining Compton scattering. In Compton scattering, the total cross section with massless electrons is IR divergent. With an small electron mass, the total cross section diverges as $\sigma \sim \frac{e^4}{32\pi Q^2} \ln \frac{Q^2}{m_e^2}$. The singularity is from the kinematic region where the outgoing photon is collinear to the incoming electron. The large logarithm is canceled by the process $e^-\gamma \rightarrow e^-\gamma$ at 1-loop interfered with the disconnected forward scattering amplitude, so the outgoing photon is collinear to the incoming *photon*. This says that if a cosmic ray electron comes in at ultra-high energy, and scatters off a photon in the atmosphere one should not be able to distinguish a high-energy photon coming towards us from a high-energy electron. Only the sum of the two cross sections is IR finite (or free of large logarithms at high energy). We presented a physical justification for the indistinguishability: at very high energy, there is negligible energetic cost to the photon converting to an e^+e^- pair. If the positron produced is soft and the electron goes in the photon direction, then effectively the photon has transformed into an electron. From a practical point of view, since the electron is in fact massive and clearly distinguishable from a photon when it

is slow, the criterion for distinguishability must depend on some experimental resolution to identifying a conversion or charged tracks.

More broadly, we must question when the charge flowing into a certain direction is observable or only the net (global) charge. It seems that in addition to experimental limits on the energy and angles that can be resolved, there must also be an experimental limit on how well the momentum of a charged particle can be measured. That is, the notion of infrared-and-collinear safety might need to be extended to a restriction on charge measurement when massless initial states are involved (of if large logarithms are to cancel when initial state charged particles have mass).

Part of the reason we began investigating the KLN theorem was to gain a handle on the intricate subject of asymptotic states and the S -matrix. In particular, there are proposals that the S -matrix might be rendered IR finite if initial and final states are dressed as coherent states. While the original proposals focused on QED with massive electrons [8–10, 56] there have been extensions to the cases with massless charged particles [57–59], QCD [60, 61] and gravity [52]. While this coherent state approach is intuitively appealing – it certainly makes sense in the context of the $e^+e^- + \text{photons} \rightarrow Z + \text{photons}$ case we studied here – our observations indicate that the cancellations observed may be accidental. For example, we discussed photon scattering in Sec. 2.7. We showed there that the cross section for $\gamma\gamma \rightarrow \text{photons}$ is infinite in a theory with massless electrons. This IR divergence is canceled by the process $\gamma\gamma \rightarrow e^+e^-$. In a coherent state approach, one would attempt to achieve the cancellation at the amplitude level, but this would involve dressing the photons with electrons. While such a dressing is not inconceivable, it deviates from the the Faddeev-

Kulish idea that the IR divergences originate from the long-range interactions in the Hamiltonian. If it is possible to dress states so that the S -matrix is finite, the integrals involved in the dressing are likely to be closely related to integrals involved in achieving finite cross sections, like those we have studied here.

Finally, it is worth ruminating on how to connect infrared finiteness forward scattering to experimentally testable predictions. The key may be to understand better the initial state sums. Indeed, although we have shown that one *can* achieve IR finiteness with just a final state sum, it is not clear that this is the most physical way to proceed. A case in point is massive-electron QED, where the Bloch-Nordsieck theorem holds. In QED a final-state sum is sufficient in any process, such as for $e^+e^- \rightarrow Z + \text{photons}$. However, our analysis in Section 2.5 demonstrated that combining final state emission of $e^+e^- \rightarrow Z\gamma$ with the virtual $e^+e^- \rightarrow Z$ loop is morally equivocal:⁵ these contributions do not come from cuts of the same graph and their cancellation is accidental (a consequence of Abelian exponentiation). Indeed, in the massless-electron case, the cancellation does not work without also including $e^+e^- \rightarrow e^+e^-$ forward scattering. Alternatively, one can cancel the $e^+e^- \rightarrow Z$ loop against $\gamma e^+e^- \rightarrow Z$ graphs. Doing so not only cancels the IR divergences, but also the large logarithms of $\frac{m_e}{Q}$. Thus we actually have 3 different ways to compute e^+e^- annihilation: 1) with a final state sum, à la Bloch-Nordseick, whereby a large logarithm results 2) with a final state sum, including $e^+e^- \rightarrow e^+e^-$ whereby the inclusive cross section is zero or 3) with an initial state sum, where a finite cross section with no large logarithms results. Of these, the 3rd option may be the most appealing. However, to actually connect initial-

⁵A counterargument based on spacetime symmetries [62] can be found in [63].

state jets and disconnected diagrams to experiment will require understanding initial state sums in greater detail, to higher order, and in a more complicated yet more experimentally accessible theory, QCD.

3

A Finite S -matrix

One of the most fundamental objects in high energy physics is the scattering- or S -matrix. Not only is it a bridge between a definition of a quantum theory and data from particle colliders, but the study of the S -matrix itself has led to deep insights into the mathematical and physical pillars of quantum field theory itself. The idea behind the S -matrix is that it gives the amplitude for a set of particles in an “in” state $|\psi_{\text{in}}\rangle$ at $t = -\infty$ to turn into a different set of particles in an “out” state $\langle\psi_{\text{out}}|$ at $t = +\infty$. To go from this intuitive picture to a mathematically rigorous definition of the S -matrix has proven remarkably challenging. For example, suppose we take the in and out states to be eigenstates of the Hamiltonian H with energy E . Then they would evolve in time only by a phase rotation and the S -matrix elements would all have the form $\lim_{t \rightarrow \infty} e^{-2iEt} \langle\psi_{\text{out}}|\psi_{\text{in}}\rangle$. Such an S -matrix would be both ill-defined (because of the limit) and trivial (because of the projection). In non-relativistic quantum mechanics, one avoids this infinitely oscillating phase by subtracting from H the free Hamiltonian $H_0 = \frac{\vec{p}^2}{2m}$. More precisely, one looks for states $|\psi\rangle$ which, when evolved with the full Hamiltonian, agree with in and out states evolved with the free Hamiltonian: $e^{-iHt}|\psi\rangle \rightarrow e^{-iH_0t}|\psi_{\text{in}}\rangle$ as $t \rightarrow -\infty$ and $e^{-iHt}|\psi\rangle \rightarrow e^{-iH_0t}|\psi_{\text{out}}\rangle$ as $t \rightarrow +\infty$. Then the projection of in states onto out states is given by matrix elements $\langle\psi_{\text{out}}|S|\psi_{\text{in}}\rangle$ of the operator $S = \Omega_+^\dagger \Omega_-$ where the Møller operators are defined as $\Omega_\pm = e^{iHt_\pm} e^{-iH_0t_\pm}$, with t_\pm shorthand for the $t \rightarrow \pm\infty$ limit. In this way, the free evolution, which is responsible for the infinite phase, is removed. Note that $\lim_{t \rightarrow \pm\infty} e^{-iHt}|\psi\rangle$ is not a well-defined state, so the in and out states should be thought of as either Heisenberg picture states or as Schrödinger picture states at $t = 0$ not at $t = \pm\infty$ (see Fig. 3.1). Defining the S -matrix this way gives sensible results and a pleasing physi-

cal picture: particles we scatter are free when not interacting. Their freedom means they should have momentum defined by the free Hamiltonian and the S -matrix encodes the effects of interactions impinging on this freedom.

In quantum field theory, a similar construction is fraught with complications. The Møller operators which convert from the Heisenberg picture to the interaction picture, do not exist as unitary operators acting on a Fock space (Haag's theorem [64]). So one must work entirely in the Heisenberg picture without reference to H_0 . The matching of the states at $t \rightarrow \pm\infty$ is then replaced with an asymptotic condition on the matrix elements of fields. In the Haag-Ruelle construction [65–67], a mass gap is required to isolate the few-particle asymptotic states as limits of carefully constructed wave packets. From there, one can derive the LSZ reduction theorem, relating elements of the S -matrix to time-ordered products of fields [68, 69].

While it is satisfying to know that the S -matrix can be rigorously defined, its existence requires a theory with a mass gap, a unique vacuum state, and fields whose two-point functions vanish exponentially at spacelike separation. None of these requirements hold in any real-world theory. The practical resolution to this impasse is to ignore Haag's theorem, ignore that charged particles cannot be isolated and other assumptions, and simply use the LSZ reduction theorem as if it were true, computing S -matrix elements with Feynman diagrams. Although the resulting matrix elements are singular (due to infrared divergences) as long as one combines S -matrix elements computed this way into observable cross sections, the singularities will drop out. This is guaranteed by the KLN theorem [4, 5] which says that infrared divergences will cancel when initial and final states are summed over, or

by its stronger version, that the cancellation occurs when initial *or* final states are summed over [70]. Despite the success of this pragmatic approach, it remains deeply unsettling that the underlying object we compute, the S -matrix, has no formal definition even in QED.

There has been intermittent progress on constructing an S -matrix for QED (and QCD) over the last 50 or so years. The infrared divergence problem of the S -matrix can be seen already in non-relativistic scattering off a Coulomb potential. Because of its $\frac{1}{r}$ behavior, the Coulomb potential is not square integrable, and the asymptotic states do not exist. This complication was observed by Dollard [71], and resolved by using a modified Hamiltonian $H_{\text{as}}(t)$ that appends the dominant large-distance behavior of the Coulomb interaction to the free Hamiltonian. Chung [8], independently, observed that if instead of scattering single-particle Fock-state elements, one scatters linear combinations of these elements, similar to coherent states used in quantum optics (and to an early attempt by Dirac [72]), finite amplitudes would result. In Chung’s construction, the IR divergent phase space integrals from cross-section calculations are moved into the definition of the states. Faddeev and Kulish [10] subsequently redefined the S -matrix to include the dominant long-distance interactions of QED in its asymptotic Hamiltonian (similar to Dollard), and identified Chung’s coherent states as arising during the asymptotic evolution. Over the years, various subtleties in the coherent-state approach to soft singularities in QED have been explored [56, 73, 74], and attempts have been made to construct a finite S -matrix for theories like QCD with massless charged particles and hence collinear singularities [57–59, 61].

Remarkably, in all this literature, there are very few explicit calculations of what a finite S -matrix looks like. Indeed almost all of the papers concentrate on the singularities alone.

Doing so sidesteps the challenge how to handle finite parts of the amplitudes and precludes the possibility of actually calculating anything physical. With an explicit prescription, you have to contend with questions such as: what quantum numbers do the dressed states have? They cannot have well-defined energy and momentum outside of the singular limit, since they are superpositions of states with different numbers of non-collinear finite-energy particles.

The basic aspiration of much of this literature is that when there are long-range interactions, the S -matrix should be defined through asymptotic Møller operators $\Omega_{\pm}^{\text{as}} = e^{iHt_{\pm}} e^{-iH_{\text{as}}t_{\pm}}$ with some kind of asymptotic Hamiltonian H_{as} replacing the free Hamiltonian H_0 . Despite the simple summary, working out the details and establishing a productive calculational framework has proved a resilient challenge.

In this paper we continue the quest for a finite S -matrix by folding into the previous analysis insights from the modern understanding of scattering amplitudes and factorization. We argue that the principle by which the asymptotic Hamiltonian is to be defined is not that the dominant long-distance interactions be included (which allows for $H_{\text{as}} = H$ and $S = \mathbb{1}$), but that the evolution of the states be independent of how they scatter.

In gauge theories, infrared divergences can be either soft or collinear in origin. Both soft and collinear interactions are universal and can be effectively separated from the remainder of the scattering process. Factorization has been understood from many perspectives [49, 50, 75–81]. A precise statement of factorization can be found in [49, 50], where it is proven that the IR divergences of any S -matrix in QCD are reproduced by the product of a hard factor, collinear factors for each relevant direction, and a single soft factor. A useful

language for understanding factorization is Soft-Collinear Effective Theory (SCET) [77–83].

The SCET Lagrangian is

$$\mathcal{L}_{\text{SCET}} = -\frac{1}{4}(F_{\mu\nu}^s)^2 + \sum_n -\frac{1}{4}(F_{\mu\nu}^{c,n})^2 + \sum_n \bar{\psi}_n^c \frac{\not{n}}{2} \left[in \cdot D + i\not{D}_{c\perp} \frac{1}{i\bar{n} \cdot D_c} i\not{D}_{c\perp} \right] \psi_n^c + \mathcal{L}_{\text{Glauber}} , \quad (3.0.1)$$

where s and c, n are soft and collinear labels respectively; these act like quantum numbers for the fields. The derivation of the SCET Lagrangian and more details on the notation can be found in the reviews [82, 83]. The Glauber interactions denoted by $\mathcal{L}_{\text{Glauber}}$ are discussed in [84]; when they are included, the SCET Lagrangian can reproduce all of the IR singularities of any non-Abelian gauge theory. The main relevant features of the SCET Lagrangian are that 1) there are no interactions between fields with different collinear-direction labels (up to Glauber effects) and 2) collinear particles going in different directions only interact through soft photons or gluons with eikonal interactions. We define the asymptotic Hamiltonian H_{as} as the SCET Hamiltonian appended with the free Hamiltonians for massive particles.

In collider physics applications, one typically adds to the SCET Hamiltonian a set of operators necessary to reproduce the hard scattering of interest. For example, one might add $\Delta\mathcal{H} = C\bar{\psi}\gamma^\mu\psi$ for jet physics applications in e^+e^- collisions. Then one determines the Wilson coefficient C by choosing it such that matrix elements computed using SCET agree with matrix elements computed in the full theory. Importantly, the infrared divergences cancel in the difference, so that C is IR-finite order-by-order in perturbation theory. Motivated by such cancellations, we define hard Møller operators as $\Omega_{\pm}^H = e^{iHt_{\pm}}e^{-iH_{\text{as}}t_{\pm}}$ and the

hard S -matrix as $S_H = \Omega_+^{H\dagger} \Omega_-^H$. Because H_{as} reproduces the IR-divergence-generating soft and collinear limits of H , we expect the hard S -matrix will be IR-finite.

To evaluate matrix elements of S_H in perturbation theory, one could attempt to work out Feynman rules in an interaction picture based on H_{as} instead of H_0 . A propagator would then be a Green's function for H_{as} , which has no known closed-form expression. Alternatively, we can write S_H suggestively as (cf. [10, 57])

$$S_H = \Omega_+^{H\dagger} \Omega_-^H = \Omega_+^{\text{as}\dagger} \Omega_+^\dagger \Omega_- \Omega_-^{\text{as}\dagger}, \quad (3.0.2)$$

where $\Omega_\pm^{\text{as}} = e^{iH_0 t_\pm} e^{-iH_{\text{as}} t_\pm}$. This encourages us to define

$$|\psi_{\text{in}}^d\rangle = \Omega_-^{\text{as}\dagger} |\psi_{\text{in}}\rangle \quad \text{and} \quad |\psi_{\text{out}}^d\rangle = \Omega_+^{\text{as}\dagger} |\psi_{\text{out}}\rangle \quad (3.0.3)$$

as dressed in and out states. Then,

$$\langle \psi_{\text{out}} | S_H | \psi_{\text{in}} \rangle = \langle \psi_{\text{out}}^d | S | \psi_{\text{in}}^d \rangle. \quad (3.0.4)$$

We will take $|\psi_{\text{in}}\rangle$ and $|\psi_{\text{out}}\rangle$ to be eigenstates of the free momentum operator P_0^μ with a few (finite number of) particles in them. Thus we can think of S_H as computing either projections among few-particle states with the hard Møller operators or projections of dressed states with the original S -matrix Møller operators. For example, in the process $e^+e^- \rightarrow Z$ in QED, $|\psi_{\text{in}}\rangle$ would be an e^+e^- state of definite momentum and $|\psi_{\text{in}}^d\rangle$ a superposition of $|e^+e^- \rangle$, $|e^+e^-\gamma\rangle$, $|e^+e^-\gamma\gamma\rangle$, and so on.

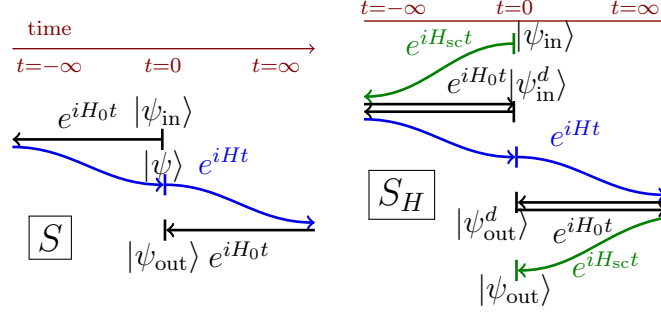


Figure 3.1: (Left) The traditional S -matrix is computed from Fock states evolved using H_0 and H . (Right) The hard S -matrix is computed either using Fock states evolved with H_{as} and H or using dressed states evolved with H_0 and H .

More explicitly, we can relate $|\psi_{\text{in}}^d\rangle$ to $|\psi_{\text{in}}\rangle$ using time-ordered perturbation theory (TOPT). For example, if $|\psi_{\text{in}}\rangle$ is the state of an electron with momentum \vec{p} , then in QED

$$|\psi_{\text{in}}^d\rangle = |\bar{u}_s(p)\rangle + e \sum_{\epsilon} \int \frac{d^{d-1}k}{(2\pi)^{d-1}} \frac{1}{2\omega_p} \frac{1}{2\omega_k} \frac{2p \cdot \epsilon}{\omega_k - \frac{\vec{p} \cdot \vec{k}}{\omega_p} - i\epsilon} |\bar{u}_s(p-k), \epsilon(k)\rangle + \dots \quad (3.0.5)$$

The denominator factor comes from the soft expansion of the TOPT propagator $(\omega_{p-k} + \omega_k - \omega_p - i\epsilon)^{-1}$. Note that the states in the expansion of $|\psi_{\text{in}}^d\rangle$ have different energies. Although electric charge and 3-momentum are conserved, energy is not as we evolve with Ω_-^{as} in TOPT. Due to the IR-divergent integral over \vec{k} , dressed states do not exist (in contrast to $|\psi_{\text{in}}\rangle$ and $|\psi_{\text{out}}\rangle$), but they do provide a useful qualitative handle on scattering.

As a concrete example, we now compute S_H for deep-inelastic scattering, $e^- \gamma^* \rightarrow e^-$ in QED with massless fermions at momentum transfer $Q = \sqrt{-q^2}$ in the Breit frame. At order e^2 , the loop contribution to the S -matrix element is, in $\overline{\text{MS}}$ and $d = 4 - 2\epsilon$ dimen-

sions [85],

$$\begin{aligned}
\mathcal{M}_A &= \text{Diagram: A wavy line (photon) connecting two vertices. The left vertex is a solid circle, and the right vertex is a solid circle. Two vertical dashed red lines are on either side of the vertices, labeled $t=-\infty$ and $t=\infty$ in red.} \\
&= \mathcal{M}_0 \frac{\alpha}{4\pi} \left[\frac{1}{\epsilon_{\text{UV}}} - \frac{2}{\epsilon_{\text{IR}}^2} - \frac{2 \ln \frac{\mu^2}{Q^2} + 4}{\epsilon_{\text{IR}}} - \ln^2 \frac{\mu^2}{Q^2} - 3 \ln \frac{\mu^2}{Q^2} - 8 + \frac{\pi^2}{6} \right], \tag{3.0.6}
\end{aligned}$$

with \mathcal{M} defined by $S_H = \mathbb{1} + (2\pi)^d \delta^d(q + p_1 - p_2) i \mathcal{M}$ and $\mathcal{M}_0 = -e \bar{u}(p_2) \gamma^\mu u(p_1)$ is the tree-level amplitude.

While this S -matrix element is IR-divergent, there are other contributions to S_H at the same order. These can be thought of as S -matrix elements for the $e^- \gamma$ components of $|\psi_{\text{in}}^d\rangle$ or $|\psi_{\text{out}}^d\rangle$. We can represent the new graphs as cuts through a broader graph, going from $0 \rightarrow -\infty \rightarrow \infty \rightarrow 0$. The first and last transitions go backward in time and represent the dressing and undressing of the state in the asymptotic regions. For example, the graph with both photon vertices coming from soft-collinear interaction in H_{as} is

$$\begin{aligned}
\mathcal{M}_B &= \text{Diagram: A wavy line (photon) connecting two vertices. The left vertex is a solid circle, and the right vertex is a solid circle. Two vertical dashed red lines are on either side of the vertices, labeled $t=-\infty$ and $t=\infty$ in red. Arrows indicate momentum flow: p_1 and p_2 are incoming, q is outgoing, and k is the photon momentum.} \\
&= \mathcal{M}_0 e^2 \mu^{4-d} \int \frac{d^{d-1}k}{(2\pi)^{d-1}} \\
&\quad \times \frac{1}{2\omega_k} \frac{1}{2\omega_1} \frac{1}{2\omega_2} \frac{8\omega_1\omega_2}{\omega_k - \frac{\vec{p}_1 \cdot \vec{k}}{\omega_1} - i\epsilon} \frac{1}{\omega_k - \frac{\vec{p}_2 \cdot \vec{k}}{\omega_2} - i\epsilon}. \tag{3.0.7}
\end{aligned}$$

To derive this integrand, we have power-expanded in the soft limit as in the method-of-regions approach [86], rather than using $\mathcal{L}_{\text{SCET}}$ directly. Although energy is not conserved in the asymptotic regions, the central region gives $\delta(\omega_k + \omega_{p_1-k} - \omega_k - \omega_{p_2-k}) \cong \delta(\omega_1 - \omega_2)$ which is factored out in the definition of \mathcal{M} .

This integral is scaleless and vanishes. Although we cannot easily separate all the UV and IR poles, the double soft/collinear pole in this amplitude is

$$\mathcal{M}_B = \mathcal{M}_0 \frac{\alpha}{4\pi} \left[-\frac{2}{\epsilon_{\text{IR}}^2} + \dots \right] \quad (3.0.8)$$

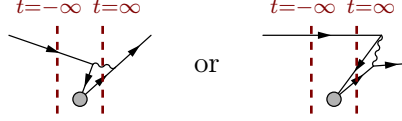
Focusing on the double pole also lets us restrict to just the soft graphs, as they contain the complete soft-collinear singularity. There are also graphs with one vertex coming from H_{as} and one coming from H :

$$\mathcal{M}_C + \mathcal{M}_D = \begin{array}{c} t=-\infty \quad t=\infty \\ \text{diagram 1} \end{array} + \begin{array}{c} t=-\infty \quad t=\infty \\ \text{diagram 2} \end{array} = \mathcal{M}_0 \frac{\alpha}{4\pi} \left[\frac{4}{\epsilon_{\text{IR}}^2} + \dots \right]. \quad (3.0.9)$$

The double IR pole from the S -matrix element cancels exactly in the sum $\mathcal{M}_A + \mathcal{M}_B + \mathcal{M}_C + \mathcal{M}_D$, as anticipated.

It is worth emphasizing the even the double-pole calculation is not trivial and requires careful manipulation of the distributions involved (cf. Ref.[70]). Moreover, the cancellation is different in nature from the cancellation in the computation of a Wilson coefficient. There the soft exchange graph (the analog of \mathcal{M}_B) is *subtracted* from \mathcal{M}_A ; here the graphs add, with the cancellation coming from graphs $\mathcal{M}_C + \mathcal{M}_D$ with one soft and one regular vertex.

The other TOPT diagrams involving soft-collinear vertices in H_{as} , such as


(3.0.10)

are not infrared divergent. In fact, the second diagram is zero, because there is no electron-positron annihilation vertex in H_{as} . Similarly, there are no diagrams with the hard vertex in the asymptotic regions, as H_{as} has only soft and collinear interactions.

To see the subleading IR poles cancel, we need a regulator other than pure dimensional regularization, such as offshellness (see [85, 87]), or explicit phase space restrictions. One should also then include graphs involving the collinear interactions in H_{as} as well as a zero-bin subtraction to avoid overcounting [87]. Using pure dimensional regularization is simplest, since all of the graphs other than \mathcal{M}_A are scaleless. Thus, after removing UV poles with renormalization, we find

$$\begin{aligned} \langle e^- | S_H | \gamma^* e^- \rangle &= (2\pi)^4 \delta^4(q + p_1 - p_2) \bar{u}(p_2) \gamma^\mu u(p_1) \\ &\times (-ie) \left[1 + \frac{\alpha}{4\pi} \left(-\ln^2 \frac{\mu^2}{Q^2} - 3 \ln \frac{\mu^2}{Q^2} - 8 + \frac{\pi^2}{6} \right) \right]. \end{aligned} \quad (3.0.11)$$

To confirm that the IR divergences cancel in S_H , without invoking scaleless-integral magic, we can impose physical cutoffs on the degrees of freedom that interact in H_{as} , such as including only photons with energy less than δ or within angle R of an electron [32]. Then the diagrams like \mathcal{M}_B are no longer scaleless. We have checked that all of the IR divergences cancel in S_H using this approach. Although S_H comes out IR finite, it retains sen-

sitivity to the scales R and δ ; in pure dimensional regularization, these cutoff scales are replaced by the single scale μ .

With a new definition of the S -matrix, it is natural to ask what are its predictions for observables? Consider an infrared-finite observable, such as the total cross section in $Z \rightarrow$ hadrons. To compute it, note that the total cross section for $Z \rightarrow \textit{anything}$, at order α_s is zero, since the forward scattering $Z \rightarrow Z$ cross section exactly cancels the cross section to everything else. This follows from unitarity, whether using S_H or S . Now, the Z has no soft or collinear interactions, so $|Z^d\rangle = |Z\rangle$. Thus $\langle Z|S_H|Z\rangle = \langle Z|S|Z\rangle$ to all orders in perturbation theory. Therefore the $Z \rightarrow Z$ forward-scattering cross section is the same with S_H and S , and so is the $Z \rightarrow$ hadrons cross section.

More generally, if we consider an observable less inclusive than the total cross section, such as a jet rate, then the details of the asymptotic dynamics will be important to determining the differential cross section. When we include this dynamics by evolving the final state with an $e^{-iH_{\text{as}}t+}$ factor, we would effectively be computing $\sum_X |\langle X|e^{-iH_{\text{as}}t}S_H|Z\rangle|^2 = \sum_X |\langle X|S|Z\rangle|^2$, so the differential cross section will agree *exactly* with one computed using S . Since infrared-safe cross sections computed using S are incontrovertible agreement with data, this is reassuring: we have not created more problems than we have solved with a finite S -matrix. On the other hand, there are also issues where physical predictions using S are ambiguous, such as with charged particles in the initial states. S_H could possibly shed light on these processes.

Having a finite S -matrix is perhaps most appealing in situations where the S -matrix is of interest for its own sake, for example, for its mathematical properties. One popular play-

ground for studying the mathematics of the S -matrix is $\mathcal{N} = 4$ super-Yang-Mills theory. This theory is a conformal gauge field theory. Although its S -matrix is UV finite, it is still IR-divergent. Moreover, its mathematical properties depend on how these IR divergences are removed. For example, the simplest approach is simply to drop the $\frac{1}{\epsilon_{\text{IR}}}$ terms, $\overline{\text{MS}}$ -style. Doing so for the planar 2-loop 6 particle amplitude, for example, gives a complicated function of the 9 kinematical invariants. If instead one employs the BDS-Ansatz, taking the ratio of the S -matrix element to the exponentiation of the 1-loop result [88, 89], then the result is a relatively simple “remainder function” of only the three dual-conformally invariant cross-ratios [90, 91]. While dual-conformal invariance is preserved by the BDS-Ansatz, the BDS remainder functions have unappealing analytic properties, such as violating the Steinmann relations [92]. A BDS-like Ansatz might preserve these [93]. A minimal normalization is another option [94]. In the computation of S_H , the IR divergences cancel automatically: the analog of the BDS subtraction comes naturally from multiplying $1/\epsilon$ counterterms for S_H with the finite $\mathcal{O}(\epsilon)$ parts S_H -matrix elements. Thus S_H -matrix elements provide some of the benefits of IR-finite remainder functions, without the arbitrariness of a ratio. Moreover, as the S_H operator is unitary, properties that follow from unitarity (perhaps including the Steinmann relations) should be automatically satisfied. This is in contrast to remainder functions which are quotients of S -matrix elements to other quantities.

In this paper, we have argued that there is nothing sacred about the traditional S -matrix. Its non-perturbative definition is absurdly complicated, and its interaction-picture definition involves an admixture of free and full-theory time evolution. In a theory with massless particles, it is natural to replace the free evolution with universal soft and collinear

evolution. Unlike S , whose matrix elements are either infinite (IR divergent) or zero (after exponentiation of the IR divergences), matrix elements of this new object S_H are IR finite to all orders.

In summary, this paper provides the first explicit construction of a S -matrix for non-Abelian gauge theories with no collinear or soft divergences; it provides rules (see also [95]) for computing S_H beyond just the cancellation of the singularities, allowing the mathematical properties of the S -matrix to be explored with the IR-divergence problem removed in a natural way; it connects to previous literature on dressed/coherent states, but also argues that such non-normalizable states are not needed for S_H or to compute observables; finally, it connects S_H -matrix elements to SCET and to remainder functions in $\mathcal{N} = 4$ SYM theory for the first time. While there is much still to be understood about S_H , it provides a solid starting point for an improved understanding of scattering in theories with massless particles.

4

An S -matrix for Massless Particles

4.1 Introduction

The scattering matrix, or S -matrix, is a fundamental object in physics. Intuitively, the S -matrix is meant to transform an “in” state $|\psi_{\text{in}}\rangle$ at $t = -\infty$ into an “out” state $\langle\psi_{\text{out}}|$ at $t = +\infty$. Unfortunately, constructing an operator in quantum field theory which achieves this projection is far from trivial. To begin, one might imagine that $S = \lim_{t \rightarrow \infty} e^{-iHt}$. However, this operator does not exist, even in a free theory. For example, acting on states with energies E_i , matrix elements of this operator would be infinitely oscillating phases. The proper resolution in quantum mechanics was first understood by Wheeler [96], who defined the S -matrix to project from a basis of metastable asymptotic states $|\psi_{\text{in}}\rangle$ (a nucleus) to other states (other nuclei) $|\psi_{\text{out}}\rangle$. This idea was expanded for use in quantum field theory by Heisenberg, Feynman, and Dyson [97–99] for calculations in quantum electrodynamics (QED). In modern language, one must factor out the evolution due to the free Hamiltonian H_0 to make S well-defined.

In the Wheeler-Heisenberg-Feynman-Dyson (henceforth “traditional”) approach, one assumes that in the far past, the “in” state is well-approximated with a freely evolving state, i.e. a state that evolves with the free Hamiltonian H_0 : $e^{-iHt}|\psi\rangle \rightarrow e^{-iH_0t}|\psi_{\text{in}}\rangle$ as $t \rightarrow -\infty$. The interaction is assumed to occur during some finite time interval so that in the far future, the time evolution is again nearly free: $e^{-iHt}|\psi\rangle \rightarrow e^{-iH_0t}|\psi_{\text{out}}\rangle$ as $t \rightarrow +\infty$. The state $|\psi\rangle$ is then related to the in and out states by Møller operators

$$\Omega_{\pm} = \lim_{t \rightarrow \pm\infty} e^{iHt} e^{-iH_0t} \tag{4.1.1}$$

as $|\psi\rangle = \Omega_+|\psi_{\text{out}}\rangle = \Omega_-|\psi_{\text{in}}\rangle$ and so $|\psi_{\text{out}}\rangle = S|\psi_{\text{in}}\rangle$ where the traditional S -matrix is defined as

$$S = \Omega_+^\dagger \Omega_- . \quad (4.1.2)$$

Unfortunately, this textbook approach has problems too: bare S -matrix elements computed this way are both ultraviolet (UV) and infrared (IR) divergent.¹ Ultraviolet divergences are by now completely understood: they are an artifact of computing S -matrix elements using unphysical fields in terms of unphysical (bare) parameters. When S -matrix elements are computed with physical, renormalized, fields in terms of physical observable parameters, the UV divergences disappear. IR divergences, however, are not as well understood and remain an active area of research. In theories with massless charged particles, such as QCD, S -matrix elements have IR divergences of both soft and collinear origin. Historically, three approaches have been explored to ameliorate the problem: the “cross section method”, the “dressed-state method” and the “modification-of- S method”.

The first way of dealing with IR divergences, referred to as the *cross section method* (following [57, 100]) is the most common. It argues that S -matrix elements themselves are not physical; only cross sections, determined by the squares of S -matrix elements integrated over sufficiently inclusive phase space regions, correspond to observables. Importantly, in this method, IR divergences cancel between virtual contributions and real emission contributions to *different* final states. The cancellation in QED was demonstrated definitively by Bloch and Nordsieck [2] in 1937. They showed that cross sections in QED (with massive

¹In this paper, we use “IR divergences” to refer to any divergence that is not of short-distance origin. So IR divergences come from both soft and collinear regions.

fermions) are IR finite order-by-order in perturbation theory when processes with all possible numbers of final state photons with energies less than some cutoff δ are summed over. The proof of Bloch-Nordsieck cancellation [24–26] crucially relies on Abelian exponentiation [24]: the soft singularities at any given order in α in QED are given by the exponential of the 1-loop soft-singularities. For theories with massless charged particles, such as QCD, Bloch-Nordsieck fails [3].

In non-Abelian gauge theories, the theorem of Kinoshita, Lee and Nauenberg (KLN) [4, 5] is often invoked to establish IR finiteness. The KLN theorem states that for any given process a finite cross section can be obtained by summing over all possible initial and final states for processes whose energy E lies within some compact energy window around a reference energy E_0 , i.e. $|E - E_0| < \delta$ for a given δ . In fact, the KLN theorem is weaker and its proof more complicated than required. First of all, energy is conserved, so the cancellation must occur without the energy window. Second of all, one does not need to sum over initial and final states; the sum over only final states for a fixed initial state will do, as will the sum over initial states for a fixed final state. This stronger version of the KLN theorem was proven recently by Frye et al. [70]. The proof is one line: for a given initial state, the probability of it becoming *anything* is 1, which is finite to all orders in perturbation theory. Importantly, both the KLN theorem and its stronger version by Frye et al. generically require the sum of diagrams to include the forward scattering contribution, which is usually excluded from a cross section definition. Unless they happen to be IR finite on their own, the forward scattering diagrams are crucial to achieve IR finiteness. Multiple illustrative examples can be found in [70]. If one wants the cross section to be finite when summing over

only a restricted set of final states, insights beyond Block-Nordsieck, KLN, and Frye et al. are required, such as those coming from factorization (e.g. [49, 50, 75–79, 81, 101]).

In the second approach to remedy IR divergences, the *dressed-state method*, the S -matrix is defined in the traditional way, but it is evaluated between states $|\psi^d\rangle$ that are not the usual few-particle Fock states $|p_1, \dots, p_n\rangle$. One of the first proposals in this direction was by Chung [8] (see also [102]), who argued that in QED one should replace single-particle electron states $|p\rangle$ with dressed states of the form $|p^d\rangle = e^R|p\rangle$ with R defined as

$$R|p\rangle = e \sum_{j=1}^2 \int \frac{d^{d-1}k}{(2\pi)^{d-1} \sqrt{2\omega_k}} \frac{p \cdot \epsilon_j(k)}{p \cdot k} a_k^{j\dagger} |p\rangle, \quad (4.1.3)$$

where ϵ_j is a photon polarization vector and $a_k^{j\dagger}$ is its corresponding creation operator.

The idea behind this dressing is that the eikonal factors $\frac{p \cdot \epsilon}{p \cdot k}$ give the real emission amplitude in the singular (soft) limit, which is then canceled by virtual contributions, so that $\langle p_3^d \dots p_n^d | S | p_1^d p_2^d \rangle$ is IR finite. The exponentiation of the eikonal interaction is the same mechanism (Abelian exponentiation) as invoked in the Bloch-Nordsieck cancellation. Indeed, the proof of the IR finiteness of these dressed states in QED is essentially the same as in the proof of the Bloch-Nordsieck theorem. This cloud of photons in the dressing has the same form as Glauber’s coherent states [103] used in quantum optics [104, 105] (these are, roughly speaking, $e^R|0\rangle$), and so the dressed states in this case are commonly called *coherent states*.

While the coherent state approach is in some ways appealing, it has drawbacks. The main problem is that the IR divergences are just moved from the amplitudes to the states.

That is, the coherent states themselves are IR divergent and therefore not normalizable elements of a Fock space (although they may be understood as living in a non-separable von Neumann space, as explained in a series of papers by Kibble [9, 73, 106, 107]). The IR divergence problem is therefore still present in this construction; it has merely been moved from the S -matrix elements to the states of the theory. Additionally, generalizing beyond massive QED to theories like QCD with collinear divergences and color factors has remained elusive [108, 109]. In particular, no prescription is given for how to go beyond the singular points (zero energy or exactly collinear). For example, the coherent states are sums over particles with different momenta, so they do not have well-defined momenta themselves. Is momentum then conserved by the S -matrix in the coherent-state basis? How does one integrate over coherent states to produce an observable cross section? These problems are not commonly discussed in the literature. As far as we know, no one has explicitly computed an S -matrix element between coherent states. This defect gives the coherent-state literature a rather formal aspect.

The third approach to removing the IR divergences in scattering theory is to redefine the S -matrix rather than the states. That the traditional S -matrix inaccurately captures the asymptotic dynamics arises already in non-relativistic scattering of a charged particle off a Coulomb potential in non-relativistic quantum mechanics. The standard assumption that particles move freely at asymptotic times is not justified for non-square-integrable potentials, like the $\frac{1}{r}$ Coulomb potential, and leads to ill-defined S -matrix elements. In modern

language, the S -matrix element for non-relativistic Coulomb scattering has the form

$$\langle \vec{p}_f | S | \vec{p}_i \rangle \sim \frac{\alpha}{(\vec{p}_i - \vec{p}_f)^2} e^{-i\alpha \frac{m}{|\vec{p}_i - \vec{p}_f|} \frac{1}{2\epsilon_{\text{IR}}}}. \quad (4.1.4)$$

We see that the leading term of order α , corresponding to the first Born approximation, is not problematic: except in the exactly forward limit, there are no divergences in the tree-level scattering process. The logarithmic IR divergence (showing up as a $\frac{1}{\epsilon_{\text{IR}}}$ pole in $d = 4 - 2\epsilon$ dimensions) first appears in the second Born approximation, where it is seen to be purely imaginary. Moreover, the IR divergent part exponentiates (as do all IR divergences in QED), into the *Coulomb phase*. Thus, in non-relativistic quantum mechanics, one can apply the cross section ideology even without the inclusive phase space integrals: the cross section for the scattering of a single electron off a Coulomb potential is well-defined. However, the S -matrix is not.

One of the first attempts to define an S -matrix for potentials that are not square-integrable was made by Dollard [71] in 1971. He noted that when incoming momentum eigenstates are evolved to late times with the Coulomb interaction $H = H_0 + \frac{\alpha}{r}$, there is a residual logarithmic time dependence for large t :

$$e^{-i \int^t H(t') dt'} |p\rangle \cong e^{-i \left(\frac{p^2}{2m} t + \frac{m\alpha}{|p|} \ln t \right)} |p\rangle \quad (4.1.5)$$

The intuition for this form is that at large t , the particle moves approximately on a classical trajectory with $r = \frac{pt}{m}$, which gives the logarithmic dependence on t when integrated up to infinity. While the $e^{-i \frac{p^2}{2m} t}$ is removed by Wheeler's $e^{iH_0 t}$ factor, the other term is not

and persists to generate the $\frac{1}{\epsilon_{\text{IR}}}$ divergences in the S -matrix. Dollard then proposed to replace to the $e^{iH_0 t}$ factor with a $e^{iH_{\text{as}}(t)}$ factor, with $H_{\text{as}}(t)$ defined with exactly the logarithmic time dependence needed to cancel the time dependence in Eq. (4.1.5). He then showed that a modified S -matrix, defined with his asymptotic Hamiltonian replacing H_0 , exists for Coulomb scattering.

When the electron is relativistic, the IR divergence in the second Born approximation has a real part that does not cancel at the cross section level. So first-quantized quantum mechanics is insufficient to produce an IR-finite cross section: QED is needed. Faddeev and Kulish [10] combined the aforementioned work of Chung in QED and Dollard's in non-relativistic quantum mechanics. They observed that in QED, infrared divergences have both a real part (as Chung observed) and an imaginary part (the relativistic generalization of the Coulomb phase). These can be combined into a modified S -matrix of the form

$$S_{\text{FK}} = \lim_{t_{\pm} \rightarrow \pm\infty} e^{-R(t_+)} e^{-i\Phi(t_+)} S e^{-i\Phi(t_-)} e^{R(t_-)}, \quad (4.1.6)$$

where

$$\Phi(t) = \frac{\alpha}{2} \int \frac{d^3 p}{(2\pi)^3} \frac{d^3 q}{(2\pi)^3} : \rho(p) \rho(q) : \frac{p \cdot q}{\sqrt{(p \cdot q)^2 - m^4}} \ln |t| \quad (4.1.7)$$

corresponds to the Coulomb phase (compare to the $\ln t$ dependence in Dollard's form, Eq. (4.1.5)). The factor R is similar to Chung's in Eq. (4.1.3) but with a power-expanded phase, and annihilation operators included as well:

$$R(t) = e \sum_{j=1}^2 \int \frac{d^3 p}{(2\pi)^3} \frac{d^3 k}{(2\pi)^3 \sqrt{2\omega_k}} \left[\frac{p \cdot \epsilon_j^*(k)}{p \cdot k} a_k^{j\dagger} e^{i \frac{p \cdot k}{\omega_p} t} - \frac{p \cdot \epsilon_j(k)}{p \cdot k} a_k^j e^{-i \frac{p \cdot k}{\omega_p} t} \right] \rho(\vec{p}) \quad (4.1.8)$$

where

$$\rho(p) = \sum_s \left(a_p^{s\dagger} a_p^s - b_p^{s\dagger} b_p^s \right) \quad (4.1.9)$$

is the electron-number operator. Acting on states, it pulls out the direction p of each fermion and multiplies the contribution by 1 for electrons or -1 for positrons:

$$\rho(p)|q_1 \cdots q_n\rangle = \sum \pm (2\pi)^3 \delta^3(\vec{p} - \vec{q}_j) |q_1 \cdots q_n\rangle \quad (4.1.10)$$

Faddeev and Kulish proceed to argue that S_{FK} has finite matrix elements between coherent states in QED. They argued that one should include the phase factors in a redefinition of the S -matrix while including the e^R factors in dressing the states. Although there are some suspicious orders-of-limit and signs in Faddeev and Kulish's paper (see [57]), we believe their construction is essentially valid. Indeed, one goal of our paper is to translate this classic work in QED to modern language. As we will show in Section 4.2.3, both the real and imaginary parts in the factor $e^{i\Phi(t-)}e^R$ are reproduced by the action of a single Wilson line.

In the 50 odd years since Faddeev and Kulish's work, there has been intermittent progress on generalizing the coherent state construction from QED to non-Abelian theories. Early work [108, 110, 111] focused on trying to use coherent states to salvage the Bloch-Nordsieck theorem, following the QCD counterexamples given by Doria et al. [3, 30]. Although soft divergences in QCD do not exponentiate into a compact form as they do in QED [112, 113], they still have a universal form and factorize off of the hard scattering [50, 75]. Using this observation, it has been argued using a frequency-ordered for-

malism that soft-finite dressed states can be constructed between S -matrix elements in QCD [61, 108]. Collinear divergences and the soft-collinear overlap in gauge theories were explored in [57, 58, 114, 115]. An explicit check of the dressed formalism was performed by Forde and Signer [59] who used explicit cutoffs to separate the regions and showed that the cross section for $e^+e^- \rightarrow \text{jets}$ can be reproduced at leading power at order α_s through finite S -matrix elements. Ref. [58] argued that if soft-collinear factorization holds in QCD, then the dressed state formalism should allow one to construct a finite S -matrix in QCD to all orders. Collinear factorization was proven diagrammatically at large N a decade later [116] and a full proof of collinear factorization and soft/collinear factorization for QCD to all orders in perturbation theory was given in [49, 50], inspired by [75, 76, 78, 80, 81, 101]. One goal of the current paper is to combine these various insights to provide, for the first time, an explicit construction of an IR-finite S -matrix for QCD.

In all of this literature, there are a number of unresolved issues. First, there are essentially no results about the finite parts of a finite S -matrix. Showing the cancellation of the IR singularities is one thing, but to evaluate S one needs to deal with complications of momentum conservation, cutoffs, UV divergences, and to actually be able to compute the resulting integrals. A prescription to determine the finite parts of the modified S -matrix is required if we are to explore the S -matrix's properties. While some authors have suggested criteria such as that the dressed states should be gauge [74] or BRST invariant [117], or have asymptotic charges [62, 63], or be compatible with decoherence [118–120], the necessity of these choices is unclear. Certainly nothing goes wrong at the level of cross sections if we proceed using the cross section method. After the finite part is fixed, one must

further explain how to relate modified S -matrix elements to observables: what is the measure for integration over momenta in the von Neumann space of dressed states (if one goes that route)? To agree with data, the predictions had better reduce to what one calculates using the IR-divergent S , but how that will happen in any of the approaches to dressed states is rarely discussed. In this paper, we attempt to raise the bar for constructing a finite S -matrix by providing a motivated, calculable scheme, and give explicit expression for S -matrix elements and observables in a number of cases in QED, QCD, and $\mathcal{N} = 4$ super Yang-Mills theory.

The organization of this paper is as follows. We start by motivating and defining a “hard” S -matrix in Section 4.2. We show how to get finite answers, and connect to the previous work on QED using dressed states in Section 4.2.1. In Section 4.2.2, we discuss how to compute observables and show that the same predictions for infrared-safe differential cross sections results from S_H as from the traditional S . In Section 4.2.3 we connect our construction to the expressions of Faddeev and Kulish in QED. We then proceed to explicit calculations, working out the Feynman rules and some toy examples in Section 4.3. In Section 4.4 we demonstrate IR finiteness in the process $\gamma^* e^- \rightarrow e^-$ in QED using cutoffs, and illustrate the relative simplicity when pure dimensional regularization is invoked. In Section 4.5 we discuss $Z \rightarrow e^+ e^-$ including the connection to the Coulomb phase and the Glauber operator as well as an explicit calculation of the thrust distribution, both exactly at NLO and to the leading logarithmic level using the asymptotic interactions. Section 4.5.2 makes explicit some of the general observations about exclusive measurements from Section 4.2.2. Section 4.6 gives some examples in $\mathcal{N} = 4$ super Yang-Mills theory, connecting to obser-

vations about remainder functions, renormalization and subtractions schemes. Concluding remarks and a summary of our main results are given in Section 4.7.

4.2 The hard S -matrix

The intuition behind scattering is that one starts with some initial state, usually well-approximated as a superposition of momentum eigenstates, which then evolves with time into a region of spacetime where it interacts, and then a new state emerges. The S -matrix is meant to be a projection of this emergent final state onto a basis of momentum eigenstates. For scattering off a local (square-integrable) potential, this picture works fine. The S -matrix is then defined as $S = \Omega_+^\dagger \Omega_-$ as in Eq. (4.1.2) with the Møller operators Ω_\pm defined in Eq. (4.1.1). However, when the interactions cannot be confined to a finite-volume interaction region, as in Coulomb scattering or in a quantum field theory with massless particles, this picture breaks down: the states at early and late times continue to interact, so the momentum-eigenstate approximation is no longer valid.

As mentioned in the introduction, the simplest example with the traditional definition of S breaks down is for non-relativistic scattering off a Coulomb potential. In this case, the Møller operators acting on momentum eigenstates generate an infrared divergent “Coulomb” phase. While the infrared divergence is a problem for a formal definition of the S -matrix, it is not a problem for cross section calculations that depend only on squares of S -matrix elements. In relativistic Coulomb scattering, or in QED, S has both an infrared divergent Coulomb phase and an infrared divergent real part. A convenient feature (Abelian exponen-

tiation [24]) of QED is that a closed form expression is known for the IR-divergent contribution to all orders in perturbation theory for any process. Indeed, the 1-loop divergences are given by $S \sim \frac{\gamma_{\text{cusp}}}{\epsilon_{\text{IR}}}$ where the cusp-anomalous dimension is (see [121])

$$\gamma_{\text{cusp}} = -\frac{\alpha}{\pi} [(\beta - i\pi) \coth \beta - 1] , \quad (4.2.1)$$

with the cusp angle defined by $\cosh \beta = \frac{v_1 \cdot v_2}{|v_1||v_2|}$ and $v_1^\mu = \frac{p_1^\mu}{E_1}$ and $v_2^\mu = \frac{p_2^\mu}{E_2}$ are the 4-velocities of the incoming and outgoing electrons. To all orders, the IR divergences exponentiate as $S \sim \exp \frac{-\gamma_{\text{cusp}}}{2\epsilon_{\text{IR}}}$ [122]. Thus, it is possible to factor out IR-divergent parts from the S -matrix and redefine a new S -matrix that is IR-finite order-by-order. This was done by Chung and Faddeev and Kulish, as discussed in the introduction. Note that the non-relativistic limit corresponds to $\beta \rightarrow 0$ in which case $\gamma_{\text{cusp}} = i\alpha \frac{1}{\beta}$ becomes the purely imaginary Coulomb phase.

When the charged particles are also massless, as in QED with $m_e = 0$, new IR divergences appear associated with collinear divergences. Soft-collinear divergences appear as double IR-poles. Indeed, in the $m_e \rightarrow 0$ limit, v_i^μ becomes lightlike, so $\beta \rightarrow \infty$. At large β in the cusp angle $\gamma_{\text{cusp}} \sim -\frac{\alpha}{\pi}\beta$ diverges linearly with β , so the S -matrix now has double, $\frac{1}{\epsilon_{\text{IR}}^2}$ poles. In QCD, or other non-Abelian theories, the cusp angle gets corrections beyond one loop and the IR divergences do not exponentiate into a closed form expression [112, 113, 123]. These complications have made it difficult to come up with a complete formulation of an IR-finite S -matrix in general quantum field theories [57–59].

The approach we take in this paper is to construct an S -matrix that is IR finite by re-

placing the free Hamiltonian H_0 in the definition of the traditional S -matrix with an appropriate asymptotic Hamiltonian H_{as} . That is, we can define new **hard Møller operators**

$$\Omega_{\pm}^H = \lim_{t_{\pm} \rightarrow \pm\infty} e^{iHt_{\pm}} e^{-iH_{\text{as}}t_{\pm}}, \quad (4.2.2)$$

and a **hard S -matrix** as

$$S_H = \Omega_+^{H\dagger} \Omega_-^H. \quad (4.2.3)$$

Ideally, we would want to choose H_{as} so that the hard Møller operators exist, as unitary operators on the Hilbert space. Proving their existence is challenging, as even in a mass-gapped theory, where we can take $H_{\text{as}} = H_0$, they do not exist by Haag's theorem [64].

From a practical point of view, we can be less ambitious and aim to choose H_{as} so that the hard S -matrix is free of IR divergences at each order in perturbation theory. If this was our only criteria, we could choose $H_{\text{as}} = H$, so that $S_H = \mathbb{1}$.

A better criteria for defining H_{as} is that, in addition to capturing long-distance interactions, the asymptotic Hamiltonian should be defined so that the asymptotic evolution of the states is independent of how they scatter. It is possible to define H_{as} this way due to universality of infrared divergences in gauge theories. Using factorization [49, 50, 75–79, 81, 101], the soft and collinear interactions can be separated from the hard scattering process: Any S -matrix element in gauge theories can be reproduced by the product of a hard factor, collinear factors for each relevant direction, and a single soft factor. See [50] for a concise statement of factorization at the amplitude level.

In order to exploit factorization, we employ methods developed in Soft-Collinear Effective

Theory (SCET). The theory provides a systematic power expansion of the QED or QCD Lagrangian, and reproduces all infrared effects. The leading power Lagrangian in SCET is [82, 83]

$$\begin{aligned} \mathcal{L}_{\text{SCET}} = & -\frac{1}{4}(F_{\mu\nu}^s)^2 + \sum_n -\frac{1}{4}(F_{\mu\nu}^{c,n})^2 \\ & + \sum_n \bar{\psi}_n^c \frac{\not{n}}{2} \left[i n \cdot D_c + g n \cdot A_s^a(x_-) T^a + i \not{D}_{c\perp} \frac{1}{i \bar{n} \cdot D_c} i \not{D}_{c\perp} \right] \psi_n^c + \mathcal{L}_{\text{Glauber}}, \end{aligned} \quad (4.2.4)$$

where s and c, n are soft and collinear labels respectively and the collinear covariant derivative is

$$iD_\mu^c = i\partial_\mu + g A_\mu^{c,a} T^a. \quad (4.2.5)$$

The last term $\mathcal{L}_{\text{Glauber}}$ describes Coulomb or Glauber gluon interactions [84] (see also [22]). Pedagogical introductions to SCET can be found in [82, 83, 124].

We define the asymptotic Hamiltonian H_{as} to be the SCET Hamiltonian appended with free Hamiltonians for massive particles. The hard S -matrix is then defined in terms of H_{as} using Eqs. (4.2.2) and (4.2.3).

Although the SCET Lagrangian looks complicated and non-local, much of the complication comes from being careful to include only leading-power interactions. In principle, for a theory to be valid at leading power, one could include any subleading power interactions one wants. Exploiting this flexibility, the collinear interactions in $\mathcal{L}_{\text{SCET}}$ can be replaced simply with the full interactions of QCD: $i\bar{\psi}_n^c \not{D}_c \psi_n^c$. The soft interactions, from the $\bar{\psi}_n^c \frac{\not{n}}{2} n \cdot A_s^a(x_-) \psi_n^c$ term, are also not that complicated: they are equivalent to treating the

collinear fermions as being infinitely energetic, with no recoil. That is, the fermions act as classical sources for radiation moving in a straight line along the n^μ direction. This leads to an alternative representation of the soft interactions as coming from Wilson lines. This connection is made more precise in Section 4.2.3.

In practice, when computing S_H elements we will not use the explicit and cumbersome interactions in $\mathcal{L}_{\text{SCET}}$. Instead, we will take the method-of-regions approach [82, 86]. We start with a particular Feynman diagram and then expand to leading power based on the collinear or soft scaling associated with particles involved. In a sense, this is the most straightforward and foolproof way to compute S_H amplitudes. Numerous examples are given in subsequent sections.

We also, in accord with the general principles of the method of regions, do not impose any hard cutoffs on the momenta of the soft and collinear particles that interact through H_{as} . Imposing cutoffs is helpful for demonstrating explicit IR-divergence cancellation, and some examples are provided in Section 4.4.1. However, cutoffs generally lead to very difficult integrals, and moreover they break symmetries like gauge-invariance that we would like S_H to respect. More precisely, it is only the finite, cutoff-dependent remainder terms that may depend on gauge – the IR divergence cancellation mechanism is gauge-independent. Since the cutoff-dependent finite parts are unphysical anyway, it is not a problem that they are also gauge-dependent. In general, however, the whole framework with cutoffs is rather unwieldy.

When using pure dimensional regularization, the diagrams involving H_{as} interactions will lead to scaleless integrals. These integrals are both UV and IR divergent. The IR diver-

gences cancel in other contributions to S_H (as we will provide ample demonstration), but the UV divergences must be removed through renormalization. As a consequence, in pure dimensional regularization, S_H -matrix elements are not guaranteed to be independent of renormalization scheme. Indeed, they are generally complex and will depend on the scale μ at which renormalization is performed. The S_H -matrix is *not* scale independent: $\frac{d}{d\mu}S_H \neq 0$, in contrast to S which does satisfy the Callan-Symanzik equation $\frac{d}{d\mu}S = 0$. This is unsatisfying, but not unsettling, as S_H elements are not themselves observable. (To be fair, if S -matrix elements are IR divergent, it is not clear what it means to say they are scale-independent). In any case, one should think of $S_H(\mu)$ like one thinks about the strong coupling constant $\alpha_s(\mu)$ in $\overline{\text{MS}}$. While $\alpha_s(\mu)$ is not observable, it is still an extraordinarily useful concept. The running coupling indeed encodes qualitatively and quantitatively a lot of important physics, such as unification and confinement. As with $\alpha_s(\mu)$, when $S_H(\mu)$ is used to compute an observable, the scale dependence will cancel. We demonstrate that in general in Section 4.2.2, and provide an explicit example in Section 4.5.

4.2.1 S_H and dressed states

The usual way of calculating S -matrix elements in perturbation theory is to work in the interaction picture, where one expands the interactions in terms of freely evolving fields. The propagators for free fields have a relatively simple form, and S -matrix elements then become integrals over these propagators. One might try to work out Feynman rules for S_H analogously, in an asymptotic interaction picture. Then propagators would correspond to non-perturbative Green's functions for the soft and collinear fields in $\mathcal{L}_{\text{SCET}}$, including all of

their interactions. Unfortunately, finding a closed-form expression for these propagators is not possible. In any case, it is not necessary, since if we want to work perturbatively in the coupling constants, we must do so consistently in both H and H_{as} .

To proceed, we note that the hard S-matrix can be written suggestively as

$$S_H = \Omega_+^{H\dagger} \Omega_-^H = \Omega_+^{\text{as}} \Omega_+^\dagger \Omega_- \Omega_-^{\text{as}\dagger} = \Omega_+^{\text{as}} S \Omega_-^{\text{as}\dagger}, \quad (4.2.6)$$

where

$$\Omega_\pm^{\text{as}} = \lim_{t \rightarrow \pm\infty} e^{iH_{\text{as}}t} e^{-iH_0t} \quad (4.2.7)$$

are asymptotic Møller operators and $\Omega_\pm = \lim_{t \rightarrow \pm\infty} e^{iHt} e^{-iH_0t}$ are the usual Møller operators. Inserting complete sets of states lets us write hard S -matrix elements between a Heisenberg picture out-state $|\psi_{\text{out}}\rangle$ and a Heisenberg picture in-state $|\psi_{\text{in}}\rangle$ as

$$\langle \psi_{\text{out}} | S_H | \psi_{\text{in}} \rangle = \int d\Pi_{\psi'_{\text{out}}} \int d\Pi_{\psi'_{\text{in}}} \langle \psi_{\text{out}} | \Omega_+^{\text{as}} | \psi'_{\text{out}} \rangle \langle \psi'_{\text{out}} | S | \psi'_{\text{in}} \rangle \langle \psi'_{\text{in}} | \Omega_-^{\text{as}\dagger} | \psi_{\text{in}} \rangle. \quad (4.2.8)$$

Here the integral is over complete sets of Fock-space states $|\psi'_{\text{in}}\rangle$ and $|\psi'_{\text{out}}\rangle$. The hard scattering matrix elements are written as a product of three terms. The middle term is the traditional S -matrix and the outer terms correspond to evolution with the asymptotic Møller operators. The Feynman rules for these contributions closely resemble those of time-ordered perturbation theory and are derived in Section 4.3.1 below.

Another interpretation of the hard matrix elements can be obtained by defining dressed

states as

$$\begin{aligned} |\psi_{\text{in}}^d\rangle &\equiv \Omega_-^{\text{as}\dagger} |\psi_{\text{in}}\rangle, \\ |\psi_{\text{out}}^d\rangle &\equiv \Omega_+^{\text{as}\dagger} |\psi_{\text{out}}\rangle. \end{aligned} \tag{4.2.9}$$

Then

$$\langle \psi_{\text{out}} | S_H | \psi_{\text{in}} \rangle = \langle \psi_{\text{out}}^d | S | \psi_{\text{in}}^d \rangle, \tag{4.2.10}$$

i.e. the matrix elements of the hard S -matrix are equivalent to matrix elements of the traditional S -matrix between dressed states. This connection was made in the context of QED in [57]. The role of the asymptotic evolution can then be viewed as transforming the in-state defined at $t = 0$ into a dressed state at $t = -\infty$ that scatters in the traditional way (with S). The role of dressed states is illustrated in Figure 4.1.

The dressed states $|\psi_{\text{in}}^d\rangle$ and $|\psi_{\text{out}}^d\rangle$ are not normalizable elements of the Fock space that $|\psi_{\text{in}}\rangle$ and $|\psi_{\text{out}}\rangle$ live in. Indeed, if we expand them perturbatively their coefficients in the Fock space basis contain infrared divergent integrals. For example, starting with an $|e^+e^-\rangle$ state

$$|\psi_{\text{in}}\rangle = |\bar{v}_s(p_1)u_{s'}(p_2)\rangle = \sqrt{2\omega_{p_1}}b_{p_1}^{s\dagger}\sqrt{2\omega_{p_2}}a_{p_2}^{s'\dagger}|0\rangle \tag{4.2.11}$$

in QED, the asymptotic Møller operator can add or remove soft photons with each factor of the coupling e . Up to order $\mathcal{O}(e^2)$ the dressed state will be a superposition of the leading

order $|e^+e^- \rangle$ state, $|e^+e^-\gamma \rangle$ states and $|e^+e^-\gamma\gamma \rangle$ Fock states. Explicitly,

$$\begin{aligned}
|\psi_{\text{in}}^d \rangle &= |\bar{v}(p_1)u(p_2) \rangle \\
&- e \int \frac{d^3k}{(2\pi)^3} \frac{1}{2\omega_k} \left[\frac{p_1 \cdot \epsilon}{p_1 \cdot k} \left| \bar{v}(p_1 - k)u(p_2)\epsilon(k) \right\rangle - \frac{p_2 \cdot \epsilon}{p_2 \cdot k} \left| \bar{v}(p_1)u(p_2 - k)\epsilon(k) \right\rangle \right] \\
&+ \frac{e^2}{2} \int \frac{d^3k_1}{(2\pi)^3} \frac{1}{2\omega_{k_1}} \int \frac{d^3k_2}{(2\pi)^3} \frac{1}{2\omega_{k_2}} \\
&\quad \times \left[\frac{p_1 \cdot \epsilon_1}{p_1 \cdot k_1} \frac{p_1 \cdot \epsilon_2}{p_1 \cdot k_2} \left| \bar{v}(p_1 - k_1 - k_2)u(p_2)\epsilon_1(k_1)\epsilon_2(k_2) \right\rangle \right. \\
&\quad + \frac{p_2 \cdot \epsilon_1}{p_2 \cdot k_1} \frac{p_2 \cdot \epsilon_2}{p_2 \cdot k_2} \left| \bar{v}(p_1)u(p_2 - k_1 - k_2)\epsilon_1(k_1)\epsilon_2(k_2) \right\rangle \\
&\quad - \frac{p_1 \cdot \epsilon_1}{p_1 \cdot k_1} \frac{p_2 \cdot \epsilon_2}{p_2 \cdot k_2} \left| \bar{v}(p_1 - k_1)u(p_2 - k_2)\epsilon_1(k_1)\epsilon_2(k_2) \right\rangle \\
&\quad \left. - \frac{p_1 \cdot \epsilon_2}{p_1 \cdot k_2} \frac{p_2 \cdot \epsilon_1}{p_2 \cdot k_1} \left| \bar{v}(p_1 - k_2)u(p_2 - k_1)\epsilon_1(k_1)\epsilon_2(k_2) \right\rangle \right] \\
&- e^2 \int \frac{d^3k}{(2\pi)^3} \frac{1}{2\omega_k} \frac{p_1 \cdot p_2}{p_1 \cdot k p_2 \cdot k} \left| \bar{v}(p_1 - k)u(p_2 + k) \right\rangle + \dots
\end{aligned} \tag{4.2.12}$$

Let us make a few observations about these dressed states. First, note that the Fock states being added have different 3-momenta. When k has exactly zero momentum (the case almost exclusively considered in the literature), momentum is conserved. But if one really wants to take these dressed states seriously, k must be allowed to have finite energy too, and then $|\psi_{\text{in}}^d \rangle$ is not a momentum eigenstate.

Second, the coefficient at order e^2 is a UV and IR divergent integral. The IR divergence is expected; it is exactly the IR divergence that cancels the IR divergence in elements of S to make elements of S_H IR finite. Nevertheless, it makes $|\psi_{\text{in}}^d \rangle$ hard to deal with as a state. The divergence requires an excursion from the traditional Fock space to a von Neumann space [9, 73, 106, 107]. The UV divergence is due to the fact a soft momentum is not sensi-

tive to any hard scale in the problem, so there is no natural cutoff on the k integrals. One could, of course, put in explicit hard cutoffs on the soft momenta, however, it is easier to simply renormalize the UV divergence by rescaling $|\psi_{\text{in}}^d\rangle$.

Third, it is not each separate electron that is being dressed. Rather it is the combination. Indeed, the IR divergence in the example above comes from loops connecting the two electrons. These loops are critical to cancelling the IR divergences in S_H . In Chung's original formulation (cf. Eq. (4.1.3)), a picture can be sketched for a coherent state as an electron moving with a cloud of photons around it. But this picture is too naive: the cloud depends on all the charged particles. This is even clearer in QCD, where the soft factors come with non-Abelian color matrices so one cannot rely on the crutch of Abelian exponentiation to move the dressing factors from state to state at will. A discussion of additional complications in QCD and the failure of Bloch-Nordsieck mechanism, can be found in [108].

In conclusion, although the dressed state picture fits in naturally with the construction of S_H we have presented, we doubt that thinking of the dressed states as physical states will ultimately be profitable.

We emphasize that for the purpose of having finite matrix elements, neither the in- and out-states $|\psi_{\text{in}}\rangle$ and $|\psi_{\text{out}}\rangle$, nor the dressed states $|\psi_{\text{in}}^d\rangle$ and $|\psi_{\text{out}}^d\rangle$, need to be eigenstates of the asymptotic Hamiltonian. In the examples to follow we will take $|\psi_{\text{in}}\rangle$ and $|\psi_{\text{out}}\rangle$ to be eigenstates of the free momentum operator P_0^μ with a finite number of particles, but in principle they can be taken to be any sensible linear combination of states in the relevant Hilbert space, i.e. with finite coefficients, in contrast to the usual coherent states which are an infinite linear superposition of Fock state elements. The S_H -matrix elements between

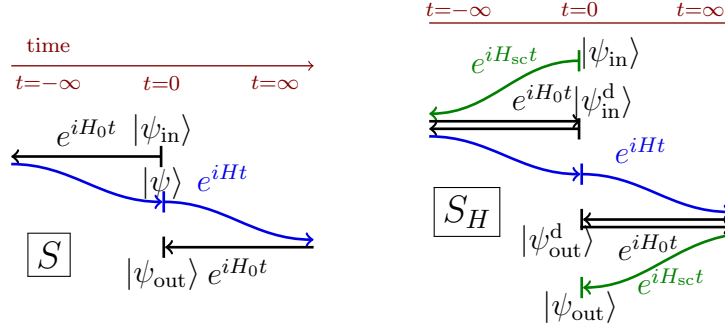


Figure 4.1: (Left) The traditional S -matrix is computed from Fock states evolved using H_0 and H . (Right) The hard S -matrix is computed either using Fock states evolved with H_{sc} and H or using dressed states evolved with H_0 and H .

any such states are always finite.

4.2.2 Computing observables using S_H

To compute an observable using S_H , one must specify what is to be included in the measurement and what is not. As a concrete example, consider computing the inclusive decay rate of the Z boson in perturbation theory. Since the Z does not couple to massless gauge bosons, it has no interactions in H_{as} and therefore $\Omega_{\pm}^{as}|Z\rangle = |Z\rangle$. The rate is then (up to kinematic factors)

$$\Gamma_Z \propto \sum_{X \neq Z} |\langle X|S_H|Z\rangle|^2 = \sum_{X \neq Z} \langle Z|\Omega_-^{as} S^\dagger \Omega_+^{as\dagger}|X\rangle \langle X|\Omega_+^{as} S \Omega_-^{as\dagger}|Z\rangle. \quad (4.2.13)$$

The sum is over all states in the theory except the Z itself, since $Z \rightarrow Z$ does not contribute to the rate and includes an implicit integral over the phase space for $|X\rangle$. Now we

write $\sum_{X \neq Z} |X\rangle\langle X| = \mathbb{1} - |Z\rangle\langle Z|$ to get

$$\Gamma_Z \propto \langle Z|Z\rangle - \langle Z|S^\dagger \Omega_+^{\text{as}\dagger}|Z\rangle\langle Z|\Omega_+^{\text{as}}S|Z\rangle = \sum_{X \neq Z} |\langle X|S|Z\rangle|^2, \quad (4.2.14)$$

where $\Omega_+^{\text{as}\dagger}|Z\rangle = |Z\rangle$ was used in the last step. So the sum over final states gives the same decay rate using S_H as it would using S . The key here was that there are no asymptotic interactions for Z . If there were, then the derivation would not hold. But in that case, the $Z \rightarrow Z$ forward scattering amplitude would be infrared divergent using S so it is not clear what physical result we should expect.

Suppose we wanted to compute something less inclusive than the total decay rate. The observable has to be infrared safe. For example, we could consider a 2-jet rate in $e^+e^- \rightarrow$ hadrons. Such a rate depends on the jet definition, which depends on exactly how the soft and collinear momenta are handled. In other words, it depends not only on the hard process, which is roughly speaking the jet-production amplitude, but also on the evolution of the jets after the hard scattering occurs. For this evolution, we need to include the dynamics induced by $e^{-iH_{\text{as}}t_+} \equiv \lim_{t \rightarrow \infty} e^{-iH_{\text{as}}t}$, as the state evolves from $t = 0$ to $t = \infty$ after the hard scattering. That is, we should define our exclusive cross section as

$$\sigma_{2\text{-jet}} = \sum_X \sum_Y |\langle X|e^{-iH_{\text{as}}t_+}|Y\rangle\langle Y|S_H|Z\rangle|^2 \delta \left[N_{\text{jets}}(X) - 2 \right]. \quad (4.2.15)$$

Here $N_{\text{jets}}(X)$ is the measurement function which takes as input the momenta of the particles in the final state X and returns the number of jets according to some jet definition. The factor $\langle Y|S_H|Z\rangle$ gives the amplitude to produce the jets and $\langle X|e^{-iH_{\text{as}}t_+}|Y\rangle$ gives

the amplitude for those jets to evolve into a state with the particles in $|X\rangle$ at asymptotic times. The sum over Y can be as restrictive as desired. For example, if Y is taken to be only $|q\bar{q}\rangle$ quark-antiquark states, the distribution will be valid to leading power. To get the jet mass distribution exactly right, including subleading power effects, one should extend the sum from over $|\bar{q}q\rangle$ states to anything that could possibly evolve into a state X with $N_{\text{jets}}(X) = 2$. For example, $|\bar{q}qg\rangle$ should be included. If all states are allowed then one can replace $\sum_Y |Y\rangle\langle Y|$ with $\mathbb{1}$. In that case, the rate reduces to

$$\sigma_{2\text{-jet}} = \sum_X |\langle X | e^{iH_0 t_+} S | Z \rangle|^2 \delta [N_{\text{jets}}(X) - 2]. \quad (4.2.16)$$

The $e^{iH_0 t_+}$ factor generates a phase $e^{iE_X t}$ which is constant for all X by energy conservation and therefore drops out of the absolute value. Thereby the exclusive cross section reduces to the same thing one would compute using S (in agreement with a century of theory/experiment comparisons). A cartoon of the reduction of the cross section to the one computed with S for this process is shown in Fig. 4.2.

Just because one *can* reduce cross section calculations using S_H to those using S , does not mean one should. Additional physical insight is gained by maintaining the separation into a calculation of S_H first and then of the evolution using $e^{-iH_{\text{as}} t_+}$ or equivalently Ω_{as}^+ . In particular, since H_{as} is independent of the hard scattering, the separation leads to the physical picture of a short-distance amplitude for jet production followed by an evolution from short-to-long distances where the jets are resolved into their constituents. For example, in the computation of thrust in e^+e^- events, when the events comprise pencil-like jets,

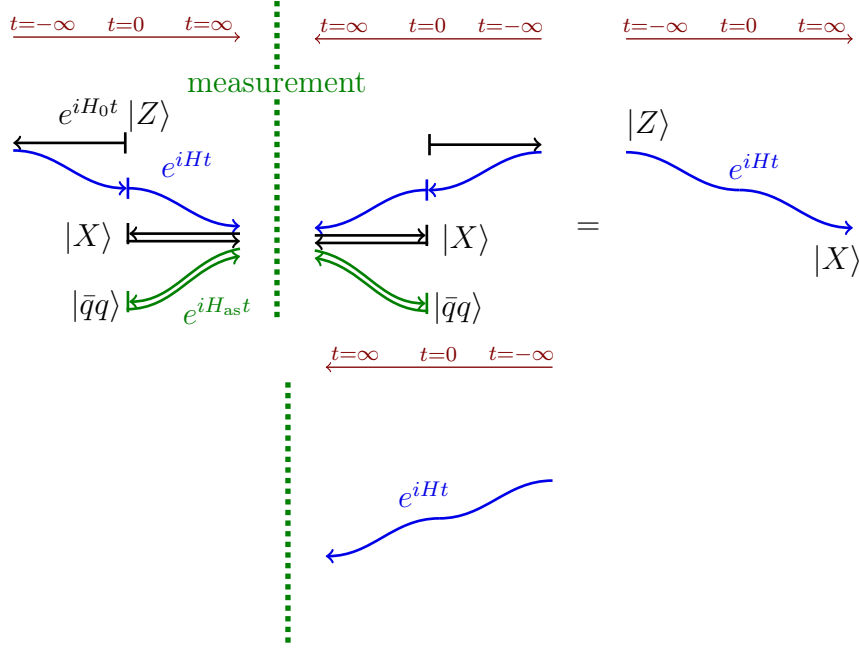


Figure 4.2: An observable is computed by integrating the square of an amplitude against a measurement function, inserted at $t = \infty$. In computing an exclusive observable sensitive to the asymptotic dynamics, one must evolve the dressed states to $+\infty$ using the asymptotic Hamiltonian. The example $Z \rightarrow \text{jets}$ is illustrated on the left. The result is equivalent to evolving the initial state $|Z\rangle$ at $t = -\infty$ with the full Hamiltonian to the set of states $|X\rangle$ on which the measurement is performed at $t = +\infty$ (right).

the structure of the distribution is almost completely determined by the asymptotic evolution alone. This example, and the utility of the separation will be discussed more in Section 4.5.

The above discussion of observables also helps clarify how one should think of assigning hard or soft/collinear labels to the particles in the states. Consider, for example, the process $Z \rightarrow \bar{q}qg$. In what circumstances should one consider the gluon momentum to be collinear to the quark or antiquark momenta, or soft?

On the one hand, if one declares the gluon momentum to be soft or collinear, then there are necessarily interactions in H_{as} that can produce the gluon through a real emission. Due

to factorization, the amplitude for this emission from H_{as} will approach that from H , but with an opposite sign. So the two will cancel in the exact soft/collinear limits. In other words, if the gluon momentum is soft/collinear, then the hard matrix element $\langle \bar{q}qg|S_H|Z \rangle$ will vanish in soft/collinear limits. In this case, there is also a contribution to a $\bar{q}qg$ final state from the hard $\bar{q}q$ production $\langle \bar{q}q|S_H|Z \rangle$ and then an emission of g through the asymptotic interactions. This additional contribution is not power suppressed and adds to the $\langle \bar{q}qg|S_H|Z \rangle$ amplitude to produce the full distribution, in agreement with $\langle \bar{q}qg|S|Z \rangle$. Such a deconstruction corresponds to the picture of matching onto a 2-jet operator $C_2\mathcal{O}_2$ and then matching on to a 3-jet operator $C_3\mathcal{O}_3$ in SCET. [125, 126]. In such matching, the Wilson coefficient C_3 vanishes in soft and collinear limits.

On the other hand, it does not really make sense to compute $\langle \bar{q}qg|S_H|Z \rangle$ when the gluon is soft or collinear. The hard S -matrix is meant to give amplitudes for production of hard particles. The evolution of those hard particles into jets with soft/collinear substructure is subsequently determined by H_{as} . Thus, a more sensible convention is to consider only matrix elements $\langle \bar{q}qg|S_H|Z \rangle$ when all 3 final state particles are considered hard. In this case, these particles have no interactions with each other in H_{as} and there are no contributions to $\langle \bar{q}qg|S_H|Z \rangle$ that have real emissions from the asymptotic region. Thus, all the contributions to S_H involving the asymptotic region are virtual (and give scaleless integrals in pure dimensional regularization). In other words, if one is interested in 3-jet production, one should study $\langle \bar{q}qg|S_H|Z \rangle$ and if one is interested in 2-jet production, one should study $\langle \bar{q}q|S_H|Z \rangle$. Although the final predictions for IR-safe differential cross sections are independent of what convention we take for assigning labels to the final state particles

(and always agree with the result from S), the hard S -matrix should always be thought of as giving the amplitudes for producing hard particles. With this convention $\langle \bar{q}qg|S_H|Z\rangle$ no longer vanishes in soft or collinear limits. Instead in these limits, it factorizes into $\langle \bar{q}qg|e^{-iH_{\text{ast}+}}|\bar{q}q\rangle\langle \bar{q}q|S_H|Z\rangle$. Since the splitting amplitudes $\langle \bar{q}qg|e^{-iH_{\text{ast}+}}|\bar{q}q\rangle$ are universal [49, 50, 116], this restricts the possible form that $\langle \bar{q}qg|S_H|Z\rangle$ could have. Implications of these restrictions have been discussed extensively (see [127, 128]) and are one instance of the deep structure present in S_H -matrix elements.

In summary, one has two choices:

- Allow states in which S_H matrix elements are taken to have soft or collinear momenta. Observables computed this way will only be valid to leading power, but can be computed efficiently exploiting factorization.
- Insist that all states in which S_H matrix elements are taken have only hard momenta. Then all the contributions from the asymptotic regions are virtual, and scaleless in dimensional regularization. Observables agree *exactly* with their computation using S .

We emphasize that with either choice, S_H matrix elements are IR finite. The general observations in this section are backed up with explicit calculations in Section 4.5.2.

4.2.3 Soft Wilson lines

To connect our framework to previous work, we consider the QED case with massive electrons. In this case, there are only soft interactions in the asymptotic Hamiltonian. The interaction in the SCET Hamiltonian between soft photons and collinear fermions has the

form (see Eq. (4.2.4))

$$H_{\text{soft}}^{\text{int}}(t) = e \sum_n \int d^3x n \cdot A(x_-) \bar{\xi}_n(x) \frac{\not{n}}{2} \xi_n(x), \quad (4.2.17)$$

where n^μ is a lightlike 4-vector labeling the fermion, \bar{n}^μ is the direction backwards to n^μ , and $x_- = \bar{n} \cdot x$. For simplicity, we take $n^\mu = (1, 0, 0, 1)$ so $\bar{n}^\mu = (1, 0, 0, -1)$ and $x_- = t + z$. The dependence of the interaction only on x_- follows from the multipole expansion². The collinear fields have only half the degrees of freedom of fields in QED: they only describe electrons in this case, as pair-creation is power-suppressed. So we can write

$$\xi_n(x) = \int \frac{d^3p}{(2\pi)^3} \frac{1}{\sqrt{2\omega_p}} u(p) a_p e^{-ipx}, \quad \bar{\xi}_n(x) = \int \frac{d^3q}{(2\pi)^3} \frac{1}{\sqrt{2\omega_q}} \bar{u}(q) a_q^\dagger e^{iqx}. \quad (4.2.18)$$

The field expansion for the soft photon is as usual, but the phase is power expanded,

$$A_\mu(x_-) = \sum_{j=1}^2 \int \frac{d^3k}{(2\pi)^3} \frac{1}{\sqrt{2\omega_k}} \left[\epsilon_\mu^j(k) a_k^j e^{-i\frac{1}{2}k^+x_-} + \epsilon_\mu^{j*}(k) a_k^{j\dagger} e^{i\frac{1}{2}k^+x_-} \right]. \quad (4.2.19)$$

²A collinear momentum scales as $(p^-, p^+, p_\perp) \sim (\lambda^2, 1, \lambda)$ so x scales like $(x_-, x_+, x^\perp) \sim (1, \lambda^{-2}, \lambda^{-1})$. Then since a soft momentum scales homogeneously like $k \sim \lambda^2$, only the k^+x_- component is relevant at leading power. See [82] for more details.

Inserting these field expansions and integrating over d^3x gives

$$\begin{aligned}
H_{\text{soft}}^{\text{int}}(t) = e \sum_n \int \frac{d^3p}{(2\pi)^3 \sqrt{2\omega_p}} \frac{d^3q}{(2\pi)^3 \sqrt{2\omega_q}} \frac{d^3k}{(2\pi)^3 \sqrt{2\omega_k}} (2\pi)^3 \delta^2(\vec{p}_\perp - \vec{q}_\perp) \bar{u}(q) \frac{\not{q}}{2} u(p) a_q^\dagger a_p \\
\times \sum_{j=1}^2 \left[n \cdot \epsilon^j(k) a_k^j \delta\left(q^z - p^z - \frac{1}{2}k^+\right) e^{i(\omega_q - \omega_p - \frac{1}{2}k^+)t} \right. \\
\left. + n \cdot \epsilon^{j*}(k) a_k^{j\dagger} \delta\left(q^z - p^z + \frac{1}{2}k^+\right) e^{i(\omega_q - \omega_p + \frac{1}{2}k^+)t} \right]. \quad (4.2.20)
\end{aligned}$$

Since $k^+ \ll p^z$ after doing the q integral, we can replace $a_q^\dagger \cong a_p^\dagger$ at leading power and write

$$\frac{1}{\sqrt{2\omega_p}} \frac{1}{\sqrt{2\omega_q}} \bar{u}(q) \frac{\not{q}}{2} u(p) \cong \frac{1}{2\omega_p} p \cdot \bar{n} \cong 1. \quad (4.2.21)$$

Power expanding the energy ω_q gives

$$\omega_q = \sqrt{\vec{p}_\perp^2 + \left(p^z \pm \frac{1}{2}k^+\right)^2} \cong \omega_p \pm \frac{p^z}{2\omega_p} k^+ \cong \omega_p \mp \frac{1}{2}k^+, \quad (4.2.22)$$

and hence the argument of the exponential becomes $i(\omega_q - \omega_p \mp \frac{1}{2}k^+)t \cong \mp i k^+ t$. So we get

$$H_{\text{soft}}^{\text{int}}(t) = e \sum_n A_\mu(t n^\mu) \int \frac{d^3p}{(2\pi)^3} a_p^\dagger a_p. \quad (4.2.23)$$

Then we find that the asymptotic Møller operator acting on a single electron state gives

$$\Omega_+^{\text{soft}}|p\rangle = T \left\{ \exp \left[-i \int_0^\infty dt H_{\text{soft}}^{\text{int}}(t) \right] \right\} |p\rangle = P \left\{ \exp \left[-ie \int_0^\infty ds n \cdot A(sn^\mu) \right] \right\} |p\rangle, \quad (4.2.24)$$

with P a path-ordered product. The path ordering is actually superfluous in QED, but is important in the non-Abelian case. The soft Wilson line in QED is defined as

$$Y_n^\dagger = \exp \left[-ie \int_0^\infty ds \, n \cdot A(sn^\mu) e^{-\varepsilon s} \right], \quad (4.2.25)$$

where the factor $e^{-\varepsilon s}$ ensures convergence near $s = \infty$. Then, the action of the asymptotic soft Møller operator is the same as that of a product of soft Wilson lines:

$$\Omega_+^{\text{soft}} |p_1 \cdots p_j\rangle = T \left\{ Y_{n_1}^\dagger \cdots Y_{n_j}^\dagger \right\} |p_1 \cdots p_j\rangle \quad (4.2.26)$$

For antiparticles, one would have Y_n factors instead, and for incoming particles, one would have factors of $\overline{Y_n}$, defined as Y_n^\dagger but with an integral from $-\infty$ to 0 [49].

We can combine the time-ordered product of exponential into a single exponential using the Magnus expansion [129],

$$\begin{aligned} T \left\{ \exp \left[\int_0^\infty dt \, \mathcal{O}(t) \right] \right\} &= \exp \left\{ \int_0^\infty dt \, \mathcal{O}(t) + \frac{1}{2} \int_0^\infty dt \int_t^\infty ds \, [\mathcal{O}(s), \mathcal{O}(t)] \right. \\ &\quad \left. \times \frac{1}{6} \int_0^\infty dt \int_t^\infty ds \int_s^\infty du \, \left([\mathcal{O}(u), [\mathcal{O}(s), \mathcal{O}(t)]] + [\mathcal{O}(t), [\mathcal{O}(s), \mathcal{O}(u)]] \right) + \cdots \right\}, \end{aligned} \quad (4.2.27)$$

where the higher order terms are sums of nested commutators. The commutators of two fields in Feynman gauge can be computed directly from the field expansions in Eq. (4.2.19),

$$\left[n_1 \cdot A(sn_1^\mu), n_2 \cdot A(tn_2^\mu) \right] = - \int \frac{d^3 k}{(2\pi)^3} \frac{n_1 \cdot n_2}{2\omega_k} \left[e^{-i(sn_1^\mu - tn_2^\mu)k_\mu} - e^{i(sn_1^\mu - tn_2^\mu)k_\mu} \right]. \quad (4.2.28)$$

Since the commutator in Eq. (4.2.28) is a c -number, additional commutators vanish. This is the essence of Abelian exponentiation. Then, we can combine all the time-ordered exponentials into a single exponential:

$$T \left\{ Y_{n_1}^\dagger \cdots Y_{n_j}^\dagger \right\} = \exp \left[-ie \sum_j \int_0^\infty ds n_j \cdot A(sn_j^\mu) \right] \exp \left[i \sum_{ij} \Phi_{ij} \right] \quad (4.2.29)$$

where

$$i\Phi_{ij} \equiv -e^2 \frac{1}{2} \int_0^\infty dt \int_t^\infty ds \left[n_i \cdot A(sn_i^\mu), n_j \cdot A(tn_j^\mu) \right] e^{-\varepsilon(s+t)}. \quad (4.2.30)$$

When acting on states with electrons, this combination is exactly of the form $e^R e^{i\Phi}$ that Faddeev and Kulish write (see Eq. (4.1.6)), with R the expression in Eq. (4.1.8).

The electron-number operator $\rho(\vec{p})$ from Eq. (4.1.9) is of the same origin as the $a_p^\dagger a_p$ in Eq. (4.2.23).

Consider the case of an outgoing electron and positron in QED, where we want to simplify the time-ordered product of two Wilson lines $T\{Y_{n_1}^\dagger Y_{n_2}\}$. Then

$$\mathcal{O}(t) = -ie [n_1 \cdot A(tn_1^\mu) - n_2 \cdot A(tn_2^\mu)]. \quad (4.2.31)$$

To see the connection to the Coulomb phase, let us do the integrations over s and t in Eq. (4.2.30) using Eq. (4.2.28),

$$i\Phi_{ij} = ie^2 \int \frac{d^3k}{(2\pi)^3} \frac{1}{2\omega_k} \text{Im} \frac{n_1 \cdot n_2}{(n_1 \cdot k - i\varepsilon)((n_1 - n_2) \cdot k - 2i\varepsilon)}. \quad (4.2.32)$$

Taking $n_1 = (1, 0, 0, 1)$ and $n_2 = (1, 0, 0, -1)$ we can simplify this to

$$\Phi = e^2 \int \frac{d^3k}{(2\pi)^3} \frac{1}{2\omega_k} \text{Im} \frac{2}{(\omega_k - k_z - i\varepsilon)(-2k_z - 2i\varepsilon)} = -\frac{e^2}{16\pi^2} \int \frac{d^2k_\perp}{k_\perp^2}. \quad (4.2.33)$$

This is the usual divergent integral appearing in the Coulomb phase (cf. Eq. (4.5.9)). When one of the electrons is incoming, the $\int_0^\infty ds$ gets replaced with $\int_{-\infty}^0 ds$ in Eq. (4.2.30) and we get

$$\Phi = -e^2 \int \frac{d^3k}{(2\pi)^3} \frac{1}{2\omega_k} \text{Im} \frac{2}{(\omega_k + k_z + i\varepsilon)(\omega_k - k_z - i\varepsilon)} = 0, \quad (4.2.34)$$

which is consistent with the Coulomb phase vanishing for timelike kinematics.

In this way, we have shown that our framework agrees with previous work in the case of QED, where there are soft but not collinear singularities and the gauge boson is Abelian. Note that both the Coulomb phase and the real part of the exponent emerge from the single soft-collinear interaction in H_{as} .

In the non-Abelian case, one cannot combine the path-ordered exponentials into the exponential of a single closed-form expression as in Eq. (4.2.29): the gauge generators do not commute. There is an analog of Abelian exponentiation, called non-Abelian exponentiation [112, 113, 123] but one must include higher order commutators, and no closed form expression is known. Thus, a Faddeev-Kulish type formulation of the dressed states is impossible for QCD. The Wilson-line description of the soft interactions is still valid, however, and the soft interactions in QCD still factorize off of the scattering operator into soft Wilson lines.

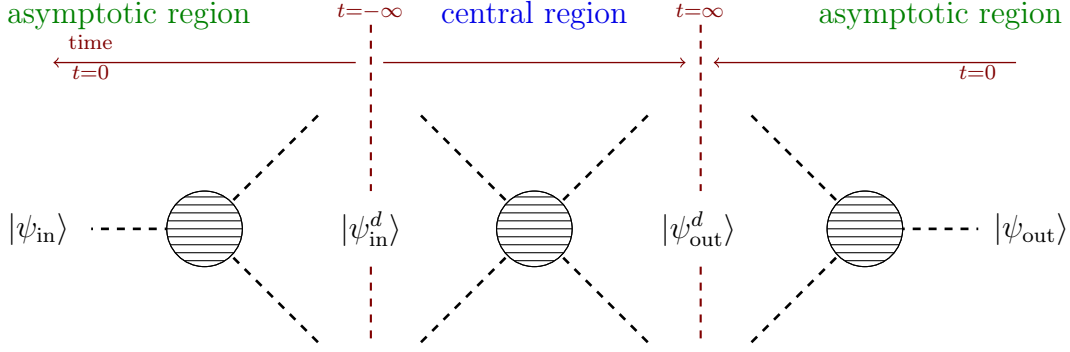


Figure 4.3: In order to facilitate calculations in perturbation theory, we divide the matrix elements of S_H into three parts. In the two outer parts, the asymptotic evolution Møller operators Ω_{\pm}^{as} work to dress the in- and out-states. The middle part corresponds to a calculation of traditional S -matrix elements between dressed states.

4.3 Computing the hard S -matrix

In this section, we show how to compute S_H -matrix elements perturbatively. We will use the formula in Eq. (4.2.8):

$$\langle \psi_{\text{out}} | S_H | \psi_{\text{in}} \rangle = \int d\Pi_{\psi'_{\text{out}}} \int d\Pi_{\psi'_{\text{in}}} \underbrace{\langle \psi_{\text{out}} | \Omega_+^{\text{as}} | \psi'_{\text{out}} \rangle}_{\text{asymptotic region}} \underbrace{\langle \psi'_{\text{out}} | S | \psi'_{\text{in}} \rangle}_{\text{central region}} \underbrace{\langle \psi'_{\text{in}} | \Omega_-^{\text{as}\dagger} | \psi_{\text{in}} \rangle}_{\text{asymptotic region}}. \quad (4.3.1)$$

We call the two outer matrix elements the *asymptotic region* and the part involving $\langle \psi'_{\text{out}} | S | \psi'_{\text{in}} \rangle$ the *central region*. The asymptotic regions go from $0 > t > -\infty$ and $\infty > t > 0$, both backward in time. The central region calculation is just that of an ordinary S -matrix. A cartoon of the division is shown in Fig 4.3. In this section we establish the Feynman rules for the asymptotic regions, which are similar to those in old-fashioned, time-ordered perturbation theory with a few changes. We also give an example calculation in ϕ^3 theory that clarifies some of the subtleties. Calculations for physical process in QED, QCD and $\mathcal{N} = 4$ SYM theories are given in subsequent sections.

4.3.1 Asymptotic region Feynman rules

We have reduced the problem of computing matrix elements of S_H to calculating matrix elements of S and matrix elements of the form

$$\langle \psi_{\text{out}} | \Omega_+^{\text{as}} | \psi'_{\text{out}} \rangle \quad \text{and} \quad \langle \psi'_{\text{in}} | \Omega_-^{\text{as}\dagger} | \psi_{\text{in}} \rangle \quad (4.3.2)$$

in perturbation theory. To evaluate these matrix elements, we separate the asymptotic Hamiltonian into a free part and an interaction part:

$$H_{\text{as}} = H_0 + V_{\text{as}} \quad (4.3.3)$$

Defining the operator $U_+^{\text{as}}(t)$ by the equation $\Omega_+^{\text{as}} = \lim_{t \rightarrow \infty} U_+^{\text{as}}(t)$, it satisfies the differential equation

$$\begin{aligned} -i\partial_t U_+^{\text{as}}(t) &= U_+^{\text{as}}(t) V_{\text{as}}^I(t), \\ U_+^{\text{as}}(0) &= 1, \end{aligned} \quad (4.3.4)$$

where the superscript I indicates that V_{as}^I is the interaction picture potential, i.e. the asymptotic potential $V_{\text{as}}[\phi_0] = -\int d^3x \mathcal{L}_{\text{as}}[\phi_0]$ expressed in terms of freely-evolving interaction picture fields ϕ_0 , and where \mathcal{L}_{as} is the Lagrangian density corresponding to the

asymptotic interactions. This differential equation has the solution

$$\begin{aligned}
U_+^{\text{as}}(t) &= 1 + i \int_0^t dt' V_{\text{as}}^I(t') + i^2 \int_0^t dt' \int_0^{t'} dt'' V_{\text{as}}^I(t'') V_{\text{as}}^I(t') + \dots \\
&= \overline{T} \left\{ \exp \left[i \int_0^t dt' \int d^3 \vec{x} V_{\text{as}}^I(t') \right] \right\} ,
\end{aligned} \tag{4.3.5}$$

where \overline{T} denotes an anti time-ordered product.

To see how to evaluate matrix elements of this operator, consider the following diagram in scalar ϕ^3 theory:

$$S_A^{+(2)} = \tag{4.3.6}$$

The free fields are given by

$$\phi_0(x) = \int \frac{d^3 p}{(2\pi)^3} \frac{1}{\sqrt{2\omega_p}} \left(a_p e^{-ipx} + a_p^\dagger e^{ipx} \right) . \tag{4.3.7}$$

One-particle states in the free theory are

$$|p\rangle = \sqrt{2\omega_p} a_p^\dagger |0\rangle . \tag{4.3.8}$$

Up to renormalization, which will be discussed later, the external states are as usual taken

to be free creation operators acting on the free vacuum. We therefore aim to calculate

$$S^+ = \langle p_1 p_2 | \Omega_+^{\text{as}}(t) | p'_1 p'_2 \rangle = \langle p_1 p_2 | \bar{T} \left\{ \exp \left[-i \int_0^\infty dt' \int d^3 \vec{x} \mathcal{L}_{\text{as}}[\phi_0] \right] \right\} | p'_1 p'_2 \rangle . \quad (4.3.9)$$

The second order term in g is

$$S^{+(2)} = \langle 0 | \sqrt{2\omega_{p_1}} a_{p_1} \sqrt{2\omega_{p_2}} a_{p_2} \int_0^\infty dt_x \int_{t_x}^\infty dt_y \int d^3 \vec{x} \int d^3 \vec{y} \\ \times \frac{-ig}{3!} \phi_0^3(x) \frac{-ig}{3!} \phi_0^3(y) \sqrt{2\omega_{p'_1}} a_{p'_1}^\dagger \sqrt{2\omega_{p'_2}} a_{p'_2}^\dagger | 0 \rangle . \quad (4.3.10)$$

Inserting Eq. (4.3.7) and commuting creation and annihilation operators, gives the following expression corresponding to the diagram above:

$$S_A^{+(2)} = (-ig)^2 \langle 0 | \sqrt{2\omega_{p_1}} a_{p_1} \sqrt{2\omega_{p_2}} a_{p_2} \int_0^\infty dt_x \int_{t_x}^\infty dt_y \int d^3 \vec{x} \int d^3 \vec{y} \\ \times \int \frac{d^3 q_1}{(2\pi)^3 \sqrt{2\omega_{q_1}}} a_{q_1}^\dagger e^{iq_1 x} \int \frac{d^3 q_2}{(2\pi)^3 \sqrt{2\omega_{q_2}}} a_{q_2}^\dagger e^{iq_2 y} \int \frac{d^3 k}{(2\pi)^3 \sqrt{2\omega_k}} a_{k'} e^{-ik' x} \\ \times \int \frac{d^3 k'}{(2\pi)^3 \sqrt{2\omega_{k'}}} a_k^\dagger e^{ik y} \int \frac{d^3 q'_2}{(2\pi)^3 \sqrt{2\omega_{q'_2}}} a_{q'_2} e^{-iq'_2 y} \int \frac{d^3 q'_1}{(2\pi)^3 \sqrt{2\omega_{q'_1}}} a_{q'_1} e^{-iq'_1 x} \\ \times \sqrt{2\omega_{p'_1}} a_{p'_1}^\dagger \sqrt{2\omega_{p'_2}} a_{p'_2}^\dagger | 0 \rangle \quad (4.3.11)$$

Integrating over \vec{x} and \vec{y} gives δ -function. Integrating over these δ -functions and the additional δ -functions coming from the creation and annihilation operators reduces the expres-

sion to

$$S_A^{+(2)} = (2\pi)^3 \delta^3(\vec{p}_1 + \vec{p}_2 - \vec{p}_1' - \vec{p}_2') (-ig)^2 \\ \times \frac{1}{2\omega_k} \int_0^\infty dt_x \int_{t_x}^\infty dt_y e^{i(\omega_1 - \omega_1' - \omega_k)t_x} e^{i(\omega_2 - \omega_2' + \omega_k)t_y}. \quad (4.3.12)$$

Finally, the integrals over t_x and t_y give

$$S_A^{+(2)} = (2\pi)^3 \delta^3(\vec{p}_1 + \vec{p}_2 - \vec{p}_1' - \vec{p}_2') (-ig)^2 \\ \times \frac{1}{2\omega_k} \frac{-i}{\omega_1' + \omega_2' - \omega_1 - \omega_2 - i\varepsilon} \frac{-i}{\omega_2' - \omega_2 - \omega_k - i\varepsilon}. \quad (4.3.13)$$

More generally, the Feynman rules for the asymptotic regions are the same as those in ordinary relativistic time-ordered perturbation theory (see [130] for example) with two differences: 1) Since the outermost integral goes from 0 to ∞ instead of $-\infty$ to ∞ , the overall energy-conserving δ -function; and $2\pi\delta(E_f - E_i)$ is replaced by a propagator $\frac{i}{E_f - E_i + i\varepsilon}$ 2) the evolution is backwards in time ($e^{iH_{\text{as}}t}$ instead of $e^{-iH_{\text{as}}t}$) so the whole amplitude is complex conjugated. This means $ig \rightarrow -ig$ and $\frac{i}{E + i\varepsilon} \rightarrow \frac{-i}{E - i\varepsilon}$.

For explicit computations and consistency checks, one has to be very careful about the $i\varepsilon$ prescription. It is important to keep in mind that the propagators $\frac{-i}{E - i\varepsilon}$ are distributions, only defined under integration. The $i\varepsilon$ comes from an integral representation of the θ func-

tion,

$$\begin{aligned} \int_0^\infty dt e^{-i\omega t} &= \int_{-\infty}^\infty dt \theta(t) e^{-i\omega t} \\ &= \int_{-\infty}^\infty dt \left[\int_{-\infty}^\infty \frac{dE}{2\pi} e^{iEt} \frac{-i}{E - i\varepsilon} \right] e^{-i\omega t} = \int_{-\infty}^\infty dE \delta(E - \omega) \frac{-i}{E - i\varepsilon} = \frac{-i}{\omega - i\varepsilon}, \end{aligned} \quad (4.3.14)$$

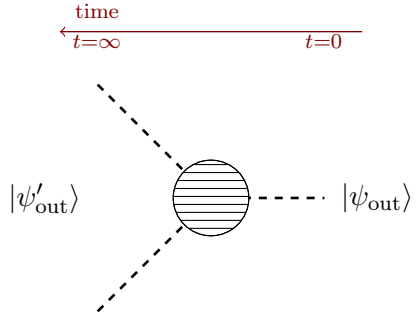
so it really should be associated with the shift $\omega \rightarrow \omega - i\varepsilon$ for any integral ending at $t = +\infty$ or $\omega \rightarrow \omega + i\varepsilon$ for any integral starting at $t = -\infty$. When we have a nested integral, like Eq. (4.3.12), we get

$$\int_0^\infty dt_2 \int_{t_1}^\infty dt_2 e^{i\omega_1 t_2} e^{i\omega_2 t_2} \rightarrow \int_0^\infty dt_2 \int_{t_1}^\infty dt_2 e^{i(\omega_1 - i\varepsilon)t_1} e^{i(\omega_2 - i\varepsilon)t_2} = \frac{-i}{\omega_2 - i\varepsilon} \frac{-i}{\omega_1 + \omega_2 - 2i\varepsilon}. \quad (4.3.15)$$

So each vertex gives another factor of ε . An example of the importance of careful treatment of these distributions is given in Section 4.3.2.

In summary, the Feynman rules for $\langle \psi_{\text{out}} | \Omega_+^{\text{as}} | \psi'_{\text{out}} \rangle$ are as follows

- Draw all relevant time-ordered diagrams between the state $|\psi_{\text{out}}\rangle$ at $t = 0$ on the right and $|\psi'_{\text{out}}\rangle$ at $t = \infty$ on the left:



$$(4.3.16)$$

- Assign momenta k_i^μ to each internal line, with $k_i^0 = \omega_k = \sqrt{m^2 + \vec{k}_i^2}$ the on-shell energy.
- Start at the far left of the diagram ($t = \infty$), and move a vertical cut rightwards in time until a vertex is crossed. After each vertex is crossed, include a factor of

$$\frac{-i}{(E'_{\text{out}} - ni\varepsilon) - E_{\text{cut}}} \quad (4.3.17)$$

where $E_{\text{cut}} = \sum \omega_{\text{cut}}$ is the total energy of the particles in the cut, $E'_{\text{out}} = \sum \omega'_{\text{out}}$ is the total energy of the particles in $|\psi'_{\text{out}}\rangle$, and n is the number of vertices that have already been crossed in the asymptotic region. Note that the $-i\varepsilon$ comes from a $+i\varepsilon$ from the $t = +\infty$ region, and is then complex conjugated.

- For each vertex, add a factor of $(2\pi)^3 \delta^3(\sum \vec{p}_i)$ to impose 3-momentum conservation and $-ig$ for the interaction (or whatever the interaction is, just as in regular Feynman rules, complex-conjugated).
- Integrate over $\prod_i \int \frac{d^3 k_i}{(2\pi)^3 2\omega_i}$ for the momentum of each internal line.

The Feynman rules for $\langle \psi'_{\text{in}} | \Omega_-^{\text{as}} | \psi_{\text{in}} \rangle$ are identical except that the diagrams go from $t =$

$-\infty$ on the right to $t = 0$ on the left

$$(4.3.18)$$

and the propagators are

$$\frac{-i}{(E'_{\text{in}} - in\varepsilon) - E_{\text{cut}}}, \quad (4.3.19)$$

where $E'_{\text{in}} = \sum \omega'_{\text{in}}$ is the total energy of the particles in $|\psi'_{\text{in}}\rangle$.

4.3.2 Cross check in ϕ^3 theory

To validate the Feynman rules, consider the case where $H_{\text{as}} = H$. In this case, the hard S -matrix is trivial $S_H = \mathbb{1}$. Perturbatively, this means that diagrams with all vertices in the central region should be exactly canceled by diagrams involving vertices in the asymptotic regions. Moreover, the cancellation should occur for each time-ordered diagram on its own.

We can check this cancellation in any theory and any diagram, so we take ϕ^3 theory with

Lagrangian $\mathcal{L} = -\frac{1}{2}\phi\Box\phi + \frac{g}{3!}\phi^3$ for simplicity and consider the diagram

$$\begin{array}{c}
\vec{k} \\
\text{---} t_1 \quad \text{---} t_2 \text{---} \\
\vec{p} \quad \quad \vec{p} \\
\text{---} \quad \quad \text{---} \\
\vec{p} - \vec{k}
\end{array} \quad (4.3.20)$$

We sum over diagrams with t_1 and t_2 going from 0 to $-\infty$ to ∞ and back to 0. Let us call the initial energy as $\omega_i = \omega_p$, the final energy $\omega_f = \omega_p$ and the energy of the intermediate state $\omega_c = \omega_{p-k} + \omega_k$.

The usual time-ordered perturbation theory loop (i.e. the contribution from S to S_H with all vertices in the central region) is

$$\begin{aligned} \mathcal{S}_1 = & \text{(diagram)} = \frac{(ig)^2}{2} \int \frac{d^3k}{(2\pi)^3} \frac{i}{4\omega_k \omega_{p-k}} \frac{1}{\omega_i - \omega_c + i\varepsilon} 2\pi \delta(\omega_i - \omega_f). \\ & (4.3.21) \end{aligned}$$

To see this cancel other diagrams, it is helpful to break this diagram down further, into the contribution into 3 regions: first, $-\infty < t_1 < t_2 < 0$ then $-\infty < t_1 < 0 < t_2 < \infty$ and finally $0 < t_1 < t_2 < \infty$:

$$\mathcal{S}_1 = \frac{(ig)^2}{2} \int \frac{d^3k}{(2\pi)^3 4\omega_k \omega_{p-k}} \left[\frac{i}{\omega_i - \omega_c + i\epsilon} \frac{i}{\omega_i - \omega_f + 2i\epsilon} + \frac{i}{\omega_i - \omega_c + i\epsilon} \frac{i}{\omega_f - \omega_c + i\epsilon} + \frac{i}{\omega_f - \omega_c + i\epsilon} \frac{i}{\omega_f - \omega_i + 2i\epsilon} \right] \quad (4.3.22)$$

In this decomposition, we have employed the careful treatment of the distributions dis-

cussed around Eq. (4.3.15).

Contributions from the loop in the asymptotic region are given by

$$\begin{aligned}
S_2 &= \text{(diagram)} \\
&= \frac{(-ig)^2}{2} \int \frac{d^3 k}{(2\pi)^3} \frac{-i}{4\omega_k \omega_{p-k}} \frac{-i}{\omega_i - \omega_c - i\varepsilon} \frac{-i}{\omega_i - \omega_f - 2i\varepsilon}, \\
\\
S_3 &= \text{(diagram)} \\
&= \frac{(-ig)^2}{2} \int \frac{d^3 k}{(2\pi)^3} \frac{-i}{4\omega_k \omega_{p-k}} \frac{-i}{\omega_f - \omega_c - i\varepsilon} \frac{-i}{\omega_f - \omega_i - 2i\varepsilon},
\end{aligned} \tag{4.3.23}$$

and contributions from the loop divided between the two asymptotic regions is

$$\begin{aligned} \mathcal{S}_4 &= \text{diagram} \\ &= \frac{(-ig)^2}{2} \int \frac{d^3k}{(2\pi)^3} \frac{1}{4\omega_k \omega_{p-k}} \frac{-i}{\omega_c - \omega_i - i\varepsilon} \frac{-i}{\omega_c - \omega_f - i\varepsilon}. \end{aligned} \quad (4.3.24)$$

Lastly, there are contributions from diagrams with one vertex in the asymptotic region:

$$\begin{aligned} \mathcal{S}_5 &= \text{---} \text{---} \text{---} \text{---} \text{---} \text{---} \\ &= \frac{(ig)(-ig)}{2} \int \frac{d^3k}{(2\pi)^3} \frac{-i}{4\omega_k \omega_{p-k} \omega_c - \omega_f - i\varepsilon} (2\pi) \delta(\omega_c - \omega_i), \quad (4.3.25) \\ \mathcal{S}_6 &= \text{---} \text{---} \text{---} \text{---} \text{---} \text{---} \\ &= \frac{(ig)(-ig)}{2} \int \frac{d^3k}{(2\pi)^3} \frac{-i}{4\omega_k \omega_{p-k} \omega_c - \omega_i - i\varepsilon} (2\pi) \delta(\omega_c - \omega_f). \end{aligned}$$

Adding these contributions up, we find

$$\sum_{i=1}^6 \mathcal{S}_i = 0. \quad (4.3.26)$$

Similarly, all the contributions to the other time ordering of the diagram in Eq. (4.3.20) sum up to zero. Note that for the cancellation to occur, it was important to keep track of the distributional nature of the diagrams as encoded in the factors of ε .

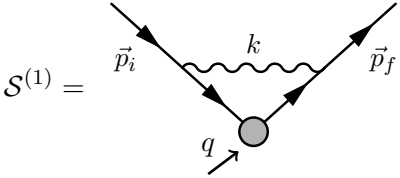
4.4 QED: Deep Inelastic Scattering

As a first real application, we consider the $e^- \gamma^* \rightarrow e^-$ in QED with a massless electron. We call this deep inelastic scattering (DIS) in reference to the analogous process in QCD at the parton level, although obviously there is nothing inelastic about this scattering. We want

to establish two facts about this process: that the hard S -matrix is IR-finite and what its value is. To compute the value for S_H it is most sensible to use dimensional regularization. In dim reg, all the diagrams with interactions in the asymptotic region give scaleless integrals that formally vanish, so the bare S_H -matrix element is determined by the S -matrix element alone. However, in pure dimensional regularization, it is difficult to separate UV from IR singularities. Therefore to check the cancellation of IR divergences, we use explicit cutoffs in the asymptotic regions.

4.4.1 S_H using cutoffs on H_{as}

In this section, we look at the diagram where a photon is exchanged between the two electron legs. The Feynman diagram in Feynman-'t Hooft gauge is given by [85]

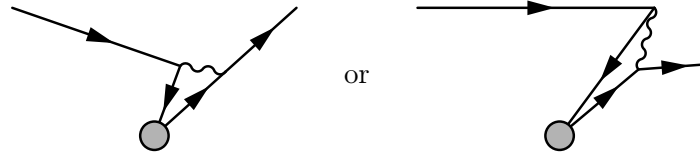


$$S^{(1)} = i\mathcal{M}_0(2\pi)^d \delta^d(p_i + q - p_f) \frac{\alpha}{4\pi} \left[\frac{1}{\epsilon_{\text{UV}}} - \frac{2}{\epsilon_{\text{IR}}^2} - \frac{4}{\epsilon_{\text{IR}}} - \frac{2 \ln \frac{\tilde{\mu}^2}{Q^2}}{\epsilon_{\text{IR}}} - \ln^2 \frac{\tilde{\mu}^2}{Q^2} - 3 \ln \frac{\tilde{\mu}^2}{Q^2} - 8 + \frac{\pi^2}{6} \right], \quad (4.4.1)$$

with $\tilde{\mu}^2 = 4\pi e^{-\gamma_E} \mu^2$ and $\mathcal{M}_0 = -e\bar{u}_f \gamma^\alpha u_i$ the tree-level matrix element. To get a cancellation of the IR divergent terms, we need to add contributions to S_H from graphs with vertices in the asymptotic regions. We would like to avoid the possible double counting of the soft and collinear degrees of freedom in H_{as} . Working in pure dimensional regularization,

the soft-collinear overlap always gives scaleless integrals that vanish. Indeed, the method-of-regions approach is to simply discount the overlap region all together. If one works with regulators that separate the UV from IR, one can explicitly remove the overlap through a zero-bin subtraction procedure [87]. In SCET, this is done by computing the soft contribution and the collinear contribution then subtracting the soft-collinear overlap through a soft-collinear power expansion at the diagram level. If one formulates SCET in terms of operators with full theory fields, as in [50], the zero-bin subtraction appears as an operator-level subtraction. In this section, we take the pragmatic approach of [50]: we exclude by hand the soft-collinear region in H_{as} . So we compute soft contributions from H_{as} by power expanding in the soft limit, and then integrating photon momenta up to some ω^{max} . We compute the collinear contributions by power expanding in the collinear limit and including only those photons with energy greater than ω^{max} that are within θ^{max} of one of the collinear directions. Similar calculations showing IR divergence cancellations for thrust and jet broadening can be found in [131].

To check IR divergence cancellations, we only need to look at a subset of time-ordered perturbation theory diagrams. For example, the diagrams


(4.4.2)

are not IR divergent. Although these diagrams give finite contributions to S_H , they do not need to be analyzed for the purposes of demonstrating IR finiteness.

It is natural to work in the Breit or “brick-wall” frame, where the off-shell photon has no energy, $q^\mu = (0, 0, 0, Q)$ and p_i and p_f are back to back. Defining θ as the angle between \vec{k} and Q , we have

$$p_i^\mu = (\omega_i, 0, 0, \omega_i), \quad p_f^\mu = (\omega_f, 0, 0, -\omega_f), \quad k^\mu = (\omega_k, 0, \omega_k \sin \theta, \omega_k \cos \theta) \quad (4.4.3)$$

and

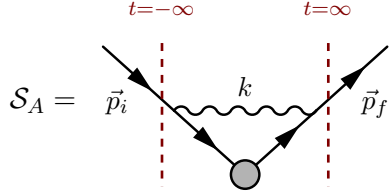
$$\omega_{i-k} = \sqrt{\omega_i^2 - 2\omega_i\omega_k \cos \theta + \omega_k^2}, \quad \omega_{f-k} = \sqrt{\omega_i^2 + 2\omega_i\omega_k \cos \theta + \omega_k^2}. \quad (4.4.4)$$

If we were to impose overall energy conservation, then we would also have $\omega_i = \omega_f = \frac{Q}{2}$. However, in time-ordered perturbation theory graphs involving vertices in the asymptotic regions, energy conservation is not guaranteed, so for those diagrams we leave ω_i and ω_f more general until energy conservation can be established. With these kinematics the phase space integral becomes

$$\int \frac{d^{d-1}k}{(2\pi)^{d-1}} = \frac{\Omega_{d-2}}{(2\pi)^{d-1}} \int d\omega_k \omega_k^{d-2} \int_{-1}^1 d\cos \theta (1 - \cos^2 \theta)^{\frac{d-4}{2}}, \quad (4.4.5)$$

where $\Omega_{d-2} = 2\pi^{\frac{d-2}{2}}/\Gamma(\frac{d-2}{2})$ is the $d-2$ -dimensional solid angle.

The graph with all the vertices in the central region is



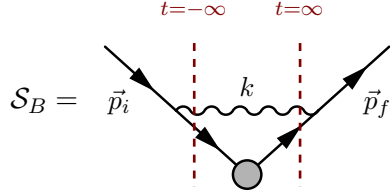
$$\begin{aligned}
\mathcal{S}_A &= (-ie)^3 \mu^{4-d} \int \frac{d^{d-1}k}{(2\pi)^{d-1}} \frac{1}{2\omega_k} \frac{1}{2\omega_{i-k}} \frac{1}{2\omega_{f-k}} \frac{i}{\omega_i - \omega_{i-k} - \omega_k + i\epsilon} \frac{i}{\omega_f - \omega_{f-k} - \omega_k + i\epsilon} \\
&\quad \times \bar{u}_f \gamma^\mu u_{f-k} \bar{u}_{f-k} \gamma^\alpha u_{i-k} \bar{u}_{i-k} \gamma^\nu u_i (-g^{\mu\nu}) (2\pi)^d \delta^d(p_i + q - p_f) .
\end{aligned} \tag{4.4.6}$$

This graph is UV and IR divergent. But since this is the only IR-divergent time ordering, we know its result must reproduce the IR divergences of the sum over all time orderings, i.e. the Feynman diagram in the full theory. So we can then read the IR divergences directly off of Eq. (4.4.1):

$$\mathcal{S}_A = i\mathcal{M}_0 (2\pi)^d \delta^d(p_i + q - p_f) \frac{\alpha}{4\pi} \left[-\frac{2}{\epsilon_{\text{IR}}^2} - \frac{4}{\epsilon_{\text{IR}}} - \frac{2 \ln \frac{\tilde{\mu}^2}{Q^2}}{\epsilon_{\text{IR}}} + \text{IR-finite} \right] \tag{4.4.7}$$

The contribution with both interactions in an asymptotic region is given by soft photon exchange alone; there are no collinear photons that couple to both the incoming and outgoing electrons since these are back-to-back. Thus, we need to power expand the integrand in Eq. (4.4.1) at small ω_k and restrict to $\omega_k < \omega^{\text{max}}$. *Before* power expanding, the time-

ordered perturbation theory amplitude has the form



$$\begin{aligned}
\mathcal{S}_B = & (-ie)(ie)^2 \mu^{4-d} \int \frac{d^{d-1}k}{(2\pi)^{d-1}} \frac{1}{2\omega_k} \frac{1}{2\omega_{i-k}} \frac{1}{2\omega_{f-k}} \theta(\omega^{\max} - \omega_k) \\
& \times \frac{-i}{\omega_{i-k} + \omega_k - \omega_i - i\varepsilon} \frac{-i}{\omega_{f-k} + \omega_k - \omega_f - i\varepsilon} \\
& \times \bar{u}_f \gamma^\mu u_{f-k} \bar{u}_{f-k} \gamma^\alpha u_{i-k} \bar{u}_{i-k} \gamma^\nu u_i (-g^{\mu\nu}) (2\pi)^d \delta^{d-1}(\vec{p}_i + \vec{q} - \vec{p}_f) \delta(\omega_{i-k} - \omega_{f-k}).
\end{aligned} \tag{4.4.8}$$

Note that the overall energy-conserving δ -function $\delta(\omega_i - \omega_f)$ from Eq. (4.4.1) is replaced with $\delta(\omega_{i-k} - \omega_{f-k})$. In Eq. (4.4.8), however at leading power the two δ -functions agree. In the soft limit, the energies of the intermediate electrons are

$$\omega_{i-k} = \sqrt{\omega_i^2 - 2\omega_i\omega_k \cos\theta + \omega_k^2} \cong \omega_i - \omega_k \cos\theta, \tag{4.4.9}$$

$$\omega_{f-k} = \sqrt{\omega_f^2 + 2\omega_f\omega_k \cos\theta + \omega_k^2} \cong \omega_f + \omega_k \cos\theta, \tag{4.4.10}$$

and the numerators are expanded as

$$\bar{u}_f \gamma^\mu u_{f-k} \bar{u}_{f-k} \gamma^\alpha u_{i-k} \bar{u}_{i-k} \gamma^\nu u_i (-g^{\mu\nu}) \cong -4p_i \cdot p_f \bar{u}_f \gamma^\alpha \bar{u}_i = -8\omega_i\omega_f \bar{u}_f \gamma^\alpha \bar{u}_i. \tag{4.4.11}$$

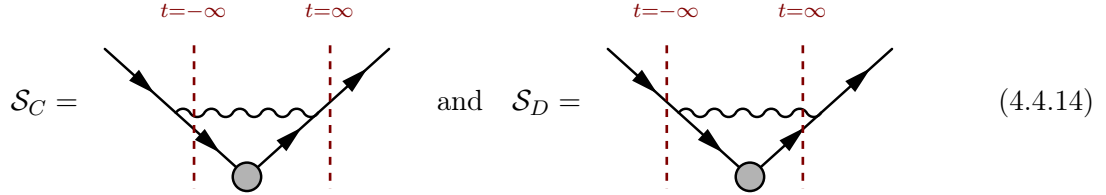
Inserting the power expansion, the amplitude reduces to

$$\mathcal{S}_B = -i\mathcal{M}_0(2\pi)^d \delta^d(p_i + q - p_f) \frac{\Omega_{d-2}}{(2\pi)^{d-1}} \mu^{2\epsilon} \times \int_0^{\omega^{\max}} d\omega_k \omega_k^{1-2\epsilon} \int_{-1}^1 dx (1-x^2)^{-\epsilon} \frac{1}{\omega_k(1-x) - i\varepsilon} \frac{1}{\omega_k(1+x) - i\varepsilon}, \quad (4.4.12)$$

where $x = \cos \theta$. Performing the integrals gives

$$\mathcal{S}_B = i\mathcal{M}_0(2\pi)^d \delta^d(p_i + q - p_f) \frac{\alpha}{4\pi} \left[-\frac{2}{\epsilon_{\text{IR}}^2} + \frac{2 \ln \frac{(2\omega^{\max})^2}{\tilde{\mu}^2}}{\epsilon_{\text{IR}}} + \frac{\pi^2}{2} - \ln^2 \frac{(2\omega^{\max})^2}{\tilde{\mu}^2} \right]. \quad (4.4.13)$$

The remaining two graphs are



$$\mathcal{S}_C = \quad \text{and} \quad \mathcal{S}_D = \quad (4.4.14)$$

These have one vertex in the asymptotic region and one in the central region. In the first graph, the asymptotic vertex forces the exchanged photon to either be soft or collinear to the direction of the outgoing electron. In the second graph, the photon can be soft or collinear to the incoming electron. We must therefore power expand each in soft and collinear limits separately.

Before doing any expansion the first graph is

$$\mathcal{S}_C = (-ie)^2 (ie) \mu^{4-d} \int \frac{d^{d-1}k}{(2\pi)^{d-1}} \frac{1}{2\omega_k} \frac{1}{2\omega_{i-k}} \frac{1}{2\omega_{f-k}} \frac{-i}{\omega_{i-k} + \omega_k - \omega_i - i\varepsilon} \frac{i}{\omega_f - \omega_{f-k} - \omega_k + i\varepsilon} \\ \bar{u}_f \gamma^\mu u_{f-k} \bar{u}_{f-k} \gamma^\alpha u_{i-k} \bar{u}_{i-k} \gamma^\nu u_i (-g^{\mu\nu}) (2\pi)^d \delta^{d-1}(\vec{p}_i + \vec{q} - \vec{p}_f) \delta(\omega_{i-k} + \omega_k - \omega_f). \quad (4.4.15)$$

In the soft limit, this reduces to the same integral as in \mathcal{S}_B up to a sign flip since only one vertex is anti-time ordered. \mathcal{S}_D is similar. So we get

$$\mathcal{S}_C^{\text{soft}} = \mathcal{S}_D^{\text{soft}} = i\mathcal{M}_0 (2\pi)^d \delta^d(p_i + q - p_f) \frac{\alpha}{4\pi} \left[\frac{2}{\epsilon_{\text{IR}}^2} - \frac{2 \ln \frac{(2\omega^{\text{max}})^2}{\tilde{\mu}^2}}{\epsilon_{\text{IR}}} - \frac{\pi^2}{2} + \ln^2 \frac{(2\omega^{\text{max}})^2}{\tilde{\mu}^2} \right]. \quad (4.4.16)$$

These will cancel the double poles of $\mathcal{S}_A + \mathcal{S}_B$.

The graph \mathcal{S}_C has a collinear singularity when $\theta \rightarrow 0$. For the collinear graphs, as mentioned above, we consider collinear photons to be collinear but not soft, so they have energies $\omega_k > \omega^{\text{max}}$ and angles $0 < \theta < \theta^{\text{max}}$. In the collinear limit, $k \parallel p_i$, the energies expand to

$$\omega_{i-k} = \sqrt{\omega_i^2 - 2\omega_i\omega_k \cos \theta + \omega_k^2} \cong \omega_i - \omega_k + \frac{\omega_i\omega_k}{\omega_i - \omega_k} (1 - \cos \theta), \quad (4.4.17)$$

$$\omega_{f-k} = \sqrt{\omega_f^2 + 2\omega_f\omega_k \cos \theta + \omega_k^2} \cong \omega_f + \omega_k. \quad (4.4.18)$$

Since these expansions are only valid in the regime where the electron does not recoil against the photon, i.e. for $\omega_k < \omega_i$, we put ω_i as an upper cutoff on the photon energy. The spinors in the numerator are on-shell, so in the collinear limit the numerator can be

approximated using $p_{i-k} \cong \frac{\omega_{i-k}}{\omega_i} p_i$ and $p_{f-k} \cong \frac{\omega_{f-k}}{\omega_f} p_f$, and hence

$$\bar{u}_f \gamma^\mu u_{f-k} \bar{u}_{f-k} \gamma^\alpha u_{i-k} \bar{u}_{i-k} \gamma^\nu u_i (-g^{\mu\nu}) \cong -4 p_i \cdot p_f \frac{\omega_{i-k} \omega_{f-k}}{\omega_i \omega_f} \bar{u}_f \gamma^\alpha u_i \cong -8 \omega_{i-k} \omega_{f-k} \bar{u}_f \gamma^\alpha u_i. \quad (4.4.19)$$

Then $\mathcal{S}_C^{\text{coll}}$ reduces to

$$\begin{aligned} \mathcal{S}_C^{\text{coll}} &= -ie^2 \mathcal{M}_0 (2\pi)^d \delta^d(p_i + q - p_f) \frac{\Omega_{d-2}}{(2\pi)^{d-1}} \mu^{2\epsilon} \\ &\quad \times \int_{\omega^{\max}}^{\omega_i} d\omega_k \omega_k^{1-2\epsilon} \int_0^{\theta^{\max}} d\theta \sin^{1-2\epsilon} \theta \frac{1 - \frac{\omega_k}{\omega_i}}{\omega_k(1 - \cos \theta) - i\epsilon - 2\omega_k + i\epsilon} \frac{1}{\omega_k(1 - \cos \theta) - i\epsilon - 2\omega_k + i\epsilon} \\ &= i \mathcal{M}_0 (2\pi)^d \delta^d(p_i + q - p_f) \frac{\alpha}{4\pi} \left[\frac{2}{\epsilon_{\text{IR}}} + \frac{\ln \frac{(2\omega^{\max})^2}{Q^2}}{\epsilon_{\text{IR}}} \right. \\ &\quad \left. + \left(2 + \ln \frac{(2\omega^{\max})^2}{Q^2} \right) \left(2 - \ln \frac{(\theta^{\max} \omega^{\max})^2}{\tilde{\mu}^2} \right) + \frac{1}{2} \ln^2 \frac{(2\omega^{\max})^2}{Q^2} \right]. \end{aligned} \quad (4.4.20)$$

Note that this graph has a single $\frac{1}{\epsilon}$ pole corresponding to the collinear-but-not-soft region.

The amplitude $\mathcal{S}_D^{\text{coll}}$ is the same as $\mathcal{S}_C^{\text{coll}}$.

In summary, extracting just the IR poles,

$$\mathcal{S}_A = i\mathcal{M}_0 (2\pi)^d \delta^d(p_i + q - p_f) \frac{\alpha}{4\pi} \left[-\frac{2}{\epsilon_{\text{IR}}^2} - \frac{4}{\epsilon_{\text{IR}}} - \frac{2 \ln \frac{\tilde{\mu}^2}{Q^2}}{\epsilon_{\text{IR}}} + \text{IR-finite} \right] \quad (4.4.21)$$

$$\mathcal{S}_B = i\mathcal{M}_0 (2\pi)^d \delta^d(p_i + q - p_f) \frac{\alpha}{4\pi} \left[-\frac{2}{\epsilon_{\text{IR}}^2} + \frac{2 \ln \frac{(2\omega^{\text{max}})^2}{\tilde{\mu}^2}}{\epsilon_{\text{IR}}} + \text{IR-finite} \right] \quad (4.4.22)$$

$$\mathcal{S}_C = i\mathcal{M}_0 (2\pi)^d \delta^d(p_i + q - p_f) \frac{\alpha}{4\pi} \left[\frac{2}{\epsilon_{\text{IR}}^2} + \frac{2}{\epsilon_{\text{IR}}} + \frac{\ln \frac{(2\omega^{\text{max}})^2}{Q^2}}{\epsilon_{\text{IR}}} - \frac{2 \ln \frac{(2\omega^{\text{max}})^2}{\tilde{\mu}^2}}{\epsilon_{\text{IR}}} + \text{IR-finite} \right] \quad (4.4.23)$$

$$\mathcal{S}_D = i\mathcal{M}_0 (2\pi)^d \delta^d(p_i + q - p_f) \frac{\alpha}{4\pi} \left[\frac{2}{\epsilon_{\text{IR}}^2} + \frac{2}{\epsilon_{\text{IR}}} + \frac{\ln \frac{(2\omega^{\text{max}})^2}{Q^2}}{\epsilon_{\text{IR}}} - \frac{2 \ln \frac{(2\omega^{\text{max}})^2}{\tilde{\mu}^2}}{\epsilon_{\text{IR}}} + \text{IR-finite} \right] \quad (4.4.24)$$

with $\mathcal{M}_0 = -e\bar{u}_f \gamma^\alpha u_i$ the tree-level matrix element. Summing these graphs, the IR divergences all cancel.

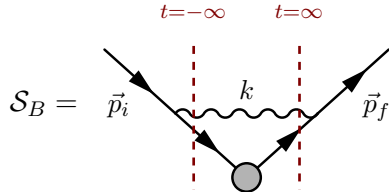
Note that this is a different mechanism from the way the cancellation happens in a matching calculation for the DIS Wilson coefficient in SCET [85]. There, the soft graph is subtracted from the full theory graph ($\mathcal{S}_A - \mathcal{S}_B$) to achieve the cancellation. Here those graphs are added, and additional graphs come in to effect the cancellation.

4.4.2 S_H in dimensional regularization

Imposing cutoffs on the asymptotic Hamiltonian is useful for showing the cancellation of IR divergences. In practice, however, the calculations are much simpler using pure dimensional regularization. Dimensional regularization respects both Lorentz and gauge invari-

ance, while explicit cutoffs do not. Moreover all 1PI graphs involving vertices in the asymptotic region are scaleless and formally vanish. This follows from simple power counting arguments: in the soft limit, we take all hard scales to infinity so there are no scales left for the amplitude to depend on. In collinear limits, only lightlike momenta in one direction are relevant and no Lorentz-invariant scale can be constructed from collinear lightlike momenta.

For an explicit example, consider the soft graph \mathcal{S}_B , from Eq. (4.4.12)



$$\mathcal{S}_B = -i\mathcal{M}_0(2\pi)^d \delta^d(p_i + q - p_f) \frac{\Omega_{d-2}}{(2\pi)^{d-2}} \mu^{2\epsilon} \int_0^\infty d\omega_k \omega_k^{-1-2\epsilon} \int_{-1}^1 dx (1-x^2)^{-1-\epsilon}. \quad (4.4.25)$$

The integral over ω_k is scaleless and formally vanishes in dimensional regularization. Note that there is also a IR divergence in this case in the angular, x , integral, so the final result has an overlapping UV/IR $\frac{1}{\epsilon_{\text{UV}}} \frac{1}{\epsilon_{\text{IR}}}$ singularity. Such singularities never occur in renormalizable theory, but they do occur in SCET. However, since when one adds up all the diagrams we know that the IR divergences cancel, the overlapping UV/IR divergences must cancel as well. These cancellations have been studied extensively in SCET (see the reviews [82, 83]).

Thus the only non-vanishing graphs in pure dimensional regularization are those with all vertices in the central region. In the central region, hard interactions are present, and these are associated with particular scales. In $d = 4 - 2\epsilon$ dimensions, in Feynman gauge, the result for the loop is given in Eq. (4.4.1). For this diagram, the UV and IR divergences can

be unambiguously separated since the UV divergences are known separately to be cancelled by the ordinary QED counterterms. For S_H diagrams, such a separation is also possible, but much more difficult, since there can be overlapping UV and IR singularities (see [87, 131, 132] for some discussion).

In any case, since the other diagrams contributing to S_H are scaleless and since S_H is IR finite, we can immediately write down the bare S_H amplitude using Eq. (4.4.1). Writing, for $|\psi_{\text{out}}\rangle \neq |\psi_{\text{in}}\rangle$,

$$\langle \psi_{\text{out}} | S_H | \psi_{\text{in}} \rangle = (2\pi)^d \delta^d(p_{\text{in}} - p_{\text{out}}) i \widehat{M} \quad (4.4.26)$$

we then have

$$\widehat{M}_{\text{bare}} = \mathcal{M}_0 \left[1 + \frac{\alpha(\mu)}{4\pi} \left(-\frac{2}{\epsilon_{\text{UV}}^2} - \frac{2 \ln \frac{\tilde{\mu}^2}{Q^2} + 3}{\epsilon_{\text{UV}}} - \ln^2 \frac{\tilde{\mu}^2}{Q^2} - 3 \ln \frac{\tilde{\mu}^2}{Q^2} - 8 + \frac{\pi^2}{6} \right) + \mathcal{O}(\alpha^2) \right]. \quad (4.4.27)$$

The renormalized S_H -matrix element is related to the bare one by operator renormalization. To remove the UV divergences, we can rescale the S -matrix by

$$Z = 1 + \frac{\alpha(\mu)}{4\pi} \left(-\frac{2}{\epsilon_{\text{UV}}^2} - \frac{2 \ln \frac{\tilde{\mu}^2}{Q^2} + 3}{\epsilon_{\text{UV}}} \right) + \mathcal{O}(\alpha^2). \quad (4.4.28)$$

So that the renormalized matrix element in $\overline{\text{MS}}$ is then

$$\widehat{M} = \left[\frac{1}{Z_4} \widehat{M}_{\text{bare}} \right] = \mathcal{M}_0 \left[1 + \frac{\alpha(\mu)}{4\pi} \left(-\ln^2 \frac{\tilde{\mu}^2}{Q^2} - 3 \ln \frac{\tilde{\mu}^2}{Q^2} - 8 + \frac{\pi^2}{6} \right) + \mathcal{O}(\alpha^2) \right], \quad (4.4.29)$$

which is UV and IR finite.

It may seem surprising that Z can depend on the scale Q : normally Z -factors are just numbers. In fact, the Q dependence is just shorthand for a more formal dependence of the S_H -matrix elements on the labels of the collinear fields. In the label formalism, the S -matrix for $e^-(p_1)\gamma^*(q) \rightarrow e^-(p_2)$ can depend on its labels, which are the large components of the momenta of the collinear particles, $p_1^- = \bar{n}_1 \cdot p_1 \sim Q$ and $p_2^+ = \bar{n}_2 \cdot p_2 \sim Q$. These labels are non-dynamical, and so the Z -factor can depend on them. Thus, one could more pedantically write

$$Z_{p_1^- p_2^+} = 1 + \frac{\alpha(\mu)}{4\pi} \left(-\frac{2}{\epsilon_{\text{UV}}^2} - \frac{2 \ln \frac{\tilde{\mu}^2}{p_1^- p_2^+} + 3}{\epsilon_{\text{UV}}} \right) + \mathcal{O}(\alpha^2) \quad (4.4.30)$$

and $S_{p_1^- p_2^+}^{H, \text{bare}} = Z_{p_1^- p_2^+} S_{p_1^- p_2^+}^H$. But writing the dependence as on Q or more generally $s_{ij} = (p_i + p_j)^2$ is simpler.

It is perhaps worth commenting on why S_H needs to be renormalized in the first place. The traditional S -matrix is also an operator, however it does not normally get an operator renormalization: its UV divergences are cancelled by rescaling the interaction strengths in the Lagrangian and the fields. The reason S_H needs to be renormalized is due to diagrams that have both interactions in the asymptotic regions and hard momentum flowing through the graph due to interactions in the central region. The soft particles in H_{as} cannot resolve the hard scales and there are no interactions in H_{as} which could be renormalized to remove the associated UV divergences. While S -matrix elements are smooth, differentiable functions of momenta, the smoothness is lost in the soft power expansion generating S_H . Thus hard scattering, from the point of S_H looks instantaneous and non-local, like a sharp,

non-differentiable cusp at the hard vertex. In other words, the additional renormalization required in S_H is the same as the need for renormalization associated with cusps in Wilson line matrix elements. The non-locality of SCET (on hard length scales) and cusp renormalization is discussed more in [82, 83].

4.5 QCD: $e^+e^- \rightarrow \text{jets}$

To illustrate the use of S_H to compute infrared-safe observables, we will explore as an example, the computation of thrust in e^+e^- events to NLO in QCD.

The hard matrix element for $\gamma^* \rightarrow \bar{q}q$ is the same as for DIS, up to a crossing. Explicitly,

$$\widehat{M} = \mathcal{M}_0 \left[1 + \frac{\alpha_s(\mu)}{4\pi} C_F \left(-\ln^2 \frac{\tilde{\mu}^2}{-Q^2 - i\varepsilon} - 3 \ln \frac{\tilde{\mu}^2}{-Q^2 - i\varepsilon} - 8 + \frac{\pi^2}{6} \right) + \mathcal{O}(\alpha^2) \right]. \quad (4.5.1)$$

Due to the $\ln(-Q^2 - i\varepsilon)$ term, this S_H -matrix element is complex. The imaginary part is the leading order expansion of the Coulomb/Glauber phase, and is present in processes with more than one charged particle in the initial or final state.

4.5.1 Glauber graph

It is perhaps illuminating to see the origin of the imaginary part from the relevant asymptotic-region graphs. Part of the reason this question is interesting in our framework is because Glauber gluons are normally associated with purely off-shell modes, with entirely transverse momentum. In time-ordered perturbation theory one has only on-shell modes.

So how is the Glauber contribution going to be reproduced?

$$\begin{aligned}
& \sim \int \frac{d^{d-1}k}{(2\pi)^{d-1}} \frac{1}{2\omega_k} \frac{1}{2\omega_{1+k}} \frac{1}{2\omega_{2-k}} \\
& \times \frac{-i}{\omega_{1+k} - (\omega_1 + \omega_k) - i\varepsilon} \frac{-i}{\omega_{1+k} + \omega_{2-k} - (\omega_1 + \omega_2) - 2i\varepsilon}
\end{aligned}
\tag{4.5.2}$$

If we were to enforce 3-momentum and energy-conservation in the central region, this would force $\omega_{1+k} = \omega_{p_2+k} = \omega_1 = \omega_2 = \frac{Q}{2}$. Then k must have exactly zero energy, as expected for an off-shell mode, and the integrand appears ill-defined. The problem however is not that k is off-shell, but that we have not been sufficiently careful handling the product of distributions.

To properly evaluate the integral, we must be patient in enforcing the energy conservation in the central region. Recall that energy conservation comes from integrating over $-\infty < t < \infty$. If we break the central region up into a $-\infty$ to 0 region and a 0 to ∞ region, then the hard vertex can be in only one of the regions. Let us also pretend for now that H_{as} is the same as H with the exception of the hard vertex. Then, if the hard vertex is at $t < 0$, the evolution from e^{-iHt} from 0 to ∞ will be exactly cancelled by the evolution

from $t = \infty$ to 0 in the asymptotic region. That is,

$$= 0. \quad (4.5.3)$$

In equations, the cancellation occurs point-by-point in phase space as

$$\left[\frac{i}{\omega_f - \omega_i + 2i\varepsilon} \frac{i}{\omega_f - \omega_c + i\varepsilon} - \frac{i}{\omega_c - \omega_i + i\varepsilon} \frac{-i}{\omega_c - \omega_f - i\varepsilon} + \frac{-i}{\omega_i - \omega_c - i\varepsilon} \frac{-i}{\omega_i - \omega_f - 2i\varepsilon} \right] \times \frac{i}{Q - \omega_i + i\varepsilon} = 0, \quad (4.5.4)$$

where $\omega_i = \omega_{1+k} + \omega_{2-k}$, $\omega_c = \omega_{1+k} + \omega_{2-k} + \omega_k$ and $\omega_f = \omega_1 + \omega_2$. In the real case,

where H_{as} is not exactly the same as H without the hard vertex, these graphs will not sum to precisely zero, but to something that is IR finite.

The cancellation of the graphs with the hard vertex at $t < 0$ implies that the nonzero contribution of the graph in Eq. (4.5.2) comes from the region where the hard vertex is at $t > 0$. So we must look at

$$\mathcal{M}_G = \sim \int \frac{d^{d-1}k}{(2\pi)^{d-1}} \frac{1}{2\omega_k} \frac{1}{2\omega_{1+k}} \frac{1}{2\omega_{2-k}} \frac{i}{\omega_{1+k} + \omega_{2-k} - Q + i\varepsilon} \times \frac{-i}{\omega_{1+k} - (\omega_1 + \omega_k) - i\varepsilon} \frac{-i}{\omega_{1+k} + \omega_{2-k} - (\omega_1 + \omega_2) - 2i\varepsilon}. \quad (4.5.5)$$

Now we only have 3-momentum conservation, not energy conservation. So, $\vec{p}_1 + \vec{p}_2 = 0$ and thus $\omega_1 = \omega_2$, but nothing forces $\omega_1 = \frac{Q}{2}$. Defining the angle between \vec{k} and \vec{p}_1 as θ , in the soft limit $\omega_{1+k} \cong \omega_1 + \omega_k \cos \theta$ and $\omega_{2-k} \cong \omega_2 + \omega_k \cos \theta$, so performing the power expansion results in

$$\mathcal{M}_G \sim \frac{i}{\omega_1 + \omega_2 - Q + i\varepsilon} \int \frac{d^{d-1}k}{(2\pi)^{d-1}} \frac{1}{\omega_k^3} \frac{1}{\cos \theta - 1 - i\varepsilon} \frac{1}{\cos \theta - i\varepsilon} \quad (4.5.6)$$

$$\sim \frac{i}{\omega_1 + \omega_2 - Q + i\varepsilon} \int d\omega_k \omega_k^{d-5} \left(\frac{1}{\epsilon_{\text{IR}}} - i\pi + \dots \right). \quad (4.5.7)$$

The ω_k integral is scaleless, being both UV and IR divergent. The $i\pi$ in this expression corresponds to the imaginary part in Eq. (4.5.1), and is known to exponentiate into the Coulomb/Glauber phase. The third graph in Eq. (4.5.3) is similar, leading to the same result as in Eq. (4.5.7) with $\frac{i}{\omega_1 + \omega_2 - Q + i\varepsilon}$ replaced by $\frac{i}{Q - \omega_1 - \omega_2 + i\varepsilon}$. The two graphs combine to produce the expected $\delta(\omega_1 + \omega_2 - Q)$ factor.

So we see that the Glauber phase is indeed reproduced by asymptotic diagrams with on-shell modes. Moreover, energy *is* conserved in this process. The key was to carefully handle the imaginary parts of the propagators and δ distributions. There are of course many other ways to compute this imaginary part (cf. [22]), but this approach clarifies the importance of carefully treating energy conservation in S_H computations.

In more complicated processes, such as $\bar{q}q \rightarrow \bar{q}q$ in QCD at 2 loops, it is known that the Glauber contribution from the full graph (the central region) is not reproduced by the eikonal approximation [75]. Consequences of this failure include collinear-factorization violation [20] and the emergence of super-leading logarithms [19]. For S_H this means that

the IR divergences of the central region will not be canceled by an asymptotic Hamiltonian with soft and collinear gluons alone. Fortunately, it has been shown that one can add to the SCET Lagrangian a set of Glauber operators [84] and remedy the failure of the soft limit. A detailed discussion of when these operators are relevant and how they resolve issues such as collinear-factorization violation can be found in [133]. The Glauber interactions, like soft interactions, are long distance and will persist after the hard scattering. Although they violate factorization, in the sense that they are long-distance interactions that depend on multiple directions, they are still independent of the hard scattering.

To connect the Glauber graph \mathcal{M}_G to the Glauber operator, we can massage the imaginary part of the integral in Eq. (4.5.6) into a more familiar form. We first drop the $i\epsilon$ in the denominator $\frac{1}{\cos\theta-1-i\epsilon}$, since the endpoint singularity at $\cos\theta = 1$ is regulated for $\epsilon < 0$ by the $(1 - \cos^2\theta)^{-\epsilon}$ factor in the measure (see Eq. (4.4.5)). Rewriting the integral in terms of $k_z = \omega_k \cos\theta$ and \vec{k}_\perp gives

$$\mathcal{M}_G \sim \int_{-\infty}^{\infty} dk_z \int \frac{d^{d-2}\vec{k}_\perp}{(2\pi)^{d-2}} \frac{1}{\sqrt{k_z^2 + \vec{k}_\perp^2}} \frac{1}{k_z - \sqrt{k_z^2 + \vec{k}_\perp^2}} \frac{1}{k_z - 2i\epsilon}. \quad (4.5.8)$$

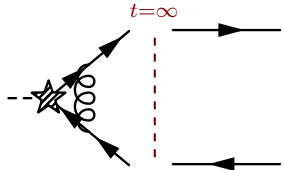
To take the imaginary part we now use $\text{Im} \left[\frac{1}{k_z - 2i\epsilon} \right] = \pi\delta(k_z)$ and integrate over k_z to get

$$\text{Im} [\mathcal{M}_G] \sim -i\pi \int \frac{d^{d-2}\vec{k}_\perp}{(2\pi)^{d-2}} \frac{1}{\vec{k}_\perp^2}. \quad (4.5.9)$$

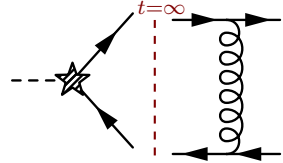
This $\frac{1}{\vec{k}_\perp^2}$ integrand is exactly what comes out of the $\bar{\xi}_{n_1} \frac{\not{k}_2}{2} \xi_{n_1} \frac{1}{\vec{p}_\perp^2} \bar{\xi}_{n_2} \frac{\not{k}_1}{2} \xi_{n_2}$ Glauber operators [22, 23, 84]. In other words, tree-level exchange in the asymptotic region corresponds

to the Glauber region expansion, except it has an opposite sign. Note that since $k_z = 0$ the on-shell energy of the Glauber gluon is $\omega_k = |\vec{k}_\perp|$. So the $\frac{1}{k_\perp^2}$ is not coming from an off-shell mode but rather from energy not being conserved in time-ordered perturbation theory. Alternative ways of understanding the Glauber phase can be found in [121, 122, 134, 135].

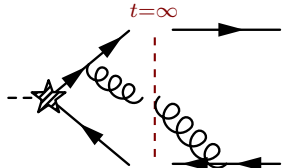
For completeness, we list the IR divergent parts of the various contributions to S_H for this process cutting off the UV divergence of the soft integrals at ω^{\max} , as in Section 4.4.1. Writing $S_H = i\widehat{M}(2\pi)^d\delta^d(q - p_1 - p_2)$, the contributions to \widehat{M} are:



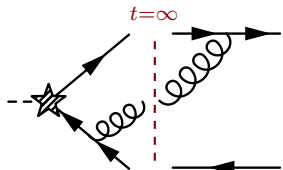
$$= \mathcal{M}_0 \frac{\alpha_s}{4\pi} C_F \left[-\frac{2}{\epsilon_{\text{IR}}^2} - \frac{4}{\epsilon_{\text{IR}}} + \frac{2 \ln \frac{Q^2}{\tilde{\mu}^2}}{\epsilon_{\text{IR}}} - \frac{2i\pi}{\epsilon_{\text{IR}}} + \text{IR-finite} \right] \quad (4.5.10)$$



$$= \mathcal{M}_0 \frac{\alpha_s}{4\pi} C_F \left[-\frac{2}{\epsilon_{\text{IR}}^2} + \frac{2 \ln \frac{(2\omega^{\max})^2}{\tilde{\mu}^2}}{\epsilon_{\text{IR}}} + \frac{2i\pi}{\epsilon_{\text{IR}}} + \text{IR-finite} \right] \quad (4.5.11)$$



$$= \mathcal{M}_0 \frac{\alpha_s}{4\pi} C_F \left[\frac{2}{\epsilon_{\text{IR}}^2} + \frac{2}{\epsilon_{\text{IR}}} + \frac{\ln \frac{(2\omega^{\max})^2}{Q^2}}{\epsilon_{\text{IR}}} - \frac{2 \ln \frac{(2\omega^{\max})^2}{\tilde{\mu}^2}}{\epsilon_{\text{IR}}} + \text{IR-finite} \right] \quad (4.5.12)$$



$$= \mathcal{M}_0 \frac{\alpha_s}{4\pi} C_F \left[\frac{2}{\epsilon_{\text{IR}}^2} + \frac{2}{\epsilon_{\text{IR}}} + \frac{\ln \frac{(2\omega^{\max})^2}{Q^2}}{\epsilon_{\text{IR}}} - \frac{2 \ln \frac{(2\omega^{\max})^2}{\tilde{\mu}^2}}{\epsilon_{\text{IR}}} + \text{IR-finite} \right] \quad (4.5.13)$$

with $\tilde{\mu}^2 = 4\pi e^{-\gamma_E} \mu^2$ and $\mathcal{M}_0 = \bar{u}_1 \gamma^\alpha v_1$ the tree level matrix element. Summing these graphs, the IR divergences all cancel.

Note that while the imaginary part of the Glauber graph, Eq. (4.5.11), cancels against the S -matrix graph, Eq. (4.5.10), the real part of the Glauber graph has the same sign as the S -matrix graph, and the sum of the two cancels against the cut graphs. This is different from how the cancellation occurs in matching to a 2-jet operator in SCET, where a single soft graph cancels both the real and imaginary parts of the divergences of the full-theory graph.

4.5.2 Thrust

Next, let us use the hard S -matrix to compute the thrust observable [136]. Thrust is a particularly simple infrared-safe e^+e^- observable. It is defined as

$$T \equiv \max_{\vec{n}} \frac{\sum_j |\vec{p}_j \cdot \vec{n}|}{\sum_j |\vec{p}_j|}. \quad (4.5.14)$$

It is convenient to use $\tau = 1 - T$ rather than T . Thrust has the property that for events that consist of two highly collimated jets $\tau \ll 1$. At small τ , thrust is approximated by the sum of the masses of these two jets $\tau \cong \frac{1}{Q^2} (m_{J1}^2 + m_{J2}^2)$, with Q the center of mass energy. Events that are more spherical have values of $\tau \sim 0.2 - 0.5$.

To compute $\frac{d\sigma}{d\tau}$ in perturbation theory using S_H , we start at lowest order, where the hard S -matrix element is

$$\widehat{\mathcal{M}}_0(\gamma^* \rightarrow \bar{q}q) = \bar{u}_i(p_q) \gamma^\mu v_j(p_{\bar{q}}). \quad (4.5.15)$$

At next to leading order we need the hard matrix element for $\bar{q}q$ final states at NLO, as given in Eq. (4.5.1):

$$\widehat{\mathcal{M}}(\gamma^* \rightarrow \bar{q}q) = \widehat{\mathcal{M}}_0 \left[1 + \frac{\alpha_s(\mu)}{4\pi} C_F \left(-\ln^2 \frac{\mu^2}{Q^2} - (3 + 2\pi i) \ln \frac{\mu^2}{Q^2} - 8 - 3\pi i + \frac{7\pi^2}{6} \right) + \mathcal{O}(\alpha_s^2) \right] \quad (4.5.16)$$

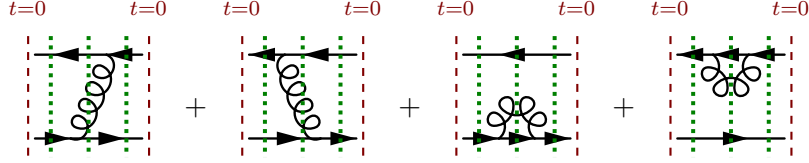
We also need the matrix elements for $\gamma^* \rightarrow \bar{q}qg$ where we treat the gluon as hard. Treating it as hard, the only contribution at order g_s is from diagrams with all vertices in the central region. Then this amplitude is identical to the S -matrix element for the same process,

$$\widehat{\mathcal{M}}(\gamma^* \rightarrow \bar{q}qg) = -g_s T_{ij}^a \bar{u}_i(p_q) \left[\gamma^\alpha \frac{1}{\not{p}_q + \not{p}_g} \gamma^\mu - \gamma^\mu \frac{1}{\not{p}_{\bar{q}} + \not{p}_g} \gamma^\alpha \right] v_j(p_{\bar{q}}) \epsilon_\alpha^*(p_g). \quad (4.5.17)$$

To compute the observable, we must then evolve these final states to $t = +\infty$ using H_{as} , as discussed in Section 4.2.2. On the formal level, this additional evolution exactly cancels the entire effect of H_{as} , so the cross section predicted is identical to that using the original S . On the practical level, however, one can gain additional insight into the distribution by actually using the S_H -matrix elements we have computed, rather than simply discarding them and starting over. To this end, it is helpful to contemplate the small τ and moderate τ regions separately.

For small τ , the gluon is necessarily soft or collinear. Thus we can disregard the hard $\widehat{\mathcal{M}}(\gamma^* \rightarrow \bar{q}qg)$ contribution. Instead, we should start with $\widehat{\mathcal{M}}_0(\gamma^* \rightarrow \bar{q}q)$ and then evolve the $\bar{q}q$ final state towards a 2-jet state with nonzero τ using H_{as} . To compute the cross sec-

tion, we need to sum the cut graphs



$$(4.5.18)$$

In these graphs each dotted green line represents a separate contribution where the measurement function at $t = \infty$ is inserted. The first two graphs have only soft contributions and the second two soft and collinear contributions (although the soft ones vanish in Feynman gauge).

The middle cuts in these graphs using soft interactions in the asymptotic regions corresponds to soft real-emission processes. The amplitude for soft emission using H_{as} is the eikonal limit of Eq. (4.5.17), with an opposite sign and without the hard matrix element,

$$\mathcal{M}_{\text{soft}} = g_s T_{ij}^a \left[\frac{p_q^\alpha}{p_q \cdot p_g} - \frac{p_{\bar{q}}^\alpha}{p_{\bar{q}} \cdot p_g} \right] \epsilon_\alpha^*(p_g). \quad (4.5.19)$$

Then the contribution to the differential thrust cross section at order α_s from these four cuts is

$$\begin{aligned} \left[\frac{d\sigma}{d\tau} \right]_{\text{soft,R}} &= \sigma_0 \int \frac{d^3 p_g}{(2\pi)^3} \frac{1}{2\omega_g} |\mathcal{M}_{\text{soft}}|^2 \\ &\times \left[\delta \left(\tau - \frac{p_{\bar{q}} \cdot p_g}{Q^2} \right) \theta(\vec{p}_g \cdot \vec{p}_{\bar{q}}) + \delta \left(\tau - \frac{p_q \cdot p_g}{Q^2} \right) \theta(\vec{p}_g \cdot \vec{p}_q) \right]. \end{aligned} \quad (4.5.20)$$

In this expression, the θ -functions project onto the appropriate hemisphere defined by the

thrust axis (which aligns with the $\bar{q} - q$ axis at leading power). The first and third cuts in all the graphs, using soft interactions, give the virtual contributions. Summing all of them, the result is the same as the contribution to thrust from the thrust soft function [133, 137]:

$$\frac{1}{\sigma_0} \left[\frac{d\sigma}{d\tau} \right]_{\text{soft,R}} + \frac{1}{\sigma_0} \left[\frac{d\sigma}{d\tau} \right]_{\text{soft,V}} = \delta(\tau) \left[1 + C_F \frac{\alpha_s}{4\pi} \left(\frac{\pi^2}{3} \right) \right] - 16 C_F \frac{\alpha_s}{4\pi} \left[\frac{\ln \frac{\tau Q}{\mu}}{\tau} \right]_+ + \mathcal{O}(\alpha_s^2) \quad (4.5.21)$$

Although the real and virtual contributions are separately infrared divergent, the final contribution to the cross section is not.

Similarly, the contribution from collinear graphs gives the jet functions. The net contribution is

$$\frac{1}{\sigma_0} \left[\frac{d\sigma}{d\tau} \right]_{\text{coll}} = \delta(\tau) + C_F \frac{\alpha_s}{4\pi} \left\{ \delta(\tau) (7 - \pi^2) + \left[\frac{-3 + 4 \ln \frac{\tau Q^2}{\mu^2}}{\tau} \right]_+ \right\} + \mathcal{O}(\alpha_s^2) . \quad (4.5.22)$$

Multiplying these by the S_H -matrix element squared, the sum is

$$\frac{1}{\sigma_0} \left[\frac{d\sigma}{d\tau} \right]_{\text{soft+coll}} = \delta(\tau) + C_F \frac{\alpha_s}{2\pi} \left\{ \delta(\tau) \left(\frac{\pi^2}{3} - 1 \right) - 3 \left[\frac{1}{\tau} \right]_+ - 4 \left[\frac{\ln \tau}{\tau} \right]_+ \right\} . \quad (4.5.23)$$

This agrees with the exact NLO thrust distribution at leading power (see [16]). Note that the μ dependence of S_H -matrix elements exactly cancels against the μ dependence of the soft and collinear contributions in the asymptotic region.

For values of τ that are not small, one should necessarily treat the gluon as hard. The measurement function in this region is therefore only sensitive to hard particles. Since there are no asymptotic interactions between hard gluons and hard quarks, the S_H -matrix el-

ement in this regime is the same as in Eq. (4.5.17). Integrating the square of this matrix element against the thrust measurement function gives for $\tau > 0$,

$$\begin{aligned} \frac{1}{\sigma_0} \left[\frac{d\sigma}{d\tau} \right]_{3\text{-jet}} &= C_F \frac{\alpha_s}{2\pi} \left\{ 3(1+\tau)(3\tau-1) \right. \\ &\quad \left. + \frac{[4+6\tau(\tau-1)] \ln(1-2\tau)}{\tau(1-\tau)} - \frac{[4+6\tau(\tau-1)] \ln \tau}{\tau(1-\tau)} \right\}. \end{aligned} \quad (4.5.24)$$

Near $\tau = 0$ this contribution coming from $\widehat{\mathcal{M}}(\gamma^* \rightarrow \bar{q}qg)$ is singular, and the phase space integral is IR divergent. However, at $\tau = 0$ there is also the contribution from $\widehat{\mathcal{M}}(\gamma^* \rightarrow \bar{q}q)$. Although we can define the measurement function so that it is not sensitive to any gluon that couples in H_{as} , we cannot remove the soft and collinear gluons from H_{as} . These gluons still contribute to the cross section through loops, and affect the thrust distribution at $\tau = 0$. The virtual graphs are the first and third cuts in all the diagrams in Eq. (4.5.18). These graphs are IR divergent. If we work in $4 - 2\varepsilon$ dimensions, the full phase space integral over the 3-jet contribution $|\widehat{\mathcal{M}}(\gamma^* \rightarrow \bar{q}qg)|^2$ generates $\frac{1}{\epsilon_{\text{IR}}^2}$ and $\frac{1}{\epsilon_{\text{IR}}}$ poles that exactly cancel the $\frac{1}{\epsilon_{\text{IR}}^2}$ and $\frac{1}{\epsilon_{\text{IR}}}$ from the virtual graphs. The result is that

$$\begin{aligned} \frac{1}{\sigma_0} \left[\frac{d\sigma}{d\tau} \right]_{3\text{-jet}} + \frac{1}{\sigma_0} \left[\frac{d\sigma}{d\tau} \right]_{2\text{-jet}} &= \delta(\tau) + C_F \frac{\alpha_s}{2\pi} \left\{ \delta(\tau) \left(\frac{\pi^2}{3} - 1 \right) \right. \\ &\quad \left. + \left[3(1+\tau)(3\tau-1) + \frac{[4+6\tau(\tau-1)] \ln(1-2\tau)}{1-\tau} \right] \left[\frac{1}{\tau} \right]_+ - \frac{4+6\tau(\tau-1)}{1-\tau} \left[\frac{\ln \tau}{\tau} \right]_+ \right\}, \end{aligned} \quad (4.5.25)$$

which is the exact NLO thrust distribution in QCD.

So we see that S_H is capable of both reproducing distributions in fixed order QCD and,

through the asymptotic expansion, reproducing just the leading-power parts of those distributions. An advantage of leading-power approach is that one is not forced to compute the S_H -matrix elements and the asymptotic evolution to the same order in α_s . Instead, one can use exponentiation properties of the soft and collinear emission to evaluate the asymptotic evolution to all orders in perturbation theory. In particular, one can perform resummation with the renormalization group, since the soft and collinear contributions are each associated with only a single scale. Doing so in this example reproduces the resummed thrust distribution computed using SCET [133, 137].

4.6 $\mathcal{N} = 4$ Super Yang-Mills

To further illustrate the features of S_H , we now consider amplitudes in $\mathcal{N} = 4$ super Yang-Mills (SYM) theory. $\mathcal{N} = 4$ SYM is a superconformal $SU(N_c)$ gauge theory in which scattering amplitudes have been studied quite extensively. To leading order in $\frac{1}{N_c}$, the only Feynman diagrams that contribute have planar topology and each loop order gives an additional factor of the 't Hooft coupling $\lambda = g_s^2 N_c$. Since only one color structure is relevant at large N_c , is convenient to factor out the group theory factors. In addition, the amplitude for n -gluon scattering is totally symmetric in the permutation of the external legs. With these observations, it is conventional to write the L -loop amplitude with n external legs as

$$\mathcal{A}_n^{(L)} = g_s^{n-2} \left[(4\pi e^{-\gamma})^\epsilon \frac{g_s^2 N_c}{8\pi^2} \right]^L \sum_{\rho} \text{Tr} (T^{a_{\rho(1)}} \dots T^{a_{\rho(n)}}) A_n^{(L)}(\rho(1), \rho(2), \dots, \rho(n)), \quad (4.6.1)$$

where the sum is over non-cyclic permutations ρ of the external legs. The arguments $\rho(1)$, etc., refer to the permutation of the momenta and helicities of the legs. It is furthermore convenient to scale out the kinematic dependence of the tree-level amplitude by defining

$$M_n^{(L)}(\epsilon) \equiv \frac{A_n^{(L)}(\epsilon)}{A_n^{(0)}(\epsilon)}. \quad (4.6.2)$$

In addition, we will find it useful to discuss the terms of each order in ϵ separately, so we write

$$M_n^{(L)}(\epsilon) = \sum \epsilon^r M_n^{(L)}(\epsilon^r), \quad (4.6.3)$$

and decompose other quantities analogously.

In general, the bare n -leg L -loop amplitude is an extremely complicated function of the external momenta, even for planar maximal-helicity violating (MHV) amplitudes. What is interesting though is that there seems to be structure in the L -loop amplitude after the 1-loop amplitude is subtracted. More precisely, the ABDK/BDS ansatz proposes that the L -loop amplitude should be expressible in terms of the 1-loop amplitude and some transcendental constants [88, 89]. More precisely, the full matrix element with n legs has the form

$$\mathcal{M}_n^{\text{BDS}} = \exp \left[\sum_L \left((4\pi e^{-\gamma})^\epsilon \frac{g_s^2 N_c}{8\pi^2} \right)^L \left(f^{(L)}(\epsilon) M_n^{(1)}(L\epsilon) + C^{(L)} + E_n^{(L)}(\epsilon) \right) \right], \quad (4.6.4)$$

where $f^{(L)}(\epsilon)$ is independent of n and related to the cusp anomalous dimension (explicitly $f^{(1)}(\epsilon) = 1$ and $f^{(2)}(\epsilon) = -\zeta_2 - \zeta_3\epsilon - \zeta_4\epsilon^2 + \dots$). The numbers $C^{(L)}$ are also independent of n

and represent the part of the L -loop amplitude not given by the exponentiation of the first term. By explicit computation it is known that $C^{(1)} = 0$ and $C^{(2)} = -\frac{1}{2}\zeta_2^2$. Finally, $E_n^{(L)}(\epsilon)$ has only positive powers of ϵ , so that $E_n^{(L)}(0) = 0$.

It turns out the BDS ansatz was not quite correct: there is more structure to the amplitudes than just the numbers $C^{(L)}$ for $n > 5$. Thus, it is common to express amplitudes as ratios of the bare amplitudes and the BDS ansatz. More precisely, the remainder function is defined as

$$R_n = \ln \left[\frac{\mathcal{M}_n}{\mathcal{M}_n^{\text{BDS}}} \right], \quad (4.6.5)$$

and one can expand R_n order-by-order in g_s .

While the remainder functions have some nice properties, such as respecting dual conformal invariance, they violate other conditions, such as the Steinmann relations [92]. To preserve the Steinmann relations, the BDS ansatz is modified to the “BDS-like” ansatz [93]. For certain amplitudes ($n = 8$ for example), it has been shown that both the Steinmann relations and dual conformal invariance cannot be satisfied simultaneously [138]. That the BDS ansatz violates the Steinmann relations is due to the additional subtraction of finite, $\mathcal{O}(\epsilon^0)$, terms in Eq. (4.6.4) in addition to the IR divergences. A more conservative ansatz is the “minimally-normalized” amplitude $\mathcal{M}_n^{\text{min}}$ defined as [94]

$$\mathcal{M}_n^{\text{min}} = \exp \left[\sum_L \left((4\pi e^{-\gamma})^\epsilon \frac{g_s^2 N_c}{8\pi^2} \right)^L \left(f^{(L)}(\epsilon) M_n^{(1,\text{div})}(L\epsilon) + C^{(L)} \right) \right], \quad (4.6.6)$$

where the IR divergences of $M_n^{(1)}$ are

$$M_n^{(1,\text{div})}(\epsilon) = -\frac{1}{2\epsilon^2} \sum_{i=1}^n \left(\frac{\mu^2}{-s_{i,i+1}} \right)^\epsilon. \quad (4.6.7)$$

The ratio $\frac{\mathcal{M}_n}{\mathcal{M}_n^{\text{min}}}$ of the full amplitude to the minimally normalized amplitude is IR finite, just like the BDS remainder function in Eq. (4.6.5), but the finite parts of $\frac{\mathcal{M}_n}{\mathcal{M}_n^{\text{min}}}$ and $\frac{\mathcal{M}_n}{\mathcal{M}_n^{\text{BDS}}}$ are different.

In this section we relate some of these observations to the hard S -matrix element. We will see that the hard S -matrix element computed in $\overline{\text{MS}}$ corresponds closely to the minimally normalized amplitude.

4.6.1 4-point amplitude

We begin by discussing the MHV amplitude with 4 external legs. The IR divergences of the 1-loop amplitude for $n = 4$ are known to agree with the divergences of

$$C_4^{(1)}(\epsilon) = -\frac{e^{\gamma\epsilon}}{\Gamma(1-\epsilon)} \frac{1}{\epsilon^2} \left[\left(\frac{\mu^2}{-s} \right)^\epsilon + \left(\frac{\mu^2}{-t} \right)^\epsilon \right], \quad (4.6.8)$$

and the divergences of the 2-loop amplitude agree with the divergences of

$$C_4^{(2)}(\epsilon) = \frac{1}{2} \left(C_4^{(1)}(\epsilon) \right)^2 + C_4^{(1)}(\epsilon) \left(M_4^{(1)}(\epsilon) - C_4^{(1)}(\epsilon) \right) - (\zeta_2 + \epsilon\zeta_3) \frac{e^{-\epsilon\gamma}\Gamma(1-2\epsilon)}{\Gamma(1-\epsilon)} C_4^{(1)}(2\epsilon). \quad (4.6.9)$$

These formulas are due to Catani [139] (see also [140]). Note that $C_4^{(2)}(\epsilon)$ depends on the complete 1-loop amplitude $M_4^{(1)}(\epsilon)$. Thus, although the quantity $M_4^{(2)}(\epsilon) - C_4^{(2)}(\epsilon)$ is IR-

finite, more terms are being subtracted this way than those determined by the universality of IR-divergences. These extra terms depend on quantities such as $M_4^{(1)}(\epsilon^2)$ which are not fixed by factorization alone. Although factorization does not determine $M_4^{(1)}(\epsilon)$, its appearance in the universal formula can be understood from the point of view of effective field theory [141]: it comes from a cross term between the non-universal 1-loop Wilson coefficient and the universal 1-loop divergences. An equivalent mechanism explains its appearance during the computation of S_H , as we now show.

With 4 legs ($n = 4$), the 1-loop amplitude is

$$M_4^1(\epsilon) = -\frac{2}{\epsilon^2} + \frac{1}{\epsilon} M_4^{(1)}(\epsilon^{-1}) + M_4^{(1)}(\epsilon^0) + \mathcal{O}(\epsilon), \quad (4.6.10)$$

where

$$M_4^{(1)}(\epsilon^{-2}) = -2, \quad (4.6.11)$$

$$M_4^{(1)}(\epsilon^{-1}) = -\ln \frac{\mu^2}{-s} - \ln \frac{\mu^2}{-t}, \quad (4.6.12)$$

$$M_4^{(1)}(\epsilon^0) = -\ln \frac{\mu^2}{-t} \ln \frac{\mu^2}{-s} + \frac{2\pi^2}{3}, \quad (4.6.13)$$

$$M_4^{(1)}(\epsilon^1) = -\frac{\pi^2}{2} \ln \frac{-s}{u} - \frac{1}{3} \ln^3 \frac{-s}{u} + \frac{\pi^2}{12} \ln \frac{\mu^2}{-s} - \frac{1}{6} \ln^3 \frac{\mu^2}{-s} + \frac{\pi^2}{4} \ln \frac{\mu^2}{u}, \quad (4.6.14)$$

$$+ \frac{1}{2} \ln^2 \frac{-s}{u} \ln \frac{\mu^2}{u} - \frac{1}{2} \ln \frac{-s}{u} \ln \frac{-t}{u} \ln \frac{\mu^2}{u} - \ln \frac{-s}{u} \text{Li}_2 \frac{-s}{u} + \text{Li}_3 \frac{-s}{u} + \frac{7}{3} \zeta_3 + (s \leftrightarrow t),$$

$$M_4^{(1)}(\epsilon^2) = \frac{5\pi^2}{24} \ln^2 \frac{-s}{u} + \frac{1}{8} \ln^4 \frac{-s}{u} + \frac{3}{8} \ln \frac{-s}{u} \ln \frac{-t}{u} + \frac{1}{6} \ln^3 \frac{-s}{u} \ln \frac{-t}{u} \quad (4.6.15)$$

$$- \frac{1}{4} \ln^2 \frac{-s}{u} \ln^2 \frac{-t}{u} + \frac{\pi^2}{24} \ln^2 \frac{\mu^2}{-s} - \frac{1}{24} \ln^4 \frac{\mu^2}{s} - \frac{\pi^2}{2} \ln \frac{-s}{u} \ln \frac{\mu^2}{u}$$

$$- \frac{1}{3} \ln^3 \frac{-s}{u} \ln \frac{\mu^2}{u} + \frac{\pi^2}{8} \ln^2 \frac{\mu^2}{u} + \frac{1}{4} \ln^2 \frac{-s}{u} \ln^2 \frac{\mu^2}{u} - \frac{1}{4} \ln \frac{-s}{u} \ln \frac{-t}{u} \ln^2 \frac{\mu^2}{u}$$

$$+ \frac{7}{3} \zeta_3 \ln^2 \frac{\mu^2}{-s} + \frac{1}{2} \ln^2 \frac{-s}{u} \text{Li}_2 \frac{-s}{u} - \ln \frac{-s}{u} \ln \frac{\mu^2}{u} \text{Li}_2 \frac{-s}{u} + \ln \frac{\mu^2}{u} \text{Li}_3 \frac{-s}{u}$$

$$- \ln \frac{-s}{u} \text{Li}_3 \frac{-t}{u} - \text{Li}_4 \frac{-s}{u} + \frac{49\pi^4}{720} + (s \leftrightarrow t).$$

In these expressions, $s = (p_1 + p_2)^2$, $t = (p_1 + p_3)^2$, $u = -t - s$ and the convention is that incoming momenta are treated as outgoing with negative energy. Note that the ϵ are all ϵ_{IR} since $\mathcal{N} = 4$ SYM is UV finite.

At 2-loops, the amplitude can be written as

$$M_4^{(2)} = \frac{2}{\epsilon^4} - \frac{2}{\epsilon^3} M_4^{(1)}(\epsilon^{-1}) + \frac{1}{\epsilon^2} \left[\frac{\pi^2}{12} + \frac{1}{2} M_4^{(1)}(\epsilon^{-1})^2 - 2M_4^{(1)}(\epsilon^0) \right]$$

$$+ \frac{1}{\epsilon} \left[-\frac{\pi^2}{12} M_4^{(1)}(\epsilon^{-1}) + M_4^{(1)}(\epsilon^{-1}) M_4^{(1)}(\epsilon^0) - 2M_4^{(1)}(\epsilon^1) + \frac{\zeta_3}{2} \right] + M_4^{(2)}(\epsilon^0) + \mathcal{O}(\epsilon), \quad (4.6.16)$$

where

$$M_4^{(2)}(\epsilon^0) = \frac{1}{2} \left[M_4^{(1)}(\epsilon^0) \right]^2 - \frac{\pi^2}{6} M_4^{(1)}(\epsilon^0) - \frac{\pi^4}{120} \\ + M_4^{(1)}(\epsilon^{-1}) \left[M_4^{(1)}(\epsilon^1) - \frac{\zeta_3}{2} \right] + M_4^{(1)}(\epsilon^{-2}) M_4^{(1)}(\epsilon^2). \quad (4.6.17)$$

Although there is some hint of exponentiation in this expression, it is not particularly simple. That is, if one defines an IR finite 2-loop amplitude by dropping all the singular terms in ϵ and then taking $\epsilon \rightarrow 0$ the result, $M_4^{(2)}(\epsilon^0)$, is complicated, with all the polylogarithms from Eqs. (4.6.14) and (4.6.15).

The appearance of the $\mathcal{O}(\epsilon^1)$ and $\mathcal{O}(\epsilon^2)$ terms from $\mathcal{M}_4^{(1)}$ in the 2-loop amplitude hints at a relationship between them. Indeed, the BDS/ABDK ansatz notes that if we subtract $C_4^{(2)}(\epsilon^0)$ in Eq. (4.6.9) from $M_4^{(2)}(\epsilon^0)$ the result is relatively simple,

$$M_4^{(2)}(\epsilon^0) - C_4^{(2)}(\epsilon^0) = \frac{1}{2} \left(M_4^{(1)}(\epsilon^0) - C_4^{(1)}(\epsilon^0) \right)^2 - \zeta_2 \left(M_4^{(1)}(\epsilon^0) - C_4^{(1)}(\epsilon^0) \right) - \frac{21}{8} \zeta_4. \quad (4.6.18)$$

Recall that $C_4^{(2)}(\epsilon)$ is not fixed by the IR structure alone, but includes additional terms.

Although this relation works well for the 4-point amplitude, it is somewhat ad hoc and requires modification for $n > 5$ legs and higher loops.

Now let us consider the hard S -matrix elements. We define them analogously to S -matrix elements, adding a hat. So $\widehat{A}_n^{(L)}$ is the color-stripped hard- S -matrix element for n legs at L loops. This amplitude is IR finite, but UV divergent. Denoting $\widehat{M}_n^{(L)}(\epsilon) \equiv \widehat{A}_n^{(L)}(\epsilon)/\widehat{A}_n^{(0)}(\epsilon)$ in analogy to Eq. (4.6.2), the 1-loop bare hard matrix element is the same

as Eq. (4.6.10) with ϵ_{IR} replaced by ϵ_{UV} . The renormalized matrix element is then

$$\widehat{M}_4^{(1)} = \left[\frac{1}{Z_4} (\widehat{\mathcal{M}}_4)_{\text{bare}} \right]_{\text{color-stripped}}^{\text{1-loop}} = -\ln \frac{\mu^2}{-t} \ln \frac{\mu^2}{-s} + \frac{2\pi^2}{3} + \mathcal{O}(\epsilon), \quad (4.6.19)$$

where, with minimal subtraction ($\overline{\text{MS}}$),

$$Z_4^{\overline{\text{MS}}} = 1 + (4\pi e^{-\gamma})^\epsilon \frac{g_s^2 N_c}{8\pi^2} \left[-\frac{2}{\epsilon^2} + \frac{1}{\epsilon} \left(-\ln \frac{\mu^2}{-s} - \ln \frac{\mu^2}{-t} \right) \right] + \mathcal{O}(g_s^4 N_c^2). \quad (\overline{\text{MS}}) \quad (4.6.20)$$

Note that $\widehat{M}_4^{(1)}$ is finite as $\epsilon \rightarrow 0$, since the IR divergences are absent in hard S -matrix elements and the UV divergences are removed through renormalization. There are nevertheless terms of $\mathcal{O}(\epsilon)$ and $\mathcal{O}(\epsilon^2)$ in the matrix elements in d dimensions. These terms are the same as the $\mathcal{O}(\epsilon)$ and $\mathcal{O}(\epsilon^2)$ terms in $M_4^{(1)}$. Then the 2-loop hard S -matrix element gets a contribution from both the 2-loop graphs, giving $M_4^{(2)}(\epsilon^0)$ after renormalization, as well as a contribution from the cross terms between the $\frac{1}{\epsilon^2}$ and $\frac{1}{\epsilon}$ terms in Z_4 and the $\mathcal{O}(\epsilon)$ and $\mathcal{O}(\epsilon^2)$ terms in $\widehat{M}_4^{(1)}$. The result is that

$$\widehat{M}_4^{(2)} = \frac{1}{2} \left[\widehat{M}_4^{(1)} - \frac{\pi^2}{6} \right]^2 - \frac{\pi^4}{45} + \frac{\zeta_3}{2} \left(\ln \frac{\mu^2}{-s} + \ln \frac{\mu^2}{-t} \right). \quad (\overline{\text{MS}}) \quad (4.6.21)$$

This matrix element is significantly simpler than $M_4^{(2)}(\epsilon^0)$ in Eq. (4.6.17), and does not require any ad-hoc subtractions.

4.6.2 Scheme choice

Dimensional regularization and minimal subtraction is the most widespread scheme in use in SCET. We must keep in mind, however, that due to renormalization there is scheme dependence in S_H . This is not a problem *per se*, since S_H itself is not directly observable.

One expects that one S_H -matrix elements are combined into observables the scheme dependence will cancel. Indeed, the cancellations that occur will be similar to the cancellations that occur in SCET. For example, Ref. [142] showed that physical observables agree when conventional dimensional regularization, four-dimensional helicity scheme, or dimensional reduction are used, despite the fact that the hard, jet and soft functions are different in the different schemes. In a normal, local field theory, the counterterms are strongly constrained: they must just be numbers. In SCET the counterterms can depend on the labels for the various collinear directions which translates to dependence of hard kinematical quantities, like s and t , as in Eq. (4.6.20). However, one cannot choose an arbitrary function of labels, as the dependence must be canceled by contributions from soft and jet functions. Roughly speaking the combination, $H \otimes J \otimes S$ must be scheme independent, where the hard function H corresponds to the square of our hard S -matrix elements. More discussion of these constraints can be found in [142].

Let us suppose that adding a finite part to the counterterm is not problematic. More precisely, suppose we can add a finite part $\delta_4(\epsilon)$ to the Z_4 renormalization constant. Then

the color-stripped hard S -matrix element at 1-loop shifts from the $\overline{\text{MS}}$ version by $\delta_4^{(1)}(\epsilon)$:

$$\widehat{M}_4^{\delta,(1)} = \widehat{M}_4^{(1)} - \delta_4^{(1)}(\epsilon) \quad (4.6.22)$$

At 2-loops, the shift picks up a cross term between $\delta_4^{(1)}$ and the divergent parts of the bare amplitude $(\widehat{\mathcal{M}}_4)_{\text{bare}}$:

$$\widehat{M}_4^{\delta,(2)} = \widehat{M}_4^{(2)} - \sum_{j=0}^2 \widehat{M}_4^{\text{bare},(1)}(\epsilon^{-j}) \delta_4^{(1)}(\epsilon^j) - \delta_4^{(2)}, \quad (4.6.23)$$

so that

$$\begin{aligned} \widehat{M}_4^{\delta,(2)} = \frac{1}{2} \left[\widehat{M}_4^{\delta,(1)} - \frac{\pi^2}{6} \right]^2 - \frac{\pi^4}{45} - \delta_4^{(1)}(\epsilon^1) \widehat{M}_4^{\text{bare},(1)}(\epsilon^{-1}) - \delta_4^{(1)}(\epsilon^2) \widehat{M}_4^{\text{bare},(1)}(\epsilon^{-2}) \\ - \frac{\pi^2}{6} \delta_4^{(1)}(\epsilon^0) - \frac{1}{2} [\delta_4^{(1)}(\epsilon^0)]^2 + \frac{\zeta_3}{2} \left(\ln \frac{\mu^2}{-s} + \ln \frac{\mu^2}{-t} \right) - \delta_4^{(2)}. \end{aligned} \quad (4.6.24)$$

This motivates choosing a “BDS” subtraction scheme, where

$$\delta_4^{(1)} = -\frac{\pi^2}{6} - \frac{\zeta_3}{2}\epsilon, \quad \delta_4^{(2)} = -\frac{\pi^4}{120} + \mathcal{O}(\epsilon) \quad (4.6.25)$$

or equivalently

$$Z_4^{\text{BDS}} = 1 + (4\pi e^{-\gamma})^\epsilon \frac{g^2 N_c}{8\pi^2} \left[-\frac{2}{\epsilon^2} - \frac{1}{\epsilon} \left(\ln \frac{\mu^2}{-s} + \ln \frac{\mu^2}{-t} \right) - \frac{\pi^2}{6} - \frac{\zeta_3}{2} \epsilon \right] + \mathcal{O}(g^4). \quad (\text{BDS scheme}) \quad (4.6.26)$$

Then we get simply

$$\widehat{M}_4^{\text{BDS},(2)} = \frac{1}{2} \left[\widehat{M}_4^{\text{BDS},(1)} - \frac{\pi^2}{6} \right]^2. \quad (\text{BDS scheme}) \quad (4.6.27)$$

There are two things to note about this result. First, it is nontrivial that one can pick pure numbers for δ_4 to cancel the explicit s and t dependence in Eq. (4.6.21). This was possible only because the $\ln \frac{\mu^2}{-s} + \ln \frac{\mu^2}{-t}$ factor in Eq. (4.6.21) is the same as in $\widehat{M}^{(1)}(\epsilon^{-1})$. Second it is impossible to choose δ_4 to remove the $\frac{\pi^2}{6}$ in Eq. (4.6.27). Thus there is a sense in which the constant term $\frac{\pi^2}{6} = \zeta_2$ of the second order amplitude is scheme independent. This term gives the constant $C_2 = \frac{1}{2}\zeta_2^2$ from Eq. (4.6.4).

The BDS ansatz implies that to all orders, the 4-gluon planar amplitude exponentiates in the BDS subtraction scheme. In the language of the hard S -matrix, this means that the finite parts of the counterterms will be pure numbers to all orders. Indeed, for dual-conformal invariance to be respected by the 4-point amplitude, we should not be adding extra dependence on s and t into the counterterms. Equivalently, we can say that the dual-conformal anomaly is manifest in the BDS subtraction scheme but somewhat obscure in $\overline{\text{MS}}$.

4.6.3 6-point amplitude

The amplitude with 6 external particles is more interesting because it can depend on more kinematic invariants. The hard MHV S -matrix element with 6 legs in $\overline{\text{MS}}$ is

$$\widehat{M}_6^{(1)}(\epsilon) = \sum_{\text{cycles}} \left[-\frac{1}{2} \ln^2(-s_{12}) - \ln \frac{-s_{12}}{-s_{123}} \ln \frac{-s_{23}}{-s_{123}} + \frac{1}{4} \ln^2 \frac{-s_{123}}{-s_{234}} \right] \\ - \text{Li}_2(1-u) - \text{Li}_2(1-v) - \text{Li}_2(1-w) + 6\zeta_2 + \mathcal{O}(\epsilon), \quad (\overline{\text{MS}}) \quad (4.6.28)$$

where the 3 dual-conformal cross ratios are

$$u = \frac{s_{12}s_{45}}{s_{123}s_{345}}, \quad v = \frac{s_{23}s_{56}}{s_{234}s_{123}}, \quad w = \frac{s_{34}s_{61}}{s_{345}s_{234}}. \quad (4.6.29)$$

The notation here is that $s_{123} = (p_1 + p_2 + p_3)^2$ and sum over cycles means sum over the 6 rotations of the labels, e.g. $s_{123} \rightarrow s_{234}$ and so on. This amplitude is simply the bare 1-loop MHV amplitude [143, 144] with IR divergences converted to UV divergences by the diagrams involving H_{as} and then removed by counterterms:

$$Z_6 = 1 + (4\pi e^{-\gamma})^\epsilon \frac{g^2 N_c}{8\pi^2} \left[-\frac{2}{\epsilon^2} - \frac{1}{\epsilon} \sum_{\text{cycles}} \left(\ln \frac{\mu^2}{-s_{12}} \right) \right] + \mathcal{O}(g_s^4) \quad (\overline{\text{MS}}) \quad (4.6.30)$$

The “BDS-like” ansatz adds to this amplitude the terms on the second line plus another cyclic sum,

$$Y_6 = \text{Li}_2(1-u) + \text{Li}_2(1-v) + \text{Li}_2(1-w) + \frac{1}{2} (\ln^2 u + \ln^2 v + \ln^2 w). \quad (4.6.31)$$

If we are free to shift the counterterm, $Z_6^{\text{BDS-like}} = Z_6 - Y_6$, then the matrix element has a somewhat simpler form:

$$\widehat{M}_6^{(1)}(\epsilon) = \sum_{\text{cycles}} \left[-\ln(-s_{12}) \ln(-s_{23}) + \frac{1}{2} \ln(-s_{12}) \ln(-s_{45}) \right] + 6\zeta_2 \quad (\text{BDS-like scheme}) \quad (4.6.32)$$

In particular, it is a function of only 2-particle invariants. This means that when the amplitude is exponentiated, it cannot violate the Steinmann relations (these require 3 particle invariants) [145, 146].

Note however, that we do not know how to specify this BDS-like subtraction scheme at higher order. More importantly, we do not know if it is consistent. As mentioned above, (see [142]) there are constraints on the scheme from self-consistency of SCET. Since we also do not know general constraints on the finite parts of the counterterms, it is safest to restrict to conventional dimensional regularization with minimal subtraction, where SCET at least is believed to be consistent. In $\overline{\text{MS}}$, the counterterm is in Eq. (4.6.30) and the hard matrix element is in Eq. (4.6.28). In $\overline{\text{MS}}$, the hard matrix elements agree with the minimally-normalized amplitudes discussed in [94] up to at least 2-loops and preserve the Steinmann relations.

4.7 Summary and Outlook

The traditional S -matrix is only well defined if time evolution of a theory is well approximated by free evolution at early or late times. Indeed, the free Hamiltonian H_0 is part of the definition of S used for perturbative calculations. When a theory has massless particles,

the interactions do not die off fast enough at asymptotic times, resulting in a poorly defined, divergent S -matrix. We argue that a sensible, finite S -matrix is obtained by replacing H_0 in its definition with an asymptotic Hamiltonian H_{as} that correctly accounts for all the asymptotic interactions. Our key principle for choosing H_{as} is that the states should evolve before and after they scatter independently of how they scatter. That such an H_{as} exists and makes the S -matrix finite is guaranteed by theorems of hard-collinear-soft factorization. Capitalizing on these theorems, we define H_{as} as the leading power expansion of the full Hamiltonian in soft and collinear limits, and call the corresponding S -matrix the hard S -matrix, S_H . S_H is finite order-by-order in perturbation theory, as we have verified through a number of explicit examples in QED, QCD and $\mathcal{N} = 4$ super Yang-Mills theory.

While the traditional S -matrix is IR divergent, it can still be used to compute IR-finite observables. This is done by summing over a broad enough set of processes so that the sum is finite even though individual contributions are divergent. With S_H , the same physical predictions result using the matrix elements of a scattering operator that are finite process-by-process.

We presented a method and Feynman rules for the perturbative calculation of S_H -matrix elements. The method involves separating S_H into three parts: An asymptotic part evolving the state from $t = 0$ to $t = -\infty$, the evolution from $t = -\infty$ to $t = \infty$ and an asymptotic part evolving from $t = \infty$ to $t = 0$. Each asymptotic part is calculated using Feynman rules similar to those in time-ordered perturbation theory but without overall energy conservation, and the middle part consists of conventional Feynman diagrams. The three part picture is presented for calculational convenience, since it breaks up calculations into es-

essentially usual time-ordered perturbation theory and Feynman diagrams, and bypasses the need to derive a new interaction picture with modified propagators.

The hard S -matrix has numerous advantages over the traditional S -matrix. The first advantage is the obvious one: S_H exists. Second, matrix elements of S_H have a rich structure with diverse interpretations. One can interpret the asymptotic evolution as dressing the states, so that a initial Fock state with a finite number of particles evolves into a dressed state with an infinite number of particles at asymptotic times. This connects our construction to previous work on coherent states, such as by Chung [8] or Faddeev and Kulish [10]. Alternatively, S_H -matrix elements can be interpreted as Wilson coefficients in Soft-Collinear Effective Theory. Finally, S_H -matrix elements are closely related to finite remainder functions studied in the amplitude community. Indeed, much of the progress in understanding scattering amplitudes over the last few decades has comprised results about an object, the S -matrix, that formally does not exist. Since there is so much interest in the S -matrix itself (as opposed to cross sections), it is logical to try to put this object on a firmer theoretical footing. Doing so was one of the main motivations of this paper.

There are a number of new ideas contained in this paper. These include:

- The first explicit calculation of a finite S -matrix in theories with massless particles.

While other authors have introduced similar concepts in QED, there are no explicit calculations in the literature of actual matrix elements. The majority of papers focuses on just the IR divergence cancellation. Issues such as regulator dependence, renormalization, subtraction schemes, phase space integrals, computation of observables, completeness of the Hilbert space, etc., are all glossed over unless one is able to

do explicit computations.

- We present a new rationale for choosing the asymptotic Hamiltonian. While others have argued that the asymptotic Hamiltonian should make the S -matrix IR finite, we argue that such a criterion is not restrictive enough: one could choose $H_{\text{as}} = H$ to satisfy that requirement. Instead we argue that one should use that the asymptotic evolution is independent of the hard scattering. That there exists an asymptotic Hamiltonian with this property in gauge theories is non-trivial and follows from factorization theorems.
- We connect the literature on coherent states to that of factorization and that of scattering amplitudes. In particular, the hard S -matrix elements can be identified as S -matrix elements of coherent states, as Wilson coefficients in SCET, and as finite remainder functions in $\mathcal{N} = 4$ SYM fields corresponding to BDS-inspired subtraction schemes.
- We provide an explicit set of Feynman rules to evaluate S_H elements in perturbation theory. These rules involve distributions and products of distributions that must be handled with some care.
- We provide a number of examples of S_H -matrix element calculations, both using pure dimensional regularization and with explicit cutoffs on H_{as} .
- We examine how the Glauber/Coulomb phase arises in asymptotic-region diagrams. In particular, energy non-conservation in the asymptotic regions allows the Glauber contribution to be reproduced (and cancelled) without off-shell modes.

- We demonstrate that infrared-safe observables computed with S_H will agree with those computed using the normal S -matrix, and, to leading power, with those computed using SCET or other factorization frameworks. We are not aware of any paper on dressed states that makes a physical prediction using them. In our framework, one can see how the dressing occurs, but also how the states get “undressed” in the final asymptotic evolution before the measurement is made.
- Although predictions using S_H reduce (almost trivially) to predictions using S , matrix elements of S_H can be studied as interesting objects on their own. These matrix elements are scheme and scale-dependent, but still have physical interpretations, just like the $\overline{\text{MS}}$ couplings $\alpha_s(\mu)$.

These last two bullets are perhaps worth some additional discussion. The incontrovertible truth is that cross sections computed with S , despite coming from IR-divergent amplitudes, are in perfect agreement with observations. Thus, no matter how one attempts to make scattering amplitudes finite, the framework must reproduce these cross sections exactly. In other words, it is foolhardy to try to make different predictions at the cross section level with a new S -matrix. That being said, there are situations, in particular those with charged initial states such as $e^+e^- + \text{photons} \rightarrow Z + \text{photons}$, where it is not entirely clear what the physical cross section is supposed to be [70]. In such situations, a finite S_H may provide some clarity.

Although we cannot expect S_H to revolutionize the computation of physical cross sections, having a finite S -matrix is still enormously beneficial for the study of scattering am-

plitudes themselves. Indeed, the majority of research of scattering amplitudes focuses on S -matrix elements themselves, not on observables. So it is this community that might benefit first from S_H . As an example, we showed that certain S_H elements in a supersymmetric theory naturally satisfy the Steinmann relations, at least to two loops. In contrast, S -matrix elements are IR divergent and, depending on how the IR divergences are subtracted, the Steinmann relations may or may not be satisfied. More broadly, because S_H corresponds to the matrix elements of a single unitary operator, rather than a ratio of such matrix elements, it should automatically satisfy any constraints that follow from unitarity. One might also imagine that properties stemming from analyticity would be more transparent in matrix elements of a single operator rather than a ratio.

Finally, let us briefly discuss how to think about S_H non-perturbatively. In this paper, we have advocated for computing S_H in dimensional regularization with $\overline{\text{MS}}$ subtraction. At each order in perturbation theory, one can compute S_H elements this way. It may seem counterintuitive, but perturbation theory has historically been the best way to orient investigation into non-perturbative physics, and a perturbative approach could be similarly successful for S_H . One can also resum S_H using renormalization group techniques to examine its all-orders behavior. Alternatively, one could (in principle) compute S_H numerically with hard cutoffs, but to compare to the perturbative results in dimensional regularization, one would have to convert between the cutoff scheme and $\overline{\text{MS}}$. Through various approaches like these, it should be possible to explore the analytic structure of S_H . It would be interesting to look at its properties in the Borel plane, for example, or whether a renormalon-free mass scheme naturally emerges. More generally, since S_H is IR finite, it resembles more closely

S in a theory with a mass gap than the IR-divergent S . Thus one might hope that when massless particles are present, the S -matrix bootstrap program might make more progress with S_H than it has on S .

5

Sequential Discontinuities of Feynman Integrals and the Monodromy Group

5.1 Introduction

Feynman integrals – integrals over Feynman propagators appearing in perturbative quantum field theory calculations – are primarily useful for making observable predictions about particle physics experiments. Famously, they have been used to make some of the most precise predictions in the history of science [147]. However, these integrals have also increasingly become recognized as interesting mathematical objects in their own right, exhibiting a variety of geometric, analytic, and number-theoretic properties.

One of the aspects of Feynman integrals that has become better understood in recent years is the class of transcendental functions they evaluate to in integer dimensions. In particular, at low loop order and low particle multiplicity, they can often be expressed in terms of generalized polylogarithms [148–150]. These functions are under good theoretical and numerical control, due in part to the symbol and coaction [91, 151–154], which provide a systematic way to understand their analytic structure and to exploit identities among them. In particular, arbitrarily complicated polylogarithms can be broken down into simpler building blocks such as logarithms and Riemann zeta values, at the cost of losing only integration boundary data.

Knowing the analytic structure of polylogarithms has proven especially useful in the computation of Feynman integrals and scattering amplitudes, as the branch cut structure of these quantities is constrained by physical principles such as locality and causality. For example, in the Euclidean region where all Mandelstam invariants are negative, Feynman integrals can only have logarithmic branch points at the vanishing loci of sums of external

momenta. This places strong constraints on the symbol and coaction of the polylogarithms these integrals produce.¹

That Feynman integrals have branch cut singularities has been known since the early days of quantum field theory. In a seminal paper by Landau [45], these branch cuts were shown to be associated with regions of external momenta where the poles in Feynman propagators coalesce around the integration contour, so that the contour is pinched between the singularities (see also [162, 163]). Cutkosky subsequently gave a general formula relating the discontinuity across these branch cuts to “cut graphs” in which some Feynman propagators are replaced by delta functions [6]. ’t Hooft and Veltman later gave a simple diagrammatic derivation of Cutkosky’s cutting rules [54, 164]. However, these works are mostly confined to the study of a *single* discontinuity of Feynman integrals.

In this paper, we are interested in studying the cutting rules for discontinuities of discontinuities: is there a way to compute sequential discontinuities of Feynman integrals with cut diagrams, as there is for single discontinuities? Cutkosky and his contemporaries touched on this topic, but computing sequential discontinuities is significantly more complicated than computing a single discontinuity. Usually cuts are computed as differences $F(s + i\epsilon) - F(s - i\epsilon) = F(s + i\epsilon) - F(s + i\epsilon)^*$, which is appropriate in the physical region where the cut runs along the positive kinematic invariants and the amplitude satisfies a reality condition $F(s - i\epsilon) = F(s + i\epsilon)^*$. However, as can be observed in explicit examples, when taking cuts of cuts the situation is more complicated. In particular, one generically

¹This constraint on the symbol can also be extended to Feynman integrals that evaluate to elliptic polylogarithms [155–157]; however, no coaction has been worked out for the types of worse-than-elliptic integrals that appear in Feynman integrals in integer dimensions (see for instance [158–161]).

encounters complex branch points, and the same reality condition does not hold (in fact, this reality condition clashes with holomorphy).

Some progress on the study of sequential discontinuities was made in [165], where a formula relating sequential discontinuities in different channels to a sum over cuts was conjectured. Drawing inspiration from this work, we make use of time-ordered perturbation theory (TOPT) to derive more general relations between the sequential discontinuities of Feynman integrals and cut integrals. In particular, our method clarifies the role of the $\pm i\varepsilon$ prescription in cut integrals, and emphasizes the importance of considering monodromies around branch points rather than discontinuities across branch cuts. In this approach, we can analytically continue from the physical region, along a path that goes around a branch point and returns to the physical region. Now the function can be thought as being evaluated on a different sheet and the discontinuity is the difference between the initial value and the final value.² While this monodromy-based approach does not appear to be widely used, it goes back at least as far as [166]. Our approach was also influenced more by recent mathematical literature on polylogarithms, where the general theory simplifies in several ways. One of the simplifications is that the monodromy group is represented by numerical matrices whose entries are integer (or rational) multiples of powers of $i\pi$. Using these methods, we are able to compute discontinuities in any sequence of channels, including discontinuities in the same channel.

Sequential cuts of Feynman integrals can also be computed using the multivariate residue calculus of Leray [167]. This has been worked out explicitly at one loop [168]. While this

²Note that in this approach the choice for where to place the branch cut is not important.

approach is both general and mathematically rigorous, it quickly becomes computationally onerous. More Hodge-theoretic approaches have also been developed in [169, 170]; we comment on the difference between these approaches and the methods of the present paper at the end of Section 5.4.2. In this paper, our main goal was to come up with a prescription for computing sequential discontinuities that was more computationally tractable than existing approaches.

One set of constraints on sequential discontinuities are the Steinmann relations. As originally studied by Steinmann [7], these relations follow from causality and express linear relations between vacuum expectation values of certain types of operator products called R -products (as defined in [171]). Steinmann originally studied these relations for the case of four local gauge-invariant operators; they were subsequently generalized to higher multiplicity [172–175]. Later, it was shown that the Steinmann relations imply scattering amplitudes cannot have double discontinuities in partially overlapping momentum channels [176]. The Steinmann relations have also been studied directly from the point of view of S -matrix theory, without reference to local fields and their commutators; for a review, see [177].

Steinmann-type constraints have proven extremely useful for the modern amplitude bootstrap program, which attempts to determine the functional forms of Feynman integrals or scattering amplitudes from their general properties (such as symmetries, analytic properties, and factorization in certain kinematic limits). So far, these methods have been mostly applied to processes in planar $\mathcal{N} = 4$ supersymmetric Yang-Mills theory [92, 144, 146, 178–186], where there is rich theoretical data available and integrability-based computations provide crucial consistency checks [187–190]. However, similar analytic constraints and

bootstrap techniques are expected to extend to non-supersymmetric quantities as well (see for instance [191–193]).

Scattering amplitudes in Yang-Mills theories necessarily involve massless particles, so the Steinmann relations, originally derived in field theories with a mass gap, do not necessarily apply. Indeed, massless particles engender infrared divergences in these theories. In planar $\mathcal{N} = 4$, instead of studying the amplitude itself, one typically studies finite Feynman integrals (see for instance [186, 194–198]) or remainder functions, defined as ratios of amplitudes or ratios of amplitudes to the exponentiation of lower-order amplitudes. It is to these types of remainder functions that Steinmann-type constraints are often applied [92, 94, 138, 184].³ While there has been some progress in systematically extracting the infrared-finite content of the S -matrix (for example, through the construction of an infrared-finite S -matrix [95, 200]), there remains some uncertainty over how and when constraints like Steinmann relations should hold. One goal of this paper is to pry away some of the strong assumptions used in the axiomatic field theory approach. Thus, rather than full scattering amplitudes in mass-gapped theories, we study Feynman integrals directly.

More broadly, in this paper we set out to provide some clarity on how to think about and compute sequential discontinuities of Feynman integrals, and to study the types of constraints these sequential discontinuities satisfy. We treat this problem both at the level of cut integrals and at the level of polylogarithmic functions. In particular, we make use of time-ordered perturbation theory (TOPT) to prove new relations between the sequential

³The Steinmann relations were first used to analyze these amplitudes in the multi-Regge limit, where it was also pointed out that normalizing by the BDS ansatz did not preserve these relations [199].

discontinuities of Feynman integrals and their cuts. We also describe how these discontinuities can be computed systematically from polylogarithmic representations of these integrals with the use of variation matrices and the monodromy group, both of which we describe in some detail.

The main practical results of this paper take the form of relations between discontinuities of Feynman integrals and cuts of those integrals. For example, we show that the m^{th} discontinuity of the Feynman integral \mathcal{M} in a momentum channel corresponding to the Mandelstam invariant s satisfies the relation

$$[\text{Disc}_s^m \mathcal{M}]_{R^s} = m! \sum_{k=m}^{\infty} \begin{Bmatrix} k \\ m \end{Bmatrix} (-1)^{m-k} [\mathcal{M}_{k\text{-cuts}}]_{R_+^s}, \quad (5.1.1)$$

where $\{k_m\} = \frac{1}{m!} \sum_{\ell=1}^m (-1)^{m-\ell} \binom{m}{\ell} \ell^k$ are the Stirling numbers of the second kind. On the left side of the equation, we compute m discontinuities in the s channel by taking m monodromies around a branch point in s . We write this as

$$\text{Disc}_s^m \mathcal{M} = (\mathbb{1} - \mathcal{M}_{\text{cut}_s})^m \mathcal{M} \quad (5.1.2)$$

These monodromies are taken by analytically continuing along a closed contour that goes between the region R^s , which we define to be the region in which all Mandelstam invariants are real and negative, except for s which is real and positive, and the Euclidean region R^\star , where all invariants are negative. On the right-hand side, $\mathcal{M}_{k\text{-cuts}}$ denotes the sum over all ways of cutting the Feynman integral k times, with positive energy flowing across all cuts.

These cuts must be computed in the region R_+^s , where the $+$ subscript indicates that all the Feynman propagators in these cut diagrams should be assigned $+i\varepsilon$. A careful treatment of the $\pm i\varepsilon$ in the cut diagrams is essential to have a sensible (and correct) formula relating discontinuities and cuts. Eq. (5.1.1) is derived in Section 5.5. We also derive similar relations between cuts and discontinuities in different channels.

One thing that our analysis makes clear is that sequential discontinuities can only be nonzero when there exists at least one TOPT diagram that depends on the energies corresponding to each cut momentum channel. When one of these energies is not present, the cut in this channel vanishes. Since the energies that appear in TOPT diagrams always take the form of sums of external energies $\sum_{i \in J} E_i$, where the sets of summed-over external particle indices J that appear in a given diagram are strict subsets or supersets of each other, TOPT graphs never have sequential discontinuities in partially-overlapping momentum channels. This amounts to a new proof of the Steinmann relations in perturbation theory. We emphasize that the relations we derive between sequential discontinuities and cuts hold for individual Feynman integrals, and as such the Steinmann relations must also be obeyed by individual Feynman integrals.

This is a long paper, partly because we wanted to give a pedagogical introduction to various subjects relevant for the main results in a uniform language. We begin in Sections 5.2 and 5.3 by reviewing first the cutting rules and then the discontinuities of integrals in both covariant perturbation theory and TOPT. These sections essentially review what is needed to understand and prove the relation between single discontinuities and cuts, as in the optical theorem. We proceed in Section 5.4 to introduce the main mathematical tools we use

for computing sequential discontinuities. Here, we discuss the maximal analytic continuation of polylogarithmic functions and introduce the formalism of variation matrices. We then show how the discontinuities of polylogarithms can be computed using the action of the monodromy group. Our treatment of these topics draws heavily from [201, 202], but is intended to be introductory since these topics have not featured prominently in the physics literature. In Section 5.5 we use these tools to prove our main results for sequential discontinuities and cuts of Feynman integrals. A corollary is a new integral-by-integral proof of the Steinmann relations. In Section 5.6 we work through some explicit examples that illustrate these new relations between the cuts and discontinuities of Feynman integrals, including bubble, triangle, and box diagrams up to L -loop order. A summary and discussion of some possible implications of our work and future directions are given in Section 5.7.

We also include in this paper a number of appendices with some technical details not needed for the main results of the paper. Appendix B.1 discusses the relation between the variation matrix and the coproduct. Appendix B.2 discusses the relationship between the monodromy group associated with a polylogarithm and the fundamental group of the manifold on which it is defined, and explicitly works out the relation between these groups in the case of the triangle and box ladder integrals. Appendix B.3 shows how single-valued functions can be easily constructed in the variation matrix formalism. In Appendix B.4, we provide details on how the permutation symmetry of the one-loop triangle integral acts on its rational and transcendental parts. Appendix B.5 presents the variation matrix for the transcendental function $\Phi_2(z, \bar{z})$ appearing in the two-loop ladder triangle and box diagrams. Finally, Appendices B.6 and B.7 give some details of the calculation of cuts of the

three-loop and L -loop triangle diagrams.

5.2 Cutting rules: a review

The branch points and branch cuts of Feynman integrals have been studied since the early days of S -matrix theory. Landau described how to compute the location of these branch hypersurfaces[45], and later Cutkosky described how to compute discontinuities across these hypersurfaces, using Feynman integrals with cut propagators [6]. In this section we review the cutting rules and the relationship between cuts, discontinuities, and the imaginary part of a scattering amplitude.

5.2.1 Cutkosky, 't Hooft and Veltman

We begin with the generalized optical theorem, which states that the imaginary part of a scattering amplitude \mathcal{A} is given by a sum over intermediate states X ,

$$\text{Im}\mathcal{A}(A \rightarrow B) = i \sum_X \int d\Pi_X (2\pi)^4 \delta^4(p_A - p_X) \mathcal{A}(A \rightarrow X) \mathcal{A}^*(X \leftarrow B). \quad (5.2.1)$$

This optical theorem is non-perturbative and follows from the unitarity of the S -matrix.

By expanding each side order-by-order in any coupling, the theorem implies a constraint on the sum of all Feynman diagrams contributing to \mathcal{A} at any order. However, it does not provide any constraints on individual diagrams. Some nontrivial checks on the optical theorem, including examples where disconnected diagrams play a crucial role, can be found in [70].

One can derive stronger results than the optical theorem by directly studying individual

Feynman integrals. These integrals are Lorentz-invariant integrals over Feynman propagators, and take the form

$$\mathcal{M}(p) = \int \prod_{\ell} \frac{d^d k_{\ell}}{i(2\pi)^d} \prod_j \frac{1}{[q_j(k, p)]^2 - m_j^2 + i\varepsilon} . \quad (5.2.2)$$

In our notation, the integer ℓ indexes L loop momenta k_{ℓ} , and j indexes the internal lines. The variables k and p denote the collective set of loop and external momenta, respectively, while $q_j(k, p)$ and m_j denote the momentum and mass of the j^{th} internal line. We do not include factors of i in the numerators of the propagators, but include a factor of $1/i$ per loop integral in anticipation of the i 's generated by the k_{ℓ}^0 integrals. Throughout this paper, we take incoming particles to have positive energy.

Feynman integrals are defined in terms of external four-momenta p^{μ} , but since they are Lorentz invariant they depend only on invariants of the form $s_I = P_I^2$, where $P_I^{\mu} \equiv \sum_{i \in I} p_i^{\mu}$ denotes a sum of external momenta. These invariants cannot all be independent. For instance, in four dimensions a Feynman integral $\mathcal{M}(p)$ depends on n external momenta and hence (at most) $4n$ independent quantities, while there are 2^n invariants s_I . The number of independent invariants is further reduced by momentum conservation and the on-shell condition for each external particle. Thus, the s_I are highly interdependent. The constraints on the s_I are easiest to derive using their expression in terms of four-momenta.

The integral \mathcal{M} may become singular as $i\varepsilon \rightarrow 0$ in the propagators. For physical momenta the Mandelstam invariants s_I are real, but we can analytically continue \mathcal{M} to be a function of complex s_I . Then the singularities as $i\varepsilon \rightarrow 0$ can be thought of as the endpoints

of branch cuts on a Riemann surface (more generally a hypersurface of maximal analytic continuation) associated to \mathcal{M} . In 1959, Landau derived a set of equations whose solutions indicate the regions of momenta where these singularities may reside, collectively known as the *Landau surface* [45]. The Landau surface may be disconnected, but each connected component corresponds to some set of propagators becoming singular: $[q_j(k, p)]^2 = m_j^2$.

Cutkosky

Shortly after Landau's paper, Cutkosky gave a prescription for computing the discontinuity across one region of the Landau surface [6]. If the singularity is associated with the region \mathcal{L}_J where the propagators $j \in J$ go on-shell, then the discontinuity is given by

$$\text{Disc}_{\mathcal{L}_J} \mathcal{M} = \int \prod_{\ell} \frac{d^d k_{\ell}}{i(2\pi)^d} \left[\prod_{j \in J} (-2\pi i) \delta(q_j^2 - m_j^2) \Theta(q_j^0) \right] \prod_{k \notin J} \frac{1}{q_k^2 - m_k^2}. \quad (5.2.3)$$

Cutkosky also considered the singularities of $\text{Disc}_{\mathcal{L}_J} \mathcal{M}$. He argued that the discontinuity across a region of the Landau surface associated with a set of propagators K (that are in the complement of J) going on shell is given by

$$\text{Disc}_{\mathcal{L}_K} \text{Disc}_{\mathcal{L}_J} \mathcal{M} = \text{Disc}_{\mathcal{L}_{J \cup K}} \mathcal{M}. \quad (5.2.4)$$

This is the type of sequential discontinuity we focus on in this paper.

Unfortunately, Cutkosky's results are phrased entirely in terms of discontinuities across regions of the Landau surface where particular propagators go on-shell. However, it is generally not possible to isolate a region corresponding to the singularity locus of (just) a given

set of propagators in the space of independent invariants. For example, a string of bubbles depends only on a single external kinematic invariant p^2 , but the Landau equations identify a different branch hypersurface when the propagators in different bubbles are cut. Thus, Cutkosky's formula gives no constraint for sequential discontinuities in the same channel, a central focus of this paper.

't Hooft and Veltman

A simplified treatment of cuts and discontinuities was provided in the 1970's by 't Hooft and Veltman [54, 164]. Their approach sidestepped the Landau equations and analytic continuation entirely, to provide a constraint on \mathcal{M} directly. They start with the Feynman graph associated with the Feynman integral \mathcal{M} , and consider all possible colorings of the vertices of this graph as either black or white. The following rules are then assigned to the edges between these colored vertices:

$$\bullet \text{---} \bullet \equiv \frac{1}{p^2 - m^2 + i\varepsilon} \quad \circ \text{---} \circ \equiv \frac{1}{p^2 - m^2 - i\varepsilon} \quad \bullet \text{---} \rightarrow \circ \equiv -2\pi i \delta(p^2 - m^2) \Theta(p^0) \quad (5.2.5)$$

The graph with all black vertices is the original time-ordered Feynman integral \mathcal{M} , with all $+i\varepsilon$ propagators, while the graph with all white vertices corresponds to $-\overline{\mathcal{M}}$, where $\overline{\mathcal{M}}$ is defined by

$$\overline{\mathcal{M}}(p) = \int \prod_{\ell} \frac{d^d k_{\ell}}{-i(2\pi)^d} \prod_j \frac{1}{[q_j(k, p)]^2 - m_j^2 - i\varepsilon}. \quad (5.2.6)$$

Propagators connecting black and white vertices are said to be cut, meaning these lines are on-shell and positive energy flows from black to white. Using the position-space version of these rules, 't Hooft and Veltman showed that the sum over all possible assignments of

white and black vertices is zero. This implies what we call the *covariant* cutting rules

$$\begin{aligned} \mathcal{M} - \overline{\mathcal{M}} = \sum_{\bullet, \circ} (-1)^{L_\circ} \int \prod_\ell \frac{d^d k_\ell}{i(2\pi)^d} \prod_{\bullet-\bullet} \frac{1}{q_j^2 - m_j^2 + i\varepsilon} \\ \times \prod_{\bullet-\circ} (-2\pi i) \delta(q_j^2 - m_j^2) \Theta(q_0) \prod_{\circ-\circ} \frac{1}{q_j^2 - m_j^2 - i\varepsilon}, \quad (5.2.7) \end{aligned}$$

where the sum is over all diagrams with mixed black and white vertices and L_\circ is the number of loops connecting exclusively white vertices.

There are a few important aspects of this equation to note. First, the covariant cutting rules (like Cutkosky's rules) do not require unitarity. Eq. (5.2.7) is derived algebraically, as a constraint among integrals over propagators and delta functions. In a unitary theory, $\overline{\mathcal{M}}$ is related to the complex-conjugated integral \mathcal{M}^* (where the numerators and vertices are complex conjugated in addition to $+i\varepsilon \rightarrow -i\varepsilon$), and the numerators of cut propagators correspond to a sum over physical spins. Then the sum over cuts gives the total scattering cross section, and the generalized optical theorem in Eq. (5.2.1) results.

Second, even in a non-unitary theory the covariant cutting rules relate an integral with all $+i\varepsilon$ propagators to an integral with all $-i\varepsilon$ propagators. Since the Feynman integrals we consider have all the other sources of imaginary parts stripped out, the cutting rules directly compute $\text{Im}\mathcal{M}$. Although we would like to view \mathcal{M} as an analytic function, so that $\text{Im}\mathcal{M}$ is related to the discontinuity of \mathcal{M} around a branch point, this has to be done with some care. The covariant cutting rules directly let us compute only $\mathcal{M} - \overline{\mathcal{M}}$.

Third, if we compare to Cutkosky's formula in Eq. (5.2.3) we note that the covariant cutting rules involve mixed $+i\varepsilon$ and $-i\varepsilon$ propagators, while Eq. (5.2.3) is agnostic to the

pole positions of the propagators. This does not make the two equations inconsistent, since left-hand-side of Eq. (5.2.3) is the discontinuity across a Landau surface defined by the cut propagators while the left-hand side of Eq. (5.2.7) is the imaginary part of \mathcal{M} . It does however make it difficult to explicitly verify Cutkosky's equation. In contrast, Eq. (5.2.7) can be verified in a straightforward manner in any number of examples.

Finally, because the 't Hooft-Veltman derivation of the cutting rules builds on a single constraint among all the diagrams (the largest time equation), it is hard to break it down further to derive constraints on individual Feynman diagrams. Although such a dissection might be possible, we find it more transparent to work in time-ordered perturbation theory where the cutting rules can be derived in a way that makes generalizations to sequential cuts and discontinuities more straightforward.

5.2.2 Time-ordered perturbation theory

To prove the cutting rules in time-ordered perturbation theory (TOPT) we exploit the following simple mathematical identity. If some functions A_j, B_j and C_j are related by

$$A_j - B_j = C_j, \quad (5.2.8)$$

then

$$A_1 \cdots A_n - B_1 \cdots B_n = C_1 B_2 \cdots B_n + A_1 C_2 B_3 \cdots B_n + \cdots + A_1 \cdots A_{n-1} C_n. \quad (5.2.9)$$

For $n = 1$, there are no A_j or B_j on the right hand side, and so Eq. (5.2.9) reduces to Eq. (5.2.8).

For example, if we take $A_j = \frac{1}{p_j^2 + i\varepsilon}$, $B_j = \frac{1}{p_j^2 - i\varepsilon}$ and $C_j = -2\pi i \delta(p_j^2)$, then Eq. (5.2.8) corresponds to the familiar relation

$$\frac{1}{p_j^2 + i\varepsilon} - \frac{1}{p_j^2 - i\varepsilon} = -2\pi i \delta(p_j^2). \quad (5.2.10)$$

To be clear, this is an identity in the sense of distributions; it is the cutting equation for $\mathcal{M} = \frac{1}{p_j^2 + i\varepsilon}$. In general, with this choice of A_j, B_j and C_j , the left hand side of Eq. (5.2.9) corresponds to the difference between an integral with all $+i\varepsilon$ propagators and one with all $-i\varepsilon$ propagators, which is either $\mathcal{M} - \overline{\mathcal{M}}$ or $\mathcal{M} + \overline{\mathcal{M}}$ depending on the number of loops. For an even number of loops, Eq. (5.2.9) can be applied, but even then it produces some combination of propagators with $+i\varepsilon$ propagators, some $-i\varepsilon$ propagators and delta functions with no clear relation to Eq. (5.2.7).

To derive the cutting rules using Eq. (5.2.9), we use TOPT. Recall that covariant Feynman rules are derived for time-ordered products by inserting extra energy integrals to generate expressions in terms of Feynman propagators. In TOPT, the time-orderings are kept separate and no energy integral is introduced. Each Feynman diagram is the sum of $v!$ TOPT diagrams, with v the number of vertices. In a time-ordered diagram, the internal lines are on-shell (meaning $q_0 = \omega_q = \sqrt{\vec{q}^2 + m^2}$) and three-momentum is conserved at each vertex, but energy is in general *not* conserved at each vertex. The positive sign is always taken for the energy (in front of the square root), because all intermediate states are

Fock-state elements of physical on-shell positive-energy particles. Each intermediate state contributes an energy denominator to the TOPT amplitude M , with its energy subtracted from the initial energy, along with a $+i\varepsilon$. M also acquires a $\frac{1}{2\omega_q}$ from each propagator with momentum q , and loop three-momenta are integrated over. A detailed discussion, and a derivation of the TOPT Feynman rules, is given in [130] (see also [95]). For example, the scalar loop can be written as

$$\begin{aligned}
\text{Diagram: } & \text{A circle with momentum } k \text{ at the top and } p-k \text{ at the bottom. Two external lines with momentum } p \text{ enter from the left and exit to the right.} \\
& = \int \frac{d^4k}{i(2\pi)^4} \frac{1}{k^2 - m_1^2 + i\varepsilon} \frac{1}{(p-k)^2 - m_2^2 + i\varepsilon} \\
& = - \int \frac{d^3k}{(2\pi)^3} \frac{1}{2\omega_k} \frac{1}{2\omega_{p-k}} \left[\frac{1}{E_p - (\omega_k + \omega_{p-k}) + i\varepsilon} + \frac{1}{E_p - (\omega_k + \omega_{p-k} + 2\omega_p) + i\varepsilon} \right],
\end{aligned} \tag{5.2.11}$$

where $E_p = \omega_p = \sqrt{\vec{p}^2 + m^2}$ is the energy of p^μ and

$$\omega_k = \sqrt{\vec{k}^2 + m_1^2}, \quad \omega_{p-k} = \sqrt{(\vec{p} - \vec{k})^2 + m_2^2} \tag{5.2.12}$$

are the energies of the virtual particles. Eq. (5.2.11) can be verified by performing the k_0 integral, which picks up two of the four poles. In terms of diagrams, we have

$$\begin{aligned}
\text{Diagram 1: } & \text{A circle with momentum } k \text{ at the top and } p-k \text{ at the bottom. Two external lines with momentum } p \text{ enter from the left and exit to the right.} \\
& = \text{Diagram 2: A circle with momentum } \vec{k} \text{ at the top and } \vec{p}-\vec{k} \text{ at the bottom. Two external lines with momentum } \vec{p} \text{ enter from the left and exit to the right.} \\
& \quad + \text{Diagram 3: A lens-shaped diagram with two vertices. The top vertex has momentum } \vec{p} \text{ entering. The bottom vertex has momentum } \vec{p} \text{ entering. The left side of the lens has momentum } \vec{k} \text{ entering. The right side of the lens has momentum } \vec{p}-\vec{k} \text{ entering.}
\end{aligned} \tag{5.2.13}$$

The intermediate state in the TOPT diagrams changes as each vertex is passed in time (where time flows to the right). In the first diagram this state includes only the k and $p-k$ lines, so its energy is $\omega_k + \omega_{p-k}$; in the second diagram, the intermediate state includes also

the energy of the initial and final states, and thus its energy is $\omega_k + \omega_{p-k} + 2\omega_p$.

It is often difficult to perform the k_0 integrals to reduce Feynman diagrams to TOPT diagrams. Their equivalence is easiest to show from more general principles of quantum field theory, since both compute the same time-ordered products (cf. [124, 130]). Keep in mind that although the $+i\varepsilon$ is necessary to determine the k_0 integration contour, it cannot be removed after the integration is done. Indeed the $+i\varepsilon$ originates from the fact that particles move forward in time with positive energy and is an essential part of the Lippmann-Schwinger propagator in TOPT.

Now for each term in the TOPT decomposition we can apply the identity in Eq. (5.2.9), using the TOPT analog of Eq. (5.2.10):

$$\frac{1}{E_j - \omega_j + i\varepsilon} - \frac{1}{E_j - \omega_j - i\varepsilon} = -2\pi i \delta(E_j - \omega_j). \quad (5.2.14)$$

The sum of all TOPT diagrams with a given topology and all $+i\varepsilon$ propagators gives the Feynman diagram \mathcal{M} , while the sum of these diagrams with all $-i\varepsilon$ propagators gives $\overline{\mathcal{M}}$. The remaining terms have $\delta(E_j - \omega_j)$ factors which impose energy conservation at an intermediate time. These diagrams neatly split in two along the cut, with positive energy automatically flowing across the cut (because TOPT diagrams have positive energy at any intermediate time) and where all cut particles are on-shell (since all particles are on-shell in TOPT). By Eq. (5.2.9) all the propagators before the cut have $+i\varepsilon$ and those after the cut have $-i\varepsilon$. Thus the cut TOPT diagram is one particular time-ordering of a white/black partition, which is one time-ordering of a cut Feynman diagram. The sum of all possible

cut TOPT diagrams gives all the possible time-orderings of the black and white vertices, and therefore reproduces the full cut Feynman diagram and confirms the cutting rules.

For example, when we apply Eq. (5.2.9) to Eq. (5.2.13), there is only one intermediate state in each diagram to cut (in contrast to the Feynman diagram, which has two intermediate propagators to cut). Cutting the first diagram gives

$$\text{Diagram 1} = \int \frac{d^3k}{(2\pi)^3} \frac{1}{2\omega_k} \frac{1}{2\omega_{p-k}} (-2\pi i) \delta(E_p - (\omega_k + \omega_{p-k})) \quad (5.2.15)$$

$$\begin{aligned} &= \int \frac{d^4k}{i(2\pi)^4} \int \frac{d^4k'}{i(2\pi)^4} (-2\pi i) (2\pi)^3 \delta^4(p - k - k') \\ &\quad \times (-2\pi i) \delta(k^2 - m_1^2) \Theta(k_0) (-2\pi i) \delta(k'^2 - m_2^2) \Theta(k'_0). \end{aligned} \quad (5.2.16)$$

So this diagram alone gives the cut of the Feynman diagram. The cut of the other diagram is zero, since energy conservation at the cut is impossible to satisfy:

$$\text{Diagram 2} = \int \frac{d^3k}{(2\pi)^3} \frac{1}{2\omega_k} \frac{1}{2\omega_{p-k}} (-2\pi i) \delta(E_p - (2\omega_p + \omega_k + \omega_{p-k})) = 0. \quad (5.2.17)$$

This is typical of TOPT graphs: when one time-ordering can be cut, the same diagram with vertices in reversed time order cannot be cut.

More broadly, the key reason why the cutting rules can be derived diagrammatically in TOPT is that cuts in TOPT are associated with internal multiparticle states, not individual particles. So a cut, which replaces a TOPT propagator by a delta function, splits the diagram in two, ordered by time, in contrast to Feynman diagrams, where using Eq. (5.2.9)

just opens up a loop.

In fact, we have derived something stronger than the covariant cutting rules: the constraint on the amplitude holds for each time-ordered Feynman diagram separately *and* it holds point-by-point in phase space. Indeed, the equation that we use to prove it, Eq. (5.2.9) holds at the integrand level. Let us define an individual TOPT integrand for fixed loop momenta as

$$M \equiv \frac{1}{E_1 - \omega_1 + i\varepsilon} \cdots \frac{1}{E_n - \omega_n + i\varepsilon}, \quad \overline{M} \equiv \frac{1}{E_1 - \omega_1 - i\varepsilon} \cdots \frac{1}{E_n - \omega_n - i\varepsilon}. \quad (5.2.18)$$

Then, by putting in the explicit form of the TOPT propagators, Eq. (5.2.9) gives what we call the *time-ordered cutting rules*:

$$M - \overline{M} = \sum_j \frac{1}{E_1 - \omega_1 + i\varepsilon} \cdots \frac{1}{E_{j-1} - \omega_{j-1} + i\varepsilon} (-2\pi i) \delta(E_j - \omega_j) \times \frac{1}{E_{j+1} - \omega_{j+1} - i\varepsilon} \cdots \frac{1}{E_n - \omega_n - i\varepsilon}. \quad (5.2.19)$$

When the loop momenta are integrated over, this equation implies the cutting rules, but this equation holds for any E_j and ω_j .

5.3 Discontinuities

Having understood the cutting rules in covariant perturbation theory and in time-ordered perturbation theory, we can now proceed to connect cuts to the discontinuities of amplitudes. As discussed above, the Feynman integral \mathcal{M} , viewed as an analytic function of

Mandelstam invariants, is a multi-valued function on a complex manifold. Cutkosky showed that one can compute the discontinuity of \mathcal{M} across some region of its Landau surface by summing over integrals in which different sets of propagators have been cut. However, to provide practical constraints on amplitudes we need a prescription much more explicit than Cutkosky's. For example, how do we identify what region of the surface we are probing from knowledge of which Feynman propagators have been cut? And how do we actually perform the analytic continuation around the relevant branch points?

There are two related concepts that we will discuss, and which we want to connect. The first is the total discontinuity of a Feynman integral in a particular *region*, which is computed by the covariant cutting rules. A region in this context is the specification of the signs of the Mandelstam invariants, and the signs of the energies (which particles are incoming and which are outgoing), if necessary. Once the signs are specified, we can compute the total discontinuity using Eq. (5.2.7). The second concept is the discontinuity of a Feynman integral with respect to a particular kinematic invariant s . More specifically, we define $\text{Disc}_s \mathcal{M}$ as the difference between \mathcal{M} before and after analytic continuation along a path that encircles the branch point in s (but no other branch points). Since Mandelstam invariants are not all independent, this has to be done with some care.

5.3.1 Covariant approach

We begin with the total discontinuity $\mathcal{M} - \overline{\mathcal{M}}$, which can be computed using the covariant cutting rules in Eq. (5.2.7). As defined in Eqs. (5.2.2) and (5.2.6), \mathcal{M} is a Feynman integral with all $+i\varepsilon$ propagators and $\overline{\mathcal{M}}$ is the same integral with all $-i\varepsilon$ propagators, multiplied

by a factor of $(-1)^L$. At any real phase-space point, \mathcal{M} and $\overline{\mathcal{M}}$ are complex conjugates of each other for finite values of $i\varepsilon$. From this point of view, \mathcal{M} and $\overline{\mathcal{M}}$ are separated by a branch cut at $i\varepsilon = 0$, and may have a finite difference as $i\varepsilon \rightarrow 0$ from the positive or negative direction. In contrast, viewed as an analytic function of the momenta, \mathcal{M} and $\overline{\mathcal{M}}$ are evaluations of the same function at different points on a complex manifold. Thus the finite difference between \mathcal{M} and $\overline{\mathcal{M}}$ can be thought of as the discontinuity of a single function \mathcal{M} . We would like to understand the analytic continuation contour along which \mathcal{M} can be transformed into $\overline{\mathcal{M}}$, as this will allow us to connect the total discontinuity computed by the covariant cutting rules to the notion of discontinuities with respect to particular Mandelstam invariants.

The branch cut between \mathcal{M} and $\overline{\mathcal{M}}$ starts at a branch point (more generally, a branch hypersurface) somewhere in the space of Mandelstam invariants on which \mathcal{M} depends. As such, the discontinuity can be computed by analytically continuing \mathcal{M} around this branch point to the other side of the branch cut. To do this, we can continue \mathcal{M} into a regime where it is analytic, and then to the region where it matches $\overline{\mathcal{M}}$. For example, suppose $\mathcal{M} = \ln(-s + i\varepsilon)$ and $\overline{\mathcal{M}} = \ln(-s - i\varepsilon)$, and take $s > 0$. Then we can continue \mathcal{M} along the path $s \rightarrow e^{i\alpha}s$ with $0 \leq \alpha \leq \pi$ to the region where $s < 0$. From this region we can either go back and reproduce \mathcal{M} using $s \rightarrow e^{-i\alpha}s$, or keep going to arrive at $\overline{\mathcal{M}}$ on the other side of the branch cut using $s \rightarrow e^{i\alpha}s$. We can also continue increasing the phase of s in this manner: as α increases, we end up on higher and higher sheets of the Riemann surface of $\ln(-s)$. A single discontinuity corresponds to the single monodromy around the branch point at $s = 0$. In equations, for the logarithm we have $\text{Disc}_s \mathcal{M} = \text{Disc}_{\text{tot}} \mathcal{M} = \mathcal{M} - \overline{\mathcal{M}} =$

$2i\text{Im } \mathcal{M}$ in the region where $s > 0$.

A useful concept for studying the analytic properties of Feynman integrals is the Euclidean region. In this region, all Mandelstam invariants are negative and \mathcal{M} is analytic.⁴ To see that integrals are analytic in the Euclidean region, it is helpful to write a general Feynman integral in the Symanzik representation [203]. This is done by using Feynman parameters and then integrating over the loop momenta. The result is that a Feynman amplitude as in Eq. (5.2.2) can be written as

$$\mathcal{M}(p) = \int \prod_{x_j \geq 0} dx_j \delta(1 - \sum x_j) \frac{\mathcal{U}^{n-2L-2}}{\mathcal{F}^{n-2L}}. \quad (5.3.1)$$

Here, the first Symanzik polynomial \mathcal{U} is

$$\mathcal{U} = \sum_{\mathcal{T}_1} \left[\prod_{j \notin \mathcal{T}_1} x_j \right], \quad (5.3.2)$$

where the sum is over all 1-trees \mathcal{T}_1 , which correspond to tree chapters/diagrams that connect all vertices in the graph. The second Symanzik polynomial \mathcal{F} is

$$\mathcal{F} = \sum_{\mathcal{T}_2} \left[\prod_{j \notin \mathcal{T}_2} x_j \right] (-s_{P(\mathcal{T}_2)}) + \mathcal{U} \sum_{j=1} x_j m_j^2 - i\varepsilon, \quad s_{P(\mathcal{T}_2)} = \left[\sum_{j \notin \mathcal{T}_2} p_j \right]^2, \quad (5.3.3)$$

where m_j are the masses of the internal lines and the sum is over 2-trees \mathcal{T}_2 , which them-

⁴Due to momentum conservation, the Euclidean region will not exist for all Feynman integrals. However, while the Euclidean region is useful for motivating the relations we derive in Section 5.5, these relations will not depend on the existence of this region. In particular, any number of Mandelstam invariants will be allowed to be positive in these relations in addition to the Mandelstams with respect to which discontinuities are being taken.

selves correspond to pairs of disconnected tree chapters/diagrams that involve all vertices of the original graph. The nice thing about this parametrization is that \mathcal{M} is now manifestly a function of Mandelstam invariants.

Singularities in \mathcal{M} can only arise when $\mathcal{F} = 0$. Since the integration region corresponds to $x_i \geq 0$ and in the Euclidean region $s_{P(\mathcal{T}_2)} < 0$ for all \mathcal{T}_2 and $m_j^2 \geq 0$ for all j , the denominator will never vanish and the result will be analytic in the external momenta. Note that the Euclidean regime is identified with a stronger requirement than that \mathcal{M} is analytic; it requires that all Mandelstam invariants are negative, not just those associated with 2-trees from a particular graph. We denote the Euclidean region by R^\star .

We denote generic regions, in which kinematic invariants can be positive or negative, by R . We use the more precise notation R_+ to indicate a region in which all positive invariants are slightly above the associated branch cut, i.e. all propagators have $+i\varepsilon$. The region in which all positive invariants are instead below the associated branch cut, and all propagators have $-i\varepsilon$, will be denoted R_- .⁵ Thus, we write

$$[\text{Disc}_{\text{tot}}\mathcal{M}]_R = [\mathcal{M} - \overline{\mathcal{M}}]_R = \mathcal{M}_{R_+} - \mathcal{M}_{R_-} . \quad (5.3.4)$$

To compute the right hand side, we would like to understand how to analytically continue the amplitude between R_+ , R^\star , and R_- . There are many ways to do this. The precise path should not affect the answer for the discontinuity. It is nevertheless important to know that the path exists, and having an explicit path can help determine which branch points are

⁵With Feynman propagators, the amplitude in this region also has a $(-1)^L$ in R_- due to the additional rotation of the energies in the loop integrals. With TOPT, we simply flip $i\varepsilon \rightarrow -i\varepsilon$ as there are no energies in the loop integrals.

encircled.

Since \mathcal{M} is Lorentz invariant, it may seem most natural to continue the invariants themselves. For example, we can rotate all the positive invariants to negative values via $s_I \rightarrow e^{i\alpha}s_I$ with $0 < \alpha < \pi$ while leaving the negative invariants stationary. This puts us in R^\star , where all $s_I < 0$ and the amplitude is nonsingular. We can then keep going, and analytically continue all the invariants that were originally positive further by extending $0 < \alpha < 2\pi$, to end up in R_- . Unfortunately, since the invariants are not all independent, this procedure can be ambiguous. For example, in massless four-particle kinematics, if we want to rotate s from being positive to negative while holding the other invariants fixed, we could try the above analytic continuation path. But if we rewrite our amplitude or integral to depend just on the other invariants using the relation $s = -t - u$, this rotation would seem to have no effect. Thus, one must be careful to do the rotation in a manner that respects the reparameterization invariance of the integrals.

In this paper, we will restrict ourselves to analytic continuations in external energies that respect overall energy conservation and leave all external three-momenta fixed. In addition to avoiding the issue described above, this choice facilitates our derivation of relations between sequential discontinuities and cut integrals, and leads to unambiguous predictions. In addition, rotating the energies while respecting four-momentum conservation ensures that we always satisfy any Gram determinant constraints.

In general, there are many ways to rotate external energies to get from a region R to the Euclidean region. For example, if the momenta in R all take non-exceptional values, one can uniformly lower the energies $E_j \rightarrow \alpha E_j$ with $\alpha < 1$. Eventually, at some point α_{\min}

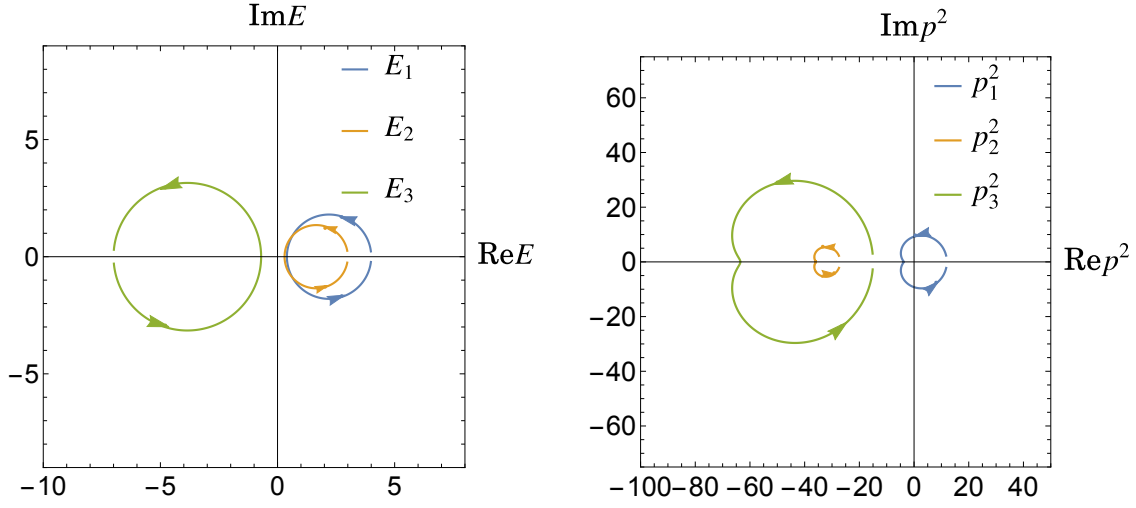


Figure 5.1: Example analytic continuation involving three external energies. We start at the kinematic point $p_1 = (4, 2, 0, 0)$, $p_2 = (3, 6, 0, 0)$, and $p_3 = (-7, -8, 0, 0)$, where we have $p_1^2 = 12$, $p_2^2 = -27$, and $p_3^2 = -15$. We rotate the energies by $E_j \rightarrow [0.1 + 0.9e^{i\pi s} \cos(\pi s)]E_j$ with $0 \leq s < 1$. During this rotation the positive invariant p_1^2 circles its branch point at $p_1^2 = 0$, thus taking us from $R^1 \rightarrow R^* \rightarrow R^1$, but changing the sign of the corresponding $i\varepsilon$. The small gaps at the beginning and end of the paths represent the $\pm i\varepsilon$.

all the invariants become negative. One can then rotate the energies in the complex plane around $\alpha_{\min} E_j$ and return to $\alpha = 1$ on the opposite side of the real energy axis. This procedure respects energy-momentum conservation everywhere along the path. One only has to be careful that the invariants do not encircle their branch points twice. A concrete example involving three momenta that follows a path homotopic to the one described in this paragraph is shown in Fig 5.1. We construct a number of similar paths for the examples we consider in Section 5.6.

Let us now assume that an appropriate analytic continuation in the energies has been chosen, which takes us from a region R_+ to the corresponding region R_- (where all Mandelstam invariants have the same sign, but each $+i\varepsilon$ has been changed to $-i\varepsilon$). Then the difference between \mathcal{M} before and after this analytic continuation should match the total

discontinuity of a Feynman integral in the region R using the covariant cutting rules:

$$[\text{Disc}_{\text{tot}}\mathcal{M}]_{R_+} = \mathcal{M}_{R_+} - \mathcal{M}_{R_-} = \sum_{\text{cuts}} \mathcal{M}_{R_{+|-}} \quad (5.3.5)$$

We emphasize the right side of this equation involves a sum over all cuts (in all channels), as explicitly given in Eq. (5.2.7). When we cut a set of propagators, we replace each one by

$$\text{cut} : \frac{1}{p^2 - m^2 + i\varepsilon} \rightarrow -2\pi i \delta(p^2 - m^2) \theta(p_0) \quad (5.3.6)$$

and use $+i\varepsilon$ for all propagators before the cut and $-i\varepsilon$ for all propagators after the cut, as implied by the subscript on $R_{+|-}$.

We would now like to derive a concrete relation between $\text{Disc}_s\mathcal{M}$, and the cuts of \mathcal{M} . The discontinuity of \mathcal{M} with respect to s corresponds to analytically continuing \mathcal{M} from being evaluated at $s + i\varepsilon$ to being evaluated at $s - i\varepsilon$, while the other invariants remain unchanged. Let us denote the region in which $s > 0$ and all other kinematic invariants are negative by R^s . As only the invariant s is positive in this region, all the nonzero cuts in the sum in Eq. (5.3.5) are in the s -channel. As a result, we have

$$[\text{Disc}_{\text{tot}}\mathcal{M}]_{R^s} = \mathcal{M}_{R_+^s} - \mathcal{M}_{R_-^s} = \sum_{\text{cuts}} \mathcal{M}_{R_{+|-}} = \sum_{\text{cuts in } s} \mathcal{M}_{R_{+|-}} . \quad (5.3.7)$$

To further connect this sum of cut integrals to $\text{Disc}_s\mathcal{M}$, we must show that the analytic continuation corresponding to Disc_{tot} in this region encircles a branch point in only in s , and in no other invariants. This turns out to be easiest to see in TOPT, which we turn to

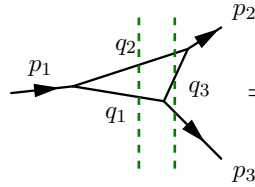
now.

5.3.2 Discontinuities in TOPT

In TOPT all internal lines are on-shell with positive energy and real masses ($q_0 > 0$ and $q^2 \geq 0$). External lines, however, have no such restriction; they can have $p^2 < 0$ if the diagram is meant to be embedded in a larger diagram (for example, the off-shell photon in deep-inelastic scattering is spacelike), and incoming external particles can have negative energy if they correspond to outgoing particles.

Because we are ultimately interested in the analytic properties of Feynman integrals as functions of external energies, it is helpful to separate out the contributions to TOPT propagators from internal and external lines. In particular, we can put each TOPT propagator in the form $1/(E_P - \omega_q + i\varepsilon)$, where E_P corresponds to a sum over external energies, and $\omega_q = \sum_j \omega_j$ is a sum over particles in internal lines, where $\omega_j = \sqrt{\vec{q}_j^2 + m_j^2}$.

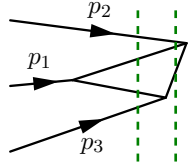
Consider for example the one-loop TOPT graph, with all $E_i > 0$:



$$= \frac{1}{E_1 - (\omega_1 + \omega_2) + i\varepsilon} \frac{1}{(E_1 - E_3) - (\omega_2 + \omega_3) + i\varepsilon} . \quad (5.3.8)$$

In the first propagator, $E_P = E_1 = E_2 + E_3$ and $\omega_q = \omega_1 + \omega_2$ while in the second propagator $E_P = E_1 - E_3 = E_2$ and $\omega_q = \omega_2 + \omega_3$. If we had drawn p_2 and p_3 as incoming lines with negative energy, the diagram would have been more awkward to draw, but we would have

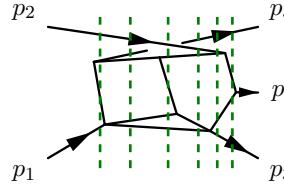
found an equivalent expression:



$$= \frac{1}{E_1 - (\omega_1 + \omega_2) + i\varepsilon} \frac{1}{(E_1 + E_3) - (\omega_2 + \omega_3) + i\varepsilon}. \quad (5.3.9)$$

The value of the diagram is the same since we have flipped $E_3 \rightarrow -E_3$. We use the convention that all lines have positive energies, as in Eq. (5.3.8).

For a general TOPT graph, the energies E_I appearing in the amplitude have a natural sequence. We begin with the total initial-state energy on the far left. Each time a vertex connecting to an external momentum is passed, the external energy is either added, if it is incoming, or subtracted, if it is outgoing. If the vertex is purely internal, then E_P does not change. For example, consider this graph:



$$(5.3.10)$$

The initial energy is $E_1 + E_2$ and the energy past the first vertex (i.e. the energies of the states that cross the first vertical dashed line) is $E_1 + E_2 + E_5 + \omega_1 + \omega_2 + \omega_3$ for some internal energies ω_j ; this first propagator depends on the difference between these energies $E_P = -E_5$. The sequence of E_P values as we move forward in time is

$$-E_5, \quad E_1 - E_5, \quad E_1 - E_5, \quad E_1 - E_5, \quad E_1 - E_5 - E_3, \quad E_1 - E_5 - E_3 + E_2. \quad (5.3.11)$$

If we took all momenta to be incoming, then we would flip the sign of E_3, E_4 , and E_5 so all the signs in Eq. (5.3.11) would be positive. The corresponding sequence is

$$5 \rightarrow 1 \rightarrow 3 \rightarrow 2 \rightarrow 4. \quad (5.3.12)$$

If we use energy conservation to rewrite the energy sum (i.e. $E_5 = -E_4 - E_3 - E_2 - E_1$), the sequence would be the same, in the opposite direction: $5 \leftarrow 1 \leftarrow 3 \leftarrow 2 \leftarrow 4$. The fact that the energies appearing in each successive propagator are a subset of the energies that appeared the preceding propagators (or vice versa) will be important to proving the Steinmann relations in Section 5.5.

Each energy E_I is the energy of a four-vector $P_I^\mu = \sum_{i \in I} \pm p_i^\mu$. Thus there is a one-to-one correspondence between invariants $s_I = P_I^2$ and these energies. A TOPT propagator $1/(E_I - \omega_q + i\varepsilon)$ can only become singular when $E_I = \omega_q$, which only happens if $s_I > 0$. To check this claim, note that the three-momentum \vec{P}_I is the same as the sum of the three-momenta of all the internal particles contributing to ω_q , namely $\vec{P}_I = \sum_j \vec{q}_j$. So we have two four-vectors, $P_I^\mu = (E_I, \vec{P}_I)$ and $q^\mu = (\omega_q, \vec{P}_I)$, with the same three-momentum. Recall that ω_q is the sum of the (positive) energies of the on-shell internal lines. Thus, the four-vector q^μ must be timelike, $q^2 > 0$, since it corresponds to the sum of four-momenta of physical on-shell particles. Therefore, P_I^μ must be timelike when $E_I = \omega_q$. So if $s_I = P_I^2 < 0$ then $E_I \neq \omega_q$. Thus the TOPT propagators can go on-shell only in the kinematical regions where there are singularities in the full Feynman integral, namely when $s_I > 0$. As a corollary, we can drop the $+i\varepsilon$ in any TOPT propagator corresponding to a negative

invariant.

Now let us discuss how to take the discontinuity of a TOPT graph. A TOPT graph is a product of propagators of the form $1/(E_I - \omega_q + i\varepsilon)$. To take the discontinuity in the channel s_I associated with E_I , we want to analytically continue E_I around the pole of this propagator. More precisely, we want to continue E_I around the branch point E_I^* at the end of the line of possible values of ω_q for a given external momentum. This branch point E_I^* is at least as large as the magnitude of the momentum in the channel, $E_I^* \geq |\vec{P}_I|$ but can be strictly larger, for example, if the internal lines are massive. The analytic continuation between R_+ and R_- should have all the energies pass around their branch points, holding the three-momenta fixed and respecting energy conservation.⁶

Taking the difference between a single TOPT propagator before and after this analytic continuation gives

$$\text{Disc}_{\text{tot}} \frac{1}{E_I - \omega_q + i\varepsilon} = \frac{1}{E_I - \omega_q + i\varepsilon} - \frac{1}{E_I - \omega_q - i\varepsilon} = -2\pi i \delta(E_I - \omega_q), \quad (5.3.13)$$

as expected. Similarly, taking the difference between a generic TOPT graph M before and after analytically continuing from $R \rightarrow R^* \rightarrow R$ using a path that encircles the branch points in all of the energies, we get

$$\text{Disc}_{\text{tot}} M = \sum_j \frac{1}{E_{I_1} - \omega_1 + i\varepsilon} \cdots (-2\pi i) \delta(E_{I_j} - \omega_j) \cdots \frac{1}{E_{I_n} - \omega_n - i\varepsilon}. \quad (5.3.14)$$

⁶Note that we do not require the external masses to remain fixed, so in general external particles will not remain on-shell during this analytic continuation.

If we sum over all TOPT chapters/diagrams with the same topology, this reproduces the covariant cutting rules for the total discontinuity of the corresponding Feynman integral \mathcal{M} . That is, we have shown that Eq. (5.2.7) holds with the left-hand side explicitly written as a discontinuity, and have thereby rederived Eq. (5.3.5) using TOPT.

Eqs. (5.3.5) and (5.3.14) hold in any region R . Let us now focus on the region R^s , where only the Mandelstam invariant s is positive, and all other invariants are negative. Since we have shown that singularities in TOPT chapters/diagrams only arise when the energy and corresponding invariant are positive ($P^0 > 0$ and $s = P^2 > 0$), in the region R^s there can only be singularities associated with s . In other words, as we continue from R_+^s to R^* and back to R_-^s , we can only pass around branch points associated with s . This is what we set out to show at the end of the last subsection. As a result, we can now write

$$[\text{Disc}_s \mathcal{M}]_{R^s} = [\text{Disc}_{\text{tot}} \mathcal{M}]_{R^s} = \sum_{\text{all cuts}} \mathcal{M}_{R_{+|-}} = \sum_{\text{cuts in } s} \mathcal{M}_{R_{+|-}}. \quad (5.3.15)$$

Stated more formally, what we have shown is that the analytic continuation used to compute $\text{Disc}_s \mathcal{M}$ is homotopic to the path used to compute $\text{Disc}_{\text{tot}} \mathcal{M}$ in the region R^s .

We would next like to generalize this formula to the case of sequential discontinuities, in the same or different channels. Unfortunately, we cannot simply repeat the procedure that allowed us to compute the first discontinuity. The problem is that this first discontinuity takes the difference of two functions on the branch cut, and thus seems to be only defined on the branch cut itself. For example, $\text{Disc} \ln^2(s) = 4\pi i \ln |s| \Theta(-s)$ is only defined for negative real s , where the branch cut is. In addition, when we take a cut, we turn all

the propagators beyond the cut from $+i\varepsilon$ to $-i\varepsilon$. What is then the right way to cut a $-i\varepsilon$ propagator? To proceed, we will now describe a more sophisticated set of mathematical tools that will allow us to analytically continue Feynman integrals beyond the cut plane. This will allow us to take sequential discontinuities of Feynman integrals.

5.4 Discontinuities as monodromies

The $\pm i\varepsilon$ notation in Feynman propagators is sufficient to compute single discontinuities of Feynman integrals, because this first discontinuity computes the difference between the value of the integral on different sides of a branch cut. For sequential discontinuities, we must explore a larger swath of the analytic structure of the various polylogarithmic functions that appear in a given Feynman integral.⁷ The $\pm i\varepsilon$ notation is not sufficient to describe this structure. Thus, in this section we review how polylogarithmic functions can be analytically continued beyond the principal branch, and how the resulting functions can be related back to the $\pm i\varepsilon$ prescription. We also discuss how these types of analytic continuations can be carried out on TOPT propagators.

5.4.1 Warm-up: the natural logarithm

Consider first the natural logarithm. It can be defined in the region $|s - 1| < 1$ by the sum

$$\ln s \equiv - \sum_{n=1}^{\infty} \frac{1}{n} (1-s)^n \quad \text{for } |s-1| < 1. \quad (5.4.1)$$

⁷While more general types of functions are known to appear in Feynman integrals, we leave these generalizations to future work.

To define $\ln s$ outside the region $|s - 1| < 1$, one can series expand Eq. (5.4.1) around points other than $s = 1$ that are within the original region of convergence to find sum representations that are valid beyond this region. Iterating this procedure, one can extend the function $\ln s$ to the entire complex plane, excluding a curve going from the origin to infinity (the branch cut). This is called the cut complex plane. Since the cut complex plane is simply connected, this analytic continuation is uniquely defined, once the location of the branch cut has been chosen. While the shape of this branch cut is in principle arbitrary, some of this arbitrariness can be removed if we ask that the continued logarithm satisfy the reality property $f(\bar{s}) = \overline{f(s)}$. The standard branch cut choice for the logarithm, going from 0 to $-\infty$ along the real s axis, is consistent with this requirement. We call $\ln s$ with this choice of branch cut the *principal branch* of the logarithm.

With the standard placement of the branch cut for $\ln s$ along the negative real axis, the value of $\ln s$ for negative real s is usually defined to mean the function produced by analytic continuation going counterclockwise from the positive real axis. Moreover, the discontinuity of the logarithm, which computes the difference between the value of this function just above and below the negative real axis, is given by

$$\text{Disc}_s \ln s = \ln(s + i\varepsilon) - \ln(s - i\varepsilon) \xrightarrow{pd} 2\pi i \Theta(-s), \quad (5.4.2)$$

where \xrightarrow{pd} indicates that this value for $\text{Disc}_s \ln s$ only makes sense when one restricts to infinitesimally-separated points that are both in the principle domain. The fact that this discontinuity is nonzero for negative values of s illustrates the ambiguity in defining $\ln s$

directly on the branch cut. This is consistent with the way discontinuities were calculated in the previous section, as the only way to analytically continue a function back to the same point in the cut complex plane is if we start and end on the cut.

The $\pm i\varepsilon$ notation is sufficient for indicating which side of a branch cut we are on when we restrict ourselves to the principal branch of a function. However, when taking additional discontinuities, the $\pm i\varepsilon$ notation and the associated non-analytic theta function are problematic. The single logarithm is a bit too simple, but already $\ln^2 s$ demonstrates the problem. Its discontinuity in the cut complex plane picture is

$$\text{Disc}_s \ln^2 s = \ln^2(s + i\varepsilon) - \ln^2(s - i\varepsilon) \xrightarrow{pd} 4\pi i \ln |s| \Theta(-s). \quad (5.4.3)$$

As with $\ln s$, the discontinuity of $\ln^2 s$ is only nonzero for real $s < 0$, since otherwise $\ln^2(s + i\varepsilon)$ and $\ln^2(s - i\varepsilon)$ agree. But if this discontinuity is only nonzero on the negative real axis, further analytic continuations are ambiguous, and correspondingly so are sequential discontinuities.

To proceed, we note that an alternative way to define the logarithm (other than Eq. (5.4.1)) is through the contour integral

$$\ln s = \int_1^s \frac{dx}{x}. \quad (5.4.4)$$

The integration is to be performed along any contour within the cut complex plane that goes from 1 to s . This definition agrees with the series definition and analytic continuation.

The discontinuity across the branch cut can then be computed as

$$\text{Disc}_s \ln s = \int_1^{s+i\varepsilon} \frac{dx}{x} - \int_1^{s-i\varepsilon} \frac{dx}{x} = \int_{\mathcal{C}_0} \frac{dx}{x} = 2\pi i, \quad (5.4.5)$$

where \mathcal{C}_0 is the infinitesimal contour that wraps around the origin once counterclockwise. For other functions, like $\ln^3 s$ or the dilogarithm $\text{Li}_2(s)$, the discontinuity will not be constant. In such cases we can consider further discontinuities. To do so we need to consider the maximal analytic continuation of our functions, in which we do not restrict their domain to the cut complex plane.

A clue to how to proceed is given by the closed contour \mathcal{C}_0 in Eq. (5.4.5), which apparently passes right through the cut. Indeed, although the integral computation agrees with the discontinuity across the cut, what it is actually computing is the difference between the value of the function on two sheets of a Riemann surface; the location of the branch cut is immaterial. The only invariant is the location of the branch point, at $s = 0$ for the logarithm. This is the unmovable singularity of the integrand.

We can extend the definition of the logarithm in Eq. (5.4.4) beyond the cut complex plane by simply writing

$$\ln_\gamma s = \int_\gamma \frac{dx}{x}, \quad (5.4.6)$$

where the integration contour γ can be any path from 1 to s that does not pass through the origin. This is the maximal analytic continuation of $\ln s$. The domain of the maximal analytic continuation in this case is an infinite number of copies of the complex plane with a branch point at $s = 0$. These additional copies can be accessed by integration contours

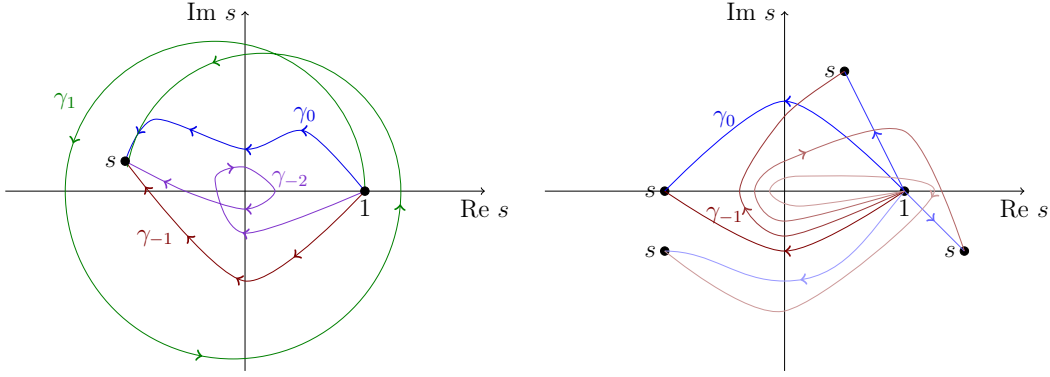


Figure 5.2: The logarithm can be defined as an integral along a path $\ln_\gamma s = \int_\gamma \frac{dx}{x}$, where the paths begin at $x = 1$ and end at $x = s$. The value of $\ln_\gamma s$ depends on the number of times the integration contour wraps around the branch point at the origin. We define families of paths by γ_n where n denotes the number of times the path circles the origin. The family labelled γ_0 is defined to give the principal branch of the logarithm. On the negative real s axis $\ln_{\gamma_0} s = \ln(s + i\varepsilon)$ and $\ln_{\gamma_{-1}} s = \ln(s - i\varepsilon)$.

that wrap around this branch point a given number of times. By considering all such paths, we obtain an infinite number of values for $\ln s$ that differ by multiples of $2\pi i$. This is illustrated in Fig. 5.2, where we denote by γ_n equivalence classes of paths that end at s after wrapping around the origin n times in the counterclockwise direction. The principal branch of the logarithm corresponds to paths that never cross the negative real s axis.

The infinite tower of values associated with $\ln s$ can be thought of as being generated by the closed integration contour around the branch point at the origin. This integral is referred to as the monodromy of $\ln s$ around the origin, and constitutes the only element of the natural logarithm's monodromy group. The discontinuities of polylogarithms can be computed in terms of their monodromies; for instance, in our new notation the discontinuity across the branch cut of $\ln s$ becomes

$$\ln_{\gamma_0} s - \ln_{\gamma_{-1}} s = \int_{\mathcal{C}_0} \frac{dx}{x} = 2\pi i, \quad (5.4.7)$$

where the integral over \mathcal{O}_0 is the monodromy. To connect the monodromy picture to the cut-plane picture, we now identify

$$\ln(s + i\varepsilon) = \ln_{\gamma_0} s, \quad \ln(s - i\varepsilon) = \ln_{\gamma_{-1}} s. \quad (5.4.8)$$

To be clear, $\ln(s \pm i\varepsilon)$ on the left side of these equations means we approach the real s axis from above or below on the principal branch of the logarithm on the cut complex plane.

The logarithms on the right hand side are defined through contours and have no branch cut — the function $\ln_{\gamma} s$ is analytic on the negative real s axis (and everywhere else) as long as the path γ is deformed smoothly to change s . With this identification, Eq. (5.4.7) then agrees with Eq. (5.4.2) up to the theta function. Indeed, the discontinuity defined in terms of the monodromy is an analytic function, while the difference using the principal branch of the logarithm comes with a non-analytic $\Theta(-s)$.

If we adopt the relations in Eq. (5.4.8) as analytic generalizations of $\ln(s + i\varepsilon)$ and $\ln(s - i\varepsilon)$, we can easily compute discontinuities of powers of logarithms by simply substituting in Eq. (5.4.7). For instance,

$$\text{Disc}_s \ln^2(s + i\varepsilon) = \ln^2(s + i\varepsilon) - \ln^2(s - i\varepsilon) = (2\pi i) [2 \ln(s + i\varepsilon) - 2\pi i] \quad (5.4.9)$$

and

$$\begin{aligned}\text{Disc}_s \ln^3(s + i\varepsilon) &= \ln^3(s + i\varepsilon) - \ln^3(s - i\varepsilon) \\ &= (2\pi i) [3 \ln^2(s + i\varepsilon) - 6\pi i \ln(s + i\varepsilon) - 4\pi^2] .\end{aligned}\tag{5.4.10}$$

We can now proceed to take additional discontinuities by subtracting from the function its value with all $+i\varepsilon$ switched to $-i\varepsilon$. We then find

$$\text{Disc}_s \text{Disc}_s \ln^3(s + i\varepsilon) = (2\pi i)^2 [6 \ln(s + i\varepsilon) - 12\pi i]\tag{5.4.11}$$

and

$$\text{Disc}_s \text{Disc}_s \text{Disc}_s \ln^3(s + i\varepsilon) = 6(2\pi i)^3 .\tag{5.4.12}$$

If we take any further discontinuities of $\ln^3(s + i\varepsilon)$ we get zero. It is worth emphasizing here that Disc does *not* in general satisfy the product rule

$$\text{Disc}(AB) \neq A \text{Disc} B + B \text{Disc} A .\tag{5.4.13}$$

The discontinuity operator computes a finite difference around a branch point, which is not an infinitesimal differential in any sense.

In summary, we have seen that for powers of logarithms, we can compute sequential discontinuities by identifying the $\pm i\varepsilon$ prescription with integration contours that end on different Riemann sheets, and the discontinuity across the cut with the monodromy around

the branch point. In general, the transcendental functions that show up in scattering amplitudes are more complicated than logarithms, and depend on many Mandelstam invariants with many branch points. Understanding the monodromy group of these more complicated functions will help us untangle their analytic structure, and thereby help us compute their sequential discontinuities. Correspondingly, we now turn to a systematic procedure for computing the generators of the monodromy group associated with a general polylogarithmic function.

5.4.2 The monodromy group

Given a function defined by a contour integral, we can determine the effect of an analytically continuing around one of its branch points by integrating along a closed contour that encircles this branch point. The integrals along these closed contours are referred to as the monodromies of the function, and form a group. By computing an explicit representation of this group, we can compute the value of this function anywhere in its maximally analytically continued domain. We illustrate how this group can be systematically computed, by working through some examples.

One branch point

Let us first return to the example of powers of logarithms $\ln^n s$, for any positive integer n . As the discontinuities of $\ln^n s$ involve lower powers of $\ln s$, we consider all powers up to n

simultaneously. The total differential of these functions is

$$d\left(\frac{\ln^n s}{n!}\right) = \left(\frac{\ln^{n-1} s}{(n-1)!}\right) \frac{ds}{s}, \quad (5.4.14)$$

where we have normalized $\ln^n s$ by a factor of $n!$ for convenience. Let's take $n = 3$ for concreteness and collect the functions that appear in the derivatives of $\ln^3 s$ into a vector

$$\mathcal{V} \equiv \begin{pmatrix} 1 & \ln s & \frac{1}{2} \ln^2 s & \frac{1}{3!} \ln^3 s \end{pmatrix}. \quad (5.4.15)$$

The differential relations in Eq. (5.4.14) can then be put in the matrix form

$$d\mathcal{V} = \mathcal{V} \cdot \omega, \quad (5.4.16)$$

where the *connection* ω is an $(n+1) \times (n+1)$ matrix defined on $\mathbb{C}^* \equiv \mathbb{C} \setminus \{0\}$ whose entries are one-forms:

$$\omega = \begin{pmatrix} 0 & \frac{ds}{s} & 0 & 0 \\ 0 & 0 & \frac{ds}{s} & 0 \\ 0 & 0 & 0 & \frac{ds}{s} \\ 0 & 0 & 0 & 0 \end{pmatrix}. \quad (5.4.17)$$

As we analytically continue $\ln^n s$ around $s = 0$, the vector of functions \mathcal{V} will mix with other functions that, like \mathcal{V} , satisfy the differential equation in Eq. (5.4.16). These other functions have lower transcendental weight, and the mixing coefficients will be proportional to powers of $i\pi$. Thus, general solutions to Eq. (5.4.16) will contain all the possible informa-

tion about the monodromies of the function.

As there are $n + 1$ independent solutions to Eq. (5.4.16), we can group these solutions into an upper-triangular matrix \mathcal{M} called the *variation matrix*, which we normalize to have 1's along the diagonal. The variation matrix on the principal branch of the logarithm for $n = 3$ can be written as

$$\mathcal{M}_{\gamma_0}(s) = \begin{pmatrix} 1 & \ln s & \frac{1}{2} \ln^2 s & \frac{1}{3!} \ln^3 s \\ 0 & 1 & \ln s & \frac{1}{2} \ln^2 s \\ 0 & 0 & 1 & \ln s \\ 0 & 0 & 0 & 1 \end{pmatrix}. \quad (5.4.18)$$

Variation matrices have a close connection to the coproduct structure often utilized in Feynman integral calculations. Further discussion of this connection is given in Appendix B.1.

To extend the variation matrix in Eq. (5.4.18) beyond the cut complex plane, we need to determine the effect of deforming the integration contour defining its entries around their branch points. Although this extension changes the value of the function at s , the differentials of the function will still be related by the differential equation Eq. (5.4.16). Since the general solution to this differential equation are linear combinations of the rows of the variation matrix, we can interpret the action of the monodromy as multiplication of the variation matrix by another matrix, the monodromy matrix.

The most general solution to the differential equation in Eq. (5.4.16) is given by

$$\mathcal{M}_\gamma(s) = \mathcal{P} \exp \left(\int_\gamma \omega \right), \quad (5.4.19)$$

where $\mathcal{P} \exp(\int_\gamma \omega)$ is a path-ordered exponential along the path γ starting at 1 and ending at s . For a given contour from a to b , the path-ordered exponential is defined by

$$\mathcal{P} \exp \left(\int_a^b \omega \right) = 1 + \int_a^b \omega + \int_a^b \omega \circ \omega + \cdots \quad (5.4.20)$$

where $\int_a^b \omega \circ \omega$ denotes an iterated integral. Since ω is a matrix, $\omega \circ \omega$ implies matrix multiplication:

$$\int_a^b \omega \circ \omega = \int_{a \leq t_1 \leq t_2 \leq b} \omega_{ik}(t_1) \omega_{kj}(t_2). \quad (5.4.21)$$

Here we have made the matrix indices explicit for clarity, and k is to be summed over. Note that the expansion in powers of ω is finite since ω is nilpotent.

For differential forms in several variables $x = (x_1, \dots, x_n)$, these iterated integrals are defined as follows. First, we choose a path γ parametrized by $t \in [0, 1]$ and defined by $(x_1(t), \dots, x_n(t))$. Then, given some differential forms $\xi_1(x), \dots, \xi_k(x)$ in the variables x , we can pull them back to the path γ , whereupon they become differential forms $\gamma^* \xi_i(t)$ in the variable t parameterizing the path. The iterated integral of these forms along γ is defined as

$$\int_\gamma \xi_1(x) \circ \cdots \circ \xi_k(x) = \int_{0 \leq t_1 \leq \cdots \leq t_k \leq 1} \gamma^* \xi_1(t_1) \cdots \gamma^* \xi_k(t_k). \quad (5.4.22)$$

We discuss how to evaluate integrals of this type in more detail in Appendix B.5.

Given an integration contour γ that ends at x , the value of $\mathcal{M}_\gamma(x)$ can be computed by integrating the path-ordered exponential in Eq. (5.4.19). We can split up any path γ between the basepoint (where the integration starts) and x into a contour γ_0 that goes from the basepoint to x without encircling any of branch points (the poles in ω), and a series of contours $\{\gamma'_j\}$ that each begin and end at x and encircle one of the branch points of ω . That is, we have $\gamma = \gamma_0 \circ \gamma'_{i_1} \circ \cdots \circ \gamma'_{i_n}$, where $\gamma_a \circ \gamma_b$ denotes the composition of paths in which we first run along the path γ_a and then along the path γ_b . A very useful feature of defining matrices as path-ordered exponentials is that composing two paths corresponds to matrix multiplication. So

$$\mathcal{M}_\gamma = \mathcal{M}_{\gamma_0 \circ \gamma'_{i_1} \circ \cdots \circ \gamma'_{i_n}} = \mathcal{M}_{\gamma_0} \cdot \mathcal{M}_{\gamma'_{i_1}} \cdots \mathcal{M}_{\gamma'_{i_n}} \quad (5.4.23)$$

Now, this contour can also be written $\gamma = (\gamma_0 \circ \gamma'_{i_1} \circ \gamma_0^{-1}) \circ \cdots \circ (\gamma_0 \circ \gamma'_{i_n} \circ \gamma_0^{-1}) \circ \gamma_0$, where $\gamma_{i_k} = \gamma_0 \circ \gamma'_{i_k} \circ \gamma_0^{-1}$ encircles the same poles as γ'_{i_k} but starts and ends at the same basepoint as γ_0 rather than starting and ending at x . Hence, we can also write

$$\mathcal{M}_\gamma = \mathcal{M}_{\gamma_{i_1} \circ \cdots \circ \gamma_{i_n} \circ \gamma_0} = \mathcal{M}_{\gamma_{i_1}} \cdots \mathcal{M}_{\gamma_{i_n}} \cdot \mathcal{M}_{\gamma_0} \quad (5.4.24)$$

where we are now *prepending* closed contour integrals from a common basepoint onto the integration contour before we arrive at x . This convention ensures that the monodromy matrices are independent of the endpoint x . In summary, to compute the monodromy from x along the path γ'_1 followed by γ'_2 we first multiply on the left by \mathcal{M}_{γ_2} followed by multiplication on the left of the result by \mathcal{M}_{γ_1} where the paths γ_1 and γ_2 start and end at the

basepoint independent of x .

In the case of $\ln^n s$ there is only a single branch point at the origin. The contour γ_0 can be taken to be the straight path from 1 to s , except when s lies on the negative real axis, in which case we deform the path γ_0 to go just above the branch point at zero. Then one can check that the variation matrix \mathcal{M}_{γ_0} in Eq. (5.4.18) is exactly $\mathcal{P} \exp \int_{\gamma_0} \omega$ along this path (see Eq. (5.4.29) below). Since there is only one branch point, we define paths γ_+ and γ_- that encircle the origin counterclockwise or clockwise with unit radius. We can thus decompose a general path γ into some number of iterations of γ_+ or γ_- , followed by γ_0 , namely $\gamma = \gamma_+ \circ \cdots \circ \gamma_+ \circ \gamma_0$ or $\gamma = \gamma_- \circ \cdots \circ \gamma_- \circ \gamma_0$.

Given a member γ_k of the equivalence class of contours that encircle the origin k times clockwise and end at s , we have

$$\mathcal{M}_{\gamma_k}(s) = (\mathcal{M}_{\gamma_-})^k \cdot \mathcal{M}_{\gamma_0}(s). \quad (5.4.25)$$

The matrix \mathcal{M}_{γ_-} can thus be seen to be a generator of the monodromy group, since it maps $\ln^n s$ to its value after encircling the branch point $s = 0$ one more time.

Since we have specified their integration contours, $\mathcal{M}_{\gamma_0}(s)$ and \mathcal{M}_{γ_-} can be computed directly. To calculate \mathcal{M}_{γ_-} , we parametrize the path γ_- by $s = \exp(-i\theta)$ for $\theta \in [0, 2\pi]$.

This gives us $\frac{ds}{s} = -id\theta$, and thus

$$\int_{\gamma_-} \underbrace{\frac{ds}{s} \circ \cdots \circ \frac{ds}{s}}_j = \frac{(-2\pi i)^j}{j!}. \quad (5.4.26)$$

The analogous set of integrals over γ_0 just return the logarithms we started with, namely

$$\int_{\gamma_0} \underbrace{\frac{ds}{s} \circ \dots \circ \frac{ds}{s}}_j = \frac{\ln^j(s)}{j!}. \quad (5.4.27)$$

Expanding the path-ordered exponentials and evaluating the iterated integrals as described above on the connection in Eq. (5.4.17) for $\ln^3 s$, we find

$$\mathcal{M}_{\gamma_-}(s) = \mathbb{1} + \int_{\gamma_-} \omega + \int_{\gamma_-} \omega \circ \omega + \int_{\gamma_-} \omega \circ \omega \circ \omega = \begin{pmatrix} 1 & -2\pi i & \frac{1}{2}(-2\pi i)^2 & \frac{1}{3!}(-2\pi i)^3 \\ 0 & 1 & -2\pi i & \frac{1}{2}(-2\pi i)^2 \\ 0 & 0 & 1 & -2\pi i \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (5.4.28)$$

and

$$\mathcal{M}_{\gamma_0}(s) = \mathbb{1} + \int_{\gamma_0} \omega + \int_{\gamma_0} \omega \circ \omega + \int_{\gamma_0} \omega \circ \omega \circ \omega = \begin{pmatrix} 1 & \ln s & \frac{1}{2} \ln^2 s & \frac{1}{3!} \ln^3 s \\ 0 & 1 & \ln s & \frac{1}{2} \ln^2 s \\ 0 & 0 & 1 & \ln s \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad (5.4.29)$$

in agreement with Eq. (5.4.18).

Using Eq. (5.4.25), we can then compute the effect of going around the branch point by

multiplying these matrices. For example, we can calculate the first discontinuity by

$$(\mathbf{1} - \mathcal{M}_{\gamma_-}) \cdot \mathcal{M}_{\gamma_0}(s) = - \begin{pmatrix} 0 & -2\pi i & -2\pi i \ln s + \frac{(-2\pi i)^2}{2} & \frac{-2\pi i}{2} \ln^2 s + \frac{(-2\pi i)^2}{2} \ln s + \frac{(-2\pi i)^3}{3!} \\ 0 & 0 & -2\pi i & -2\pi i \ln s + \frac{(-2\pi i)^2}{2} \\ 0 & 0 & 0 & -2\pi i \\ 0 & 0 & 0 & 0 \end{pmatrix}. \quad (5.4.30)$$

The discontinuity of $\ln^3 s$ is then $3!$ times the top-right entry of this matrix, in agreement with Eq. (5.4.10).

More generally, under the action of \mathcal{M}_{γ_-} the entry in the first row and last column of $\mathcal{M}_{\gamma_0}(s)$ transforms as

$$\mathcal{M}_{\gamma_-} \frac{\ln^n(s)}{n!} = \sum_{k=0}^n \frac{\ln^{n-k}(s)}{(n-k)!} \frac{(-2\pi i)^k}{k!}. \quad (5.4.31)$$

Here we are generalizing notation slightly by having \mathcal{M}_{γ_-} act on a function rather than the variation matrix in which it is the upper-right entry. Thus, the discontinuity is

$$\text{Disc}_s \ln^n s = (\mathbf{1} - \mathcal{M}_{\gamma_-}) \ln^n s = - \sum_{k=1}^n n! \frac{\ln^{n-k}(s)}{(n-k)!} \frac{(-2\pi i)^k}{k!}. \quad (5.4.32)$$

This agrees with what we get using the substitution $\ln(s - i\varepsilon) = \ln(s + i\varepsilon) - 2\pi i$, as we did

for instance in Eq. (5.4.9), which gives us

$$\text{Disc}_s \ln^n(s + i\varepsilon) = \ln^n(s + i\varepsilon) - \ln^n(s - i\varepsilon) = - \sum_{k=1}^n \binom{n}{k} \ln^{n-k}(s + i\varepsilon) (-2\pi i)^k \quad (5.4.33)$$

for arbitrary n .

Further discontinuities can be computed by acting with the same operator $\mathbb{1} - \mathcal{M}_{\gamma_-}$.

For later reference, we list here some general formulas that can be derived either using the substitution method or with the use of monodromy matrices:

$$\text{Disc}_s \text{Disc}_s \ln^n(s + i\varepsilon) = \ln^n(s + i\varepsilon) - 2[\ln(s + i\varepsilon) - 2\pi i]^n + [\ln(s + i\varepsilon) - 4\pi i]^n \quad (5.4.34)$$

$$= \sum_{k=1}^n (2^k - 2) \binom{n}{k} \ln^{n-k}(s + i\varepsilon) (-2\pi i)^k. \quad (5.4.35)$$

Similarly, the formula for m discontinuities is

$$\text{Disc}_s^m \ln^n(s + i\varepsilon) = \sum_{\ell=0}^m (-1)^\ell \binom{m}{\ell} [\ln(s + i\varepsilon) - \ell 2\pi i]^n \quad (5.4.36)$$

$$= (-1)^m m! \sum_{k=1}^n \left\{ \begin{matrix} k \\ m \end{matrix} \right\} \binom{n}{k} \ln^{n-k}(s + i\varepsilon) (-2\pi i)^k \quad (5.4.37)$$

where

$$\left\{ \begin{matrix} k \\ m \end{matrix} \right\} = \frac{1}{m!} \sum_{\ell=1}^m (-1)^{m-\ell} \binom{m}{\ell} \ell^k \quad (5.4.38)$$

are the Stirling numbers of second kind. These numbers have a useful combinatorial interpretation: $\{m^k\}$ is the number of ways of partitioning a set of k elements into m non-empty sets.

Multiple branch points

Let us now consider an example involving two branch points, the dilogarithm

$$\text{Li}_2(s) \equiv \sum_{n=1}^{\infty} \frac{s^n}{n^2} \quad \text{for } |s| < 1. \quad (5.4.39)$$

Similar to the definition of the logarithm in Eq. (5.4.1), this power series definition is only convergent in the region $|s| < 1$, but can be uniquely continued to the rest of the cut complex plane, where the branch cut is usually placed on the positive real axis running from 1 to ∞ . The dilogarithm can also be given by an integral definition,

$$\text{Li}_2(s) \equiv \int_0^s \frac{dx}{x} \text{Li}_1(x), \quad \text{with} \quad \text{Li}_1(s) \equiv \int_0^s \frac{dx}{1-x} = -\ln(1-s), \quad (5.4.40)$$

We write the integral in terms of $\text{Li}_1(s)$ rather than $-\ln(1-s)$ to make the singularities more transparent, as $\text{Li}_1(s)$ and $\text{Li}_2(s)$ both have branch points at $s = 1$, with a branch cut conventionally going from 1 to ∞ along the positive real s axis. The standard placement of the branch cut for $\text{Li}_n(s)$, from $1 < s < \infty$ is consistent with the standard branch cut for the logarithm, $s < 0$.

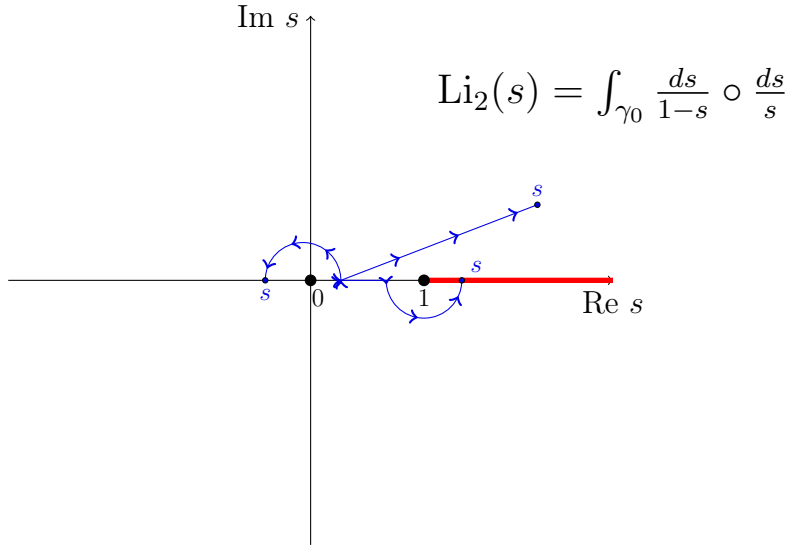


Figure 5.3: $\text{Li}_2(s)$ has branch points at $s = 0$ and $s = 1$. The principal branch of $\text{Li}_2(s)$ has a branch cut on the real line from $s = 1$ to $+\infty$. The standard contour γ_0 in the analytic integral definition of $\text{Li}_2(s)$ begins at a basepoint at $s = \varepsilon > 0$ and proceeds in a straight line to s , diverting in a counterclockwise path around the branch points when necessary.

Using equation (5.4.40), we have

$$d\text{Li}_2(s) = -\frac{ds}{s} \ln(1-s), \quad d\ln(1-s) = -\frac{ds}{1-s}. \quad (5.4.41)$$

We can again put these relations in a matrix form

$$d \begin{pmatrix} 1 & \text{Li}_1(s) & \text{Li}_2(s) \end{pmatrix} = \begin{pmatrix} 1 & \text{Li}_1(s) & \text{Li}_2(s) \end{pmatrix} \cdot \omega, \quad (5.4.42)$$

where

$$\omega = \begin{pmatrix} 0 & \frac{ds}{1-s} & 0 \\ 0 & 0 & \frac{ds}{s} \\ 0 & 0 & 0 \end{pmatrix} \quad (5.4.43)$$

is defined on $\mathbb{C} \setminus \{0, 1\}$.

For $\text{Li}_2(s)$, we take the basepoint to be $s = 0$ and the path γ_0 defining its principal branch to be the straight line from 0 to s , which avoids the branch points at 0 and 1 with a counterclockwise detour if necessary. This is shown in Fig. 5.3. Note that this contour is problematic for the differential $\frac{ds}{s}$, which diverges at the lower integration bound. This can be dealt with using tangential basepoint regularization, which amounts to introducing a cutoff ε on the lower integration limit and dropping the powers of $\ln \varepsilon$ that result (see for instance [204]).⁸ For example,

$$\int_0^s \frac{ds}{s} \circ \frac{ds}{1-s} = \int_\varepsilon^s \frac{ds'}{1-s'} \int_\varepsilon^{s'} \frac{ds''}{s''} = -\text{Li}_2(s) - \ln(1-s) \ln \frac{s}{\varepsilon} \quad (5.4.44)$$

$$\cong -\text{Li}_2(s) - \ln(1-s) \ln s, \quad (5.4.45)$$

where \cong means terms divergent in ε are dropped and then $\varepsilon \rightarrow 0$. Then it is straightfor-

⁸In more detail, this regularization implies a choice of parametrization for the path in which the tangent vector to the path at the basepoint is of length one. The monodromy group is then defined with respect to the paths that satisfy this constraint. In other words, we consider homotopy classes of paths which can be continuously deformed into one another with the tangent at the basepoint being kept constant.

ward to compute the variation matrix by integrating ω along γ_0 :

$$\mathcal{M}_{\gamma_0}(s) = \mathcal{P} \exp \left(\int_{\gamma_0} \omega \right) = \begin{pmatrix} 1 & \text{Li}_1(s) & \text{Li}_2(s) \\ 0 & 1 & \ln s \\ 0 & 0 & 1 \end{pmatrix}. \quad (5.4.46)$$

Note that the this variation matrix encodes precisely the coproduct structure of $\text{Li}_2(s)$,

$$\Delta \text{Li}_2(s) = 1 \otimes \text{Li}_2(s) + \text{Li}_1(s) \otimes \ln s + \text{Li}_2(s) \otimes 1, \quad (5.4.47)$$

as discussed further in Appendix B.1.

We would now like to extend this construction to the maximal analytic continuation of $\text{Li}_2(s)$. As there are multiple branch points, we should in general be careful to distinguish between infinitesimal contours that encircle these branch points, and the full contours that not only wrap around these points but also start and end at our chosen basepoint of integration. For $\ln^n s$ we took the basepoint to be 1, but for all the other functions we study in this paper we will take the basepoint to be 0 (or a small value ϵ on the positive real axis, when regularization is required). We denote the infinitesimal contour in a variable x that encircles the point p counterclockwise by \mathfrak{C}_p^x . In contrast, we denote the path around $x = p$ that starts and ends at the basepoint by \mathfrak{X}_p^x . When the function under study only depends on a single variable x , we will often drop the index indicating which variable the contour is taken in.

The contribution from moving along any contour is computed by evaluating the path-

ordered exponential $\mathcal{P} \exp(\int_{\gamma} \omega)$ on the contour. For the monodromy around 0, we find

$$\mathcal{M}_{\mathbb{S}^1_0} = \mathcal{M}_{\mathbb{D}_0} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 2\pi i \\ 0 & 0 & 1 \end{pmatrix}. \quad (5.4.48)$$

To compute the monodromy matrix associated with the branch point at 1, we first use Eq. (5.4.46) to determine the contribution from the path between 0 and 1, and compute the infinitesimal contour around 1 as before. We find

$$\mathcal{M}_{0 \rightarrow 1} = \begin{pmatrix} 1 & 0 & \frac{\pi^2}{6} \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \mathcal{M}_{\mathbb{D}_1} = \begin{pmatrix} 1 & -2\pi i & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad (5.4.49)$$

where we have dropped all logarithmically-divergent terms in accordance with tangential basepoint regularization. The complete path thus gives

$$\mathcal{M}_{\mathbb{S}^1_1} = \mathcal{M}_{0 \rightarrow 1} \cdot \mathcal{M}_{\mathbb{D}_1} \cdot (\mathcal{M}_{0 \rightarrow 1})^{-1} = \begin{pmatrix} 1 & -2\pi i & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \quad (5.4.50)$$

We highlight again that the action of the monodromy matrices proceeds from left to right; Eq. (5.4.50) computes the effect of moving from 0 to 1 along the real line, rotating counter-clockwise around an infinitesimal contour centered at 1, and then moving back to 0.

Acting with these matrices on \mathcal{M}_{γ_0} allows us to compute any sequence of monodromies

on the functions appearing in \mathcal{M}_{γ_0} . For instance, prepending a monodromy around 0 to the path γ_0 gives

$$\mathcal{M}_{\mathbb{S}^{\circlearrowleft}_0} \cdot \mathcal{M}_{\gamma_0} = \begin{pmatrix} 1 & \text{Li}_1(s) & \text{Li}_2(s) \\ 0 & 1 & \ln s + 2\pi i \\ 0 & 0 & 1 \end{pmatrix}, \quad (5.4.51)$$

while prepending a contour around 1 gives

$$\mathcal{M}_{\mathbb{S}^{\circlearrowleft}_1} \cdot \mathcal{M}_{\gamma_0} = \begin{pmatrix} 1 & \text{Li}_1(s) - 2\pi i & \text{Li}_2(s) - 2\pi i \ln s \\ 0 & 1 & \ln s \\ 0 & 0 & 1 \end{pmatrix}. \quad (5.4.52)$$

These matrices imply that $\text{Li}_1(s)$ and $\text{Li}_2(s)$ only have a monodromy around $s = 1$ while $\ln s$ only has a monodromy around $s = 0$, as expected. We can also now compute the sequential discontinuity of $\text{Li}_2(s)$ by first taking the monodromy around $s = 1$ and then around $s = 0$. As we prepend these contours, this corresponds to

$$(\mathbb{1} - \mathcal{M}_{\mathbb{S}^{\circlearrowleft}_1}) \cdot (\mathbb{1} - \mathcal{M}_{\mathbb{S}^{\circlearrowleft}_0}) \cdot \mathcal{M}_{\gamma_0} = \begin{pmatrix} 0 & 0 & -(2\pi i)^2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad (5.4.53)$$

which tells us that $\text{Disc}_0 \text{Disc}_1 \text{Li}_2(s) = -(2\pi i)^2$. Similarly, we can compute that $(\mathbb{1} - \mathcal{M}_{\mathbb{S}^{\circlearrowleft}_1}) \cdot (\mathbb{1} - \mathcal{M}_{\mathbb{S}^{\circlearrowleft}_0}) \cdot \mathcal{M}_{\gamma_0} = 0$, consistent with the fact that $\text{Li}_2(s)$ does not have a discontinuity around $s = 0$.

Multiple variables

Let us finally turn to an example involving multiple variables. We consider the two-variable function

$$\Phi_1(z, \bar{z}) = 2\text{Li}_2(z) - 2\text{Li}_2(\bar{z}) - \ln(z\bar{z}) \left[\text{Li}_1(z) - \text{Li}_1(\bar{z}) \right]. \quad (5.4.54)$$

This function arises in the one-loop triangle and one-loop box integrals (see Section 5.6.2 below). Here we treat z and \bar{z} as independent variables, so this function is analytic for $|z - \frac{1}{2}| < \frac{1}{2}$ and $|\bar{z} - \frac{1}{2}| < \frac{1}{2}$. Following the same steps as in our previous examples, we first compute

$$d\Phi_1 = \left(\frac{dz}{z} - \frac{d\bar{z}}{\bar{z}} \right) (\text{Li}_1(z) + \text{Li}_1(\bar{z})) - \left(\frac{dz}{1-z} - \frac{d\bar{z}}{1-\bar{z}} \right) \ln(z\bar{z}). \quad (5.4.55)$$

This can be put in the matrix form $d\mathcal{M}_{\gamma_0} = \mathcal{M}_{\gamma_0} \cdot \omega$, where

$$\omega = \begin{pmatrix} 0 & \frac{dz}{z} + \frac{d\bar{z}}{\bar{z}} & \frac{dz}{1-z} + \frac{d\bar{z}}{1-\bar{z}} & 0 \\ 0 & 0 & 0 & -\frac{dz}{1-z} + \frac{d\bar{z}}{1-\bar{z}} \\ 0 & 0 & 0 & \frac{dz}{z} - \frac{d\bar{z}}{\bar{z}} \\ 0 & 0 & 0 & 0 \end{pmatrix}. \quad (5.4.56)$$

The connection ω is well-defined in $\mathbb{C}^2 \setminus \{z = 0, z = 1, \bar{z} = 0, \bar{z} = 1\}$, so there are now four codimension-one branching varieties.

We can define a path γ_0 between the basepoint $(0, 0)$ and (z, \bar{z}) in the same way we did

for $\text{Li}_2(s)$, namely we use straight line paths, except when z or \bar{z} are on the real line outside of $(0, 1)$, when we go counterclockwise around the branch points. Integrating along this path gives the variation matrix on the principal branch. The result is

$$\mathcal{M}_{\gamma_0}(z, \bar{z}) = \mathcal{P} \exp \left(\int_{\gamma_0} \omega \right) = \begin{pmatrix} 1 & \ln(z\bar{z}) & \text{Li}_1(z) + \text{Li}_1(\bar{z}) & \Phi_1(z, \bar{z}) \\ 0 & 1 & 0 & -\text{Li}_1(z) + \text{Li}_1(\bar{z}) \\ 0 & 0 & 1 & \ln(z/\bar{z}) \\ 0 & 0 & 0 & 1 \end{pmatrix}. \quad (5.4.57)$$

Note that the antisymmetry of $\Phi_1(z, \bar{z})$ in its arguments is encoded in the matrices \mathcal{M}_{γ_0} and ω by the action of conjugation by $\text{diag}(1, 1, 1, -1)$, namely

$$\text{diag}(1, 1, 1, -1) \cdot \mathcal{M}_{\gamma_0}(z, \bar{z}) \cdot \text{diag}(1, 1, 1, -1) = \mathcal{M}_{\gamma_0}(\bar{z}, z). \quad (5.4.58)$$

Further, it can be checked that the connection ω is closed ($d\omega = 0$) and flat ($d\omega - \omega \wedge \omega = 0$). These requirements were trivially satisfied in the preceding one-variable examples, but guarantee that the functions appearing in $\mathcal{P} \exp \int_{\gamma} \omega$ only depend on the homotopy class of γ . Further discussion of this point can be found in Appendix B.5.

We now compute the monodromy matrices associated with the branch points at 0 and 1 in both z and \bar{z} by evaluating the path-ordered exponential (5.4.20) on cycles that encircle

each of these four poles. First, we compute

$$\mathcal{M}_{\mathbb{S}^{\circlearrowleft}_0^z} = \mathcal{M}_{\mathbb{S}_0^z} = \begin{pmatrix} 1 & 2\pi i & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 2\pi i \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad \mathcal{M}_{\mathbb{S}^{\circlearrowleft}_0^{\bar{z}}} = \mathcal{M}_{\mathbb{S}_0^{\bar{z}}} = \begin{pmatrix} 1 & 2\pi i & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -2\pi i \\ 0 & 0 & 0 & 1 \end{pmatrix}. \quad (5.4.59)$$

To compute the monodromy matrices associated with contours around 1 we need

$$\mathcal{M}_{0 \rightarrow 1}^z = \begin{pmatrix} 1 & 0 & 0 & 2\zeta_2 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad \mathcal{M}_{\mathbb{S}_1^z} = \begin{pmatrix} 1 & 0 & -2\pi i & 0 \\ 0 & 1 & 0 & 2\pi i \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad (5.4.60)$$

$$\mathcal{M}_{0 \rightarrow 1}^{\bar{z}} = \begin{pmatrix} 1 & 0 & 0 & -2\zeta_2 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad \mathcal{M}_{\mathbb{S}_1^{\bar{z}}} = \begin{pmatrix} 1 & 0 & -2\pi i & 0 \\ 0 & 1 & 0 & -2\pi i \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \quad (5.4.61)$$

Putting these paths together, we find

$$\mathcal{M}_{\searrow\circlearrowleft_1^z} = \mathcal{M}_{0\rightarrow 1}^z \cdot \mathcal{M}_{\circlearrowleft_1^z} \cdot (\mathcal{M}_{0\rightarrow 1}^z)^{-1} = \begin{pmatrix} 1 & 0 & -2\pi i & 0 \\ 0 & 1 & 0 & 2\pi i \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad (5.4.62)$$

$$\mathcal{M}_{\searrow\circlearrowleft_1^{\bar{z}}} = \mathcal{M}_{0\rightarrow 1}^{\bar{z}} \cdot \mathcal{M}_{\circlearrowleft_1^{\bar{z}}} \cdot (\mathcal{M}_{0\rightarrow 1}^{\bar{z}})^{-1} = \begin{pmatrix} 1 & 0 & -2\pi i & 0 \\ 0 & 1 & 0 & -2\pi i \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \quad (5.4.63)$$

Note that the matrices that encode monodromies in the variable z commute with the matrices that encode monodromies in the variable \bar{z} .

These matrices allow us to compute monodromies of $\Phi_1(z, \bar{z})$ and the other functions appearing in Eq. (5.4.57) anywhere in their domain, and therefore to compute sequential discontinuities in z or \bar{z} (and correspondingly the kinematic invariants of the triangle or box diagrams). For example, to compute a sequential discontinuity in z around 1 and then 0, we would evaluate

$$(\mathbb{1} - \mathcal{M}_{\searrow\circlearrowleft_1^z}) \cdot (\mathbb{1} - \mathcal{M}_{\searrow\circlearrowleft_0^z}) \cdot \mathcal{M}_{\gamma_0} = \begin{pmatrix} 0 & 0 & 0 & -(2\pi i)^2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \quad (5.4.64)$$

Taking these discontinuities in a different order, we get a different result

$$(\mathbb{1} - \mathcal{M}_{\mathbb{S} \odot_0^z}) \cdot (\mathbb{1} - \mathcal{M}_{\mathbb{S} \odot_1^z}) \cdot \mathcal{M}_{\gamma_0} = \begin{pmatrix} 0 & 0 & 0 & (2\pi i)^2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \quad (5.4.65)$$

It is also possible to take a discontinuity around both branch points by considering the monodromy matrix associated with ∞ . We construct this monodromy matrix and discuss the full monodromy group in Appendix B.2.

As long as we analytically continue along paths which are fully contained in the Euclidean region, we never encounter branch singularities and the functions we consider are single-valued. The variation matrix approach lends itself well to the description of single-valued functions, and in Appendix B.3 we describe a general construction that builds a single-valued version of any generalized polylogarithm from its variation matrix.

Finally, let us highlight that the variation matrices we associate with Feynman integrals using the above construction are not the same as the matrices considered in [169, 170]. The matrices considered in those works look similar, insofar as they are lower-diagonal matrices in which a Feynman integral sits in the bottom-left entry and various cuts of the Feynman integral appear in other entries. However, in that approach *several* lower-diagonal matrices are associated with each graph. For example, they associate six such matrices with the triangle diagram, labeled by spanning trees whose edges are ordered. The variation matrices we consider are useful for determining how the transcendental functions appearing in them

transform under monodromies. It is unclear to us if the matrices described in [169, 170] can be used in the same way, and if they can, what representations of the monodromy group they furnish.

5.4.3 Monodromies of propagators

We have seen that the $\pm i\varepsilon$ notation is good for describing where we are on the principal branch of multivalued functions, where they describe being on opposite sides of a branch cut. We have also seen that discontinuities across the branch cut can be recast using monodromies around the branch point where the cut begins.

In the case of the logarithm, we recall that this amounts to identifying

$$\ln(s + i\varepsilon) \equiv \ln_{\gamma_0} s, \quad \ln(s - i\varepsilon) \equiv \ln_{\gamma_{-1}} s, \quad (5.4.66)$$

where γ_0 is homotopic to the straight path from 1 to s , and γ_{-1} is given by a path that first crosses the real negative axis before ending at s , as shown in Fig. 5.2. With these identifications, we have that $\ln(s - i\varepsilon) = \ln(s + i\varepsilon) - 2\pi i$ for all values of s . Using this identity, we can compute the discontinuity of not only $\ln(s + i\varepsilon)$, but also $\ln(s - i\varepsilon)$, finding

$$\text{Disc}_s \ln(s - i\varepsilon) = \text{Disc}_s [\ln(s + i\varepsilon) - 2\pi i] = 2\pi i. \quad (5.4.67)$$

This can be rewritten in a more suggestive manner:

$$\text{Disc}_s \ln(s - i\varepsilon) = \ln(s - i\varepsilon) - [\ln(s - i\varepsilon) - 2\pi i] = \ln_{\gamma_{-1}} s - \ln_{\gamma_{-2}} s. \quad (5.4.68)$$

Thus, when we take the discontinuity of $\ln(s - i\varepsilon)$, we are not computing the difference between its value and the value of $\ln(s + i\varepsilon)$. Rather, we are computing the difference between analytically continuing $\ln_{\gamma_0}s$ around the origin of s once versus twice.

For sequential discontinuities, the contour definitions are particularly helpful as they allow us to migrate away from the principal branch where $\pm i\varepsilon$ is applicable. Recall however that all the $\pm i\varepsilon$ displacements in Feynman integrals originate in the $\pm i\varepsilon$ displacement of the poles in TOPT propagators. Thus, just as we were able to identify higher winding number versions of $\ln(s \pm i\varepsilon)$ using different integration contours, we should be able to identify higher winding number versions of propagators. To do so, recall that propagator comes originally from a semi-infinite integral over time

$$\frac{\textcolor{blue}{1}}{\textcolor{blue}{E} + i\varepsilon} = -i \int_0^{\infty} dt e^{iEt}, \quad \frac{\textcolor{red}{1}}{\textcolor{red}{E} - i\varepsilon} = -i \int_0^{-\infty} dt e^{iEt}, \quad (5.4.69)$$

so that

$$\frac{\textcolor{blue}{1}}{\textcolor{blue}{E} + i\varepsilon} - \frac{\textcolor{red}{1}}{\textcolor{red}{E} - i\varepsilon} = -i \int_{-\infty}^{\infty} dt e^{iEt} = \textcolor{green}{-2\pi i \delta(E)}. \quad (5.4.70)$$

Thus, for the propagator the integration path goes from $t = 0$ to $t = \pm\infty$ and the $\pm i\varepsilon$ is shorthand for this integration path. We can correspondingly take sequential discontinuities of products of propagators in the same way as for logarithms. To do so, we introduce the notation

$$\textcolor{green}{D_j} = \textcolor{green}{-2\pi i \delta(E_j - \omega_j)} \quad (5.4.71)$$

and

$$P_j^{(n)} = \frac{1}{E_j - \omega_j + i\varepsilon} + nD_j, \quad (5.4.72)$$

where we call this distribution a propagator with winding number n . The propagators we are used to seeing correspond to $P_j^{(0)} = \frac{1}{E_j - \omega_j + i\varepsilon}$ and $P_j^{(-1)} = \frac{1}{E_j - \omega_j - i\varepsilon}$.

In this notation, a TOPT amplitude and its conjugate are

$$M = \prod_{j=1}^n P_j^{(0)}, \quad \overline{M} = \prod_{j=1}^n P_j^{(-1)}. \quad (5.4.73)$$

The TOPT cutting rules in Eq. (5.2.19) become

$$\text{Disc}_{\text{tot}} M = M - \overline{M} = \sum_{j=1}^n \left(\prod_{k=1}^{j-1} P_k^{(0)} \right) D_j \left(\prod_{k=j+1}^n P_k^{(-1)} \right). \quad (5.4.74)$$

Since Disc_{tot} is a linear operator, this can also be generalized to products of propagators with arbitrary winding numbers:

$$\text{Disc}_{\text{tot}} \prod_{j=1}^n P_j^{(l_j)} = \sum_{j=1}^n \left(\prod_{k=1}^{j-1} P_k^{(l_k)} \right) D_j \left(\prod_{k=j+1}^n P_k^{(l_k-1)} \right). \quad (5.4.75)$$

To take further discontinuities, we just use Eq. (5.4.72) to express propagators with nonzero

winding number in terms of propagators with winding number 0. Then, as in Eq. (5.4.35),

$$\begin{aligned}
\text{Disc}_{\text{tot}}^2 M &= \text{Disc}_{\text{tot}} M - \text{Disc}_{\text{tot}} \overline{M} \\
&= (P_1 \cdots P_n) - 2(P_1 - D_1) \cdots (P_n - D_n) + (P_1 - 2D_1) \cdots (P_n - 2D_n) \\
&= \sum_k (-1)^k (2^k - 2) \left[D_1 \cdots D_k P_{k+1} \cdots P_n + \text{perms} \right]
\end{aligned} \tag{5.4.76}$$

where the sum over permutations in the last bracket corresponds to the $\binom{n}{k}$ choices for which k propagators to replace with delta functions. The analog of Eq. (5.4.37) is

$$\begin{aligned}
(\text{Disc}_{\text{tot}})^m P_1 \cdots P_n &= \sum_{\ell=0}^m (-1)^\ell \binom{m}{\ell} \left[(P_1 - \ell D_1) \cdots (P_n - \ell D_n) \right] \\
&= \sum_{\ell=0}^m (-1)^\ell \binom{m}{\ell} \sum_{k=1}^n \left[P_1 \cdots P_k (\ell D_{k+1}) \cdots (\ell D_n) + \text{perms} \right]. \tag{5.4.77}
\end{aligned}$$

Although the winding numbers have been left implicit in Eq. (5.4.76) and Eq. (5.4.77), these equations are valid for any assignment of winding numbers.

Let us try to briefly summarize this section. We found that to take sequential discontinuities the $\pm i\varepsilon$ language was insufficient. For a single discontinuity, one can compare a function on two sides of a branch cut on the principal branch. However, to take additional discontinuities, one needs an analytic function defined away from the cut itself. A natural way to do that is to treat the discontinuity as a monodromy around the branch point. In the monodromy language, there is no branch cut at all (the branch cut is an artifact of projecting onto a complex plane) and the discontinuity is automatically an analytic function.

Moreover, monodromies can be computed in an algebraic way using a variation matrix and a connection. Finally, we saw that the monodromy picture led to a natural generalization of the $+i\varepsilon$ propagator to a family of propagators with additional winding numbers. These propagators will be used in the derivation of the relation between multiple cuts and sequential discontinuities, to which we now return.

5.5 Sequential discontinuities

We saw in Section 5.3 that an advantage of TOPT over the covariant formalism is that one can directly identify the origin of singularities in a particular channel. Propagators in a given TOPT diagram depend on a sequence of energies, $E_{I_1} \rightarrow \cdots \rightarrow E_{I_n}$, and each propagator will only lead to a singularity in the integration region if the corresponding energy and invariant are non-negative ($E_I \geq 0$ and $s_I = P_I^2 \geq 0$). We then saw in Section 5.4 that, while the $\pm i\varepsilon$ notation is sufficient to identify the two sides of a branch cut for taking a single discontinuity, for sequential discontinuities it proves useful to think in terms of branch points and monodromies. We now make use of these tools to derive formulas for the sequential discontinuities of Feynman integrals in terms of cuts.

If we work in a region R^s , where only a single invariant $s = s_I = P_I^2$ with $P_I^\mu = \sum_{i \in I} P_i^\mu$ is positive, then we can drop the $i\varepsilon$ in all TOPT propagators not involving the energy associated with this invariant. To make the equations in this section more transparent, we denote the energy and momentum associated with the s channel by $E_s = E_{P_I}$ and $P_s = P_I$.

In this notation, a generic TOPT diagram in R^s takes the form

$$M = \prod_{P_i \neq P_s} \frac{1}{E_{P_i} - \omega_i} \left[\frac{1}{E_s - \omega_1 + i\varepsilon} \cdots \frac{1}{E_s - \omega_n + i\varepsilon} \right]. \quad (5.5.1)$$

In this region, the discontinuity in the s channel is the same as the total discontinuity:

$$\begin{aligned} [\text{Disc}_s M]_{R^s} &= [\text{Disc}_{\text{tot}} M]_{R^s} \\ &= \prod_{P_i \neq P_s} \frac{1}{E_{P_i} - \omega_i} \sum_j \frac{1}{E_s - \omega_1 + i\varepsilon} \cdots (-2\pi i) \delta(E_s - \omega_j) \cdots \frac{1}{E_s - \omega_n - i\varepsilon}, \end{aligned} \quad (5.5.2)$$

The second equality comes from applying the TOPT cutting rules in Eq. (5.2.19) to all propagators, or equivalently just to the propagators involving E_s , as all the delta functions involving other sums of energies evaluate to zero. Summing over all TOPT graphs with a given topology then gives the discontinuity of the corresponding Feynman integral, $[\text{Disc}_s \mathcal{M}]_{R^s}$ from Eq. (5.3.15).

Before taking further discontinuities, let us pause to clarify the role being played by the region R^s in Eq. (5.5.2). In principle, the discontinuity operator Disc_s that appears in this equation can be applied anywhere in the maximal analytic domain of the function M . On the other hand, the relation between $\text{Disc}_s M$ and cut integrals in Eq. (5.5.2) only holds in regions where these cuts are allowed, and only when appropriate analytic continuation paths from R^s to R^\star are used to take this discontinuity. This requirement, that the analytic continuation path starts in the region where the cuts are being computed and only passes through an adjacent region, will become even more important when we compute sequential

discontinuities below. For instance, in the triangle and box ladder integrals we will consider in Section 5.6.2, we will see there are multiple ways of encircling branch points in the z and \bar{z} variables used there that correspond to encircling the branch point in a given Mandelstam invariant; however, only some of these monodromies in z and \bar{z} can be accessed via paths that pass through the appropriate regions. Thus, while we can compute the discontinuities of \mathcal{M} in arbitrary regions, these discontinuities must be evaluated in the appropriate region and using appropriate contours to be related to cuts. For instance, $\text{Disc}_s \mathcal{M}$ can be computed (and will in general be nonzero) in the Euclidean region, where the cuts of \mathcal{M} are zero. However, it is perfectly valid for us to analytically continue the discontinuity that has been computed using the right monodromy matrices in the Euclidean region to the region R^s , where it must satisfy Eq. (5.5.2).

5.5.1 Sequential discontinuities in the same channel

We are now ready to consider discontinuities of discontinuities. To take a second discontinuity of Eq. (5.5.2) in the s channel we can simply rotate all the energies around the same path as for the first discontinuity. This gives

$$\begin{aligned}
[\text{Disc}_s^2 M]_{R^s} = & \prod_{P_i \neq P_s} \frac{1}{E_{P_i} - \omega_i} \sum_j \sum_k P^{(0)}(E_s - \omega_1) \cdots P^{(0)}(E_s - \omega_{j-1}) \\
& \times (-2\pi i) \delta(E_s - \omega_j) P^{(-1)}(E_s - \omega_{j+1}) \cdots P^{(-1)}(E_s - \omega_{k-1}) \\
& \times (-2\pi i) \delta(E_s - \omega_k) P^{(-2)}(E_s - \omega_{k+1}) \cdots P^{(-2)}(E_s - \omega_n). \quad (5.5.3)
\end{aligned}$$

In words, the first cut turns the $+i\varepsilon$ propagators, denoted $P^{(0)}$, to $-i\varepsilon$ propagators, denoted $P^{(-1)}$. The second cut turns the $P^{(-1)}$ propagators into $P^{(-2)}$ ones.

To make sense of the $P^{(-2)}(E)$ propagators, we rewrite them using Eq. (5.4.72),

$$P^{(-2)}(E) = \frac{1}{E + i\varepsilon} - 2(-2\pi i)\delta(E). \quad (5.5.4)$$

To avoid any ambiguity, we also substitute $P^{(-1)}(E) = \frac{1}{E + i\varepsilon} - (-2\pi i)\delta(E)$. The result is a sum over cutting different numbers of s -channel propagators, in which each non-cut propagator is in the region corresponding to $+i\varepsilon$. Explicitly, we get

$$\begin{aligned} [\text{Disc}_s^2 M]_{R^s} &= \prod_{P_i \neq P} \frac{1}{E_{P_i} - \omega_j} \sum_{k=2}^n (-1)^k (2^k - 2) \\ &\times \left[(-2\pi i)\delta(E_s - \omega_1) \cdots (-2\pi i)\delta(E_s - \omega_k) \frac{1}{E_s - \omega_{k+1} + i\varepsilon} \cdots \frac{1}{E_s - \omega_n + i\varepsilon} + \text{perms} \right], \end{aligned} \quad (5.5.5)$$

similar to Eq. (5.4.76).

Summing over the double discontinuities of all TOPT diagrams with the same topology, we get the double discontinuity of the associated Feynman integral. Recall that each delta function in a TOPT diagram directly corresponds to a Feynman diagram cut. As such, we can extract the combinatorial factor from Eq. (5.5.5) and directly compute the cut Feynman diagram with all $+i\varepsilon$ propagators. Doing so, we get

$$[\text{Disc}_s^2 \mathcal{M}]_{R^s} = \left[\sum_{k=2} (-1)^k (2^k - 2) \mathcal{M}_{k\text{-cuts}} \right]_{R_+^s}, \quad (5.5.6)$$

where $\mathcal{M}_{k\text{-cuts}}$ is the sum over all possible ways to cut \mathcal{M} exactly k times, and R_+^s indicates

that all uncut propagators have $+i\varepsilon$. Each cut should split the diagram in two, and the sum of momenta flowing across it should be P_s , as the cuts in all other channels vanish in R^s .

The formula for the triple discontinuity can be computed the same way, giving

$$[\text{Disc}_s^3 \mathcal{M}]_{R^s} = \left[\sum_{k=3} (-1)^k (-3^k + 3 \cdot 2^k - 3) \mathcal{M}_{k\text{-cuts}} \right]_{R_+^s}, \quad (5.5.7)$$

and the generalization to m cuts is as in Eq. (5.4.37):

$$[\text{Disc}_s^m \mathcal{M}]_{R^s} = (1 - \mathcal{M}_{\mathbb{S} \circ_0^s})^m \mathcal{M} = m! \sum_{k=m}^{\infty} \left\{ \begin{matrix} k \\ m \end{matrix} \right\} (-1)^{m-k} [\mathcal{M}_{k\text{-cuts}}]_{R_+^s}, \quad (5.5.8)$$

where $\left\{ \begin{matrix} k \\ m \end{matrix} \right\} = \frac{1}{m!} \sum_{\ell=1}^m (-1)^{m-\ell} \binom{m}{\ell} \ell^k$ are the Stirling numbers of the second kind. We emphasize again that this relation holds when all non-cut propagators in $\mathcal{M}_{k\text{-cuts}}$ are taken to be in the region corresponding to $+i\varepsilon$. We have also included the definition of the discontinuity operator in terms of $\mathcal{M}_{\mathbb{S} \circ_0^s}$, which returns the monodromy around $s = 0$. More precisely, this monodromy matrix acts on the variation matrix \mathcal{M}_{γ_0} , which should be computed along paths from the basepoint to R^s . Examples are given in Section 5.6.

5.5.2 Sequential discontinuities in different channels

Next, let us consider how to take sequential discontinuities in different channels. Unlike the case of sequential discontinuities in the same channel, we must now analytically continue along at least two different paths. As before, we insist on using paths that rotate the

external energies while leaving the external three-momenta fixed and respecting energy-momentum conservation. This gives us $n - 1$ independent parameters that we can vary along each analytic continuation path, where n is the number of external particles. One also must make sure that the relevant invariants only encircle their branch points once. In the examples we have explored (see Section 5.6), we have not found these constraints to be overly restrictive. Nevertheless, choosing paths has to be done carefully. While Cauchy's residue theorem guarantees that normal contour integrals only depend on the homology class of the integration contour, iterated integrals in general depend on the homotopy class of the integration path. This means that one can in general find multiple discontinuity operators that give the same first discontinuity, but different sequential discontinuities. This highlights the importance of our prescription for taking discontinuities by analytically continuing through specific kinematic regions. We discuss this ambiguity in more detail in Appendix B.2.

To fix our notation, suppose we want to compute $\text{Disc}_s \text{Disc}_t \mathcal{M}$, where $s = s_I = (P_I)^2$ and $t = s_J = (P_J)^2$ for sets I and J are different momentum invariants. We abbreviate the associated energies and momenta with $E_s = E_{P_I}$, $P_s = P_I$, $E_t = E_{P_J}$, and $P_t = P_J$. We also denote by $R^{\{s,t\}}$ the region in which $s > 0$, $t > 0$, and all other Mandelstam invariants are real and negative. A general TOPT amplitude with n_s propagators in the s channel and n_t propagators in the t channel in the region $R^{\{s,t\}}$ has the form

$$[M]_{R^{\{s,t\}}} = \prod_{P_i \notin \{P_s, P_t\}} \frac{1}{E_{P_i} - \omega_i} \prod_{k=1}^{n_s} \frac{1}{E_s - \omega_k + i\varepsilon} \prod_{\ell=1}^{n_t} \frac{1}{E_t - \omega_\ell + i\varepsilon}. \quad (5.5.9)$$

We have dropped the $i\varepsilon$ from all propagators in channels other than s or t , since these will never go on shell.

To take the discontinuity in the t channel, we want to pass around the branch point at $t = 0$ and no other branch points. We can do this by passing through the region R^s , where only $s > 0$ and then back to $R^{\{s,t\}}$ on the other side of the $t = 0$ branch cut. Thus we must find a path rotating the energies, respecting energy conservation, to go from $R^{\{s,t\}} \rightarrow R^s$ (some examples are given in Section 5.6). Let us assume such a path exists. This path will encircle the branch point for E_t , located at the smallest value of ω_k appearing in any E_t propagator, but will not encircle the branch point for E_s . The difference between M before and after analytic continuation along this path is thus

$$\begin{aligned}
[\text{Disc}_t M]_{R^{\{s,t\}}} = & \prod_{P_i \notin \{P_s, P_t\}} \frac{1}{E_{P_i} - \omega_i} \prod_{k=1}^{n_s} \frac{1}{E_s - \omega_k + i\varepsilon} \\
& \times \sum_{\ell=1}^{n_t} \frac{1}{E_t - \omega_1 + i\varepsilon} \cdots (-2\pi i) \delta(E_t - \omega_\ell) \cdots \frac{1}{E_t - \omega_{n_t} - i\varepsilon}. \quad (5.5.10)
\end{aligned}$$

Again, the propagators not in the t channel will remain unaffected since our analytic continuation path has gone from $R^{\{s,t\}} \rightarrow R^s \rightarrow R^{\{s,t\}}$.

We can take a discontinuity in the s channel in an analogous way, using an analytic continuation path in energy that encircles the branch point for E_s while going from $R^{\{s,t\}} \rightarrow$

$R^t \rightarrow R^{\{s,t\}}$. This allow us to compute

$$\begin{aligned}
[\text{Disc}_s \text{Disc}_t M]_{R^{\{s,t\}}} &= \prod_{P_i \neq P_s, P_t} \frac{1}{E_{P_i} - \omega_i} \\
&\times \sum_{k=1}^{n_s} \frac{1}{E_s - \omega_1 + i\varepsilon} \cdots (-2\pi i) \delta(E_s - \omega_k) \cdots \frac{1}{E_s - \omega_{n_s} - i\varepsilon} \\
&\times \sum_{\ell=1}^{n_t} \frac{1}{E_t - \omega_1 + i\varepsilon} \cdots (-2\pi i) \delta(E_t - \omega_\ell) \cdots \frac{1}{E_t - \omega_{n_t} - i\varepsilon}. \quad (5.5.11)
\end{aligned}$$

Like before, when we take the s -channel discontinuity, the t -channel propagators are unaffected since we have not gone around the branch point at $t = 0$.

We cannot immediately sum over TOPT diagrams in Eq. (5.5.11) to get a Feynman integral, since it is not clear which Feynman propagators should get $+i\varepsilon$ and which should get $-i\varepsilon$. To remedy the problem, we rewrite each diagram in terms of all $+i\varepsilon$ propagators as we did for the sequential discontinuities in Section 5.5.1. This gives

$$\begin{aligned}
[\text{Disc}_s \text{Disc}_t M]_{R^{\{s,t\}}} &= \prod_{P_i \neq P_s, P_t} \frac{1}{E_{P_i} - \omega_i} \\
&\times \sum_{k=1}^{n_s} (-1)^k \left[(-2\pi i)^k \delta(E_s - \omega_1) \cdots \delta(E_s - \omega_k) \frac{1}{E_s - \omega_{k+1} + i\varepsilon} \cdots \frac{1}{E_s - \omega_{n_s} + i\varepsilon} + \text{perms} \right] \\
&\times \sum_{\ell=1}^{n_t} (-1)^\ell \left[(-2\pi i)^\ell \delta(E_t - \omega_1) \cdots \delta(E_t - \omega_\ell) \frac{1}{E_t - \omega_{\ell+1} + i\varepsilon} \cdots \frac{1}{E_t - \omega_{n_t} + i\varepsilon} + \text{perms} \right] \quad (5.5.12)
\end{aligned}$$

After summing over all TOPT diagrams with the same topology, we get

$$\boxed{[\text{Disc}_s \text{Disc}_t \mathcal{M}]_{R^{\{s,t\}}} = \left[\sum_{k=1}^{\infty} \sum_{\ell=1}^{\infty} (-1)^{k+\ell} \mathcal{M}_{\{k \text{ cuts in } s, \ell \text{ cuts in } t\}} \right]_{R_+^{\{s,t\}}}} \quad (5.5.13)$$

where the sum is over all diagrams with $k \geq 1$ cuts in the s -channel and $\ell \geq 1$ cuts in the t channel, and all propagators are assigned $+i\varepsilon$.

One should think of Eq. (5.5.13) as applying at an implicit phase-space point in the physical region where the cuts are to be computed. One can analytically continue the resulting cut graphs to any region one wants, such as the Euclidean region, but the result will not be the same as evaluating the cut graphs at a phase-space point in the Euclidean region. This is because the theta functions associated with the original region determine whether the cut vanishes, rather than by the kinematics of the new region. In other words, one cannot evaluate some of the cuts at a phase space point in R^s and others at a phase space point in R^t . Thus, our formula is derived assuming we want to relate cuts and discontinuities at a single phase space point in $R^{\{s,t\}}$. You can use a region other than $R^{\{s,t\}}$ (such as $R^{\{s,t,u\}}$), as long as the paths in analytic continuation between these regions exist.

In terms of monodromy matrices, this sequential discontinuity can be computed as

$$[\text{Disc}_s \text{Disc}_t \mathcal{M}]_{R^{\{s,t\}}} = (\mathbb{1} - \mathcal{M}_{\varpi_0^t})(\mathbb{1} - \mathcal{M}_{\varpi_0^s})\mathcal{M}, \quad (5.5.14)$$

where we recall that the action of these monodromy matrices should be read left to right (unlike discontinuity operators). The variation matrix \mathcal{M} should be evaluated along paths

from the basepoint to the region $R^{\{s,t\}}$. The monodromy matrices are computed from the basepoint and the monodromies are prepended to the path γ ending in $R^{\{s,t\}}$. Alternatively, one can apply the monodromy matrices in some other region, such as R^\star and then continue to $R^{\{s,t\}}$; since we are prepending the monodromies, whether we continue before or after we prepend them gives the same answer. However, we highlight again that the same is *not* true of cuts—for instance, all cuts evaluate to zero in R^\star .

One can generalize this formula to apply to m_i discontinuities in channel i without additional complication:

$$\boxed{\begin{aligned} & [(\text{Disc}_{s_1})^{m_1} \cdots (\text{Disc}_{s_n})^{m_n} \mathcal{M}]_{R^{\{s_1, \dots, s_n\}}} \\ &= (-1)^{N_{\text{discs}} - N_{\text{cuts}}} m_1! \cdots m_n! \sum_{k_1=m_1}^{\infty} \begin{Bmatrix} k_1 \\ m_1 \end{Bmatrix} \cdots \sum_{k_n=m_n}^{\infty} \begin{Bmatrix} k_n \\ m_n \end{Bmatrix} \left[\mathcal{M} \begin{Bmatrix} k_1 \text{ cuts in } s_1 \\ \vdots \\ k_n \text{ cuts in } s_n \end{Bmatrix} \right]_{R_+^{\{s_1, \dots, s_n\}}} \end{aligned}} \quad (5.5.15)$$

where

$$N_{\text{discs}} = m_1 + \cdots + m_n \quad \text{and} \quad N_{\text{cuts}} = k_1 + \cdots + k_n. \quad (5.5.16)$$

This is the master formula for computing any number of sequential discontinuities in any channels. Note that the right side of this equation does not depend on the order in which discontinuities are taken on the left side, which points to a non-obvious set of identities that the discontinuities of these functions must satisfy.

One can even go one step farther and generalize from s_i being individual invariants to being sets of invariants. For example, we might have a set $\mathcal{S}_i = \{s, t\}$. Then the discontinuity in \mathcal{S}_i is computed by taking the monodromy from a region $R^{\mathcal{S}_i}$ where the invariants in

\mathcal{S}_i are positive through the Euclidean region and back. Then

$$[\text{Disc}_{\mathcal{S}_i} \mathcal{M}]_{R^{\mathcal{S}_i}} = (\mathbb{1} - \mathcal{M}_{\mathcal{S}_i}) \mathcal{M} = \sum_j [\mathcal{M}_{\text{cuts in } s_j \in \mathcal{S}_i}]_{R_+^{\mathcal{S}_i}} \quad (5.5.17)$$

The generalization to multiple sets and multiple discontinuities is

$$\begin{aligned} & \frac{1}{m_1!} \cdots \frac{1}{m_n!} [(\text{Disc}_{\mathcal{S}_1})^{m_1} \cdots (\text{Disc}_{\mathcal{S}_n})^{m_n} \mathcal{M}]_{R^{\cup \mathcal{S}_i}} \\ &= (-1)^{N_{\text{discs}} - N_{\text{cuts}}} \sum_{k_1=m_1}^{\infty} \begin{Bmatrix} k_1 \\ m_1 \end{Bmatrix} \cdots \sum_{k_n=m_n}^{\infty} \begin{Bmatrix} k_n \\ m_n \end{Bmatrix} [\mathcal{M}_{k_j \text{ cuts from set } \mathcal{S}_j}]_{R_+^{\cup \mathcal{S}_i}} \end{aligned} \quad (5.5.18)$$

where $\text{Disc}_{\mathcal{S}_j}$ is taken between the region $R^{\cup \mathcal{S}_i}$ where all invariants in any set \mathcal{S}_i are positive to a region $R^{\cup \mathcal{S}_i / \mathcal{S}_j}$ where all the invariants have the same sign as in $R^{\cup \mathcal{S}_i}$ except for those in \mathcal{S}_j , which are negative. An example of this type of set discontinuity is given in Eq. (5.6.43) below.

In [165], a different prescription for calculating sequential discontinuities in different channels was proposed. Their proposal was that $\text{Disc}_s \text{Disc}_t \mathcal{M}$ should be computed by first calculating $\text{Disc}_t \mathcal{M}$ in R^t , and then analytically continue to $R^{\{s,t\}}$ before computing Disc_s . They defined these discontinuities as the difference between a function on different sides of a branch cut. Using the language of monodromies around a branch point rather than discontinuities across branch cuts, this can be interpreted to mean first prepending a monodromy matrix around $t = 0$ to a path going into R^t and then extending the path into $R^{\{s,t\}}$. Since the monodromy matrix is independent of the endpoint of the integration, this is the same as simply computing the discontinuity in t in the region $R^{\{s,t\}}$ to begin with.

No details were given in [165] for how to choose paths for analytic continuation.

As for the cuts, the prescription given in [165] for how to compute sequential cuts involves an algorithm with tuples of black and white dots that determines whether $+i\varepsilon$ or $-i\varepsilon$ should be chosen. For the examples they considered, this algorithm worked. However, in more complicated cases, it may not correctly account for the discontinuity of $-i\varepsilon$ propagators that appear after a first discontinuity. The main difference, however, is that [165] excluded from consideration cases where sequential discontinuities were taken in the same channel. Our formulas allow for any number of discontinuities in any channels, with no restrictions.

5.5.3 Steinmann relations

Finally, let us connect to the Steinmann relations. One of the important implications of Eq. (5.5.13) is that $[\text{Disc}_s \text{Disc}_t \mathcal{M}]_{R\{s,t\}}$ can only be nonzero when there exists at least one TOPT diagram in which both E_s and E_t appear. However, it is a general feature of TOPT that whenever two energies E_t and E_s appear in the propagators of a single diagram, one must depend on a subset of the energies that appear in the other (e.g. $E_s = E_1 + E_2 + E_3$ and $E_t = E_1 + E_2$). It follows that $[\text{Disc}_s \text{Disc}_t \mathcal{M}]_{R\{s,t\}}$ will vanish whenever s and t involve partially overlapping sets of energies. More precisely, recall from the beginning of this section that $s = (\sum_{i \in I} P_i)^2$ and $t = (\sum_{i \in J} P_i)^2$. Then,

$$[\text{Disc}_s \text{Disc}_t \mathcal{M}]_{R\{s,t\}} = 0 \quad \text{if} \quad I \not\subset J \quad \text{and} \quad J \not\subset I. \quad (5.5.19)$$

This is a version of the Steinmann relations, which state that the double sequential discontinuity in such overlapping channels must vanish, which we have thus proven at the level of Feynman integrals.

It is worth emphasizing two conditions that are necessary for our proof of the Steinmann relations to hold. First, the region $R^{\{s,t\}}$, where all invariants other than s and t are negative and all momenta are real, must exist. The existence of such regions is consistent with the assumptions of axiomatic field theory, where all particles are massive; however, when there are massless external particles, the on-shell constraint may mean the region $R^{\{s,t\}}$ is empty. In such a case, we cannot immediately apply our formulas.

Second, we go around the poles in the TOPT propagators by continuing the external energies, holding the external three-momenta fixed. This allowed us to isolate the singularities, since the internal energies ω_k depend only on the external three-momenta, which are held fixed during the analytic continuation. If one tries to impose a constraint on some of the external momenta, such as fixing their masses to zero or some other value, then one must also rotate the external momenta to maintain the mass-shell condition. In such cases, finding the singular variety for the TOPT propagators is more complicated and our derivation also does not immediately apply.

Because of these preconditions, the Steinmann relations in Eq. (5.5.19) do not restrict all possible double discontinuities in partially-overlapping channels. In particular, they do not apply to discontinuities on sheets that are far removed from the physical sheet; they only hold at *real* kinematic points, in the physical region. This subtlety appears, for instance, in the one-loop box with massless internal and external legs. This box is infrared divergent. In

$d = 4 - 2\varepsilon$ dimensions it has the expansion [205]

$$\mathcal{M}^{0m} = \frac{1}{st} \left[\frac{4}{\varepsilon^2} - \frac{2}{\varepsilon} \left(\ln \frac{-s}{\mu^2} + \ln \frac{-t}{\mu^2} \right) + 2 \ln \frac{-s}{\mu^2} \ln \frac{-t}{\mu^2} - \pi^2 + \mathcal{O}(\varepsilon) \right] \quad (5.5.20)$$

where $s = (p_1 + p_2)^2$ and $t = (p_2 + p_3)^2$ partially overlap. The $\mathcal{O}(\varepsilon^0)$ term has a $\ln(-s) \ln(-t)$ component that has a nonzero sequential discontinuity in s and t . With massless external lines, the region $R^{\{s,t\}}$ does not exist, so there is no contradiction with our formula. This observation is consistent with results from S -matrix theory; since s and t can only simultaneously vanish outside of the physical region, the Steinmann relations do not apply [206].

If *internal* particles are massless, our sequential discontinuity formulas in Eq. (5.5.8) and Eq. (5.5.13), and correspondingly the Steinmann relations in Eq. (5.5.19), should still apply. The key problem with massless *external* particles is that the massless condition constrains the surface of maximal analytic continuation; massless internal particles impose no such constraint. Nevertheless, with massless internal particles, certain cuts also have to be treated with care when applying the Steinmann relations (as explained, for instance, in [177]). When two overlapping momentum channels only depend on a single common momentum, cutting both channels can lead to a three-point vertex in which an external state decays into a pair of internal physical states. Some discussion of these vertices is given in Appendix B.7. In S -matrix theory, external states are stable and massless three-point vertices do not appear.

Finally, let us highlight the fact that the right side of Eq. (5.5.15) does not know anything about the order of the discontinuities begin taken on the left side. This implies that

the Steinmann relations force any sequence of discontinuities involving partially-overlapping channels to vanish, even if these partially-overlapping discontinuities are separated by a long sequence of unrelated discontinuities. This is related to the fact that Eq. (5.5.15) only governs discontinuities that are computed at a phase-space point in which all the relevant cuts are accessible, and holding all other variables fixed [206]. Thus, in many cases the relevant region may not correspond to real kinematics, in which case this restriction does not immediately apply.

5.6 Examples

In this section, we consider a number of examples in which we can check the general relations between cuts and discontinuities developed in the previous sections.

5.6.1 Bubbles

The first examples we consider are sequences of bubbles. The single bubble integral with massless internal lines in $d = 4 - 2\epsilon$ dimensions evaluates to

$$\mathcal{M}_1^{\text{bare}} = \text{---} \underset{p-k}{\overset{k}{\bigcirc}} \text{---} = \mu^{4-d} \int \frac{d^d k}{i(2\pi)^d} \frac{1}{k^2 + i\epsilon} \frac{1}{(p-k)^2 + i\epsilon} \quad (5.6.1)$$

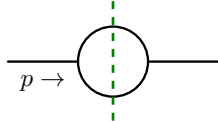
$$= -\frac{1}{16\pi^2} \left[-\frac{1}{\epsilon} + \ln \left(\frac{-s}{\tilde{\mu}^2} - i\epsilon \right) - 2 \right], \quad (5.6.2)$$

where $s = p^2$ and $\tilde{\mu}^2 = 4\pi e^{-\gamma_E} \mu^2$. The counterterm graph is analytic, so we add it to remove the UV divergence and the algebraic part of the integral (the -2 contribution), giving

a simpler answer for the renormalized amplitude:

$$\mathcal{M}_1 = -\frac{1}{16\pi^2} \ln \left(\frac{-s}{\widetilde{\mu}^2} - i\varepsilon \right) . \quad (5.6.3)$$

The cut through the bubble is finite in four dimensions:

$$\mathcal{M}_1^{\text{cut}} = \text{---} \xrightarrow{p \rightarrow} \text{---} \bigcirc \text{---} \quad (5.6.4)$$


$$= \int \frac{d^4 k}{i(2\pi)^4} (-2\pi i) \delta(k^2) \Theta(k^0) (-2\pi i) \delta[(p-k)^2] \Theta(p^0 - k^0) = \frac{i}{8\pi} \Theta(s) . \quad (5.6.5)$$

Here we have assumed $p^0 > 0$. If $p^0 < 0$, this cut vanishes but the cut with energy flowing in the opposite direction compensates and gives the same result. \mathcal{M}_1 has a branch cut on the positive real line in the s plane. The discontinuity across this branch cut is

$$\text{Disc}_s \mathcal{M}_1 = -\frac{1}{16\pi^2} (-2\pi i) \Theta(s) = \mathcal{M}_1^{\text{cut}} , \quad (5.6.6)$$

in agreement with the covariant cutting rules and the optical theorem. Similarly, the monodromy computed around the branch point at $s = 0$,

$$(\mathbb{1} - \mathcal{M}_{\varpi_0^s}) \mathcal{M}_1 = \frac{i}{8\pi} , \quad (5.6.7)$$

gives the same answer in R^s , where $s > 0$.

Sequential discontinuities in the same channel

Now we consider an example that has a nonzero sequential discontinuity in a single channel. We keep the propagators in the loops massless, but give the internal lines connecting the bubbles a mass m so that we can ignore their discontinuities for $m > \sqrt{s}$. The chain of three bubbles is given by

$$\mathcal{M}_3 = \text{---} \xrightarrow{p \rightarrow} \text{---} \bigcirc \text{---} \xrightarrow{m} \bigcirc \text{---} \xrightarrow{m} \bigcirc \text{---} \quad (5.6.8)$$

$$= \frac{1}{(-16\pi^2)^3} \left(\frac{1}{s - m^2} \right)^2 \ln^3 \left(\frac{-s}{\mu^2} - i\varepsilon \right). \quad (5.6.9)$$

Since this is just a product of logarithms, the discontinuities in s are simple to calculate using Eq. (5.4.37). We find

$$\text{Disc}_s \mathcal{M}_3 = \frac{2\pi i}{(16\pi^2)^3} \left(\frac{1}{s - m^2} \right)^2 \left[3 \ln^2 \left(\frac{-s}{\mu^2} - i\varepsilon \right) + 6\pi i \ln \left(\frac{-s}{\mu^2} - i\varepsilon \right) - 4\pi^2 \right], \quad (5.6.10)$$

$$\text{Disc}_s \text{Disc}_s \mathcal{M}_3 = -\frac{(2\pi i)^2}{(16\pi^2)^3} \left(\frac{1}{s - m^2} \right)^2 \left[6 \ln \left(\frac{-s}{\mu^2} - i\varepsilon \right) + 12\pi i \right], \quad (5.6.11)$$

and

$$\text{Disc}_s \text{Disc}_s \text{Disc}_s \mathcal{M}_3 = \frac{6(2\pi i)^3}{(16\pi^2)^3} \left(\frac{1}{s - m^2} \right)^2. \quad (5.6.12)$$

We expect these discontinuities to be related to cuts by Eq. (5.5.8).

Assuming $p^0 > 0$ and $s > 0$, and using all $+i\varepsilon$ propagators, the cut through loop A is

given by

$$\begin{aligned}
\mathcal{M}_{3A}^{\text{cut}} &= \text{Diagram: } p \rightarrow \text{circle A} \xrightarrow{m} \text{circle B} \xrightarrow{m} \text{circle C} \rightarrow \\
&= \frac{1}{(-16\pi^2)^3} \left(\frac{1}{s-m^2} \right)^2 (-2\pi i) \ln^2 \left(\frac{-s}{\mu^2} - i\varepsilon \right). \quad (5.6.13)
\end{aligned}$$

The cuts of the second and third loop give identical results since we always assign uncut propagators $+i\varepsilon$. Thus, we have $\mathcal{M}_{3B}^{\text{cut}} = \mathcal{M}_{3C}^{\text{cut}} = \mathcal{M}_{AC}^{\text{cut}}$. There are also three chapters/diagrams involving two cuts. Cutting loops A and B gives

$$\begin{aligned}
\mathcal{M}_{3AB}^{\text{cut}} &= \text{Diagram: } p \rightarrow \text{circle A} \xrightarrow{m} \text{circle B} \xrightarrow{m} \text{circle C} \rightarrow \\
&= \frac{(-2\pi i)^2}{(-16\pi^2)^3} \left(\frac{1}{s-m^2} \right)^2 \ln \left(\frac{-s}{\mu^2} - i\varepsilon \right). \quad (5.6.14)
\end{aligned}$$

The other chapters/diagrams involving two cuts give identical results: $\mathcal{M}_{3AC}^{\text{cut}} = \mathcal{M}_{3BC}^{\text{cut}} = \mathcal{M}_{3AB}^{\text{cut}}$. The triple cut is

$$\mathcal{M}_{3ABC}^{\text{cut}} = \text{Diagram: } p \rightarrow \text{circle A} \xrightarrow{m} \text{circle B} \xrightarrow{m} \text{circle C} \rightarrow \quad (5.6.15)$$

$$= \frac{(-2\pi i)^3}{(-16\pi^2)^3} \left(\frac{1}{s-m^2} \right)^2. \quad (5.6.16)$$

We can now compute the right side of Eq. (5.5.8). For $m = 1$, we get

$$-\sum_{k=1}^{\infty} (-1)^k \begin{Bmatrix} k \\ 1 \end{Bmatrix} \mathcal{M}_3^{(k\text{-cuts})} = \mathcal{M}_3^{(1\text{-cut})} - \mathcal{M}_3^{(2\text{-cuts})} + \mathcal{M}_3^{(3\text{-cuts})} \quad (5.6.17)$$

$$= (\mathcal{M}_{3A}^{\text{cut}} + \mathcal{M}_{3B}^{\text{cut}} + \mathcal{M}_{3C}^{\text{cut}}) - (\mathcal{M}_{3AB}^{\text{cut}} + \mathcal{M}_{3AC}^{\text{cut}} + \mathcal{M}_{3BC}^{\text{cut}}) + \mathcal{M}_{3ABC}^{\text{cut}} \quad (5.6.18)$$

$$= \frac{(-2\pi i)}{(-16\pi^2)^3} \left(\frac{1}{s-m^2} \right)^2 \left[3 \ln^2 \left(\frac{-s}{\mu^2} - i\varepsilon \right) + 6\pi i \ln \left(\frac{-s}{\mu^2} - i\varepsilon \right) - 4\pi^2 \right]. \quad (5.6.19)$$

This agrees with $\text{Disc}_s \mathcal{M}_3$, as expected. Similarly, for $m = 2$ and $m = 3$ we get

$$2 \sum_{k=2}^{\infty} (-1)^k \begin{Bmatrix} k \\ 2 \end{Bmatrix} \mathcal{M}_3^{(k\text{-cuts})} = 2(\mathcal{M}_{3AB}^{\text{cut}} + \mathcal{M}_{3AC}^{\text{cut}} + \mathcal{M}_{3BC}^{\text{cut}}) - 6\mathcal{M}_{3ABC}^{\text{cut}}. \quad (5.6.20)$$

and

$$-3! \sum_{k=3}^{\infty} (-1)^k \begin{Bmatrix} k \\ 3 \end{Bmatrix} \mathcal{M}_3^{(k\text{-cuts})} = 6\mathcal{M}_{3ABC}^{\text{cut}}. \quad (5.6.21)$$

It can be checked that these quantities agree with the discontinuities computed in Eq. (5.6.11) and Eq. (5.6.12).

One can similarly check that the relation in Eq. (5.5.8) holds for the m^{th} discontinuity of the n -loop bubble chain. This is not particularly surprising, since the algebra involved is essentially the same as the algebra used to derive equations like Eq. (5.5.7).

Sequential discontinuities in different channels

We now turn to an example involving discontinuities in different channels. We consider

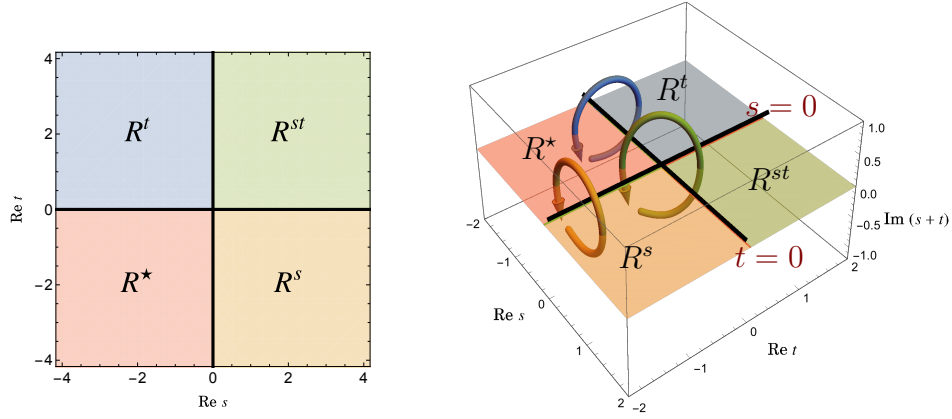


Figure 5.4: The function $\ln(-s)\ln(-t)$ has branch hypersurfaces at $s = 0$ and $t = 0$, shown in black. The Euclidean region R^* corresponds to $s < 0$ and $t < 0$. We can compute discontinuities in s and t of \mathcal{M} by rotating around the branch points as indicated by the curves on the right. These curves pass out of the real s, t plane.

the diagram

$$\mathcal{M}_{st} = \text{Diagram} = \frac{1}{256\pi^4} \ln\left(\frac{-s}{\mu^2} - i\varepsilon\right) \ln\left(\frac{-t}{\mu^2} - i\varepsilon\right), \quad (5.6.22)$$

where $s = P_s^2$ and $t = P_t^2$. This function has branch points at $s = 0$ and at $t = 0$. In the space of complex s and t , these branch points correspond to one-dimensional complex hypersurfaces. We have depicted this in Fig. 5.4.

The connection and variation matrix for this function in the Euclidean region where $s <$

0 and $t < 0$ are

$$\omega = \begin{pmatrix} 0 & \frac{ds}{s} & \frac{dt}{t} & 0 \\ 0 & 0 & 0 & \frac{dt}{t} \\ 0 & 0 & 0 & \frac{ds}{s} \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad \mathcal{M}_{\gamma_0} = \begin{pmatrix} 1 & \ln(-s) & \ln(-t) & \ln(-s)\ln(-t) \\ 0 & 1 & 0 & \ln(-t) \\ 0 & 0 & 1 & \ln(-s) \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad (5.6.23)$$

and the monodromy matrices are

$$\mathcal{M}_{\mathbb{S}^s_0} = \begin{pmatrix} 1 & 2\pi i & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 2\pi i \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad \mathcal{M}_{\mathbb{S}^t_0} = \begin{pmatrix} 1 & 0 & 2\pi i & 0 \\ 0 & 1 & 0 & 2\pi i \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \quad (5.6.24)$$

The variation matrix in a region with $s > 0$ and/or $t > 0$ is the same with $\ln(-s) \rightarrow \ln(-s - i\varepsilon)$ and/or $\ln(-t) \rightarrow \ln(-t - i\varepsilon)$.

We can compute $\text{Disc}_s \text{Disc}_t \mathcal{M}_{st}$ by computing monodromies around the branch points at $s = 0$ and $t = 0$. First, the discontinuity in s gives

$$\text{Disc}_s \mathcal{M}_{st} = (\mathbb{1} - \mathcal{M}_{\mathbb{S}^s_0}) \mathcal{M}_{st} = \frac{-2\pi i}{256\pi^4} \ln\left(\frac{-t}{\mu^2} - i\varepsilon\right). \quad (5.6.25)$$

Computing the discontinuity in t of this quantity gives

$$\text{Disc}_t \text{Disc}_s \mathcal{M}_{st} = (\mathbb{1} - \mathcal{M}_{\mathbb{S}^s_0})(\mathbb{1} - \mathcal{M}_{\mathbb{S}^t_0}) \mathcal{M}_{st} = \frac{(-2\pi i)^2}{256\pi^4}. \quad (5.6.26)$$

To compute the cuts, we must be in the region $R^{\{s,t\}}$ where neither cut vanishes. There, we find

$$[\text{Cut}_{st}\mathcal{M}_{st}]_{R^{\{s,t\}}} = \frac{(-2\pi i)^2}{256\pi^4}. \quad (5.6.27)$$

We see that the cut and the sequential discontinuity agree, as they should according to Eq. (5.5.13).

We can also compute the total discontinuity of this function in $R^{\{s,t\}}$,

$$\begin{aligned} [\text{Disc}_{\text{tot}}\mathcal{M}_{st}]_{R^{\{s,t\}}} &= \mathcal{M}_{st} - \overline{\mathcal{M}}_{st} \\ &= \frac{1}{256\pi^4} \left[\ln\left(\frac{-s}{\mu^2} - i\varepsilon\right) \ln\left(\frac{-t}{\mu^2} - i\varepsilon\right) - \ln\left(\frac{-s}{\mu^2} + i\varepsilon\right) \ln\left(\frac{-t}{\mu^2} + i\varepsilon\right) \right] \\ &= \frac{-2\pi i}{256\pi^4} \left[\ln\left(\frac{-s}{\mu^2} - i\varepsilon\right) + \ln\left(\frac{-t}{\mu^2} + i\varepsilon\right) \right] \end{aligned} \quad (5.6.28)$$

in agreement with the standard cut prescription, where the $i\varepsilon$ is flipped on the $\ln(-t)$ propagator because it comes after the s cut. We can also write this in our standardized form, where the $i\varepsilon$ are homogeneous:

$$[\text{Disc}_{\text{tot}}\mathcal{M}_{st}]_{R^{\{s,t\}}} = \frac{-2\pi i}{256\pi^4} \left[\ln\left(\frac{-s}{\mu^2} - i\varepsilon\right) + \ln\left(\frac{-t}{\mu^2} - i\varepsilon\right) + 2\pi i \right]. \quad (5.6.29)$$

According to Section 5.5.2, this should match the function returned by the operator

$\text{Disc}_{\{s,t\}}$, which corresponds to analytically continuing around both the branch points $s = 0$ and $t = 0$ along a path $R^{\{s,t\}} \rightarrow R^* \rightarrow R^{\{s,t\}}$, as depicted by the green curve in Fig. 5.4.

The result is

$$[\text{Disc}\mathcal{M}_{st}]_{R\{s,t\}} = (1 - \mathcal{M}_{\text{box}_0^s} \mathcal{M}_{\text{box}_0^t}) \mathcal{M}_{st} \quad (5.6.30)$$

$$= \frac{-2\pi i}{256\pi^4} \left[\ln \left(\frac{-s}{\mu^2} - i\varepsilon \right) + \ln \left(\frac{-t}{\mu^2} - i\varepsilon \right) + 2\pi i \right], \quad (5.6.31)$$

in agreement with Eq. (5.6.29).

5.6.2 Triangles and Boxes

Next we consider the triangle and box ladder integrals, with massless internal lines and massive external lines. These integrals are known to all loop orders [207], and can be treated simultaneously because they give rise to the same transcendental function at each order. For simplicity, we concentrate mostly on the triangle ladders, and comment on the box ladders at the end of the section. Our momentum labeling convention is shown in Fig. 5.5. All momenta are incoming, and we have $\sum p_i^\mu = 0$.

Triangle kinematics

For the triangle integrals, we follow the conventions of [165] and [208]. Since all internal lines are massless, the amplitude depends only on ratios of the invariants p_1^2 , p_2^2 , and p_3^2 . These kinematics can be parametrized using the variables u , v , z , and \bar{z} , defined as

$$u \equiv \frac{p_2^2}{p_1^2} = z\bar{z} \quad \text{and} \quad v \equiv \frac{p_3^2}{p_1^2} = (1-z)(1-\bar{z}), \quad (5.6.32)$$

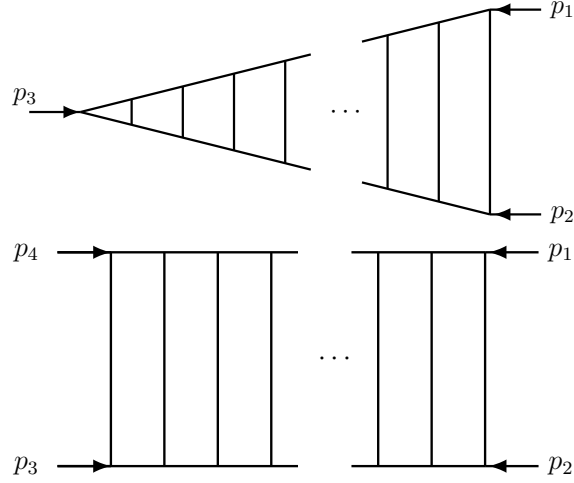


Figure 5.5: The L -loop triangle and box ladder integrals. We take all momenta incoming with p_3 along the long direction of the triangle. For the box ladders, $s = (p_1 + p_2)^2$ and $t = (p_2 + p_3)^2$.

where we choose

$$z = \frac{1 + u - v + \sqrt{1 + u^2 + v^2 - 2uv - 2u - 2v}}{2}, \quad (5.6.33)$$

$$\bar{z} = \frac{1 + u - v - \sqrt{1 + u^2 + v^2 - 2uv - 2u - 2v}}{2}. \quad (5.6.34)$$

This corresponds to the convention that $\bar{z} \leq z$ for real kinematics. The triangle ladders are invariant under the \mathbb{Z}_2 symmetry $z \leftrightarrow \bar{z}$.

For these integrals, it is possible to find real phase-space points with any pattern of signs for the invariants p_1^2 , p_2^2 , and p_3^2 . We denote the region where $p_1^2 > 0$ and $p_2^2, p_3^2 < 0$ by R^1 . In this region, z and \bar{z} are real, and $\bar{z} < 0$ while $1 < z$. Similarly, we denote the region in which $p_2^2 > 0$ and $p_1^2, p_3^2 < 0$ by R^2 , and here we have $\bar{z} < 0 < z < 1$. Finally, we denote by R^3 the region where $p_3^2 > 0$ and $p_1^2, p_2^2 < 0$, which implies $0 < \bar{z} < 1 < z$. We also consider dual regions in which two invariants are positive, such as R^{23} , where $p_1^2 < 0$ and $p_2^2, p_3^2 > 0$,

and so on. Since taking $p_j^2 \rightarrow -p_j^2$ for all j leaves u and v invariant (and therefore also z and \bar{z}), any function of u and v has the same form in a given region and the dual region in which all invariants have the opposite sign. For example, functions of u and v take the same form in R^{23} and R^1 . It is nevertheless important to distinguish a region from its dual because cuts can only be nonzero for positive invariants.

The Euclidean region, where all invariants are negative, is denoted R^* . The Euclidean region has a number of subregions, based on the relative sizes of the p_j^2 invariants (or equivalently of z and \bar{z}). Of particular importance is the region R_A^* , which corresponds to real values $0 < \bar{z} < z < 1$. The functions $\ln z$, $\ln \bar{z}$, $\text{Li}_n z$, and $\text{Li}_n \bar{z}$ are all analytic in this region. Region R_C^* corresponds to real $\bar{z} < z < 0$, and region R_B^* corresponds to real $1 < \bar{z} < z$. Finally, region R_I^* involves complex z and \bar{z} that are related by complex conjugation, namely $\bar{z} = z^*$. All of these regions correspond to two-dimensional slices of the four-dimensional space of complex z and \bar{z} , in which all the invariants p_i^2 are real. The dual of the Euclidean region, where all invariants are negative, is denoted R^{123} and also has subregions corresponding to R_A^* , R_B^* and R_C^* . A summary of the regions is shown in Fig. 5.6.

To take sequential discontinuities of Feynman integrals, we analytically continue around branch points where Mandelstam invariants vanish. This analytic continuation takes us into different kinematic regions; for example, to take $[\text{Disc}_{p_1^2}]_{R^1}$ we need to analytically continue from R^1 to R^* and back. Our formula relating cuts and discontinuities assumes that we rotate the energies while preserving $E_1 + E_2 + E_3 = 0$ and holding all three-momenta fixed. Thus, we can set $E_3 = -E_1 - E_2$ and $\vec{p}_3 = -\vec{p}_1 - \vec{p}_2$ and work in a frame where all momenta are aligned in the x direction. Then, rescaling these momenta so that $p_1^x = 1$, we can solve

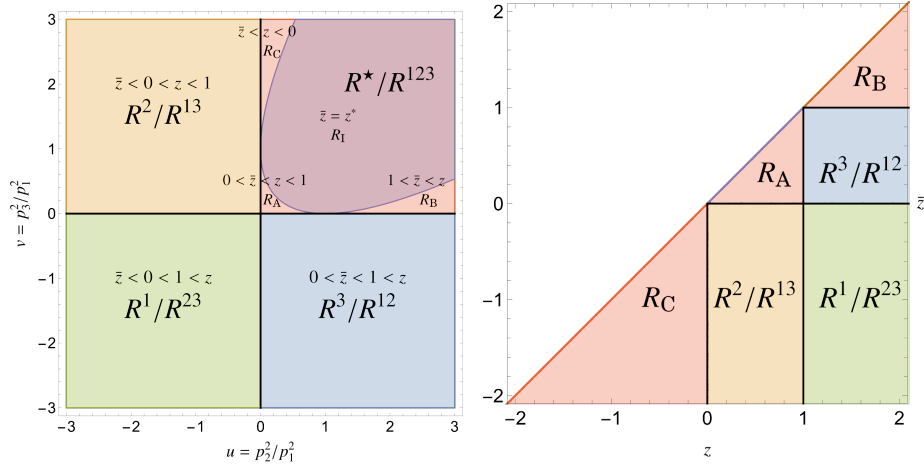


Figure 5.6: The triangle ladder integrals we consider depend only on $u = p_2^2/p_1^2$ and $v = p_3^2/p_1^2$, or equivalently on z and \bar{z} . The different regions in u, v and z, \bar{z} space correspond to regions in which the Mandelstam invariants have different relative signs. For instance, in R^1 the invariants satisfy $p_1^2 > 0$, $p_2^2 < 0$, and $p_3^2 < 0$. The Euclidean region, where $p_j^2 < 0$ for all j , has four further subregions, described in the text.

for E_1 and E_2 in terms of z , \bar{z} , and the remaining unfixed momentum component p_2^x :

$$E_1 = \frac{-2p_2^x - (z + \bar{z})}{\bar{z} - z}, \quad E_2 = \frac{2z\bar{z} + p_2^x(z + \bar{z})}{\bar{z} - z}. \quad (5.6.35)$$

One can use these equations to translate a given path in z and \bar{z} to an acceptable path in energy for a given value of p_2^x . It turns out, however, that an analytic continuation path cannot be found between any pair of regions. For example, we cannot go from R^1 to R_A^* .

To see this, note that in these coordinates, the invariants are given by

$$p_1^2 = \frac{4(p_2^x + z)(p_2^x + \bar{z})}{(z - \bar{z})^2}, \quad p_2^2 = z\bar{z}p_1^2, \quad p_3^2 = (1 - z)(1 - \bar{z})p_1^2. \quad (5.6.36)$$

In R_A^* , all the p_j^2 are negative. For a fixed value of $p_2^x > 0$, this constraint is impossible to satisfy, as $z > \bar{z} > 0$ in R_A^* , which implies $p_1^2 > 0$. In fact, we need $-1 < p_2^x < 0$ to get to

R_A^* . But then, in R^1 where $0 < \bar{z} < 1 < z$, we must have $p_2^x + z > 0$ and $p_2^x + \bar{z} < 0$, and so $p_1^2 < 0$. But this is a contradiction, since p_1^2 must be positive in R^1 . Thus, we cannot go from R^1 to R_A^* .

In addition to making sure the path exists, one must check that the path only encircles the desired branch points in the invariants once. For example, in particle j 's rest frame, $E_j \rightarrow e^{2\pi i} E_j$ would not be an acceptable path, as it would encircle the branch point in p_j^2 twice.

Some paths that satisfy all of these constraints are shown in Fig. 5.7. For example, we show a path from $R^2 \rightarrow R_A^* \rightarrow R^2$. It is also possible to construct a path from $R^2 \rightarrow R_C^* \rightarrow R^2$. Conversely, no path exists from R^2 to R_B^* , by the same type of argument that showed the impossibility of analytically continuing between R^1 and R_A^* . We also show a path that starts and ends in R^1 , after passing through R_C^* . When this path intersects the $\text{Re } z = \text{Re } \bar{z}$ plane, the branch cut in the square root that distinguishes z and \bar{z} is crossed. This path can be viewed as going around $z = 0$ and $z = 1$, or as going around $z = \infty$. The right side of this figure shows paths between other regions, such as $R^{23} \rightarrow R^3 \rightarrow R^{23}$. The existence of such paths is required to take sequential discontinuities in p_2^2 and p_3^2 .

Having constructed these paths, we can enumerate the monodromies corresponding to each of the discontinuities we're interested in computing. For sequential discontinuities in a

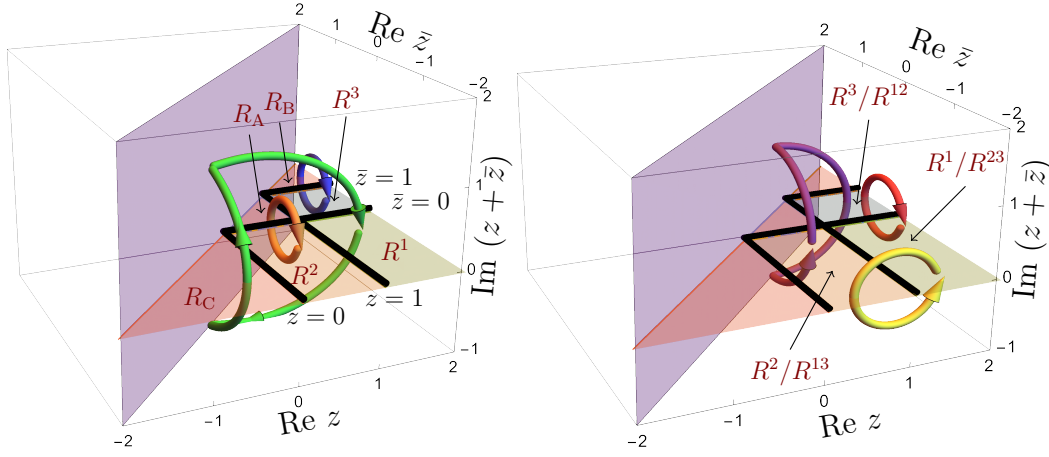


Figure 5.7: Analytic continuation of three-point and four-point ladder chapters/diagrams takes place in the four-dimensional space of complex z and \bar{z} . Sample paths of analytically continuing the energies are shown. The figure on the left depicts contours that are relevant for computing sequential discontinuities in a single channel: $R^1 \rightarrow R_C^* \rightarrow R^1$, $R^2 \rightarrow R_A^* \rightarrow R^2$ and $R^3 \rightarrow R_B^* \rightarrow R^3$. The figure on the right depicts paths relevant for computing sequential discontinuities in different channels: $R^{23} \rightarrow R^2 \rightarrow R^{23}$, $R^{13} \rightarrow R^3 \rightarrow R^{13}$ and $R^{23} \rightarrow R^3 \rightarrow R^{23}$. These paths each encircle some combination of the branch hypersurfaces shown as black lines, corresponding to where z or \bar{z} are equal to either 0 or 1.

single channel, we find

$$[\text{Disc}_{p_1^2}]_{R^1} = \mathbb{1} - \mathcal{M}_{\bar{z} \circ z_0} \cdot \mathcal{M}_{\bar{z} \circ z_1} = \mathbb{1} - \mathcal{M}_{\infty}^{\downarrow z}, \quad (5.6.37a)$$

$$[\text{Disc}_{p_2^2}]_{R^2} = \mathbb{1} - \mathcal{M}_{\bar{z} \circ z_0}, \quad (5.6.37b)$$

$$[\text{Disc}_{p_3^2}]_{R^3} = \mathbb{1} - \mathcal{M}_{\bar{z} \circ z_1}. \quad (5.6.37c)$$

In each case, there are two choices of Euclidean region that we can pass through (e.g. $R^2 \rightarrow R_A^* \rightarrow R^2$ or $R^2 \rightarrow R_C^* \rightarrow R^2$). This choice amounts to permuting $z \leftrightarrow \bar{z}$. The monodromy matrix $\mathcal{M}_{\infty}^{\downarrow z}$ corresponds to going around infinity counterclockwise, where infinity is approached along some angle that goes below the real line. This implies that the contour around infinity crosses the branch cut on the negative real axis before the one on the

positive real axis. This choice to go below the real axis corresponds to taking p_1^2 to have a small positive imaginary part, which endows z with a small negative imaginary part, as per Eq. (5.6.33). This monodromy matrix is computed in Appendix B.2.

To compute sequential discontinuities in different channels, we consider analytic continuation paths from regions with multiple positive invariants to regions in which one of these invariants has the opposite sign. To construct the discontinuity operator corresponding to each of these analytic continuations, we need to determine which branch points in z and \bar{z} the path encircles. Let us illustrate how this can be done for the path from $R^{12} \rightarrow R^2 \rightarrow R^{12}$, which computes a discontinuity in p_1^2 in the region R^{12} . We first take the differential of Eq. (5.6.32):

$$d \ln z + d \ln \bar{z} = d \ln p_2^2 - d \ln p_1^2, \quad (5.6.38)$$

$$d \ln(1 - z) + d \ln(1 - \bar{z}) = d \ln p_3^2 - d \ln p_1^2. \quad (5.6.39)$$

Since we are considering a discontinuity in p_1^2 , our path γ must satisfy

$$\oint_{\gamma} d \ln p_1^2 = 2\pi i, \quad \oint_{\gamma} d \ln p_2^2 = \oint_{\gamma} d \ln p_3^2 = 0. \quad (5.6.40)$$

Eqs. (5.6.38) and (5.6.39) then imply that

$$\oint_{\gamma} (d \ln z + d \ln \bar{z}) = -2\pi i, \quad \oint_{\gamma} (d \ln(1 - z) + d \ln(1 - \bar{z})) = -2\pi i. \quad (5.6.41)$$

We furthermore have that $0 < \bar{z} < 1 < z$ in R^{12} , while $\bar{z} < 0 < z < 1$ in R^2 . This

suggests that z should encircle 1 while \bar{z} should encircle 0 along this path. We see that this can be achieved in a manner consistent with Eq. (5.6.41) if both of these branch points are encircled clockwise. Thus, we conclude that $[\text{Disc}_{p_1^2}]_{R^{12}} = \mathbb{1} - \mathcal{M}_{\mathbb{Z}\mathcal{O}_0^{\bar{z}}} \cdot \mathcal{M}_{\mathbb{Z}\mathcal{O}_1^z}$.

Using similar reasoning, we compute the discontinuity operators in each of the regions involving two positive invariants to be

$$[\text{Disc}_{p_2^2}]_{R^{23}} = \mathbb{1} - \mathcal{M}_{\mathbb{Z}\mathcal{O}_0^{\bar{z}}} \quad , \quad [\text{Disc}_{p_3^2}]_{R^{23}} = \mathbb{1} - \mathcal{M}_{\mathbb{Z}\mathcal{O}_1^z}, \quad (5.6.42a)$$

$$[\text{Disc}_{p_1^2}]_{R^{13}} = \mathbb{1} - \mathcal{M}_{\mathbb{Z}\mathcal{O}_0^{\bar{z}}} \cdot \mathcal{M}_{\mathbb{Z}\mathcal{O}_1^z}, \quad [\text{Disc}_{p_3^2}]_{R^{13}} = \mathbb{1} - \mathcal{M}_{\mathbb{Z}\mathcal{O}_1^z}, \quad (5.6.42b)$$

$$[\text{Disc}_{p_1^2}]_{R^{12}} = \mathbb{1} - \mathcal{M}_{\mathbb{Z}\mathcal{O}_0^{\bar{z}}} \cdot \mathcal{M}_{\mathbb{Z}\mathcal{O}_1^z}, \quad [\text{Disc}_{p_2^2}]_{R^{12}} = \mathbb{1} - \mathcal{M}_{\mathbb{Z}\mathcal{O}_0^{\bar{z}}}. \quad (5.6.42c)$$

In contrast to the first discontinuity, the region that we pass through is completely fixed, so there is only a single correct monodromy matrix in each of these cases. The paths corresponding to these discontinuity operators are depicted in Fig. 5.7.

One can also consider other analytic continuation paths, such as $R_C^{123} \rightarrow R^1 \rightarrow R_C^{123}$ (not shown in the figure). Such a path exists and gives us the discontinuity with respect to the pair of invariants $\mathcal{S}_{23} = \{p_2^2, p_3^2\}$. This path encircles $z = 0$ and $z = 1$, so

$$[\text{Disc}_{\mathcal{S}_{23}}]_{R_C^{123}} = \mathbb{1} - \mathcal{M}_{\mathbb{Z}\mathcal{O}_1^z} \cdot \mathcal{M}_{\mathbb{Z}\mathcal{O}_0^z}. \quad (5.6.43)$$

Other paths that encircle the branch points of more than one invariant are also possible.

It is easiest to compute the monodromy matrices in one region and then continue the result to the other regions. The most natural region to use is R_A^* , since $0 < \bar{z} < z < 1$

so all of $\ln z, \ln \bar{z}, \text{Li}_n(z)$ and $\text{Li}_n(\bar{z})$ are analytic there. To evaluate the matrices for real values of z and \bar{z} below 0 or above 1, we need to be careful about which side of the branch cuts we are on. In the region R^i , where only $p_i^2 > 0$ and the other squared momenta are negative, we assign p_i^2 a small positive imaginary part. It can be checked using Eqs. (5.6.33) and (5.6.34) that this corresponds to giving z and \bar{z} the following small imaginary parts in these regions:

$$R^1, R^{12}, R^{13}, R^{123} : \quad z \rightarrow z - i\varepsilon, \quad \bar{z} \rightarrow \bar{z} + i\varepsilon, \quad (5.6.44a)$$

$$R^2, R^3, R^{23} : \quad z \rightarrow z + i\varepsilon, \quad \bar{z} \rightarrow \bar{z} - i\varepsilon. \quad (5.6.44b)$$

These assignments allow us to evaluate the variation matrix and monodromy matrices in the different regions.

One loop

The one-loop triangle with all massless internal lines is finite in four dimensions. In the region R_1^* , where all invariants are negative and $\bar{z} = z^*$, the Feynman integral is

$$\begin{aligned}
 T_1 = & \text{Diagram of a triangle with external momenta } p_1, p_2, p_3 \text{ and internal momenta } q_1, q_2, q_3. \\
 & = \int \frac{d^4 k}{i(2\pi)^4} \frac{1}{k^2 + i\varepsilon} \frac{1}{(p_2 - k)^2 + i\varepsilon} \frac{1}{(p_3 + k)^2 + i\varepsilon} \\
 & = \frac{1}{16\pi^2 p_1^2} \frac{1}{z - \bar{z}} \Phi_1(z, \bar{z}), \quad (5.6.45)
 \end{aligned}$$

where

$$\Phi_1(z, \bar{z}) = 2\text{Li}_2(z) - 2\text{Li}_2(\bar{z}) + \ln(z\bar{z}) \ln\left(\frac{1-z}{1-\bar{z}}\right). \quad (5.6.46)$$

In the regions R_1^\star and R_A^\star , this function is analytic.

The variation matrix for Φ_1 was given in Eq. (5.4.57):

$$\mathcal{M}_{\gamma_0}^{R_A^\star} = \begin{pmatrix} 1 & \ln z + \ln \bar{z} & \text{Li}_1(z) + \text{Li}_1(\bar{z}) & \Phi_1(z, \bar{z}) \\ 0 & 1 & 0 & -\text{Li}_1(z) + \text{Li}_1(\bar{z}) \\ 0 & 0 & 1 & \ln z - \ln \bar{z} \\ 0 & 0 & 0 & 1 \end{pmatrix}. \quad (5.6.47)$$

Here γ_0 is the straight-line path from the basepoint $(0, 0)$ to (z, \bar{z}) . In the region R_A^\star , the variation matrix is analytic. In other regions, it has the same form with z and \bar{z} on the appropriate sides of their branch cuts as determined by the displacements in Eq. (5.6.44).

Using the monodromy matrices in Eqs. (5.4.59) and (5.4.62), we can calculate the differences of paths relevant to evaluating the discontinuities in Eq. (5.6.37). We find

$$(\mathbb{1} - \mathcal{M}_{\curvearrowright_0^z})\Phi_1 = 2\pi i \left[\text{Li}_1(z) - \text{Li}_1(\bar{z}) \right], \quad (\mathbb{1} - \mathcal{M}_{\curvearrowright_1^{\bar{z}}})\Phi_1 = 2\pi i \left[\ln z - \ln \bar{z} \right], \quad (5.6.48)$$

and

$$(\mathbb{1} - \mathcal{M}_{\infty^z}^{\curvearrowright})\Phi_1 = -2\pi i \left[\text{Li}_1(z) - \text{Li}_1(\bar{z}) + \ln z - \ln \bar{z} + 2\pi i \right]. \quad (5.6.49)$$

Rewriting these results in terms of logarithms with manifestly positive arguments in the

relevant region, which in the case of $[\text{Disc}_{p_1^2} T_1]_{R^1}$ means replacing

$$\text{Li}_1(z - i\varepsilon) - \text{Li}_1(\bar{z}) + \ln z - \ln(\bar{z} + i\varepsilon) + 2\pi i = -\ln \left[\frac{(z-1)(-\bar{z})}{(1-\bar{z})z} \right], \quad (5.6.50)$$

we have

$$[\text{Disc}_{p_1^2} T_1]_{R^1} = \frac{1}{16\pi^2 p_1^2} \frac{2\pi i}{z - \bar{z}} \ln \left[\frac{(z-1)(-\bar{z})}{(1-\bar{z})z} \right], \quad (5.6.51a)$$

$$[\text{Disc}_{p_2^2} T_1]_{R^2} = \frac{1}{16\pi^2 p_1^2} \frac{2\pi i}{z - \bar{z}} \ln \left[\frac{1 - \bar{z}}{1 - z} \right], \quad (5.6.51b)$$

$$[\text{Disc}_{p_3^2} T_1]_{R^3} = \frac{1}{16\pi^2 p_1^2} \frac{2\pi i}{z - \bar{z}} \ln \left[\frac{z}{\bar{z}} \right]. \quad (5.6.51c)$$

As an initial cross check, we note that these discontinuities map to each other under the dihedral symmetry that permutes the legs of the one-loop triangle. Both the rational part and the transcendental part of these functions pick up a sign under odd permutations of the legs; for instance, under $p_2 \leftrightarrow p_3$, we have $z \rightarrow 1 - z$ and $\bar{z} \rightarrow 1 - \bar{z}$ in the logarithms, while $(z - \bar{z}) \rightarrow -(z - \bar{z})$ in the rational prefactor. The action of this symmetry is discussed in detail in Appendix B.4.

The corresponding cuts must be computed in the appropriate region. For example, the cut in p_1^2 requires $p_1^2 > 0$, and evaluates to

$$\text{Cut}_{p_1^2} T_1 = \frac{1}{16\pi^2 p_1^2} \frac{2\pi i}{z - \bar{z}} \left\{ \ln[-\bar{z}(1 - z) - i\varepsilon] - \ln[-z(1 - \bar{z}) - i\varepsilon] \right\} \Theta(p_1^2). \quad (5.6.52)$$

In region R^1 , this can be written

$$\left[\text{Cut}_{p_1^2} T_1 \right]_{R^1} = \frac{1}{16\pi^2 p_1^2} \frac{2\pi i}{z - \bar{z}} \ln \left[\frac{(z-1)(-\bar{z})}{(1-\bar{z})z} \right], \quad (5.6.53)$$

matching the discontinuity in Eq. (5.6.51a) as well as the corresponding expression in [165].

The cuts in p_2^2 and p_3^2 can similarly be computed, and agree with the discontinuities in Eqs. (5.6.51b) and (5.6.51c), and with the results of [165].

We can also compute the discontinuity in p_2 and p_3 , using Eq. (5.6.43). This gives

$$[\text{Disc}_{S_{23}} T_1]_{R_C^{123}} = \frac{1}{16\pi^2 p_1^2} \frac{2\pi i}{z - \bar{z}} [2\pi i + \text{Li}_1(z) - \text{Li}_1(\bar{z}) + \ln(z - i\varepsilon) - \ln(\bar{z} + i\varepsilon)] \quad (5.6.54)$$

$$= \frac{1}{16\pi^2 p_1^2} \frac{2\pi i}{z - \bar{z}} \ln \left[\frac{(1-\bar{z})(-z)}{(1-z)(-\bar{z})} \right]. \quad (5.6.55)$$

We should compare to the sum of the cuts in p_2 and p_3^2 which can be deduced from Eqs.(5.6.51b) and (5.6.51c):

$$\left[\text{Cut}_{p_2^2} T_1 + \text{Cut}_{p_3^2} T_1 \right]_{R_C^{123}} = \frac{1}{16\pi^2 p_1^2} \frac{2\pi i}{z - \bar{z}} \left[\ln \frac{1-\bar{z}}{1-z} + \ln \frac{z}{\bar{z}} \right]. \quad (5.6.56)$$

Again, we see the discontinuities and cuts agree.

A similar example involves going from $R_A^{123} \rightarrow R_A^* \rightarrow R_A^{123}$. A path between these regions exists that does not go around any branch points. So $[\text{Disc}_{S_{123}} T_1]_{R_C^{123}} = 0$. In R_A^{123} the sum of the cuts also vanishes, although each individual cut does not. In other words, total discontinuity in the dual Euclidean region vanishes, but the discontinuities in separate channels do not. In contrast, in the Euclidean region R_A^{123} , all the cuts vanish individually

(and the total discontinuity is still zero, using the same path).

To take sequential discontinuities in a single channel, we iterate the monodromies in Eq. (5.6.37). We find that these double discontinuities vanish in all channels,

$$\left[\text{Disc}_{p_j^2} \text{Disc}_{p_j^2} T_1 \right]_{R^j} = 0 \quad \forall j. \quad (5.6.57)$$

This is consistent with our expectations, since the triangle has at most one cut in each channel. We can also consider sequential discontinuities of the triangle in different channels, such as $\text{Disc}_{p_2^2} \text{Disc}_{p_3^2} T_1$. The corresponding double cut in p_2^2 and p_3^2 can be computed in the region R^{23} , where $p_2^2 > 0, p_3^2 > 0$, and $p_1^2 < 0$. Using the discontinuity operators defined in Eq. (5.6.42), we find

$$[\text{Disc}_{p_3^2} \text{Disc}_{p_2^2} T_1]_{R^{23}} = (\mathbb{1} - \mathcal{M}_{\curvearrowright_1^z})(\mathbb{1} - \mathcal{M}_{\curvearrowright_0^{\bar{z}}})T_1 = \frac{1}{16\pi^2 p_1^2} \frac{(2\pi i)^2}{z - \bar{z}}. \quad (5.6.58)$$

Notice that we could have equivalently taken these discontinuities in the other order, as both sequences of discontinuities are related to the same cut integrals by Eq. (5.5.13); that is, we have $[\text{Disc}_{p_3^2} \text{Disc}_{p_2^2} T_1]_{R^{23}} = [\text{Disc}_{p_2^2} \text{Disc}_{p_3^2} T_1]_{R^{23}}$. Similarly, we find

$$[\text{Disc}_{p_1^2} \text{Disc}_{p_2^2} T_1]_{R^{12}} = [\text{Disc}_{p_2^2} \text{Disc}_{p_1^2} T_1]_{R^{12}} = -\frac{1}{16\pi^2 p_1^2} \frac{(2\pi i)^2}{z - \bar{z}} \quad (5.6.59)$$

and

$$[\text{Disc}_{p_1^2} \text{Disc}_{p_3^2} T_1]_{R^{13}} = [\text{Disc}_{p_3^2} \text{Disc}_{p_1^2} T_1]_{R^{13}} = -\frac{1}{16\pi^2 p_1^2} \frac{(2\pi i)^2}{z - \bar{z}}. \quad (5.6.60)$$

Notice the additional minus sign in both of these expressions compared to Eq. (5.6.58). As discussed in Appendix B.4, these relative signs are expected from the invariance of the triangle integral under permutations of its external legs.

To illustrate the importance of using the specific operators in Eq. (5.6.42) for computing sequential discontinuities in different channels, we can see what happens if we instead use the discontinuity operators from Eq.(5.6.37). In the case of $\text{Disc}_{p_3^2}\text{Disc}_{p_2^2}T_1$ we would have found

$$\begin{aligned} [\text{Disc}_{p_3^2}[\text{Disc}_{p_2^2}T_1]_{R^2}]_{R^3} &= (\mathbb{1} - \mathcal{M}_{\mathfrak{S}\mathfrak{O}_1^{\bar{z}}})(\mathbb{1} - \mathcal{M}_{\mathfrak{S}\mathfrak{O}_0^{\bar{z}}})T_1 = \frac{1}{16\pi^2 p_1^2} \frac{(2\pi i)^2}{z - \bar{z}}, \\ [\text{Disc}_{p_2^2}[\text{Disc}_{p_3^2}T_1]_{R^3}]_{R^2} &= (\mathbb{1} - \mathcal{M}_{\mathfrak{S}\mathfrak{O}_0^{\bar{z}}})(\mathbb{1} - \mathcal{M}_{\mathfrak{S}\mathfrak{O}_1^{\bar{z}}})T_1 = -\frac{1}{16\pi^2 p_1^2} \frac{(2\pi i)^2}{z - \bar{z}}. \end{aligned} \quad (5.6.61)$$

The results differ by a sign. This highlights the importance of computing the discontinuities by analytically continuing from the region in which the cuts are being computed into adjacent regions.

Let us also reiterate that all the discontinuities we consider are computed along paths in external energies such that energy is conserved. If one tries instead to do what may seem more natural, by continuing the Lorentz invariants directly, one can run into trouble. For example, by continuing z and \bar{z} one can easily go from $R_A^{123} \rightarrow R^{23} \rightarrow R_A^{123}$ by passing around $\bar{z} = 0$ and $z = 1$. The discontinuity along this path is

$$(\mathbb{1} - \mathcal{M}_{\mathfrak{S}\mathfrak{O}_0^{\bar{z}}} \cdot \mathcal{M}_{\mathfrak{S}\mathfrak{O}_1^z})T_1 = \frac{1}{16\pi^2 p_1^2} \frac{2\pi i}{z - \bar{z}} [\text{Li}_1(z) - \text{Li}_1(\bar{z}) + \ln(z) - \ln(\bar{z}) + 2\pi i]. \quad (5.6.62)$$

This is analytic in R_A^{123} , but differs from $\text{Cut}_{p_1^2}T_1$ in R_A^{123} in Eq. (5.6.52) by the extra $2\pi i$.

Thus, specifying the regions of interest is not in general enough: one must also know how to connect them.

Two loops

Next, we consider the two-loop triangle. As before, all internal lines are taken to be massless. The Feynman integral evaluates to

$$\begin{aligned}
 T_2 = & \text{Diagram of a two-loop triangle with external momenta } p_1, p_2, p_3 \text{ and internal lines labeled 1 through 6.} \\
 & \hspace{15em} (5.6.63) \\
 & = \int \frac{d^4 k_1}{i(2\pi)^4} \int \frac{d^4 k_2}{i(2\pi)^4} \frac{1}{k_1^2 (p_3 - k_1)^2 (k_1 + p_1)^2 k_2^2 (p_3 - k_2)^2 (k_1 - k_2)^2} \\
 & = \frac{1}{(4\pi)^4} \frac{1}{p_1^2 p_3^2} \Phi_2(z, \bar{z})
 \end{aligned}$$

where in the region R_A^* the function $\Phi_2(z, \bar{z})$ takes the form

$$\Phi_2(z, \bar{z}) = 6[\text{Li}_4(z) - \text{Li}_4(\bar{z})] - 3 \ln(z\bar{z})[\text{Li}_3(z) - \text{Li}_3(\bar{z})] + \frac{1}{2} \ln^2(z\bar{z})[\text{Li}_2(z) - \text{Li}_2(\bar{z})], \quad (5.6.64)$$

and as before z and \bar{z} satisfy the relations in Eqs. (5.6.32), (5.6.33), and (5.6.34). The variation matrix for this integral is described in Appendix B.5, where the relevant monodromy matrices are also presented.

We first compute the single discontinuities, using the operators in Eq. (5.6.37):

$$\begin{aligned} \left[\text{Disc}_{p_1^2} \Phi_2 \right]_{R^1} &= (2\pi i) \times \left\{ 3\text{Li}_3(\bar{z}) - 3\text{Li}_3(z) + (\ln z + \ln \bar{z} - i\pi)(\text{Li}_2(z) - \text{Li}_2(\bar{z})) \right. \\ &\quad \left. + \frac{1}{2} \ln z (\ln z - \ln \bar{z} + 2\pi i)(\ln \bar{z} - 2\pi i) \right\}, \end{aligned} \quad (5.6.65a)$$

$$\left[\text{Disc}_{p_2^2} \Phi_2 \right]_{R^2} = (2\pi i) \times \left\{ 3\text{Li}_3(z) - 3\text{Li}_3(\bar{z}) - (\ln z + \ln \bar{z} + i\pi)[\text{Li}_2(z) - \text{Li}_2(\bar{z})] \right\}, \quad (5.6.65b)$$

$$\left[\text{Disc}_{p_3^2} \Phi_2 \right]_{R^3} = (2\pi i) \times \left\{ -\frac{1}{2} \ln z \ln \bar{z} (\ln z - \ln \bar{z}) \right\}. \quad (5.6.65c)$$

All the explicit factors of $i\pi$ in these expressions can be absorbed into polylogarithms that are manifestly real in the appropriate region (taking into account Eq. (5.6.44)). The resulting expressions agree with the cuts computed in Eqs. (5.26), (5.37) and (5.41) of [165].

The sequential discontinuities in these channels can be computed using the same monodromy matrices. We find

$$\left[\text{Disc}_{p_1^2} \text{Disc}_{p_1^2} \Phi_2 \right]_{R^1} = (2\pi i)^2 \left\{ \text{Li}_2(z) - \text{Li}_2(\bar{z}) + \frac{1}{2} (\ln z - \ln \bar{z} + 2\pi i)(\ln z + \ln \bar{z}) \right\}, \quad (5.6.66a)$$

$$\left[\text{Disc}_{p_2^2} \text{Disc}_{p_2^2} \Phi_2 \right]_{R^2} = (2\pi i)^2 \left\{ \text{Li}_2(z) - \text{Li}_2(\bar{z}) \right\}, \quad (5.6.66b)$$

$$\left[\text{Disc}_{p_3^2} \text{Disc}_{p_3^2} \Phi_2 \right]_{R^3} = 0. \quad (5.6.66c)$$

Note that the right side of Eq. (5.6.66a) can be rewritten as $\text{Li}_2(1/\bar{z}) - \text{Li}_2(1/z)$ in R^1 , and thus $[\text{Disc}_{p_1^2} \text{Disc}_{p_1^2} \Phi_2]_{R^1}$ and $[\text{Disc}_{p_2^2} \text{Disc}_{p_2^2} \Phi_2]_{R^1}$ get mapped to minus each other under the permutation $p_1 \leftrightarrow p_2$, which corresponds to $z \rightarrow 1/z, \bar{z} \rightarrow 1/\bar{z}$. This is consistent with what

we expect from Appendix B.4. The triple discontinuities all vanish,

$$\left[\text{Disc}_{p_j^2} \text{Disc}_{p_j^2} \text{Disc}_{p_j^2} \Phi_2 \right]_{R^j} = 0, \quad (5.6.67)$$

in accordance with the fact that there aren't three cuts in any of the channels.

These sequential cuts in the same channel have not been computed before to our knowledge. To do so, we regulate the IR divergence of the cuts by giving the lines labeled 4 and 5 in the figure below with a small mass m_{reg} , and work to leading power in m_{reg} . In region R^3 , we find

$$\left[T_{(12),(45)}^{\text{cut}} \right]_{R^3} = \text{Diagram} = -2T_2^{\text{cut}} \quad (5.6.68)$$

where

$$T_2^{\text{cut}} = \frac{1}{64p_1^2 p_3^2 \pi^2 (z - \bar{z})} \ln \frac{m_{\text{reg}}^2}{p_3^2} \ln \frac{z}{\bar{z}}. \quad (5.6.69)$$

The other cuts give multiples of this expression. In particular, we find

$$\left[T_{(12),(135)}^{\text{cut}} \right]_{R^3} = \text{Diagram} = 0, \quad (5.6.70)$$

$$\left[T_{(12),(234)}^{\text{cut}} \right]_{R^3} = \text{Diagram} = 0, \quad (5.6.71)$$

and

$$\left[T_{(45),(135)}^{\text{cut}} \right]_{R^3} = \text{diagram} = T_2^{\text{cut}}, \quad (5.6.72)$$

$$\left[T_{(45),(234)}^{\text{cut}} \right]_{R^3} = \text{diagram} = T_2^{\text{cut}}. \quad (5.6.73)$$

It follows that the sum of all double cuts in R^3 is exactly zero,

$$\left[\sum_{\text{double cuts}} T_2 \right]_{R^3} = 0, \quad (5.6.74)$$

which agrees with Eq. (5.6.66c). Note that the chapters/diagrams in Eq. (5.6.72) and Eq (5.6.73) both involve an isolated three-point vertex with only massless lines. For $d > 4$ such cut graphs may be zero, while they are nonzero in $d = 4$ (they contain integrals of the form $\int dx \delta(x) x^{\frac{d-4}{2}}$). If we were to set them to zero, we would get the wrong answer. This can easily be seen in the example above, as Eq. (5.6.69) would give a non-vanishing result in dimensional regularization, while the graphs in Eq. (5.6.70), Eq. (5.6.71), Eq. (5.6.72), and Eq. (5.6.73) would vanish. See Appendix B.7 for more details.

In R^2 , there is only one diagram. We find

$$\left[T_{(46),(136)}^{\text{cut}} \right]_{R^2} = \text{diagram} = -\frac{1}{256\pi^2 p_1^2 p_3^2} \frac{2\text{Li}_2(z) - 2\text{Li}_2(\bar{z})}{z - \bar{z}}. \quad (5.6.75)$$

Comparing to Eq. (5.6.65b), we then find

$$\left[\sum_{\text{double cuts}} T_2 \right]_{R^2} = \left[\text{Disc}_{p_2^2} \text{Disc}_{p_2^2} T_2(z, \bar{z}) \right]_{R^2} = 2 \left[T_{(46),(136)}^{\text{cut}} \right]_{R^2}, \quad (5.6.76)$$

in agreement with Eq. (5.5.6). The sum of double cuts in the p_1^2 channel are related by $z \leftrightarrow 1/z$, $\bar{z} \leftrightarrow 1/\bar{z}$ to the sum of double cuts in the p_2^2 channel, and thus the sum of double cuts in R^1 is related to the sequential discontinuity computed in Eq. (5.6.66a) by the same combinatorial factor. These provide highly nontrivial checks of Eq. (5.5.6).

Finally, we compute the sequential discontinuities in different channels. We find

$$\left[\text{Disc}_{p_3^2} \text{Disc}_{p_1^2} \Phi_2 \right]_{R^{13}} = (2\pi i)^2 \left\{ -\frac{1}{2} \ln^2 z + \ln z \ln(\bar{z} + i\varepsilon) - i\pi \ln z \right\}, \quad (5.6.77a)$$

$$\left[\text{Disc}_{p_2^2} \text{Disc}_{p_1^2} \Phi_2 \right]_{R^{12}} = (2\pi i)^2 \left\{ \text{Li}_2(\bar{z}) - \text{Li}_2(z - i\varepsilon) - \frac{1}{2} \ln^2 z + \ln z \ln \bar{z} - i\pi \ln z \right\}, \quad (5.6.77b)$$

$$\left[\text{Disc}_{p_2^2} \text{Disc}_{p_3^2} \Phi_2 \right]_{R^{23}} = (2\pi i)^2 \left\{ \frac{1}{2} \ln^2 z - \ln z \ln(\bar{z} - i\varepsilon) - i\pi \ln z \right\}. \quad (5.6.77c)$$

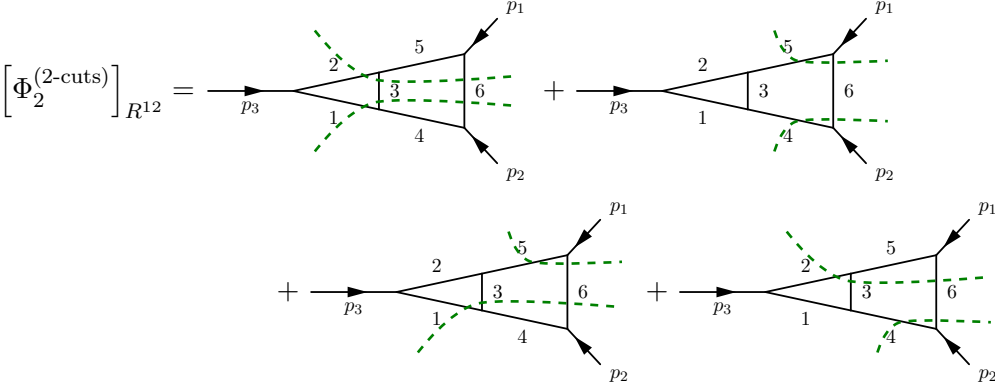
We believe these agree with the results in [165].

9

Recall that [165] uses a different cut prescription, which involves both $-i\varepsilon$ and $+i\varepsilon$ propagators, and that they use dimensional regularization and so massless three-point vertices vanish. For reasons discussed in Appendix B.7, we believe it is safer to use a mass regula-

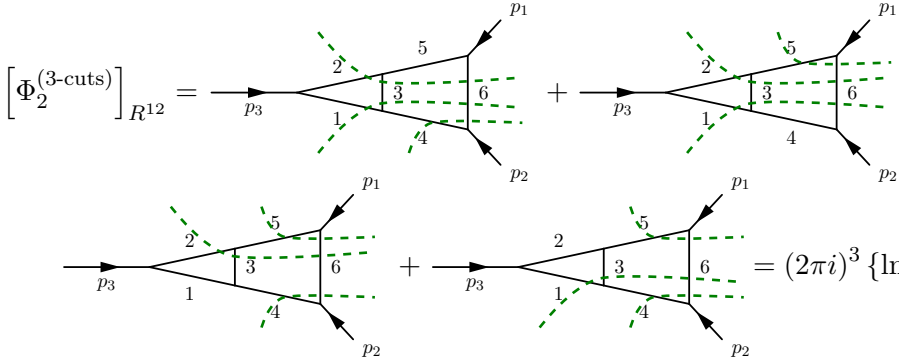
⁹These equations differ slightly from Eqs. (6.4) and (6.5) in [165]. However, summing the results from their Appendix D, we believe their (6.4) should agree with our Eq. (5.6.77a). For $\text{Disc}_{p_2^2} \text{Disc}_{p_1^2} \Phi_2$, we find that summing their expressions with some typos corrected gives $(2\pi i)^2 \{ \text{Li}_2(\bar{z}) + \text{Li}_2(1 - z) + \ln(z - 1) \ln z - \frac{1}{2} \ln^2 z + \ln z \ln \bar{z} - \frac{\pi^2}{6} \}$, which agrees with Eq. (5.6.77b).

tor. With our $+i\varepsilon$ convention, the double-cut graphs in R^{12} give



$$\begin{aligned}
\left[\Phi_2^{(2\text{-cuts})} \right]_{R^{12}} &= \text{[Diagram 1]} + \text{[Diagram 2]} \\
&\quad + \text{[Diagram 3]} + \text{[Diagram 4]} \\
&= (2\pi i)^2 \left\{ \text{Li}_2(\bar{z}) - \text{Li}_2(z - i\varepsilon) - \frac{1}{2} \ln^2 z + \ln z \ln \bar{z} + i\pi \ln z - 2\pi i \ln \bar{z} \right\}.
\end{aligned} \tag{5.6.78}$$

To match onto the discontinuity in Eq. (5.6.77b), we must in our analysis add the three-cut graphs according to Eq. (5.5.15). We find



$$\left[\Phi_2^{(3\text{-cuts})} \right]_{R^{12}} = \text{[Diagram 1]} + \text{[Diagram 2]} + \text{[Diagram 3]} + \text{[Diagram 4]} = (2\pi i)^3 \{ \ln z - \ln \bar{z} \} \tag{5.6.79}$$

Inserting into Eq. (5.5.15) the sum of all cuts gives

$$\left[\Phi_2^{(2\text{-cuts})} - \Phi_2^{(3\text{-cuts})} \right]_{R^{12}} = (2\pi i)^2 \left\{ \text{Li}_2(\bar{z}) - \text{Li}_2(z - i\varepsilon) - \frac{1}{2} \ln^2 z + \ln z \ln \bar{z} - i\pi \ln z \right\} \tag{5.6.80}$$

in agreement with the discontinuity in Eq. (5.6.77b). In particular, the three-cut chap-

ters/diagrams $\Phi_2^{(3\text{-cuts})}$ containing massless three-point vertices must be added to get the correct result. We have verified this result using a mass regulator, and the technique discussed in Appendix B.7. Note that while these chapters/diagrams add up to a finite result in this case, each diagram would naïvely be set to zero in dimensional regularization as discussed earlier, which would lead to a wrong result. Further discussion on how to calculate massless three-point vertices can be found in Appendix B.7.

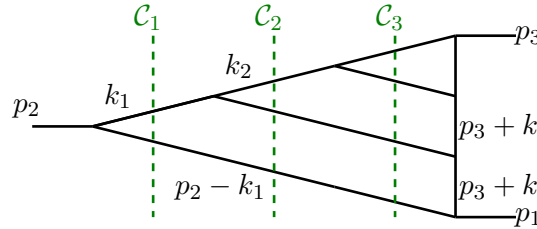
Three loops

It is instructive to continue to three loops. The most interesting case is the one in which two cuts are taken in the p_2^2 channel, where Eq. (5.5.6) tells us we should find

$$\left[\text{Disc}_{p_2^2}^2 T_3 \right]_{R^2} = \left[2T_3^{(2\text{-cuts})} - 6T_3^{(3\text{-cuts})} \right]_{R^2} \quad (5.6.81)$$

when we assign all propagators $+i\varepsilon$.

The three-loop triangle



The diagram shows a three-loop triangle. On the left, an external momentum p_2 enters from the left and splits into two lines. The upper line carries momentum k_1 and the lower line carries momentum $p_2 - k_1$. These lines meet at a vertex. From this vertex, two lines go to the right. The upper line carries momentum k_2 and the lower line carries momentum $p_2 - k_1$. These lines meet at another vertex. From this vertex, two lines go to the right. The upper line carries momentum p_3 and the lower line carries momentum $p_3 + k_2$. These lines meet at a third vertex. From this vertex, two lines go to the right. The upper line carries momentum $p_3 + k_1$ and the lower line carries momentum p_1 . Three vertical dashed green lines represent cuts, labeled \mathcal{C}_1 , \mathcal{C}_2 , and \mathcal{C}_3 from left to right. \mathcal{C}_1 is between the first and second vertices, \mathcal{C}_2 is between the second and third vertices, and \mathcal{C}_3 is between the third and fourth vertices.

(5.6.82)

is given by

$$T_3 = -\frac{1}{6(4\pi)^6 p_1^2 p_3^4 (z - \bar{z})} \left\{ [\text{Li}_3(z) - \text{Li}_3(\bar{z})] \ln^3(z\bar{z}) - 12 [\text{Li}_4(z) - \text{Li}_4(\bar{z})] \ln^2(z\bar{z}) \right. \\ \left. + 60 [\text{Li}_5(z) - \text{Li}_5(\bar{z})] \ln(z\bar{z}) - 120 [\text{Li}_6(z) - \text{Li}_6(\bar{z})] \right\}. \quad (5.6.83)$$

Taking two discontinuities in the p_2^2 channel using Eq. (5.6.37) gives

$$\text{Disc}_{p_2^2}^2 T_3 = \frac{1}{1024\pi^4 p_1^2 p_3^4 (z - \bar{z})} \left\{ [\text{Li}_3(z) - \text{Li}_3(\bar{z})] [\ln(z\bar{z}) + 2\pi i] - 4 [\text{Li}_4(z) - \text{Li}_4(\bar{z})] \right\}, \quad (5.6.84)$$

while taking three discontinuities results in

$$\text{Disc}_{p_2^2}^3 T_3 = -\frac{i}{512\pi^3 p_1^2 p_3^4 (z - \bar{z})} [\text{Li}_3(z) - \text{Li}_3(\bar{z})]. \quad (5.6.85)$$

To facilitate the cut computation, we rewrite Eq. (5.5.6) in a way that allows us to recycle results for the single cuts of the two-loop triangle. The sum of all single cuts in the p_2^2 channel of the two-loop triangle $T_2(p_1^2, p_2^2, p_3^2)$, with the traditional $i\varepsilon$ prescription involving $-i\varepsilon$'s to the right of the cut, was shown in [165] to agree with the discontinuity in p_2^2 . We can use these results if we rewrite the term corresponding to the double cut $\mathcal{C}_1\mathcal{C}_2$ in Eq. (5.6.82) to have $-i\varepsilon$'s to the right of the cut, adding a triple cut term to compensate for it. When doing so, we must be careful with the combinatorial factors that come along with massless three-point vertices, as these cut integrals involve delta functions with support only at integration endpoints. In Appendix B.7, we show that one gets an additional

factor of $\frac{1}{m!}$ compared to naïvely evaluating these delta functions to 1, where m is the number of cuts being taken. Thus, we must absorb a term $-6T_3^{(3\text{-cuts})}$ to correct the $+i\varepsilon$'s to $-i\varepsilon$'s in the term $2T_3^{\mathcal{C}_1\mathcal{C}_2}$. The result we want to verify is therefore

$$\left[\left(\text{Disc}_{p_2^2} \right)^2 T_3 \right]_{R^2} = 2 T_3^{\mathcal{C}_1, \mathcal{C}_2} \Big|_{-i\varepsilon \text{ on r.h.s.}} + 2 T_3^{\mathcal{C}_1, \mathcal{C}_3} + 2 T_3^{\mathcal{C}_2, \mathcal{C}_3}. \quad (5.6.86)$$

The first two terms in this expression correspond to cutting in \mathcal{C}_1 and summing over the one-cuts of the two-loop triangle. The details of the calculation are worked out in Appendix B.6, and the result is

$$\begin{aligned} T_3^{\mathcal{C}_1, \mathcal{C}_2} \Big|_{-i\varepsilon \text{ on r.h.s. of cut}} + T_3^{\mathcal{C}_1, \mathcal{C}_3} &= \frac{1}{2048\pi^4} \frac{1}{p_1^2 p_3^4 (z - \bar{z})} \left\{ -3 [\text{Li}_4(z) - \text{Li}_4(\bar{z})] \right. \\ &\quad \left. + \ln \left(-\frac{m^2}{p_3^2} \right) [\text{Li}_3(z) - \text{Li}_3(\bar{z})] - \frac{1}{2} [\text{Li}_2^2(z) - \text{Li}_2^2(\bar{z})] \right\}, \end{aligned} \quad (5.6.87)$$

where m is a small mass of the line labelled as k used to regulate the IR divergence of the cut graphs. The cut $T_3^{\mathcal{C}_2\mathcal{C}_3}$ is given by

$$\begin{aligned} T_3^{\mathcal{C}_2, \mathcal{C}_3} &= \frac{1}{2048\pi^4 p_1^2 p_3^4 (z - \bar{z})} \left\{ \left[-\ln \left(-\frac{m^2}{p_3^2} \right) + \ln(z\bar{z}) + 2\pi i \right] [\text{Li}_3(z) - \text{Li}_3(\bar{z})] \right. \\ &\quad \left. + \left[\frac{1}{2} \text{Li}_2(z) - \frac{1}{2} \text{Li}_2(\bar{z}) \right] - [\text{Li}_4(z) - \text{Li}_4(\bar{z})] \right\}. \end{aligned} \quad (5.6.88)$$

The sum of all cuts is therefore

$$\begin{aligned}
& 2 T_3^{\mathcal{C}_1, \mathcal{C}_2} \Big|_{-i\varepsilon \text{ on r.h.s.}} + 2 T_3^{\mathcal{C}_1, \mathcal{C}_3} + 2 T_3^{\mathcal{C}_2, \mathcal{C}_3} \\
&= \frac{1}{1024\pi^4 p_1^2 p_3^4 (z - \bar{z})} \left\{ [\text{Li}_3(z) - \text{Li}_3(\bar{z})] [\ln(z\bar{z}) + 2\pi i] - 4 [\text{Li}_4(z) - \text{Li}_4(\bar{z})] \right\} \quad (5.6.89)
\end{aligned}$$

in agreement with Eqs. (5.6.86) and (5.6.84).

L loops

Let us now consider the L -loop triangle integral,

$$T_L(p_1^2, p_2^2, p_3^2) = \text{Diagram} \quad (5.6.90)$$

$$\begin{aligned}
&= \int \frac{d^4 k_1}{i (2\pi)^4} \cdots \frac{d^4 k_L}{i (2\pi)^4} \frac{1}{k_1^2 + i\varepsilon} \frac{1}{k_2^2 + i\varepsilon} \cdots \frac{1}{k_L^2 + i\varepsilon} \frac{1}{(p_2 - k_1)^2 + i\varepsilon} \quad (5.6.91) \\
&\times \frac{1}{(k_1 - k_2)^2 + i\varepsilon} \cdots \frac{1}{(k_{L-1} - k_L)^2 + i\varepsilon} \frac{1}{(p_3 - k_1)^2 + i\varepsilon} \cdots \frac{1}{(p_3 - k_L)^2 + i\varepsilon}.
\end{aligned}$$

The result after performing the loop integration is [207]

$$T_L(p_1^2, p_2^2, p_3^2) = -\frac{1}{z - \bar{z}} \frac{1}{L! (4\pi)^{2L} p_1^2 (p_3^2)^{L-1}} \sum_{j=L}^{2L} \frac{(-1)^j j! \ln^{2L-j}(z\bar{z})}{(j-L)! (2L-j)!} [\text{Li}_j(z) - \text{Li}_j(\bar{z})], \quad (5.6.92)$$

with z and \bar{z} defined as before.

One thing we can immediately observe about this expression is that taking two or more

discontinuities along the long axis (in the p_3^2 channel) gives zero. To see this, we note that taking a discontinuity in p_3^2 corresponds to taking a monodromy around $z = 1$, which is only nonvanishing for the $\text{Li}_j(z)$ factor in Eq. (5.6.92). Using the fact that the discontinuity of $\text{Li}_n(z)$ corresponding to encircling the branch point at $z = 1$ gives $2\pi i \frac{\ln^{n-1} z}{(n-1)!}$, we get

$$\left[\text{Disc}_{p_3^2} T_L(p_1^2, p_2^2, p_3^2) \right]_{R^3} = \frac{-2\pi i}{z - \bar{z}} \frac{1}{L! (4\pi)^{2L} p_1^2 (p_3^2)^{L-1}} \sum_{j=L}^{2L} \frac{(-1)^j j! \ln^{2L-j}(z\bar{z}) \ln^{j-1}(z)}{(j-L)! (2L-j)! (j-1)!} . \quad (5.6.93)$$

In this expression, there are no longer branch points at 1 in z or \bar{z} . Thus, further discontinuities in p_3^2 vanish,

$$\left[(\text{Disc}_{p_3^2})^2 T_L \right]_{R^3} = 0 . \quad (5.6.94)$$

The sum of taking two and more cuts of the L -loop triangle along the long axis must correspondingly also vanish.

We now show that taking L discontinuities in the p_2^2 channel amounts to taking L cuts in the same channel, i.e.

$$\left[(\text{Disc}_{p_2^2})^L T_L(p_1^2, p_2^2, p_3^2) \right]_{R^2} = [L! \text{Cut}_{\mathcal{C}_1, \dots, \mathcal{C}_L} T_L(p_1^2, p_2^2, p_3^2)]_{R^2} . \quad (5.6.95)$$

We start by computing the sequential discontinuity, which amounts to taking L discontinuities of the factor $\ln^{2L-j}(z\bar{z})$ in the expression above. Only the first term in the sum over j , where $j = L$, contributes to this discontinuity. The result is

$$\left[\text{Disc}_{p_2^2}^L T_L(p_1^2, p_2^2, p_3^2) \right]_{R^2} = -\frac{i^L}{z - \bar{z}} \frac{1}{(8\pi)^L p_1^2 (p_3^2)^{L-1}} [\text{Li}_L(z) - \text{Li}_L(\bar{z})] . \quad (5.6.96)$$

Next, we calculate the cuts. Putting the lines corresponding to the cuts $\mathcal{C}_1 \cdots \mathcal{C}_L$ on-shell in the region R^2 gives the following expression:

$$\begin{aligned} \text{Cut}_{\mathcal{C}_1, \dots, \mathcal{C}_L} T_L(p_1^2, p_2^2, p_3^2) &= \int \frac{d^4 k_1}{i(2\pi)^4} \cdots \frac{d^4 k_L}{i(2\pi)^4} (-2\pi i)^{2L} \delta(k_1^2) \Theta(k_1^0) \cdots \delta(k_L^2) \Theta(k_L^0) \\ &\times \delta\left[(p_2 - k_1)^2\right] \Theta(p_2^0 - k_1^0) \delta\left[(k_1 - k_2)^2\right] \Theta(k_1^0 - k_2^0) \cdots \delta\left[(k_{L-1} - k_L)^2\right] \Theta(k_{L-1}^0 - k_L^0) \\ &\times \frac{1}{(p_3 + k_1)^2} \frac{1}{(p_3 + k_2)^2} \cdots \frac{1}{(p_3 + k_L)^2}. \quad (5.6.97) \end{aligned}$$

We perform the energy integrals using the delta functions $\delta(k_1^2) \cdots \delta(k_L^2)$, and get

$$\begin{aligned} \text{Cut}_{\mathcal{C}_1, \dots, \mathcal{C}_L} T_L(p_1^2, p_2^2, p_3^2) &= \int \frac{d^3 k_1}{(2\pi)^3 2\omega_{k_1}} \frac{d^3 k_2}{(2\pi)^3 2\omega_{k_2}} \cdots \frac{d^3 k_L}{(2\pi)^3 2\omega_{k_L}} (2\pi i)^L \delta(p_2^2 - 2p_2 \cdot k_1) \\ &\times \delta(-2k_1 \cdot k_2) \cdots \delta(-2k_{L-1} \cdot k_L) \frac{1}{(p_3 + k_1)^2} \cdots \frac{1}{(p_3 + k_L)^2}. \quad (5.6.98) \end{aligned}$$

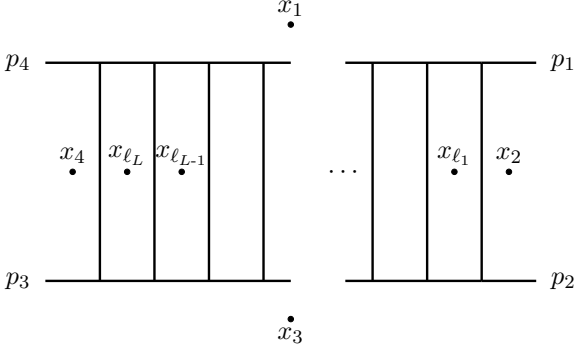
The remaining delta functions show that this cut only has support when the momenta k_1, \dots, k_L and $k_1 - k_2, \dots, k_{L-1} - k_L$ are all collinear. We therefore get a product of $L-1$ massless vertices. This configuration is singular and must be treated with care, using TOPT. As explained in Appendix B.7, evaluating the integrals over these remaining delta functions gives rise to a combinatorial factor of $\frac{1}{L!}$. The result of the integral, worked out in detail in Appendix B.7, is

$$\text{Cut}_{\mathcal{C}_1 \cdots \mathcal{C}_L} T_L(p_1^2, p_2^2, p_3^2) = -\frac{i^L}{z - \bar{z}} \frac{1}{L! (8\pi)^L p_1^2 (p_3^2)^{L-1}} [\text{Li}_L(z) - \text{Li}_L(\bar{z})]. \quad (5.6.99)$$

Comparing this result to Eq. (5.6.96), we see that Eq. (5.6.95) is indeed satisfied.

Sequential discontinuities of the L -loop box ladders

We finally comment on the sequential discontinuities of the L -loop box ladder,

$$B_L(p_1^2, p_2^2, p_3^2, p_4^2, (p_1 + p_2)^2, (p_2 + p_3)^2) =$$


$$(5.6.100)$$

These ladder integrals yield the same transcendental functions as the triangle integrals.

This is easiest to see in dual space, as first considered in [209].¹⁰ Translating to dual space, we label the dual points corresponding to loops by x_{ℓ_i} , and by x_j for external points, with $j \in \{1, 2, 3, 4\}$. The ladder integral is then given by

$$B_L \propto ((x_1 - x_3)^2)^L (x_2 - x_4)^2$$

$$\times \int \prod_{i=1}^L \frac{d^4 x_{\ell_i}}{(x_{\ell_i} - x_1)^2 (x_{\ell_i} - x_3)^2} \frac{1}{(x_2 - x_{\ell_1})^2} \prod_{i=1}^{L-1} \frac{1}{(x_{\ell_i} - x_{\ell_{i+1}})^2} \frac{1}{(x_{\ell_L} - x_4)^2}.$$

$$(5.6.101)$$

This integral is invariant under conformal transformations of the dual variables x , which can be shown using Lorentz invariance and the (less obvious) invariance under inversion

¹⁰Dual space can be defined as follows: for any planar diagram, we associate a variable x_{ℓ_i} for each loop and a variable x_i for each region between two external lines. Then, once we pick an orientation on each of the edges, we take the momentum flowing through that edge to be the difference between the dual variable on the right and the dual variable on the left. This ensures momentum conservation at each vertex. In some cases, the dual variables make manifest hidden symmetries, such as the dual conformal symmetry (see [210]).

$x^\mu \rightarrow \frac{x^\mu}{x^2}$. By a combination of translation and inversion we can send x_4 to infinity. In this limit we have $\frac{(x_2-x_4)^2}{(x_{\ell_L}-x_4)^2} \rightarrow 1$. This is precisely the triangle ladder in dual space. The box and triangle integrals therefore give the same analytic expression, and working out the exact transformation between the two, one can show that z and \bar{z} variables for the box are given in terms of the Mandelstams as

$$z\bar{z} = \frac{p_2^2 p_4^2}{p_1^2 p_3^2}, \quad (1-z)(1-\bar{z}) = \frac{st}{p_1^2 p_3^2}, \quad (5.6.102)$$

with $s = (p_1 + p_2)^2$ and $t = (p_2 + p_3)^2$. All of the analysis for the triangle integrals therefore extends to L -loop box ladders.

We can also compute the sequential discontinuity of the box ladder integrals in the s and then t channels, which is expected to vanish due to the Steinmann relations. To compute this quantity, we go to the region $R^{\{s,t\}}$, where $s, t > 0$ while all other invariants $p_i^2 < 0$. For concreteness, we consider the phase-space point

$$p_1 = (1, 5, -6, 0), \quad p_2 = (1, -6, 5, 0), \quad (5.6.103)$$

$$p_3 = (1, 7, -6, 0), \quad p_4 = (-3, -4, 7, 0). \quad (5.6.104)$$

We can analytically continue into R^t by rescaling $E_1 \rightarrow \alpha E_1$ and $E_2 \rightarrow \alpha E_2$ by $1 > \alpha > 0$, while keeping E_3 fixed and varying $E_4 = -E_1 - E_2 - E_3$ along with E_1 and E_2 . We then return to $R^{\{s,t\}}$ by the reverse path, after encircling the branch point at $s = 0$. In the z and \bar{z} variables, this corresponds to analytically continuing around $z = 1$. A similar path around

the branch point at $t = 0$ can be constructed by instead rescaling E_2 and E_3 , and also corresponds to computing a monodromy around $z = 1$. Since this sequence of discontinuity operators is identical to the sequence of operators used to compute sequential discontinuities in the p_3^2 channel of the triangle, Eq. (5.6.94) confirms that the Steinmann relations are satisfied by the box ladder integral at all loop orders. This matches the Steinmann analysis carried out in [193], where the expression that appears in Eq. (5.6.93) was also shown to reduce to a simpler functional form (given as Eq. (19) of that paper, which has slightly different rational normalization).

5.7 Conclusions

In this paper we have analyzed the discontinuities and cuts of Feynman integrals from several points of view. We first described how to compute the imaginary part of Feynman integrals in terms of cuts, reviewing the work of Cutkosky and 't Hooft and Veltman, and also described the analogous relations in non-covariant time-ordered perturbation theory. These traditional approaches are based on the idea that Feynman integrals have branch cuts in physical regions, and that integrals over propagators with $+i\epsilon$ and $-i\epsilon$ displacements are on opposite sides of these branch cuts. The main focus of this paper has been to extend these methods to sequential discontinuities. The $\pm i\epsilon$ prescription is in general insufficient for computing more than one discontinuity, but the relevant computations can be carried out by considering monodromies around the branch points of Feynman integrals. In particular, by understanding discontinuities in terms of monodromies, we are able to homogenize

the $+i\varepsilon$ and $-i\varepsilon$ propagators that appear after the first discontinuity by analytically continuing them into same cut complex plane. This allows subsequent discontinuities to be taken. For integrals that are expressible in terms of generalized polylogarithms, we have also described how discontinuities can be computed using variation matrices and the monodromy group.¹¹

The main result of this paper is a formula relating the sequential discontinuities in the same or different channels around branch points associated with invariants s_j to cuts:

$$\begin{aligned} & \frac{1}{m_1!} \cdots \frac{1}{m_n!} [(\text{Disc}_{s_1})^{m_1} \cdots (\text{Disc}_{s_n})^{m_n} \mathcal{M}]_{R^{\{s_1, \dots, s_n\}}} \\ &= \sum_{k_1=m_1}^{\infty} \begin{Bmatrix} k_1 \\ m_1 \end{Bmatrix} \cdots \sum_{k_n=m_n}^{\infty} \begin{Bmatrix} k_n \\ m_n \end{Bmatrix} (-1)^{\sum m_i - \sum k_i} \left[\mathcal{M}_{\{k_i \text{ cuts in channel } s_i\}} \right]_{R_+^{\{s_1, \dots, s_n\}}} \cdot \quad (5.7.1) \end{aligned}$$

It is crucial that these relations are understood to apply only in regions where all the cuts of interest are nonvanishing. In particular, we emphasize that these discontinuities are always taken as the difference between \mathcal{M} evaluated at the same physical value of real external momenta on different Riemann sheets.

An important consideration that we have spent considerable time exploring is that the analytic continuations by which these discontinuities are computed must be chosen with care. Paths that are homologous but not in the same homotopy class may give different answers (as discussed in Appendix B.2). In addition, the derivation of our formulas is made assuming a path exists which continues the external energies, holding the three-momenta

¹¹It should be possible to extend the variation and monodromy matrix construction to elliptic polylogarithms [155–157], which also appear in Feynman integral calculations. It would be interesting to see if it could also be used in conjunction with the diagrammatic coaction [211–213].

fixed and respecting energy conservation. We have presented many nontrivial examples of cut and discontinuity computations, and have checked that Eqs. (5.5.8) and (5.5.13) hold in these examples. For each example, we have been sure to find an explicit path in energies connecting the relevant regions, and used the path to determine which branch points are encircled. If one just picks an arbitrary path between regions, the discontinuity can still be computed, but there is no guarantee of agreement with cuts (and in fact, the agreement sometimes fails). While there is undoubtedly a more covariant way to understand the constraints on the paths, in every case where we have found an explicit path in energy we have found agreement between discontinuities and cuts according to our formulas, and conversely, in cases where our formulas seem to fail, we have not been able to find an explicit path in energy between regions (so that our formulas do not apply).

An important class of sequential discontinuities described by Eq. (5.5.13) are those in which the discontinuity channels are partially overlapping. In these cases, this equation encodes the Steinmann relations, originally derived using axiomatic quantum field theory, which state that sequential discontinuities in partially overlapping channels must vanish. In the original S -matrix program, this was shown to hold for full non-perturbative S -matrix elements in a mass-gapped scalar quantum field theory. Our analysis implies that the Steinmann relations in fact hold for individual Feynman integrals.¹² This amounts to a proof of the Steinmann relations in perturbation theory, diagram by diagram. Our proof requires only that the region where both channels can be simultaneously cut must exist, and that

¹²It had previously been observed that the Steinmann relations were obeyed by many of the Feynman integrals that appear in planar $\mathcal{N} = 4$, insofar as these integrals appear in the space of Steinmann-satisfying hexagon functions [92, 186, 214].

the external momenta are not constrained (for instance by being massless).

Of course, the constraint that all external lines be massive is a strong one, and excludes many theories of physical interest. As such, it would be good to understand the massless case in more depth. The tools we have developed should in principle apply to any Feynman integral, but a full analysis of the massless case involves an additional profusion of subtleties. For example, if we regulate the IR divergences of the massless box by going to $d > 4$ dimensions, we get a $\ln s \ln t$ contribution (see Eq. (5.5.20)), and a nonzero (and IR-finite) sequential monodromy in s and t . However, regulating the external lines with masses, as done in the four-mass box, the sequential monodromy in s and t vanishes (this follows from Eq. (5.6.57), if we use Eq. (5.6.102) to map the triangle to the box integral). Thus, this sequential discontinuity, despite being IR finite, is regulator-dependent. We leave further study of these subtleties to future work.

Time-ordered perturbation theory played an essential role in our derivation. There is a sense in which time-ordered perturbation theory is more physical than covariant perturbation theory, since particles are always on-shell. Indeed, the benefits of a non-covariant formulation in some other contexts are well-known, such as how light-cone perturbation theory is used to show factorization, and new uses are constantly being developed, such as for cosmological polytopes [215, 216]. It would be interesting to see if Steinmann-type constraints and the monodromy group could be useful as a bootstrapping technique in cosmological contexts.

The existence of IR divergences in amplitudes involving massless particles actually facilitates the study of certain aspects of these amplitudes. The IR structure of gauge the-

ories is particularly well understood: a scattering amplitude can be factorized into a hard part, a jet (collinear) part, and a soft part [49, 50, 75, 76, 79–81, 101, 139]. The hard part is IR-finite and can be interpreted as the S -matrix (the ‘hard’ S -matrix) in a computational scheme where the soft and collinear parts are included in the asymptotic Hamiltonian [95, 200]. This suggests that analytic properties of the hard part alone might be amenable to the same techniques used to study massive, IR-finite theories like we have done in this paper. Indeed, the analytic properties of scattering amplitudes in planar $\mathcal{N} = 4$ super-Yang-Mills theory are usually studied at the level of IR-finite remainder functions, which can also be interpreted as hard S -matrix elements. In fact, this connection was part of the motivation for the current work.

The soft part of the scattering amplitude in theories with massless particles can also reproduce the IR-dominated non-analytic behavior of the full S -matrix elements. The soft function, which can be represented as a matrix element of Wilson lines, satisfies a renormalization group equation and can be written as the exponential of the integral of the soft anomalous dimension [122, 217–222]. The soft anomalous dimension depends on kinematics and is a matrix in color space; it contains a dipole part, which is diagonal in color space, and a correction term with restricted kinematic dependence [128, 141]. The dipole part is determined by the cusp anomalous dimension, and is proportional to $\sum_{i < j} \ln(\frac{-p_i \cdot p_j}{\mu^2})$, where μ is the renormalization-group scale. The correction to the dipole formula depends only on the directions of the external momenta and not on their magnitudes; this implies that it can only depend on rescaling-invariant cross-ratios of the form $\rho_{ijkl} = \frac{(p_i \cdot p_j)(p_k \cdot p_l)}{(p_i \cdot p_k)(p_j \cdot p_l)}$. This constitutes a strong constraint, and in particular implies that a soft function can never

have cuts in channels with more than two particles. Since simultaneously cutting a pair of partially-overlapping two-particle channels isolates a one-particle channel, i.e. a decay, such partially-overlapping cuts are forbidden in theories with only stable particles. This is one way to understand the Steinmann relation in S -matrix theory in the soft limit. In contrast, in theories with massless particles, $1 \rightarrow n$ amplitudes do not have to identically vanish. Correspondingly, the articulation of the Steinmann relations for these theories proves to be more challenging. Nevertheless, the restriction to two-particle cuts in the soft limit gives a clue to how we might understand the analytic properties of the massless case. Also, since the soft function is an expectation value of a product of Wilson lines, one could ask what restrictions causality imposes on this expectation value.¹³

To facilitate our analysis, we have presented an introduction to the monodromies of poly-logarithmic functions, drawing inspiration from [201, 202]. A central role in this analysis is played by the connection ω and an integration contour γ . These ingredients are sufficient to determine a variation matrix via $\mathcal{M}_\gamma = \mathcal{P} \exp \int_\gamma \omega$. The variation matrix is a homotopy functional, i.e. its value depends only on the homotopy class of the integration contour γ . In typical cases, the number of homotopy classes is infinite. Nevertheless, in physical applications one rarely considers analytic continuations in the full domain of analyticity; in the examples we studied, it was sufficient to consider rotations in the phases of energies. The allowed sequences of cuts correspond to non-vanishing elements in the variation matrix, while forbidden sequences of cuts correspond to vanishing elements.

This type of reasoning, in which the vanishing of certain cuts (or sequences of cuts) is

¹³While the Steinmann relations were initially studied for correlation functions of local operators, the implications of causality on non-local operators do not seem to have been studied.

used to constrain the analytic structure of polylogarithmic scattering amplitudes and Feynman integrals, has appeared in a number of contexts (see for instance [165, 178, 179, 195, 208, 223]). These analyses are often carried out at the level of the symbol, with the resulting objects only being later upgraded to full polylogarithmic functions using the methods of [153, 154, 223] (or more implicitly, using the methods reviewed in [224]). It is important to note, however, that when such constraints are imposed directly at the level of the symbol, it is not always clear whether the corresponding cuts can arise in the physical region, or only outside of it. This could prove salient, as the Steinmann relations do not necessarily apply when the relevant cuts are not accessible within the physical region.

It would be particularly interesting to understand whether the Steinmann-type constraints that prove useful in planar $\mathcal{N} = 4$ [92] all correspond to cuts that are accessible within physical regions, or point to some further special property of these amplitudes. In particular, it has been observed that these constraints can be generalized to the *extended* Steinmann relations, which apply to sequential discontinuities at all depths in the symbol [186, 214], and that these extended constraints exhibit intriguing connections to cluster algebras [225]. The extended Steinmann relations have been used in conjunction with additional formal constraints, such as integrability (which ensures that symbols can be upgraded to genuine functions), first entry conditions (which constrain the branch cuts that are accessible on the boundary of the Euclidean region), and last entry conditions (which constrain the derivative of these amplitudes) to formulate ansätze for six- and seven-particle amplitudes in this theory, which can be further constrained in special kinematic limits to determine the amplitude at a given loop order [92, 144, 146, 179–185]. These types of constraints

can all be conveniently formulated in terms of the connection ω . The integrability condition is just the requirement that $\omega \wedge \omega = 0$, the first entry condition constrains the differentials that appear in the first row of ω , and the last entry condition constrains differentials that appear in the last column of ω .

In fact, one can consider bootstrapping Feynman integrals directly in terms of the elements of their variation matrices \mathcal{M}_γ .¹⁴ Many of the entries in the right column of \mathcal{M}_γ correspond to different (sequential) cut channels, and should therefore be expressible as integrals over the phase space of on-shell amplitudes.¹⁵ The integrability condition $\omega \wedge \omega = 0$ imposes linear constraints that relate these cut integrals to the other entries of \mathcal{M} . Moreover, when working in terms of the connection ω , one can impose additional constraints having to do with the unipotence of its monodromy matrices, namely that property that $(\mathbb{1} - \mathcal{M}_{\mathbb{S}^x_p})^k = 0$ for some integer k , where this integer k is related to the number of cuts one can take in channel corresponding to this monodromy. More generally, this unipotence property provides strong constraints on the underlying mixed Hodge structure of the polylogarithmic functions that arise from Feynman integrals, and it would be interesting to understand these constraints in more detail.

¹⁴A similar idea, of using dispersion relations to complete the coproduct of a Feynman integral, was developed in [165].

¹⁵This will not be true in channels that are only accessibly outside of physical regions.



Appendix to Chapter 2

A.1 On-shell intermediate propagators

In this appendix, we explain how to compute cut diagrams with on-shell intermediate propagators, as in Eq. (2.5.7). Consider the contribution to the total cross section for γZ scattering with a final state $e^+e^-\gamma$, interfered with a disconnected diagram:

$$\sigma_{11.A} = \text{Diagram} + \text{c.c.} \quad (\text{A.1.1})$$

This is the same as the first diagram in Eq. (2.5.7), but crossed so the γZ is incoming. We do this to separate this complication of on-shell intermediate states from that of integrating over 3-body initial-state phase space. The spin-summed cut diagram is

$$\begin{aligned} \sigma_{11.A} = \sigma_0 \int \frac{d^d k}{(2\pi)^d} \frac{d^d q}{(2\pi)^d} \frac{d^d p}{(2\pi)^d} & 2\pi\delta(k^2)\theta(k_0)2\pi\delta(p^2)\theta(p_0)2\pi\delta(q^2)\theta(q_0) \\ & \times (2\pi)^d \delta^d(q + p - p_Z) (2\pi)^{d-1} (2\omega_k) \delta^{d-1}(k' - k) \\ & \times \frac{i}{(p+k)^2 + i\varepsilon} \frac{i}{(p+k-k')^2 + i\varepsilon} \text{Tr}[\not{p}\gamma^\mu(\not{p} + \not{k})\gamma^\mu(\not{p} + \not{k} - \not{k}')\gamma^\alpha \not{q}\gamma^\alpha] + \text{c.c.} \end{aligned} \quad (\text{A.1.2})$$

The $(2\pi)^{d-1}(2\omega_k)\delta^{d-1}(k' - k) = \langle k|k' \rangle$ factor on the second line comes from projecting the incoming photon momentum onto the outgoing photon momentum in the absence of interactions. Integrating over $d^d k d^d q$ causes no problems. But once $k = k'$, the integral reduces to $\delta(p^2) \frac{i}{p^2 + i\varepsilon}$, which must be treated carefully. Integrating over all variables other

than p_0 and $\omega_p = |\vec{p}|$ gives

$$\begin{aligned} \sigma_{11.A} = \sigma_0 \frac{\Omega_{d-2} (-2(d-2)^2)}{(2\pi)^{d-2}(1-z)} \int_{-\infty}^{\infty} dp_0 \int_{\frac{z}{2}}^{\frac{1}{2}} d\omega_p \delta(p_0^2 - \omega_p^2) \left[\frac{1}{p_0^2 - \omega_p^2 + i\varepsilon} + \text{c.c.} \right] \theta(p_0) \\ \times \omega_p^{d-3} \left[\frac{[(1-p_0)^2 - \omega_p^2][\omega_p^2 - (p_0 - z)^2]}{(1-z)^2 \omega_p^2} \right]^{\frac{d-4}{2}} \\ \times \frac{p_0^4 - 2p_0 z + z^2 - (1-z)\omega_p^2 + \omega_p^4 - p_0^2(2\omega_p^2 + z - 1)}{2p_0 - z}. \quad (\text{A.1.3}) \end{aligned}$$

While $p_0^2 - \omega_p^2$ has two roots, the root with $p_0 = -\omega_p$ is off the integration contour due to the $\theta(p_0)$ in the integrand. Thus we can drop the $i\varepsilon$ term for the $p_0 + \omega_p$ factor and focus on the singularity at $p_0 = \omega$. For this singularity, it is critical to treat the product $\delta(p_0 - \omega_p) \left[\frac{1}{p_0 - \omega_p + i\varepsilon} + \text{c.c.} \right]$ as a distribution. By taking the derivative of the relation $2\pi\delta(x) = \frac{i}{x+i\varepsilon} - \frac{i}{x-i\varepsilon}$ we are led to

$$2\pi\delta'(x) = -i \left(\frac{1}{x+i\varepsilon} \right)^2 + i \left(\frac{i}{x-i\varepsilon} \right)^2. \quad (\text{A.1.4})$$

Thus we can write

$$\begin{aligned} \delta(p_0 - \omega_p) \left[\frac{1}{p_0 - \omega_p + i\varepsilon} + \text{c.c.} \right] \\ = \frac{i}{2\pi} \left(\frac{1}{p_0 - \omega_p + i\varepsilon} \right)^2 - \frac{1}{2\pi} \left(\frac{1}{p_0 - \omega_p - i\varepsilon} \right)^2 \quad (\text{A.1.5}) \\ = -\delta'(p_0 - \omega_p). \end{aligned}$$

The $\delta'(p_0 - \omega_p)$ can then be integrated by parts. So, for a test function $f(p_0)$ we have

$$\int_0^\infty dp_0 \delta(p_0^2 - \omega_p^2) \left[\frac{1}{p_0^2 - \omega_p^2 + i\varepsilon} + \text{c.c.} \right] f(p_0) = \frac{d}{dp_0} \left[\left(\frac{1}{p_0 + \omega_p} \right)^2 f(p_0) \right]_{p_0=\omega_p}. \quad (\text{A.1.6})$$

Applying this prescription to Eq. (A.1.3) gives

$$\sigma_{11.A} = \sigma_0 \Gamma_d \frac{1}{\pi} \left\{ \frac{1}{\varepsilon} - 3 - \ln z \right\}. \quad (\text{A.1.7})$$

The same technique is used to compute Eq. (2.5.7).

A highly non-trivial check on this procedure is that the cross section for $\gamma Z \rightarrow \gamma e^+ e^-$ computed this way exactly cancels the contributions from other γZ final states at the same order in perturbation theory. In particular, the other diagrams, such as the forward scattering loop and the $\gamma Z \rightarrow e^+ e^-$ process are computed without having singular intermediate propagators.

A.2 Initial state masses

In Section 2.5 we showed that the cross section for $n\gamma + e^+ e^- \rightarrow m\gamma + Z$ was IR finite for each n , summed over m . Because of its finiteness and the possible convergence of the sum over n , one might hope to connect the cross section to a physical observable. To do so, the total cross section, inclusive over all possible initial state photons, including hard non-collinear ones, is probably not the most sensible thing to try to measure. To refine the calculation to something closer to physical, we consider instead the cross section for the col-

lision of two initial-state hemisphere jets with masses less than some scale m . Using mass and hemisphere jets makes this infrared-safe cross section depend on only a single parameter, rather than say energy and angle cuts, like Steman-Weinberg jets.

The initial-state hemispheres are defined by the initial-state thrust axis. For a 2-body e^+e^- initial state, the thrust axis is the same as the collision direction and both hemisphere masses are zero. For $e^+e^-\gamma$ initial states, the thrust axis aligns with the hardest of the 3 momenta. The two softer momenta are in one hemisphere and the hemisphere mass containing the single hard particle is zero. For simplicity, we ignore the region of phase space with e^+e^- in the same hemisphere, as it is power suppressed and does not contribute an infrared divergence. Although we consider states with $m+2$ particles in the initial state, there are still at most 3 independent momenta, so we do not have to worry about the more complicated 4-body computation of the thrust axis and hemisphere masses.

We calculate the cumulant total cross section, integrated over the phase space where both hemisphere masses are less than $\sqrt{\lambda}Q$. That is, $m_{\text{hemi } 1}^2 \leq \lambda Q^2$ and $m_{\text{hemi } 2}^2 \leq \lambda Q^2$. So at leading order

$$\tilde{\sigma}_{00}(\lambda) = \sigma_0^d \delta(1-z), \quad (\text{A.2.1})$$

where $\Gamma_d = \left(\frac{4\pi e^{-\gamma_E} \mu^2}{Q^2}\right)^{\frac{4-d}{2}}$, $\sigma_0^d = \sigma_0 \frac{d-2}{2} \mu^{4-d}$, $\sigma_0 = \frac{4\pi g^2}{Q^2}$, and $z = \frac{m_Z^2}{Q^2}$ as before. The virtual correction $\tilde{\sigma}_{00}$ is the same as in Eq. (2.5.1).

$$\tilde{\sigma}_{00}(\lambda) = \sigma_0^d \frac{e^2}{\pi^2} \Gamma_d \left\{ -\frac{1}{4\varepsilon^2} - \frac{3}{8\varepsilon} + \frac{7\pi^2}{48} - 1 \right\} \delta(1-z). \quad (\text{A.2.2})$$

There is no λ dependence as the virtual contribution always contributes.

The only other contribution with no photons in the final state at this order has one photon in the initial state. The cross section is

$$\begin{aligned}\tilde{\sigma}_{10}(\lambda) = \sigma_0^d \frac{e^2}{\pi^2} \Gamma_d \Big\{ & \frac{1}{4\varepsilon^2} + \frac{3}{8\varepsilon} - \frac{1}{4} \ln^2 \lambda - \frac{3}{8} \ln \lambda - \frac{5\pi^2}{48} + \frac{7}{8} \\ & + \frac{7}{32} \ln(1-2\lambda) + \lambda \left(\frac{15}{16} + \frac{1}{2} \ln \lambda - \frac{1}{2} \ln(1-2\lambda) \right) \\ & + \lambda^2 \left(\frac{3}{16} - \frac{1}{8} \ln \lambda + \frac{1}{8} \ln(1-2\lambda) \right) - \frac{1}{2} \text{Li}_2(2\lambda) \Big\} \delta(1-z). \quad (\text{A.2.3})\end{aligned}$$

Note that the IR divergences of $\tilde{\sigma}_{00}$ and $\tilde{\sigma}_{10}$ exactly cancel, leaving $\ln \lambda$ and $\ln^2 \lambda$ terms, just as for final state jets.

For $n > 0$ photons in the final state, we have to be a little careful about the kinematics. For $Q \approx m_Z$, the jet masses can only be as large as roughly $m \lesssim Q - m_Z$, thus $\lambda \lesssim (1-z)$. For values of λ larger than this, the cumulant becomes λ independent. The precise cutoff depends on the numbers of photons in the initial state and final state. Explicitly $\tilde{\sigma}_{mn}$ becomes λ independent for $\lambda > \frac{m}{2n}(1-z)$. For simplicity, we also take $z > \frac{1}{2}$ as we want the Born process to be $e^+e^- \rightarrow Z$ not $e^+e^- \rightarrow \gamma Z$.

The various contributions for $\lambda > \frac{m}{2n}(1-z)$ are, for $m = n-1$,

$$\begin{aligned}\tilde{\sigma}_{n-1,n} = \sigma_0^d \frac{e^2}{\pi^2} \Gamma_d \Big\{ & \delta(1-z) \left(\frac{1}{4\varepsilon^2} + \frac{\ln n}{\varepsilon} - \frac{\pi^2}{16} + \frac{\ln^2 n}{2} \right) \\ & - \frac{1}{\varepsilon} \frac{(2n^2 - 2n + 1)z^2 + 2(n-1)z + 1}{4n^2} \left[\frac{1}{1-z} \right]_+ \\ & + \frac{(2n^2 - 2n + 1)z^2 + 2(n-1)z + 1}{4n^2} \left(\ln \left(\frac{(n-1)z + 1}{n^3} \right) \left[\frac{1}{1-z} \right]_+ + 2 \left[\frac{\ln(1-z)}{1-z} \right]_+ \right) \Big\}, \quad (\text{A.2.4})\end{aligned}$$

for $m = n$,

$$\begin{aligned} \tilde{\sigma}_{n,n} = \sigma_0^d \frac{e^2}{\pi^2} \Gamma_d \left\{ \delta(1-z) \left(-\frac{1}{2\varepsilon^2} - \frac{2\ln n}{\varepsilon} + \frac{\pi^2}{8} - \ln^2 n \right) \right. \\ \left. + \frac{1}{\varepsilon} \frac{(2n^2+1)z^2 - 2z + 1}{2n^2} \left[\frac{1}{1-z} \right]_+ \right. \\ \left. - \frac{1-z}{2n^2} - \frac{(2n^2+1)z^2 - 2z + 1}{2n^2} \left(\ln \left(\frac{z}{n^2} \right) \left[\frac{1}{1-z} \right]_+ + 2 \left[\frac{\ln(1-z)}{1-z} \right]_+ \right) \right\}, \quad (\text{A.2.5}) \end{aligned}$$

and for $m = n + 1$,

$$\begin{aligned} \tilde{\sigma}_{n+1,n} = \sigma_0^d \frac{e^2}{\pi^2} \Gamma_d \left\{ \left(\frac{1}{4\varepsilon^2} + \frac{\ln n}{\varepsilon} - \frac{\pi^2}{16} + \frac{\ln^2 n}{2} \right) \delta(1-z) \right. \\ \left. - \frac{1}{\varepsilon} \frac{(2n^2+2n+1)z^2 - 2(n+1)z + 1}{4n^2} \left[\frac{1}{1-z} \right]_+ \right. \\ \left. + \frac{(2n^2+2n+1)z^2 - 2(n+1)z + 1}{4n^2} \left(\ln \left(\frac{(n+1)z - 1}{n^3} \right) \left[\frac{1}{1-z} \right]_+ + 2 \left[\frac{\ln(1-z)}{1-z} \right]_+ \right) \right\}. \quad (\text{A.2.6}) \end{aligned}$$

For $\lambda \leq \frac{m}{2n}(1-z)$, we find for $m = n - 1$:

$$\begin{aligned} \tilde{\sigma}_{n-1,n} = \sigma_0^d \frac{e^2}{\pi^2} \Gamma_d \left\{ -\frac{1}{\varepsilon} \frac{(2n^2 - 2n + 1)z^2 + 2(n-1)z + 1}{4n^2} \left(\frac{1}{1-z} \right) \right. \\ \left. + \frac{(n-1)(1-z) - 2n\lambda}{4n^2(n-1)} \right. \\ \left. + \frac{(2n^2 - 2n + 1)z^2 + 2(n-1)z + 1}{4n^2(1-z)} \ln \left(\frac{((n-1)z + 1)\lambda(1-z)^2}{n^2((n-1)(1-z) - n\lambda)} \right) \right\}, \quad (\text{A.2.7}) \end{aligned}$$

for $m = n$,

$$\begin{aligned} \tilde{\sigma}_{n,n} = \sigma_0^d \frac{e^2}{\pi^2} \Gamma_d \left\{ \frac{1}{\varepsilon} \frac{(2n^2 + 1)z^2 - 2z + 1}{2n^2} \left(\frac{1}{1-z} \right) \right. \\ \left. + \frac{(1-z) - 4\lambda}{2n^2} - \frac{(2n^2 + 1)z^2 - 2z + 1}{2n^2(1-z)} \ln \left(\frac{z\lambda(1-z)^2}{n^2(1-z-\lambda)} \right) \right\}, \quad (\text{A.2.8}) \end{aligned}$$

and for $m = n + 1$,

$$\begin{aligned} \tilde{\sigma}_{n+1,n} = \sigma_0^d \frac{e^2}{\pi^2} \Gamma_d \left\{ -\frac{1}{\varepsilon} \frac{(2n^2 + 2n + 1)z^2 - 2(n+1)z + 1}{4n^2} \left(\frac{1}{1-z} \right) \right. \\ \left. + \frac{(n+1)(1-z) - 2n\lambda}{4n^2(n+1)} \right. \\ \left. + \frac{(2n^2 + 2n + 1)z^2 - 2(n+1)z + 1}{4n^2(1-z)} \ln \left(\frac{((n+1)z - 1)\lambda(1-z)^2}{n^2((n+1)(1-z) - n\lambda)} \right) \right\}. \quad (\text{A.2.9}) \end{aligned}$$

As with the total cross section, the IR divergences from these contributions cancel in triplets: $\tilde{\sigma}_{n-1,n} + \tilde{\sigma}_{n,n} + \tilde{\sigma}_{n+1,n}$ is finite for any n . To see if the sum over n converges, we look at the large n behavior of the series. The asymptotic behavior for $n \gg 1$ is:

$$\begin{aligned} \tilde{\sigma}_{n-1,n} + \tilde{\sigma}_{n,n} + \tilde{\sigma}_{n+1,n} = \sigma_0^d \Gamma_d \frac{e^2}{\pi^2} \times \begin{cases} -\frac{(1-z)^3}{6z^2n^4} + \mathcal{O}(\frac{1}{n^6}), & \lambda > 1-z \\ -\frac{z(3\lambda^2 - 4\lambda(1-z) + 2z^2 - 3z + 1)}{2n^2(1-z-\lambda)^2} + \frac{3-6\lambda}{2n^2} + \mathcal{O}(\frac{1}{n^4}), & \lambda < \frac{1-z}{2} \end{cases} \end{aligned} \quad (\text{A.2.10})$$

Note that for $\lambda > 1 - z$ the asymptotic behavior is the same as the total cross section,

Eq. (2.5.11), as expected, and that the sum converges for any λ .

While it is satisfying that the sum converges, we have be careful drawing too strong con-

clusions. As pointed out in [38] for potential scattering, series like this one are not absolutely convergent. Summing in terms in a different order will give a different answer. For example, grouping by fixed number of initial state photons $\tilde{\sigma}_{m,m-1} + \tilde{\sigma}_{m,m} + \tilde{\sigma}_{m,m+1}$ the IR divergences still cancel in triplets, however there is a leftover uncanceled IR divergence $\tilde{\sigma}_{00}$. So the sum over m is also IR divergent. Besides the ordering ambiguity, it is not at all clear that the cancellations and convergence will persist at higher order in perturbation theory or in QCD rather than QED. There is clearly much more to be understood, both computationally and physically, about initial state jets.

B

Appendix to Chapter 4

B.1 The coproduct from variation matrices

Polylogarithms come equipped with a motivic coproduct [151, 226] which is sometimes usefully upgraded to a coaction [152, 227]. The coproduct or coaction can be used to systematically decompose the analytic structure of complicated functions into simpler building blocks. These mathematical notions have been used in a wide variety of Feynman integral calculations to constrain the functional form of the answer based on knowledge of the locations of its discontinuities (see for example [91, 146, 165, 186, 223, 223, 228]). In this appendix, we show how the coproduct arises naturally in the language of the variation matrices \mathcal{M} .

Let us consider again the example of the dilogarithm, which has the variation matrix

$$\mathcal{M} = \begin{pmatrix} 1 & \text{Li}_1(z) & \text{Li}_2(z) \\ 0 & 1 & \ln(z) \\ 0 & 0 & 1 \end{pmatrix}. \quad (\text{B.1.1})$$

A couple of observations can be made about the entries in the top row and the last column of this matrix. The first is that the product $\mathcal{M}_{1i}\mathcal{M}_{i3}$ has the same transcendental weight as the original function \mathcal{M}_{13} , for all i . Second, because of the differential equation this matrix satisfies, the entries \mathcal{M}_{1i} involve the iterated integral corresponding to carrying out the first $i - 1$ integrations in the definition of $\text{Li}_2(z)$ (as given in Eq. (5.4.40)), while the entries \mathcal{M}_{i3} involve the iterated integral that results from dropping the first $i - 1$ integrations. Following these observations, we can consider defining an operator Δ that maps $\text{Li}_2(z)$ to a sum

over a tensor product of these matrix entries, which we might think of as summing over the possible ways to partition the integrations in $\text{Li}_2(z)$ into an initial and a final set:

$$\Delta \mathcal{M}_{13} = \sum_{j=1}^3 \mathcal{M}_{1j} \otimes \mathcal{M}_{j3}. \quad (\text{B.1.2})$$

Plugging in the functions that appear in \mathcal{M} , this equation becomes

$$\Delta \text{Li}_2(z) = 1 \otimes \text{Li}_2(z) + \text{Li}_1(z) \otimes \ln(z) + \text{Li}_2(z) \otimes 1, \quad (\text{B.1.3})$$

which can be recognized to be precisely the coproduct of the dilogarithm, as defined in [151].

These observations, and the corresponding construction of the coproduct, can be extended to the general case. Namely, due to the fact that each row of \mathcal{M} satisfies the same differential equation, the product $\mathcal{M}_{ij}\mathcal{M}_{jk}$ has the same transcendental weight as \mathcal{M}_{ik} for all $i \leq j \leq k$. And while generic variation matrix entries \mathcal{M}_{ik} involve sums of iterated integrals, the functions \mathcal{M}_{ij} still correspond to carrying out (some linear combination of) the initial integrations entering \mathcal{M}_{ik} , while the functions \mathcal{M}_{jk} still correspond to carrying out (some linear combination of) the final integrations in \mathcal{M}_{ik} . Correspondingly, the coproduct can be defined in terms of entries of the variation matrix by

$$\Delta \mathcal{M}_{ik} = \sum_{j=i}^k \mathcal{M}_{ij} \otimes \mathcal{M}_{jk}. \quad (\text{B.1.4})$$

As indicated by the use of general indices i and k , the coproduct can be applied to any en-

try of a variation matrix; however, as in [151], the second factor in this tensor product must be interpreted modulo factors of $i\pi$. Instances of $i\pi$ that appear in the first factor can be retained using the methods of [153].

It is worth emphasizing that the coproduct (B.1.4) can be applied to entries of the variation matrix in any region, and that it commutes with the action of the monodromy matrices. For instance, recall the variation matrix for the triangle and box integral from Eq. (5.4.57),

$$\mathcal{M}_{\gamma_0} = \begin{pmatrix} 1 & \ln(z\bar{z}) & \text{Li}_1(z) + \text{Li}_1(\bar{z}) & \Phi_1(z, \bar{z}) \\ 0 & 1 & 0 & -\text{Li}_1(z) + \text{Li}_1(\bar{z}) \\ 0 & 0 & 1 & \ln(z/\bar{z}) \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad (\text{B.1.5})$$

where we recall that

$$\Phi_1(z, \bar{z}) = -\ln(z\bar{z})(\text{Li}_1(z) - \text{Li}_1(\bar{z})) + 2(\text{Li}_2(z) - \text{Li}_2(\bar{z})). \quad (\text{B.1.6})$$

Using Eq. (B.1.4), we can easily read off the coproduct of $\Phi_1(z, \bar{z})$ from Eq. (B.1.5):

$$\begin{aligned} \Delta\Phi_1(z, \bar{z}) &= 1 \otimes \Phi_1(z, \bar{z}) - \ln(z\bar{z}) \otimes \text{Li}_1(z) + \ln(z\bar{z}) \otimes \text{Li}_1(\bar{z}) \\ &\quad + \text{Li}_1(z) \otimes \ln(z/\bar{z}) + \text{Li}_1(\bar{z}) \otimes \ln(z/\bar{z}) + \Phi_1(z, \bar{z}) \otimes 1. \end{aligned} \quad (\text{B.1.7})$$

To analytically continue Eq. (B.1.7) around one of its branch points, we can replace all of

the functions in the left factor of the coproduct with the value they take after being acted on by one of the monodromy matrices. It should be clear that this results in the same coproduct that one would get from applying Eq. (B.1.4) directly to the variation matrix that results from the action of the monodromy matrix. Further details on the properties of the coproduct can be found in [229].

B.2 The monodromy and fundamental groups

As seen in Section 5.4, the complete analytic structure of a collection of polylogarithms can be encoded in a set of monodromy matrices. These matrices occur in one-to-one correspondence with the location of simple poles in the integral definition of these polylogarithms, reflecting the fact that the corresponding integration contours are always homotopic to a composition of (some sequence of) closed contours that encircle individual poles, and a contour that does not cross any branch cuts. This indicates that there should be some relation between the monodromy group and the fundamental group describing the manifold on which these polylogarithms are defined, which has punctures at precisely the loci of these simple poles.

To make this connection between the monodromy and fundamental groups more explicit, we first observe that monodromy matrices can be written as the conjugation of a matrix with rational entries by a diagonal matrix whose entries are integer powers of $2\pi i$. For instance, the monodromy matrices of the dilogarithm from Eq. (5.4.48) and Eq. (5.4.50) can be written as

$$\mathcal{M}_{\mathbb{D}_0} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 2\pi i \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2\pi i & 0 \\ 0 & 0 & (2\pi i)^2 \end{pmatrix}^{-1} \cdot \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2\pi i & 0 \\ 0 & 0 & (2\pi i)^2 \end{pmatrix} \quad (\text{B.2.1})$$

and

$$\mathcal{M}_{\varpi_1} = \begin{pmatrix} 1 & -2\pi i & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2\pi i & 0 \\ 0 & 0 & (2\pi i)^2 \end{pmatrix}^{-1} \cdot \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2\pi i & 0 \\ 0 & 0 & (2\pi i)^2 \end{pmatrix}. \quad (\text{B.2.2})$$

These conjugated matrices can be understood as furnishing a representation of the homotopy group of $\mathbb{C} - \{0, 1\}$ by matrices in $\text{GL}(3, \mathbb{Z})$. More explicitly, the homotopy group of $\mathbb{C} - \{0, 1\}$ is the free group with two generators, which are associated with the homotopy classes of paths around $z = 0$ and $z = 1$. Up to conjugation by $\text{diag}(1, 2\pi i, (2\pi i)^2)$, the monodromy matrices give us an explicit representation of this group.

Note that this connection to the fundamental group remains valid if we compactify the complex plane by considering the monodromy matrix associated with infinity. Using the connection in Eq. (5.4.43), we can compute the monodromy matrix from an infinitesimal contour encircling infinity. For instance, if we integrate the dilogarithm integrand around a circular path that starts and ends at a complex point $|R| > 1$, we have

$$\int_{\gamma_R} \frac{ds}{1-s} \circ \frac{ds}{s} = (2\pi i)^2 \int_0^1 dt \int_0^t du \frac{Re^{2\pi i u}}{1 - Re^{2\pi i u}} = 2\pi^2 + 2\pi i \ln \frac{R-1}{R}. \quad (\text{B.2.3})$$

Since $\ln \frac{R-1}{R}$ is a continuous function for large $|R|$, $\lim_{R \rightarrow \infty} \ln \frac{R-1}{R} = 0$ and we get $2\pi^2$. The

full matrix can be computed to be

$$\mathcal{M}_{\mathcal{D}_\infty} = \begin{pmatrix} 1 & -2\pi i & 2\pi^2 \\ 0 & 1 & 2\pi i \\ 0 & 0 & 1 \end{pmatrix}. \quad (\text{B.2.4})$$

Note that going around infinity clockwise corresponds to a counterclockwise contour around 0 and 1. If we compute the matrix along a straight line path between 0 and R , we get the variation matrix in Eq. (5.4.46):

$$\mathcal{M}_{0 \rightarrow R} = \begin{pmatrix} 1 & -\ln(1-R) & \text{Li}_2(R) \\ 0 & 1 & \ln R \\ 0 & 0 & 1 \end{pmatrix}. \quad (\text{B.2.5})$$

Then, if we take $R \rightarrow \infty$ with $\text{Im } R > 0$, we get

$$\mathcal{M}_{\infty}^{\text{green}} = \mathcal{M}_{0 \rightarrow R} \cdot \mathcal{M}_{\mathcal{D}_\infty} \cdot \mathcal{M}_{0 \rightarrow R}^{-1} = \begin{pmatrix} 1 & -2\pi i & 0 \\ 0 & 1 & 2\pi i \\ 0 & 0 & 1 \end{pmatrix} = \mathcal{M}_{\mathcal{D}_0}^{\text{blue}} \cdot \mathcal{M}_{\mathcal{D}_1}^{\text{red}}. \quad (\text{B.2.6})$$

This monodromy around infinity can be written as the product of a monodromy around 0 and 1, since the path around infinity is homotopic to a path around 0 then around 1, as illustrated in the left part of Fig. B.1. There, we see that the choice to take $\text{Im } R > 0$ was what determined that we encircled the branch point at 0 first, and then the branch point at 1. If we take $R \rightarrow \infty$ with $\text{Im } R < 0$ (so that the contour circles the branch point at 1

first), the monodromy matrix differs in the top-right entry

$$\mathcal{M}_{\infty}^{\uparrow} = \mathcal{M}_{0 \rightarrow R} \cdot \mathcal{M}_{\infty} \cdot \mathcal{M}_{0 \rightarrow R}^{-1} = \begin{pmatrix} 1 & 2\pi i & 4\pi^2 \\ 0 & 1 & -2\pi i \\ 0 & 0 & 1 \end{pmatrix} = \mathcal{M}_{\infty,1} \cdot \mathcal{M}_{\infty,0} \quad (\text{B.2.7})$$

The result is the product of the 0 and 1 monodromies in the opposite order. This path around infinity is illustrated on the right in Fig. B.1.

This ambiguity at $\mathcal{O}(\pi^2)$ in the monodromy matrix associated with infinity is also present for the other monodromy matrices. For example, we could have computed the monodromy around 1 using a contour that first crosses the negative real axis before going around 1, as illustrated on the right in Fig. B.1. The result would have been

$$\mathcal{M}_{\infty,0,1} = \mathcal{M}_{\infty,0}^{-1} \cdot \mathcal{M}_{\infty,1} \cdot \mathcal{M}_{\infty,0} = \begin{pmatrix} 1 & -2\pi i & 4\pi^2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \quad (\text{B.2.8})$$

The $\mathcal{O}(\pi)$ terms in this monodromy matrix are the same as for $\mathcal{M}_{\infty,1}$ in Eq. (B.2.2), as expected from Cauchy's residue theorem, but the $\mathcal{O}(\pi^2)$ terms are different.

To describe these $\mathcal{O}(\pi^2)$ ambiguities more formally, consider a codimension-one branch variety defined by an equation $f(\{s_j\}) = 0$, for some set of variables $\{s_j\}$ which we can take to be Mandelstam invariants. To compute the monodromy around this branch variety, we find a closed path γ such that $\oint_{\gamma} d \ln f(\{s_j\}) = 2\pi i$. However, as there are many paths γ that satisfy this requirement, there is some ambiguity in this choice. In particular, all the

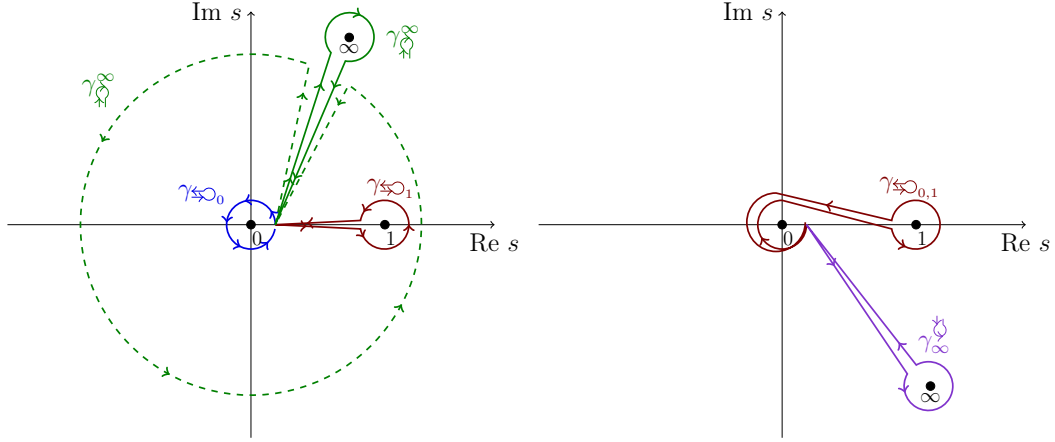


Figure B.1: Paths around 0, 1 and ∞ . We depict two possible contours that go around infinity, starting at points in the upper or lower half-plane. These are each homotopic to paths around 0 and 1, but in different orders. The two contours around infinity are not homotopically equivalent. The right panel shows that the path ambiguity is present also for paths around $s = 1$.

paths in the same homology class of γ satisfy the same relation; however, the paths in this homology class may still be in different homotopy classes. While the integral $\oint_{\gamma} d \ln f(\{s\})$ depends only on the homology class of γ , the elements of the monodromy group depend on the homotopy class of γ .

The fundamental group and first homology group are related by Hurewicz theorem, which states that the first homology group is the abelianization of the fundamental group. That is, given any two elements a and b of the fundamental group, we can quotient the fundamental group by the commutator subgroup generated by elements $aba^{-1}b^{-1}$ to obtain the homology group. The contour corresponding to the commutator $aba^{-1}b^{-1}$ is called a Pochhammer contour, and corresponds to a trivial element in homology. Thus, for every path γ which satisfies the condition $\oint_{\gamma} d \ln f(\{s_j\}) = 2\pi i$, we can find another path $\gamma aba^{-1}b^{-1}$ that also satisfies this relation. Moreover, as this new path belongs to a different homotopy class, it yields a different monodromy beyond $\mathcal{O}(\pi)$.

Despite these ambiguities, any choice of closed contours around 0, 1, and infinity will furnish us with a representation of the fundamental group on the Riemann sphere with three marked points. For instance, we can choose the rational matrices appearing in Eqs. (B.2.1), (B.2.2), and (B.2.6), which satisfy a single multiplicative identity. Note, however, that the contours used must all start at the same basepoint, so we cannot use the rational matrices corresponding to \mathcal{M}_{\odot_0} , \mathcal{M}_{\odot_1} , and $\mathcal{M}_{\odot_\infty}$.

For a multivariable function, like the function $\Phi_1(z, \bar{z})$ that appears in the one-loop triangle and box, we can carry out the same analysis for the contours in z while holding \bar{z} fixed. The contours around $z = 0$ and $z = 1$ were computed in Eqs. (5.4.59) and (5.4.62) to be

$$\mathcal{M}_{\text{triangle}_0^z} = \begin{pmatrix} 1 & 2\pi i & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 2\pi i \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad \mathcal{M}_{\text{triangle}_1^z} = \begin{pmatrix} 1 & 0 & -2\pi i & 0 \\ 0 & 1 & 0 & 2\pi i \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \quad (\text{B.2.9})$$

For the contours around infinity, a calculation analogous to the dilog case gives

$$\mathcal{M}_{\infty z}^{\infty} = \begin{pmatrix} 1 & 2\pi i & -2\pi i & -4\pi^2 \\ 0 & 1 & 0 & 2\pi i \\ 0 & 0 & 1 & 2\pi i \\ 0 & 0 & 0 & 1 \end{pmatrix} = \mathcal{M}_{\infty z_0}^z \cdot \mathcal{M}_{\infty z_1}^z, \quad (\text{B.2.10})$$

$$\mathcal{M}_{\infty}^z = \begin{pmatrix} 1 & 2\pi i & -2\pi i & 4\pi^2 \\ 0 & 1 & 0 & 2\pi i \\ 0 & 0 & 1 & 2\pi i \\ 0 & 0 & 0 & 1 \end{pmatrix} = \mathcal{M}_{\infty z_0}^z \cdot \mathcal{M}_{\infty z_1}^z. \quad (\text{B.2.11})$$

The monodromy matrices for contours in \bar{z} can be computed in a similar fashion, and commute with the monodromy matrices in z . Like for the case of the dilogarithm, each monodromy matrix gives rise to an associated rational matrix that corresponds to a generator of the fundamental group, which in this case describes the manifold corresponding to the space of complex z and \bar{z} with the points 0, 1, and infinity in each variable removed.

More generally, the monodromy group describing the discontinuity of a set of polylogarithms also furnishes us with a representation of the fundamental group describing the manifold on which these polylogarithms are defined. When we consider polylogarithms that only depend on a single variable, the relevant manifold is the Riemann sphere with n marked points and the fundamental group corresponds to the free group with $n - 1$ generators. However, the fundamental group of higher-dimensional manifolds will in general be more complicated.

B.3 Single-valued polylogarithms

Using the Knizhnik-Zamolodchikov equation, polylogarithms can be mapped to single-valued avatars of themselves [230]. In these new single-valued functions, all contributions generated by analytically continuing around branch points are systematically cancelled out by new functional dependence on variables conjugate to the variables of the original function. This type of single-valued map has proven useful in a variety of physics contexts, such as multi-Regge limits [231, 232], the infrared structure of gauge theory [192, 233], string amplitudes [234], and massless ϕ^4 theory [235]. Motivated by [201, 202] we show here that the same map can be constructed in terms of variation matrices.

We begin by considering a variation matrix \mathcal{M} that depends on any number of variables, whose discontinuities are described by a set of monodromy matrices $\{\mathcal{M}_{\varpi,k}\}$ indexed by k . In order to construct a single-valued version of the matrix \mathcal{M} , we want to find a matrix that transforms in the opposite way as \mathcal{M} when analytically continued around branch points. A natural object to consider is the inverse conjugate matrix $\overline{\mathcal{M}}^{-1}$, namely the inverse matrix of \mathcal{M} in which all variables z_j have additionally been replaced by their complex conjugates \bar{z}_j . Under the action of the monodromy group, this pair of matrices transform as

$$\mathcal{M} \rightarrow \mathcal{M}_{\varpi,k} \cdot \mathcal{M} , \tag{B.3.1}$$

$$\overline{\mathcal{M}}^{-1} \rightarrow \overline{\mathcal{M}}^{-1} \cdot \overline{\mathcal{M}}_{\varpi,k}^{-1} . \tag{B.3.2}$$

Thus, the product of these two matrices is not quite invariant under arbitrary analytic continuations, because $\overline{\mathcal{M}}_{\varpi,k}^{-1} \cdot \mathcal{M}_{\varpi,k} \neq \mathbb{1}$.

This mismatch can be fixed by decomposing the monodromy matrices as discussed in section B.2. In particular, we have

$$\mathcal{M}_{\varpi,k} = D^{-1} \cdot M_k \cdot D, \quad (\text{B.3.3})$$

where D is a diagonal matrix whose entries are integer powers of $2\pi i$, and M_k is an element of the general linear group with rational entries. Since the action of the monodromy matrices preserves transcendental weight, the matrix D (which encodes the relative weight the rows of \mathcal{M}) does not depend on k . Having made this observation, we define the single-valued matrix

$$\mathcal{M}_{\text{sv}} \equiv \overline{\mathcal{M}}^{-1} \cdot \overline{D}^{-1} \cdot D \cdot \mathcal{M}. \quad (\text{B.3.4})$$

This matrix is invariant under the action of the monodromy group, since

$$\mathcal{M}_{\text{sv}} \rightarrow \left(\overline{\mathcal{M}}^{-1} \cdot \overline{D}^{-1} \cdot M_k^{-1} \cdot \overline{D} \right) \cdot \overline{D}^{-1} \cdot D \cdot \left(D^{-1} \cdot M_k \cdot D \cdot \mathcal{M} \right) = \mathcal{M}_{\text{sv}} \quad (\text{B.3.5})$$

whenever $\bar{z}_j = z_j^*$. We note that the definition (B.3.4) is equivalent to the map defined in Eq. (3.82) of [232] using the coproduct formalism.

Let us see how this works in the case of the dilogarithm. Referring to its variation matrix

\mathcal{M} in Eq. (B.1.1), we see that $\overline{D}^{-1} \cdot D = \text{diag}(1, -1, 1)$ and

$$\overline{\mathcal{M}}^{-1} = \begin{pmatrix} 1 & -\text{Li}_1(\bar{z}) & -\text{Li}_2(\bar{z}) + \text{Li}_1(\bar{z}) \ln \bar{z} \\ 0 & 1 & -\ln \bar{z} \\ 0 & 0 & 1 \end{pmatrix}. \quad (\text{B.3.6})$$

The single-valued matrix is thus given by

$$\mathcal{M}_{\text{sv}} = \begin{pmatrix} 1 & \text{Li}_1(z) + \text{Li}_1(\bar{z}) & \text{Li}_2(z) - \text{Li}_2(\bar{z}) + \ln(z\bar{z}) \text{Li}_1(\bar{z}) \\ 0 & -1 & -\ln(z\bar{z}) \\ 0 & 0 & 1 \end{pmatrix}. \quad (\text{B.3.7})$$

It is not hard to check that all effects of analytically continuing z and \bar{z} in opposite directions around any branch point cancel out in the entries of this matrix, as expected.

B.4 Permutation symmetry of the triangle integral

The one-loop triangle integral considered in Section 5.6.2, given by

$$\frac{1}{16\pi^2 p_1^2} \frac{1}{z - \bar{z}} \Phi_1(z, \bar{z}) \quad (\text{B.4.1})$$

where $\Phi_1(z, \bar{z})$ was defined in Eq. (5.6.46), respects an S_3 symmetry under the permutation of its external legs. This symmetry turns out to be realized in an interesting way, by the collusion of this integral's rational and transcendental parts.

We first discuss the rational prefactor. To determine how the quantity $p_1^2(z - \bar{z})$ transforms under the permutation of external momenta, we consider the wedge product $p_1 \wedge p_2$. We work in the coordinate system described above Eq. (5.6.35), where $p_1 = (E_1, 1)$ and $p_2 = (E_2, p_2^x)$. In terms of a pair of basis vectors e_t and e_x , these momenta become $p_1 = E_1 e_t + e_x$ and $p_2 = E_2 e_t + p_2^x e_x$, and we have

$$p_1 \wedge p_2 = -\frac{1}{2} p_1^2 (z - \bar{z}) e_t \wedge e_x. \quad (\text{B.4.2})$$

We can correspondingly use this quantity to study the transformation properties of $p_1^2(z - \bar{z})$. Clearly, under $p_1 \leftrightarrow p_2$ the left-hand side of Eq. (B.4.2) changes sign. Similarly, under $p_1 \leftrightarrow p_3$ we have $p_1 \wedge p_2 \leftrightarrow p_3 \wedge p_2 = -p_1 \wedge p_2$. We conclude that the representation of the symmetric group S_3 when acting on $p_1^2(z - \bar{z})$ is the sign representation.

Before moving on to discuss the symmetries of $\Phi_1(z, \bar{z})$, we need to find the action of the

S_3 symmetry on z and \bar{z} . From the above, we know that

$$\sigma(p_1^2(z - \bar{z})) = p_{\sigma(1)}^2(\sigma(z) - \sigma(\bar{z})) = (-1)^{|\sigma|} p_1^2(z - \bar{z}). \quad (\text{B.4.3})$$

We also know, from Eq. (5.6.32), that under the $p_2 \leftrightarrow p_3$ permutation we have $z\bar{z} \leftrightarrow (1 - z)(1 - \bar{z})$. These constraints can be solved with the unique solution that $p_2 \leftrightarrow p_3$ corresponds to $z \leftrightarrow 1 - z$ and $\bar{z} \leftrightarrow 1 - \bar{z}$. Similarly, one can show that $p_1 \leftrightarrow p_2$ must correspond to $z \leftrightarrow \frac{1}{z}$ and $\bar{z} \leftrightarrow \frac{1}{\bar{z}}$. The action of the remaining permutations can be determined from these two transformations.

We are now ready to study the symmetry of the transcendental part of the triangle function. It turns out that this is related to the Bloch-Wigner function

$$D(z) = \Im \text{Li}_2(z) + \arg(1 - z) \ln(|z|). \quad (\text{B.4.4})$$

In particular, using

$$\Im \text{Li}_2(z) = \frac{1}{2i} (\text{Li}_2(z) - \text{Li}_2(z^*)), \quad (\text{B.4.5})$$

$$\arg(1 - z) = \frac{1}{2i} \ln \frac{1 - z}{1 - z^*}, \quad (\text{B.4.6})$$

$$\ln|z| = \frac{1}{2} \ln(zz^*), \quad (\text{B.4.7})$$

we have

$$4iD(z) = 2(\text{Li}_2(z) - \text{Li}_2(z^*)) + \ln \left(\frac{1 - z}{1 - z^*} \right) \ln(zz^*). \quad (\text{B.4.8})$$

In the region R_1^* , where $\bar{z} = z^*$, this gives precisely the transcendental part of the one-loop triangle, $\Phi_1(z, \bar{z})$.

The Bloch-Wigner function satisfies

$$D(z) = -D(1-z) = -D\left(\frac{1}{z}\right). \quad (\text{B.4.9})$$

These signs precisely compensate the signs arising from the action of the permutation group on the rational prefactor. In the other regions, where $\bar{z} \neq z^*$, the transcendental part should be thought as a function of two independent variables. Still, the same relations hold under the transformation of both z and \bar{z} .

How is the symmetry realized on the cuts? It is instructive to consider the example of a leading singularity, where the only dependence on the kinematics is in the rational prefactor, while the transcendental part is a power of $2\pi i$. By the argument above, under the action of the permutation of external legs, the rational prefactor may pick up a sign. Hence, the residue on a given leading singularity is *not* invariant under the permutation group. However, each leading singularity locus is paired with another one with opposite residue, as required by global residue theorems. It follows that the set of values the residue takes on all the leading singularities is invariant under the action of the permutation group. A similar statement holds for the rest of the cuts.

B.5 Variation matrix of the two-loop box

In this appendix we present the connection and variation matrix for the two-loop ladder triangle/box function

$$\Phi_2(z, \bar{z}) = 6[\text{Li}_4(z) - \text{Li}_4(\bar{z})] - 3 \ln(z\bar{z})[\text{Li}_3(z) - \text{Li}_3(\bar{z})] + \frac{1}{2} \ln^2(z\bar{z})[\text{Li}_2(z) - \text{Li}_2(\bar{z})]. \quad (\text{B.5.1})$$

The two-loop connection is

$$\omega = \left(\begin{array}{c|ccc|ccc|c} 0 & -\omega_1 - \bar{\omega}_1 & \omega_0 + \bar{\omega}_0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & -\bar{\omega}_0 & \omega_0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \omega_1 & -\bar{\omega}_1 & \omega_0 + \bar{\omega}_0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & \omega_0 - \bar{\omega}_0 & -\omega_0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -\bar{\omega}_0 & -\omega_0 + \bar{\omega}_0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -\omega_1 & \bar{\omega}_1 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \omega_0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \bar{\omega}_0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array} \right), \quad (\text{B.5.2})$$

where

$$\omega_0 = \frac{dz}{z}, \quad \omega_1 = \frac{dz}{z-1}, \quad (\text{B.5.3})$$

$$\bar{\omega}_0 = \frac{d\bar{z}}{\bar{z}}, \quad \bar{\omega}_1 = \frac{d\bar{z}}{\bar{z}-1}. \quad (\text{B.5.4})$$

The connection trivially satisfies $d\omega = 0$, and using the fact that $\omega_0 \wedge \omega_1 = 0$, we also have that $\omega \wedge \omega = 0$. Thus, the connection has zero curvature ($d\omega - \omega \wedge \omega = 0$).

Using this connection, we can compute the variation matrix \mathcal{M}_{γ_0} . We encounter integrals of one-forms, which are familiar, but also iterated integrals of higher weight. These can be easily evaluated by leveraging the shuffle algebra associated with iterated integrals [148, 226]. As an example, consider

$$\mathcal{M}_{1,6} = \int^{z, \bar{z}} (\omega_0 + \bar{\omega}_0) \circ (\omega_0 + \bar{\omega}_0) . \quad (\text{B.5.5})$$

While this integral can be computed using Eq. (5.4.22) along a concretely chosen contour, it is easier to use the fact that any pair of one-forms σ_1 and σ_2 satisfies

$$\int_{\gamma} (\sigma_1 \circ \sigma_2 + \sigma_2 \circ \sigma_1) = \left(\int_{\gamma} \sigma_1 \right) \left(\int_{\gamma} \sigma_2 \right) . \quad (\text{B.5.6})$$

This allows us to rewrite Eq. (B.5.5) as

$$\mathcal{M}_{1,6} = \frac{1}{2} \int^z \omega_0 \int^z \omega_0 + \int^z \omega_0 \int^{\bar{z}} \bar{\omega}_0 + \frac{1}{2} \int^{\bar{z}} \bar{\omega}_0 \int^{\bar{z}} \bar{\omega}_0 . \quad (\text{B.5.7})$$

These integrals are much simpler to evaluate, and we get

$$\mathcal{M}_{1,6} = \frac{1}{2} \ln^2 z + \ln z \ln \bar{z} + \frac{1}{2} \ln^2 \bar{z} . \quad (\text{B.5.8})$$

The relation in Eq. (B.5.6) can be iterated, to give us

$$\int_{\gamma} \sum_{\substack{\{j_1, \dots, j_n\} \\ \in \text{perms of } \{1, \dots, n\}}} \sigma_{j_1} \circ \sigma_{j_2} \circ \dots \circ \sigma_{j_n} = \int_{\gamma} \sigma_1 \int_{\gamma} \sigma_2 \dots \int_{\gamma} \sigma_n \quad (\text{B.5.9})$$

for n one-forms $\sigma_1, \dots, \sigma_n$, along with relations such as

$$\int_{\gamma} (\sigma_1 \circ \sigma_2 \circ \sigma_3 + \sigma_2 \circ \sigma_1 \circ \sigma_3 + \sigma_2 \circ \sigma_3 \circ \sigma_1) = \int_{\gamma} \sigma_1 \int_{\gamma} \sigma_2 \circ \sigma_3. \quad (\text{B.5.10})$$

Using these kinds of formulas, we can reduce the expressions in the calculation of the variation matrix to familiar integrals, such as

$$\text{Li}_n(z) = - \int_0^z \omega_1 \circ \underbrace{\omega_0 \circ \omega_0 \dots \circ \omega_0}_{n-1}, \quad (\text{B.5.11})$$

along with integrals that can easily be performed, such as

$$\int_0^z \omega_0 \circ \omega_1 \circ \omega_0 = 2\text{Li}_3(z) - \ln z \text{Li}_2(z). \quad (\text{B.5.12})$$

The iterated integrals we study have the special property that they are independent of small deformations of the integration contour which preserve its endpoints. This is a consequence of the flatness of the connection ω and is sometimes called integrability condition.

In our example, the integrability condition reads

$$(\omega_0 + \bar{\omega}_0) \wedge (\omega_0 + \bar{\omega}_0) = 0. \quad (\text{B.5.13})$$

This condition is trivial when both forms only depend on a single variable, but imposes non-trivial restrictions when two or more variables are involved.

The result of performing the integrations is

$$\mathcal{M}_{\gamma_0}(z, \bar{z}) = \left(\begin{array}{c|ccc|cc|c} 1 & \mathcal{M}_{1,2} & \mathcal{M}_{1,3} & \mathcal{M}_{1,4} & \mathcal{M}_{1,5} & \mathcal{M}_{1,6} & \mathcal{M}_{1,7} & \mathcal{M}_{1,8} & \mathcal{M}_{1,9} \\ \hline 0 & 1 & 0 & \mathcal{M}_{2,4} & \mathcal{M}_{2,5} & 0 & \mathcal{M}_{2,7} & \mathcal{M}_{2,8} & \mathcal{M}_{2,9} \\ 0 & 0 & 1 & \mathcal{M}_{3,4} & \mathcal{M}_{3,5} & \mathcal{M}_{3,6} & \mathcal{M}_{3,7} & \mathcal{M}_{3,8} & \mathcal{M}_{3,9} \\ \hline 0 & 0 & 0 & 1 & 0 & 0 & \mathcal{M}_{4,7} & \mathcal{M}_{4,8} & \mathcal{M}_{4,9} \\ 0 & 0 & 0 & 0 & 1 & 0 & \mathcal{M}_{5,7} & \mathcal{M}_{5,8} & \mathcal{M}_{5,9} \\ 0 & 0 & 0 & 0 & 0 & 1 & \mathcal{M}_{6,7} & \mathcal{M}_{6,8} & \mathcal{M}_{6,9} \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & \mathcal{M}_{7,9} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & \mathcal{M}_{8,9} \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{array} \right), \quad (\text{B.5.14})$$

where

$$\mathcal{M}_{1,2} = -\ln(1 - \bar{z}) - \ln(1 - z), \quad (\text{B.5.15})$$

$$\mathcal{M}_{1,3} = \ln z + \ln \bar{z}, \quad (\text{B.5.16})$$

$$\mathcal{M}_{1,4} = -\text{Li}_2(\bar{z}) + \ln(1 - z) (\ln z + \ln \bar{z}) + \text{Li}_2(z), \quad (\text{B.5.17})$$

$$\mathcal{M}_{1,5} = -\text{Li}_2(\bar{z}) - \ln(1 - \bar{z}) (\ln z + \ln \bar{z}) + \text{Li}_2(z), \quad (\text{B.5.18})$$

$$\mathcal{M}_{1,6} = \frac{1}{2} \ln^2 z + \ln z \ln \bar{z} + \frac{1}{2} \ln^2 \bar{z}, \quad (\text{B.5.19})$$

$$\mathcal{M}_{1,7} = \frac{1}{2} \left[6\text{Li}_3(\bar{z}) - 4\text{Li}_2(z) (\ln z + \ln \bar{z}) - 2\text{Li}_2(\bar{z}) (\ln z + \ln \bar{z}) - \ln(1 - z) \ln^2 \bar{z} \right]$$

$$- 2 \ln(1 - z) \ln z \ln \bar{z} + 6\text{Li}_3(z) - \ln(1 - z) \ln^2 z \Big], \quad (\text{B.5.20})$$

$$\begin{aligned} \mathcal{M}_{1,8} = \frac{1}{2} \Big[& - 6\text{Li}_3(z) + 4\text{Li}_2(\bar{z}) (\ln z + \ln \bar{z}) + 2\text{Li}_2(z) (\ln z + \ln \bar{z}) + \ln(1 - \bar{z}) \ln^2 z \\ & + 2 \ln(1 - \bar{z}) \ln z \ln \bar{z} - 6\text{Li}_3(\bar{z}) + \ln(1 - \bar{z}) \ln^2 \bar{z} \Big], \end{aligned} \quad (\text{B.5.21})$$

$$\mathcal{M}_{1,9} = \Phi_2, \quad (\text{B.5.22})$$

$$\mathcal{M}_{2,4} = -\ln \bar{z}, \quad \mathcal{M}_{2,5} = \ln z, \quad (\text{B.5.23})$$

$$\mathcal{M}_{2,7} = \frac{1}{2} \ln^2 \bar{z} - \ln \bar{z} \ln z, \quad \mathcal{M}_{2,8} = -\frac{1}{2} \ln^2 z + \ln z \ln \bar{z}, \quad (\text{B.5.24})$$

$$\mathcal{M}_{2,9} = \frac{1}{2} \ln z \ln^2 \bar{z} - \frac{1}{2} \ln^2 z \ln \bar{z}, \quad (\text{B.5.25})$$

$$\mathcal{M}_{3,4} = \ln(1 - z), \quad \mathcal{M}_{3,5} = -\ln(1 - \bar{z}), \quad \mathcal{M}_{3,6} = \ln z + \ln \bar{z}, \quad (\text{B.5.26})$$

$$\mathcal{M}_{3,7} = -\text{Li}_2(\bar{z}) - \ln(1 - z) (\ln z + \ln \bar{z}) - 2\text{Li}_2(z), \quad (\text{B.5.27})$$

$$\mathcal{M}_{3,8} = 2\text{Li}_2(\bar{z}) + \ln(1 - \bar{z}) (\ln z + \ln \bar{z}) + \text{Li}_2(z), \quad (\text{B.5.28})$$

$$\mathcal{M}_{3,9} = 3\text{Li}_3(\bar{z}) + \text{Li}_2(z) (\ln z + \ln \bar{z}) - \text{Li}_2(\bar{z}) (\ln z + \ln \bar{z}) - 3\text{Li}_3(z), \quad (\text{B.5.29})$$

$$\mathcal{M}_{4,7} = \ln z - \ln \bar{z}, \quad \mathcal{M}_{4,8} = -\ln z, \quad \mathcal{M}_{4,9} = \frac{1}{2} \ln^2 z - \ln z \ln \bar{z}, \quad (\text{B.5.30})$$

$$\mathcal{M}_{5,7} = -\ln \bar{z}, \quad \mathcal{M}_{5,8} = \ln \bar{z} - \ln z, \quad \mathcal{M}_{5,9} = \frac{1}{2} \ln^2 \bar{z} - \ln \bar{z} \ln z, \quad (\text{B.5.31})$$

$$\mathcal{M}_{6,7} = -\ln(1 - z), \quad \mathcal{M}_{6,8} = \ln(1 - \bar{z}), \quad \mathcal{M}_{6,9} = \text{Li}_2(z) - \text{Li}_2(\bar{z}), \quad (\text{B.5.32})$$

$$\mathcal{M}_{7,9} = \ln z, \quad \mathcal{M}_{8,9} = \ln \bar{z}. \quad (\text{B.5.33})$$

The monodromy around $z = 0$ is

$$\mathcal{M}_{\mathfrak{D}_0^z} = \left(\begin{array}{c|ccc|ccc|cc|c} 1 & 0 & 2\pi i & 0 & 0 & \frac{1}{2}(2\pi i)^2 & 0 & 0 & 0 \\ \hline 0 & 1 & 0 & 0 & 2\pi i & 0 & 0 & -\frac{1}{2}(2\pi i)^2 & 0 \\ 0 & 0 & 1 & 0 & 0 & 2\pi i & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 1 & 0 & 0 & 2\pi i & -2\pi i & \frac{1}{2}(2\pi i)^2 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & -2\pi i & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 2\pi i \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{array} \right). \quad (\text{B.5.34})$$

We note that $(\mathbb{1} - \mathcal{M}_{\mathfrak{D}_0^z})^3 = 0$. This is consistent with *three* (but not two) sequential cuts

in the p_2^2 channel of the 2-loop triangle vanishing. The monodromy around $z = 1$ is

$$\mathcal{M}_{\varpi_1^z} = \left(\begin{array}{c|ccc|ccc|c} 1 & -2\pi i & 0 & & 0 & 0 & 0 & 0 \\ \hline 0 & 1 & 0 & & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 2\pi i & 0 & 0 & & 0 \\ \hline 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & -2\pi i & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{array} \right). \quad (\text{B.5.35})$$

In this case we have $(1 - \mathcal{M}_{\varpi_1^z})^2 = 0$. This is consistent with *two* sequential cuts in the p_3^2 channel (the long direction) of the 2-loop triangle vanishing. Finally, the clockwise mon-

odromy around infinity (where we approach infinity above the real line) is

$$\mathcal{M}_{\oint_z}^\infty = \mathcal{M}_{\nearrow_0^z} \cdot \mathcal{M}_{\searrow_1^z} \quad (\text{B.5.36})$$

$$= \begin{pmatrix} 1 & -2\pi i & 2\pi i & (2\pi i)^2 & 0 & \frac{1}{2}(2\pi i)^2 & -\frac{1}{2}(2\pi i)^3 & 0 & 0 \\ 0 & 1 & 0 & 0 & 2\pi i & 0 & 0 & -\frac{1}{2}(2\pi i)^2 & 0 \\ 0 & 0 & 1 & 2\pi i & 0 & 2\pi i & -(2\pi i)^2 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 2\pi i & -2\pi i & \frac{1}{2}(2\pi i)^2 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & -2\pi i & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & -2\pi i & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 2\pi i \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

and again we get $(1 - \mathcal{M}_{\oint_z})^3 = 0$.

To compute the monodromy matrices associated with contours in \bar{z} , we can use the fact that z and \bar{z} can be exchanged in the connection from Eq. (B.5.2) via conjugation by the

matrix

$$C = \begin{pmatrix} 1 & & & & \\ & 1 & 0 & & \\ & 0 & 1 & & \\ & & & 0 & -1 & 0 \\ & & & -1 & 0 & 0 \\ & & & 0 & 0 & 1 \\ & & & & 0 & -1 \\ & & & & -1 & 0 \\ & & & & & -1 \end{pmatrix}. \quad (\text{B.5.37})$$

That is, we have $\omega(z, \bar{z}) \rightarrow \omega(\bar{z}, z) = C\omega C^{-1}$. Thus, we also have that

$$\mathcal{M}_{\nabla_0 \bar{z}} = C \mathcal{M}_{\nabla_0 z} C^{-1}, \quad \mathcal{M}_{\nabla_1 \bar{z}} = C \mathcal{M}_{\nabla_1 z} C^{-1}. \quad (\text{B.5.38})$$

These are the last monodromy matrices that are needed to construct the discontinuity operators in Eq. (5.6.42).

B.6 Cuts of the three-loop triangle

In this Appendix, we work out the details of the calculations in Section 5.6.2. We start by computing the sum of the two cuts involving \mathcal{C}_1 and write

$$\begin{aligned}
T_3^{\mathcal{C}_1, \mathcal{C}_2} \Big|_{-i\varepsilon \text{ on r.h.s.}} + T_3^{\mathcal{C}_1, \mathcal{C}_3} &= \frac{1}{2} \int \frac{d^4 k_1}{i (2\pi)^4} (-2\pi i) \delta(k_1^2 - m^2) \\
&\times \Theta(k_1^0) (-2\pi i) \delta\left[(p_2 - k_1)^2\right] \Theta(p_2^0 - k_1^0) \frac{1}{(p_3 + k_1)^2} \sum \text{Cut}_{k_1^2} T^2 \left[(p_3 + k_1)^2, k_1^2, p_3^2\right],
\end{aligned} \tag{B.6.1}$$

where $\sum \text{Cut}_{k_1^2} T^2 \left[(p_3 + k_1)^2, k_1^2, p_3^2\right]$ is the sum of cuts in k_1^2 through the two-loop triangle T_2 with masses $(p_3 + k_1)^2$, k_1^2 and p_3^2 . We take the particle with momentum k_1 to have a small mass m to regulate the IR divergences that arise in the cut calculations, and work to leading power in m^2 . The factor of $\frac{1}{2}$ arises because the mass regulator does not capture the $\frac{1}{L!}$ arising from a product of $L - 1$ massless vertices, as worked out in Appendix B.7. The sum of the cuts through the two-loop triangle is given by [165],

$$\begin{aligned}
\sum \text{Cut}_{k_1^2} T^2 \left[(p_3 + k_1)^2, m^2, p_3^2\right] &= \frac{2\pi i}{256\pi^4} \frac{1}{(p_3 + k_1)^4} \frac{1}{(1-x)(1-\bar{x})(x-\bar{x})} \\
&\times \left\{ 3 [\text{Li}_3(x) - \text{Li}_3(\bar{x})] - \ln(-x\bar{x}) [\text{Li}_2(x) - \text{Li}_2(\bar{x})] \right\}, \tag{B.6.2}
\end{aligned}$$

with

$$x\bar{x} = \frac{m^2}{(p_3 + k_1)^2}, \quad (\text{B.6.3})$$

$$(1-x)(1-\bar{x}) = \frac{p_3^2}{(p_3 + k_1)^2}. \quad (\text{B.6.4})$$

Working to leading power in $k_1^2 = m^2$, we can take either x or \bar{x} to be small. The final answer is independent of which one is picked, so we assume that \bar{x} is small, and hence $\bar{x} = \frac{m^2(1-x)}{p_3^2 x}$.

Using the delta functions, and performing the integral over the azimuthal angle, the phase space can be written as

$$\int \frac{d^4 k_1}{i(2\pi)^4} (-2\pi i) \delta(k_1^2 - m^2) \Theta(k_1^0) (-2\pi i) \delta[(p_2 - k_1)^2] \Theta(p_2^0 - k_1^0) = \frac{i}{16\pi} \int_{-1}^1 d\cos\theta. \quad (\text{B.6.5})$$

In the rest frame of p_2 , the propagator in $p_3 + k_1$ becomes

$$(p_3 + k_1)^2 = p_3^2 - m_2(\omega_3 - p\cos\theta), \quad (\text{B.6.6})$$

where p is the magnitude of the three-momentum of the outgoing particles, and where we have dropped power corrections in m^2 and hence used that $\omega_{k_1} = |\vec{k}_1| = m_2/2$. Changing variables from $\cos\theta$ to $x = 1 - \frac{p_3^2}{(p_3 + k_1)^2}$ gives a Jacobian of

$$d\cos\theta = -\frac{(p_3 + k_1)^4}{m_2 p p_3^2} dx. \quad (\text{B.6.7})$$

In this frame, the energy of p_3 is $\omega_3 = \frac{m_2^2 + p_3^2 - p_1^2}{2m_2}$ and momentum of the outgoing particles is

$p = -\frac{p_1^2}{2m_2} (z - \bar{z})$, which gives

$$(p_3 + k_1)^2 \cong \frac{1}{2} p_1^2 [(1 - z)(1 - \cos \theta) + (1 - \bar{z})(1 + \cos \theta)] , \quad (\text{B.6.8})$$

and hence we have

$$\frac{1}{1 - x} = \frac{1}{2} \frac{1 + \cos \theta}{1 - z} + \frac{1}{2} \frac{1 - \cos \theta}{1 - \bar{z}} \quad (\text{B.6.9})$$

to leading power in m^2 . This shows that $x = \bar{z}$ for $\cos \theta = -1$ and $x = z$ for $\cos \theta =$

1. Putting everything together, phase space along with the propagator in $p_3 + k_1$ can be written as

$$\begin{aligned} \int \frac{d^4 k_1}{i (2\pi)^4} (-2\pi i) \delta(k_1^2) \Theta(k_1^0) (-2\pi i) \delta[(p_2 - k_1)^2] \Theta(p_2^0 - k_1^0) \frac{1}{(p_3 + k_1)^2} \\ = -\frac{i}{16\pi} \int_{\bar{z}}^z dx \frac{1}{m_2 p (1 - x)} . \end{aligned} \quad (\text{B.6.10})$$

Then, dropping polylogarithms in \bar{x} that are subleading in the limit $m^2 \rightarrow 0$,

$$T_3^{\mathcal{C}_1, \mathcal{C}_2} \Big|_{-i\varepsilon \text{ on r.h.s. of cut}} + T_3^{\mathcal{C}_1, \mathcal{C}_3} = \frac{1}{16 (4\pi)^4} \frac{1}{m_2 p p_3^4} \int_{\bar{z}}^z \frac{dx}{x} \left\{ 3\text{Li}_3(x) - \ln(-x\bar{x}) \text{Li}_2(x) \right\} . \quad (\text{B.6.11})$$

The integration contour from \bar{z} to z in the region R^2 can be taken to be a straight line from

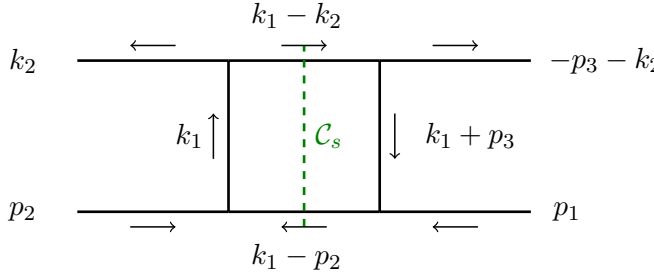
$\bar{z} < 0$ to z , with $0 < z < 1$. Integrating this expression and using $p = -\frac{p_1^2}{2m_2}(z - \bar{z})$ gives

$$T_3^{\mathcal{C}_1, \mathcal{C}_2} \Big|_{-i\varepsilon \text{ on r.h.s. of cut}} + T_3^{\mathcal{C}_1, \mathcal{C}_3} = \frac{1}{2048\pi^4} \frac{1}{p_1^2 p_3^4 (z - \bar{z})} \left\{ -3 [\text{Li}_4(z) - \text{Li}_4(\bar{z})] \right. \\ \left. + \ln \left(-\frac{m^2}{p_3^2} \right) [\text{Li}_3(z) - \text{Li}_3(\bar{z})] - \frac{1}{2} [\text{Li}_2^2(z) - \text{Li}_2^2(\bar{z})] \right\}. \quad (\text{B.6.12})$$

Next, we calculate the double cut $\mathcal{C}_2\mathcal{C}_3$, with all other propagators having a $+i\varepsilon$. We write the cut as

$$T_3^{\mathcal{C}_2, \mathcal{C}_3} = \frac{1}{2} \int \frac{d^4 k_2}{i(2\pi)^4} (-2\pi i) \delta(k_2^2 - m^2) \Theta(k_2^0) \text{Cut}_{(p_2 - k_2)^2} B(p_2^2, k_2^2, (p_3 + k_2)^2, p_1^2) \\ \times \frac{1}{(p_3 + k_2)^2} \text{Cut}_{k_2^2} T^2[(p_3 + k_2)^2, k_2^2, p_3^2], \quad (\text{B.6.13})$$

where $\text{Cut}_{(p_2 - k_2)^2} B$ is an s -channel cut through a box with one massive internal line,

$$\text{Cut}_s B(p_2^2, k_2^2, (p_3 + k_2)^2, p_1^2) =$$


$$(\text{B.6.14})$$

$$= \frac{1}{16\pi} \frac{\log \left[-\frac{m^2 p_1^2 (p_3^2 - 2\omega_{k_2} (\omega_3 - p \cos \theta))}{2m_2 (p_3^2 - m_2 (\omega_3 - p \cos \theta))^2 \omega_{k_2}} \right] + 2\pi i}{m_2 (p_3^2 - m_2 (\omega_3 - p \cos \theta)) \omega_{k_2}} \quad (\text{B.6.15})$$

where θ is now the angle between p_3 and k_2 . The cut of the three mass triangle is given by

$$\text{Cut}_{k_2^2} T^1 = \frac{i}{8\pi (\xi - \bar{\xi})} \frac{1}{(p_3 + k_2)^2} \ln \left(\frac{1 - \xi}{1 - \bar{\xi}} \right) \quad (\text{B.6.16})$$

with

$$\xi \bar{\xi} = \frac{k_2^2}{(p_3 + k_2)^2}, \quad (1 - \xi)(1 - \bar{\xi}) = \frac{p_3^2}{(p_3 + k_2)^2}, \quad (\text{B.6.17})$$

where we take $k_2^2 > 0$, $p_3^2 < 0$ and it can be shown that for these cuts, $(p_3 + k_2)^2 < 0$. As

before, we assume that $\bar{\xi}$ is small. We make a change of variables from ω_{k_2} and $\cos \theta$ to ξ

and x , defined by

$$\xi = 1 - \frac{p_3^2}{(p_3 + k_2)^2}, \quad x = 1 - \frac{p_3^2}{p_3^2 - m_2 (\omega_3 - p \cos \theta)}, \quad (\text{B.6.18})$$

with Jacobian

$$\frac{\partial(\xi, x)}{\partial(\omega_{k_2}, \cos \theta)} = \begin{pmatrix} \frac{\partial \xi}{\partial \omega_{k_2}} & \frac{\partial \xi}{\partial \cos \theta} \\ 0 & \frac{\partial x}{\partial \cos \theta} \end{pmatrix}, \quad (\text{B.6.19})$$

where

$$\frac{\partial \xi}{\partial \omega_{k_2}} = \frac{-2p_3^2 (\omega_3 - p \cos \theta)}{(p_3 + k_2)^4}, \quad \frac{\partial x}{\partial \cos \theta} = \frac{m_2 p p_3^2}{[p_3^2 - m_2 (\omega_3 - p \cos \theta)]^2}. \quad (\text{B.6.20})$$

The limits of the ξ integrals are at 0 and x , while the limits of the x integration are at z

and \bar{z} . Putting everything together, we get

$$T_3^{\mathcal{C}_2, \mathcal{C}_3} = \frac{1}{4096\pi^4 m_2 p_3^2} \int_{\bar{z}}^z \frac{dx}{x p p_3^2} \int_0^x d\xi \ln \left[\frac{-m^2 (1 - x) x p_1^2}{p_2^2 p_3^2 \xi} \right] \frac{\ln(1 - \xi)}{\xi}. \quad (\text{B.6.21})$$

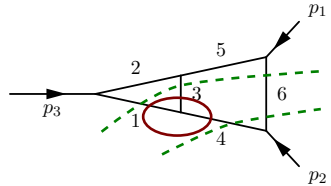
Performing the integrals in ξ and x , and using that $p = -\frac{p_1^2}{2m_2} (z - \bar{z})$ results in

$$T_3^{\mathcal{C}_2, \mathcal{C}_3} = \frac{1}{2048\pi^4 p_1^2 p_3^4 (z - \bar{z})} \left\{ \left[-\ln \left(-\frac{m^2}{p_3^2} \right) + \ln(z\bar{z}) + 2\pi i \right] [\text{Li}_3(z) - \text{Li}_3(\bar{z})] \right. \\ \left. + \left[\frac{1}{2} \text{Li}_2^2(z) - \frac{1}{2} \text{Li}_2^2(\bar{z}) \right] - [\text{Li}_4(z) - \text{Li}_4(\bar{z})] \right\}. \quad (\text{B.6.22})$$

B.7 Massless three-point vertices

When calculating cut graphs, we sometimes encounter subgraphs with cuts of massless lines on either side of a three-point vertex. This appendix discusses two important subtleties involved in computing these cut subgraphs. The first relates to evaluating the diagrams in dimensional regularization, and the second comes from delta functions evaluated at the endpoints of integration.

When evaluating diagrams with massless three-point vertices in dimensional regularization using the covariant cutting rules, one gets a delta function in the angle between the two particles multiplied with its argument raised to a power. For example, consider the graph



$$\propto \int_{-1}^1 d \cos \theta \left(1 - \cos^2 \theta\right)^{\frac{d-4}{2}} \delta(1 - \cos \theta) , \quad (\text{B.7.1})$$

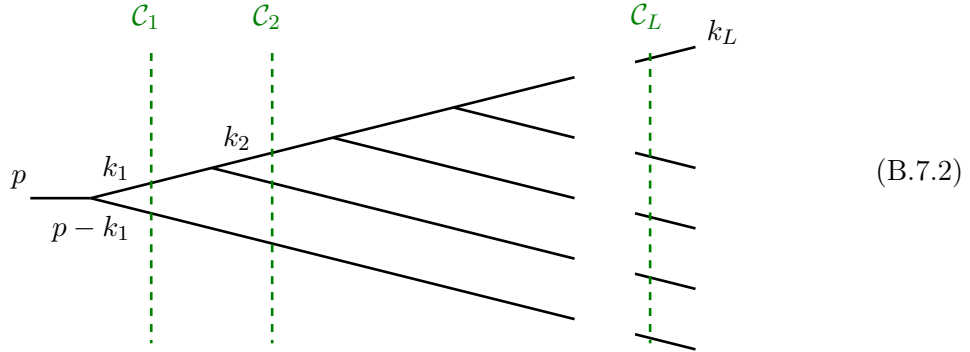
which contributes to $(\text{Disc}_{p_2})^2 \Phi_2$. The dashed lines correspond to cuts and the circled subgraph is the problematic three-point vertex. Here, θ is the angle between internal particles 1 and 3 in the diagram.

The first problem with this expression is that the limit $d \rightarrow 4$ is not smooth. For $d > 4$ the integral is zero, for $d = 4$ it is finite, and for $d < 4$ it is divergent. In [165] it was argued that one should use the $d > 4$ result and set all such graphs to zero. Indeed, such an approach seems to work in the examples considered in [165]. However, it may give results for the cut graphs that are inconsistent with the discontinuities, as discussed below

Eqs. (5.6.74) and (5.6.80). An alternative to using dimensional regularization is to give the internal lines a small mass m_{reg} and take the limit $m_{\text{reg}} \rightarrow 0$. Although masses are not great regulators in general, particularly in gauge theories where they can violate gauge invariance, for the Feynman integrals we consider in this paper they always seems to give results for the cuts consistent with the discontinuities.

The second problem is that, even if a graph or sum of graphs is IR finite in $d = 4$, the delta function of the angle between the two particles may need to be evaluated at one of the endpoints of the limits of integration. Such expressions are not generally well-defined, and more careful analysis is needed. As we will show, this ultimately results in a combinatorial factor of $\frac{1}{L!}$ compared to the naïve expectation of setting $\int_{-1}^1 \delta(1 - \cos \theta) d \cos \theta$ to 1, where $L - 1$ is the number of massless three point vertices in the cut diagram.

To see how the combinatorial factor arises, we calculate the L -loop triangle:



The incoming particle is massive with $p^2 = m^2$, and we cut the massless propagators with momentum $k_1, \dots, k_L, p - k_1$, and $k_2 - k_1, \dots, k_L - k_{L-1}$. Naïvely, using the covariant cutting

rules, one would put all the cut particles on-shell and the diagram above would be given by

$$\begin{aligned}
T = i^L \int \frac{d^4 k_1}{(2\pi)^4} \cdots \frac{d^4 k_L}{(2\pi)^4} (2\pi) \delta(k_1^2) \Theta(k_1^0) \cdots (2\pi) \delta(k_L^2) \Theta(k_L^0) \\
\times (2\pi) \delta[(p - k_1)^2] \Theta(p^0 - k_1^0) (2\pi) \delta[(k_1 - k_2)^2] \Theta(k_1^0 - k_2^0) \\
\times \cdots (2\pi) \delta[(k_{L-1} - k_L)^2] \Theta(k_{L-1}^0 - k_L^0) . \quad (\text{B.7.3})
\end{aligned}$$

We label the angle between k_i and k_j by $\theta_{i,j}$, define $\omega_i \equiv k_i^0$, and denote the angle between k_1 and the z -axis as θ . In the center of mass frame of p , the above expression can be written

$$\begin{aligned}
T = i^L \int \frac{d^3 k_1}{(2\pi)^3 2\omega_1} \cdots \frac{d^3 k_L}{(2\pi)^3 2\omega_L} (2\pi) \delta(m^2 - 2m\omega_1) (2\pi) \delta[-2\omega_1\omega_2(1 - \cos\theta_{1,2})] \\
\times \cdots (2\pi) \delta[-2\omega_{L-1}\omega_L(1 - \cos\theta_{L-1,L})] \Theta(\omega_1 > \omega_2 > \cdots > \omega_L) . \quad (\text{B.7.4})
\end{aligned}$$

Extracting the Jacobian factors results in

$$\begin{aligned}
T = \frac{i^L}{(8\pi)^L m} \int_0^\infty d\omega_1 \int_0^{\omega_1} \frac{d\omega_2}{\omega_2} \cdots \int_0^{\omega_{L-2}} \frac{d\omega_{L-1}}{\omega_{L-1}} \int_0^{\omega_{L-1}} d\omega_L \delta\left(\omega_1 - \frac{m}{2}\right) \\
\int_{-1}^1 d\cos\theta \int_{-1}^1 d\cos\theta_{1,2} \cdots \int_{-1}^1 d\cos\theta_{k_{L-1},k_L} \delta(1 - \cos\theta_{1,2}) \cdots \delta(1 - \cos\theta_{k_{L-1},k_L}) . \quad (\text{B.7.5})
\end{aligned}$$

This integral is ambiguous, since the delta functions of the angles are evaluated at the integration endpoints. To evaluate it properly, we must go back to the TOPT expression for the corresponding diagram, where we have a handle on how to make sense of these products

of delta functions. Namely, we know that they arise when using the relation

$$\lim_{\varepsilon \rightarrow 0} \left(\frac{1}{E + i\varepsilon} - \frac{1}{E - i\varepsilon} \right) = -2\pi i \delta(E). \quad (\text{B.7.6})$$

Thus, when we encounter a delta function that is evaluated at an integration endpoint, this implies we have used the distributional identity in Eq. (B.7.6) too early. For massless three-point vertices, we should instead use the expression

$$\frac{1}{E + i\varepsilon} - \frac{1}{E - i\varepsilon} = -2i \frac{\varepsilon}{E^2 + \varepsilon^2}, \quad (\text{B.7.7})$$

and only take the limit $\varepsilon \rightarrow 0$ after all the integrals have been evaluated. To shorten our equations, we define the expression that appears on the right-hand side of Eq. (B.7.7) as

$$\delta^\varepsilon \equiv \frac{1}{\pi} \frac{\varepsilon}{x^2 + \varepsilon^2}.$$

Two loops

For extra clarity, we now show how the correct combinatoric factor results in the two-loop case. The L -loop case is worked out analogously afterwards; it involves the same ideas but with longer expressions. The two-loop TOPT diagram is given by

$$T = i^2 \int \frac{d^3 k_1}{(2\pi)^3 2\omega_1} \int \frac{d^3 k_2}{(2\pi)^3 2\omega_2} \frac{1}{2\omega_1} \frac{1}{2\omega_{1-2}} (2\pi) \delta^\varepsilon(m - 2\omega_1) (2\pi) \delta^\varepsilon(m - \omega_1 - \omega_2 - \omega_{1-2}) \quad (\text{B.7.8})$$

with $\omega_{1-2} = \sqrt{\omega_1^2 + \omega_2^2 - 2\omega_1\omega_2 \cos \theta_{1,2}}$. We have already imposed three-momentum conservation. We perform the azimuthal integrals, and change variables from $\cos \theta_{1,2}$ to ω_{1-2} to

get

$$T = \frac{i^2}{(2\pi)^2 2^4} \int d\omega_1 \int d\cos\theta \int d\omega_2 \int_{\omega_1-\omega_2}^{\omega_1+\omega_2} d\omega_{1-2} \delta^\varepsilon(m-2\omega_1) \delta^\varepsilon(m-\omega_1-\omega_2-\omega_{1-2}) . \quad (\text{B.7.9})$$

We now use that

$$\int \delta^\varepsilon(x) dx = \frac{1}{\pi} \int \frac{\varepsilon}{x^2 + \varepsilon^2} dx = \frac{1}{\pi} \arctan\left(\frac{x}{\varepsilon}\right) \quad (\text{B.7.10})$$

to write

$$T = \frac{i^2}{(2\pi)^2 2^4} \int d\cos\theta \int d\omega_2 \int \frac{d\omega_1}{\omega_1} \delta^\varepsilon(m-2\omega_1) \times \frac{1}{\pi} \left[\arctan\left(\frac{m-2\omega_1}{\varepsilon}\right) - \arctan\left(\frac{m-2\omega_1-2\omega_2}{\varepsilon}\right) \right] . \quad (\text{B.7.11})$$

We can plug in $\omega_1 = \frac{m}{2}$ everywhere except at singular points, to get

$$T = \frac{1}{2^6 \pi^3} \int d\cos\theta \int d\omega_2 \int d\omega_1 \delta^\varepsilon(m-2\omega_1) \left[\arctan\left(\frac{m-2\omega_1}{\varepsilon}\right) - \arctan\left(\frac{-2\omega_2}{\varepsilon}\right) \right] . \quad (\text{B.7.12})$$

Since $\frac{d}{dx} \arctan\left(\frac{x}{\varepsilon}\right) = \pi \delta^\varepsilon(x)$, we get

$$\begin{aligned} & \int_0^\infty d\omega_1 \pi \delta^\varepsilon(m-2\omega_1) \left[\arctan\left(\frac{m-2\omega_1}{\varepsilon}\right) - \arctan\left(\frac{-2\omega_2}{\varepsilon}\right) \right] \\ &= \frac{1}{2} \left[\left(\arctan\left(\frac{m-2\omega_1}{\varepsilon}\right) - \arctan\left(\frac{-2\omega_2}{\varepsilon}\right) \right)^2 \right]_0^\infty = \frac{\pi^2}{2} , \quad (\text{B.7.13}) \end{aligned}$$

where we have taken the limit $\varepsilon \rightarrow 0^+$ when writing the last equation. The factor of $\frac{1}{2}$ in this equation, arising from the integral over the product of arctan and a δ^ε function, has

the same origin as the $\frac{1}{L!}$ factor in the L -loop case. The combinatorial factor arises because the δ^ε s in Eq. (B.7.9) only have support on the endpoint of the sequential delta function.

We plug this into Eq. (B.7.12) to get

$$T = \frac{i^2}{2^7 \pi^2 m} \int d\cos\theta \int d\omega_2. \quad (\text{B.7.14})$$

Comparing to the $L = 2$ case of Eq. (B.7.5), we learn that we must multiply the right hand side of $\int_{-1}^1 \delta(1 - \cos\theta_{1,2}) \stackrel{?}{=} 1$ by a combinatorial factor of $\frac{1}{2}$. Although this factor of $\frac{1}{2}$ could potentially be justified in the two-loop case by claiming that the delta function in Eq. (B.7.5) is only integrated up to its endpoint, and hence should be evaluated to give $\frac{1}{2}$, that argument does not generalize to the L -loop case, where we will see that we encounter a combinatorial factor of $\frac{1}{L!}$ rather than $\frac{1}{2!}$.

L loops

The L -loop TOPT diagram is given by

$$\begin{aligned} T = i^L \int \frac{d^3 k_1}{(2\pi)^3 2\omega_1} \cdots \frac{d^3 k_L}{(2\pi)^3 2\omega_L} \frac{1}{2\omega_1} \frac{1}{2\omega_{1-2}} \cdots \frac{1}{2\omega_{(L-1)-L}} (2\pi) \delta^\varepsilon(m - 2\omega_1) \\ \times (2\pi) \delta^\varepsilon(m - \omega_1 - \omega_2 - \omega_{1-2}) (2\pi) \delta^\varepsilon(m - \omega_1 - \omega_3 - \omega_{1-2} - \omega_{2-3}) \\ \times \cdots (2\pi) \delta^\varepsilon(m - \omega_1 - \omega_L - \omega_{1-2} - \cdots - \omega_{(L-1)-L}), \quad (\text{B.7.15}) \end{aligned}$$

where

$$\omega_{i-j} = \sqrt{\omega_i^2 + \omega_j^2 - 2\omega_i\omega_j \cos\theta_{i,j}}. \quad (\text{B.7.16})$$

Performing the azimuthal integrals gives

$$\begin{aligned}
T &= \frac{i^L}{2^{2L} (2\pi)^L} \int_0^\infty \omega_1 d\omega_1 \int_0^{\omega_1} \omega_2 d\omega_2 \cdots \int_0^{\omega_{L-1}} \omega_L d\omega_L \frac{1}{\omega_1} \frac{1}{\omega_{1-2}} \cdots \frac{1}{\omega_{(L-1)-L}} \\
&\quad \times \int_{-1}^1 d\cos\theta \int_{-1}^1 d\cos\theta_{1,2} \cdots \int_{-1}^1 d\cos\theta_{k_{L-1},k_L} \delta^\varepsilon(m - 2\omega_1) \\
&\quad \times \delta^\varepsilon(m - \omega_1 - \omega_2 - \omega_{1-2}) \delta^\varepsilon(m - \omega_1 - \omega_3 - \omega_{1-2} - \omega_{2-3}) \\
&\quad \times \cdots \delta^\varepsilon(m - \omega_1 - \omega_L - \omega_{1-2} - \cdots - \omega_{(L-1)-L}) . \quad (\text{B.7.17})
\end{aligned}$$

We change variables from the $\cos\theta_{i,i+1}$ variables to x_1, \dots, x_{L-1} with $x_i = \omega_{i-(i+1)}$. The

Jacobian for each i is given by

$$J_i = \left(\frac{\partial \omega_{k_i - k_{i+1}}}{\partial \cos\theta_{i,i+1}} \right)^{-1} = - \left(\frac{\omega_i \omega_{k_{i+1}}}{\omega_{k_i - k_{i+1}}} \right)^{-1} , \quad (\text{B.7.18})$$

so

$$\begin{aligned}
T &= \frac{i^L}{2^{2L} (2\pi)^L} \int_{-1}^1 \cos\theta \int_0^\infty \frac{d\omega_1}{\omega_1} \int_0^{\omega_1} \frac{d\omega_2}{\omega_2} \cdots \int_0^{\omega_{L-2}} \frac{d\omega_{L-1}}{\omega_{L-1}} \int_0^{\omega_{L-1}} d\omega_L \delta^\varepsilon(m - 2\omega_1) \\
&\quad \times \int_{\omega_1 - \omega_2}^{\omega_1 + \omega_2} dx_1 \delta^\varepsilon(m - \omega_1 - \omega_2 - x_1) \int_{\omega_2 - \omega_3}^{\omega_2 + \omega_3} dx_2 \delta^\varepsilon(m - \omega_1 - \omega_3 - x_1 - x_2) \\
&\quad \times \cdots \int_{\omega_{L-1} - \omega_L}^{\omega_{L-1} + \omega_L} dx_{L-1} \delta^\varepsilon(m - \omega_1 - \omega_L - x_1 - x_2 - \cdots - x_{L-1}) . \quad (\text{B.7.19})
\end{aligned}$$

Shifting the integrals gives

$$\begin{aligned}
T = & \frac{i^L}{2^{2L+1} (2\pi)^L} \int_{-1}^1 d\cos\theta \int_0^\infty \frac{dx_0}{x_0} \int_0^{x_0} \frac{d\omega_2}{\omega_2} \cdots \int_0^{\omega_{L-2}} \frac{d\omega_{L-1}}{\omega_{L-1}} \int_0^{\omega_{L-1}} d\omega_L \delta^\varepsilon \left(x_0 - \frac{m}{2} \right) \\
& \times \int_{-\omega_2+x_0}^{\omega_2+x_0} dx_1 \delta^\varepsilon (m - \omega_2 - x_0 - x_1) \int_{\omega_2-\omega_3+x_1}^{\omega_2+\omega_3+x_1} dx_{1,2} \delta^\varepsilon (m - \omega_3 - x_0 - x_{1,2}) \\
& \times \cdots \int_{\omega_{L-1}-\omega_L+x_{1,L-2}}^{\omega_{L-1}+\omega_L+x_{1,L-2}} dx_{1,L-1} \delta^\varepsilon (m - \omega_L - x_0 - x_{1,L-1}) , \quad (\text{B.7.20})
\end{aligned}$$

where $x_{1,i} = x_1 + \cdots + x_i$ and $x_0 = \omega_1$. We now have a product of delta functions where each is evaluated at the endpoint of the previous one. To handle this more carefully, we use the δ^ε distributions. In particular, we investigate the expression

$$\begin{aligned}
\mathcal{I} = & \int_0^\infty dx_0 \delta^\varepsilon \left(x_0 - \frac{m}{2} \right) \int_{-\omega_2+x_0}^{\omega_2+x_0} dx_1 \delta^\varepsilon (m - \omega_2 - x_0 - x_1) \\
& \times \int_{\omega_2-\omega_3+x_1}^{\omega_2+\omega_3+x_1} dx_{1,2} \delta^\varepsilon (m - \omega_3 - x_0 - x_{1,2}) \cdots \int_{\omega_{L-1}-\omega_L+x_{1,L-2}}^{\omega_{L-1}+\omega_L+x_{1,L-2}} dx_{1,L-1} \delta^\varepsilon (m - \omega_L - x_0 - x_{1,L-1}) \\
& \times F(x_0, x_1, x_{1,2}, \dots, x_{1,L-1}), \quad (\text{B.7.21})
\end{aligned}$$

where F is a test function, which we take to be a smooth function of compact support. We aim to compute the $\epsilon \rightarrow 0^+$ limit of this integral. We use the fact that if $x = a + \epsilon y$, then

$$\delta^\epsilon(x - a)dx = \frac{dy}{\pi(1 + y^2)}.$$

Using this formula repeatedly, we find

$$\begin{aligned} \mathcal{I} = & \int_{-\frac{m}{2\epsilon}}^{\infty} \frac{dy_0}{\pi(1+y_0^2)} \int_{y_0}^{y_0+\frac{2\omega_2}{\epsilon}} \frac{dy_1}{\pi(1+y_1^2)} \int_{y_1}^{y_1+\frac{2\omega_3}{\epsilon}} \frac{dy_{1,2}}{\pi(1+y_{1,2}^2)} \cdots \int_{y_{1,L-2}}^{y_{1,L-2}+\frac{2\omega_L}{\epsilon}} \frac{dy_{1,L-1}}{\pi(1+y_{1,L-1}^2)} \\ & \times F(m/2 + \epsilon y_0, m/2 - \omega_2 + \epsilon y_1, m/2 - \omega_3 + \epsilon y_{1,2}, \dots, m/2 - \omega_L + \epsilon y_{1,L-1}). \quad (\text{B.7.22}) \end{aligned}$$

Since the function F is smooth, we can series expand it around $\epsilon = 0$. We keep only the zeroth-order terms in the expansion; the higher-order terms do not contribute in the limit $\epsilon \rightarrow 0^+$.

If ω_i vanishes, then the integral over $y_{1,i-1}$ vanishes, as the upper and lower integration limits are coincident. If all of the ω_i are strictly positive, then the upper integration limits all become $+\infty$ in the $\epsilon \rightarrow 0^+$ limit. Hence, we obtain

$$\begin{aligned} \lim_{\epsilon \rightarrow 0^+} \mathcal{I} = & F(m/2, m/2 - \omega_2, \dots, m/2 - \omega_L) \\ & \times \int_{-\infty}^{\infty} \frac{dy_0}{\pi(1+y_0^2)} \int_{y_0}^{\infty} \frac{dy_1}{\pi(1+y_1^2)} \int_{y_1}^{\infty} \frac{dy_{1,2}}{\pi(1+y_{1,2}^2)} \cdots \int_{y_{1,L-2}}^{\infty} \frac{dy_{1,L-1}}{\pi(1+y_{1,L-1}^2)}. \quad (\text{B.7.23}) \end{aligned}$$

Performing the integrals one by one, we get an arctan function raised to a power each time, just as in Eq. (B.7.13). The result after performing $L - 1$ integrations is

$$\begin{aligned} \lim_{\epsilon \rightarrow 0^+} \mathcal{I} = & F(m/2, m/2 - \omega_2, \dots, m/2 - \omega_L) \\ & \times (-1)^{L-1} \int_{-\infty}^{\infty} \frac{dy_0}{2^{L-1} (L-1)! \pi^L (1+y_0^2)} (\pi - 2 \arctan(y_0))^{L-1}. \quad (\text{B.7.24}) \end{aligned}$$

The last integral evaluates to

$$\begin{aligned} \lim_{\epsilon \rightarrow 0^+} \mathcal{I} &= F(m/2, m/2 - \omega_2, \dots, m/2 - \omega_L) \\ &\times (-1)^L \left[\frac{1}{2^L L! \pi^L} (\pi - 2 \arctan(y_0))^L \right]_{-\infty}^{\infty} = \frac{1}{L!}. \end{aligned} \quad (\text{B.7.25})$$

Making use of this in Eq. (B.7.20), we get

$$T = \frac{i^L}{(8\pi)^L L! m} \int_{-1}^1 d \cos \theta \int_0^{m/2} \frac{d\omega_2}{\omega_2} \dots \int_0^{\omega_{L-2}} \frac{d\omega_{L-1}}{\omega_{L-1}} \int_0^{\omega_{L-1}} d\omega_L. \quad (\text{B.7.26})$$

In particular, this result has an extra factor of $\frac{1}{L!}$ compared to what one would get by evaluating each of the delta functions to 1. Although we can compute these integrals in TOPT, it is harder to find this combinatorial factor using covariant Feynman rules.

Bibliography

- [1] E. Noether, “Invariante variationsprobleme,” *Nachr. d. König. Gesellsch. d. Wiss. zu Göttingen, Math-phys. Klasse* (1918) 235–257.
- [2] F. Bloch and A. Nordsieck, “Note on the Radiation Field of the Electron,” *Phys. Rev.* **52** (1937) 54–59.
- [3] R. Doria, J. Frenkel, and J. C. Taylor, “Counter Example to Nonabelian Bloch-Nordsieck Theorem,” *Nucl. Phys.* **B168** (1980) 93–110.
- [4] T. Kinoshita, “Mass Singularities of Feynman Amplitudes,” *J. Math. Phys.* **3** (1962) 650–677.
- [5] T. D. Lee and M. Nauenberg, “Degenerate Systems and Mass Singularities,” *Phys. Rev.* **133** (1964) B1549–B1562.
- [6] R. E. Cutkosky, “Singularities and Discontinuities of Feynman Amplitudes,” *Journal of Mathematical Physics* **1** no. 5, (1960) 429–433.
- [7] O. Steinmann, “Über den Zusammenhang Zwischen den Wightmanfunktionen und den Retardierten Kommutatoren,” *Helvetica Physica Acta* **33** (1960) 257–298.
- [8] V. Chung, “Infrared Divergence in Quantum Electrodynamics,” *Phys. Rev.* **140** (1965) B1110–B1122.
- [9] T. W. B. Kibble, “Coherent Soft-Photon States and Infrared Divergences. ii. Mass-Shell Singularities of Green’s Functions,” *Phys. Rev.* **173** (1968) 1527–1535.
- [10] P. P. Kulish and L. D. Faddeev, “Asymptotic Conditions and Infrared Divergences in Quantum Electrodynamics,” *Theor. Math. Phys.* **4** (1970) 745.
- [11] M. Dasgupta and G. P. Salam, “Resummation of Nonglobal QCD Observables,” *Phys. Lett.* **B512** (2001) 323–330, [arXiv:hep-ph/0104277](#) [hep-ph].

- [12] A. Banfi, G. Marchesini, and G. Smye, “Away From Jet Energy Flow,” *JHEP* **08** (2002) 006, [arXiv:hep-ph/0206076](#) [[hep-ph](#)].
- [13] R. Kelley, M. D. Schwartz, and H. X. Zhu, “Resummation of Jet Mass With and Without a Jet Veto,” [arXiv:1102.0561](#) [[hep-ph](#)].
- [14] A. Hornig, C. Lee, I. W. Stewart, J. R. Walsh, and S. Zuberi, “Non-Global Structure of the Dijet Soft Function,” *JHEP* **08** (2011) 054, [arXiv:1105.4628](#) [[hep-ph](#)]. [Erratum: *JHEP*10,101(2017)].
- [15] R. Kelley, M. D. Schwartz, R. M. Schabinger, and H. X. Zhu, “The Two-Loop Hemisphere Soft Function,” *Phys. Rev.* **D84** (2011) 045022, [arXiv:1105.3676](#) [[hep-ph](#)].
- [16] M. D. Schwartz and H. X. Zhu, “Nonglobal Logarithms at Three Loops, Four Loops, Five Loops, and beyond,” *Phys. Rev.* **D90** no. 6, (2014) 065004, [arXiv:1403.4949](#) [[hep-ph](#)].
- [17] T. Becher, M. Neubert, L. Rothen, and D. Y. Shao, “Factorization and Resummation for Jet Processes,” *JHEP* **11** (2016) 019, [arXiv:1605.02737](#) [[hep-ph](#)]. [Erratum: *JHEP*05,154(2017)].
- [18] A. J. Larkoski, I. Moult, and D. Neill, “Non-Global Logarithms, Factorization, and the Soft Substructure of Jets,” *JHEP* **09** (2015) 143, [arXiv:1501.04596](#) [[hep-ph](#)].
- [19] J. R. Forshaw, A. Kyrieleis, and M. H. Seymour, “Super-Leading Logarithms in Non-Global Observables in QCD,” *JHEP* **08** (2006) 059, [arXiv:hep-ph/0604094](#) [[hep-ph](#)].
- [20] J. R. Forshaw, A. Kyrieleis, and M. H. Seymour, “Super-leading logarithms in non-global observables in QCD: Colour basis independent calculation,” *JHEP* **09** (2008) 128, [arXiv:0808.1269](#) [[hep-ph](#)].
- [21] J. Keates and M. H. Seymour, “Super-Leading Logarithms in Non-Global Observables in QCD: Fixed Order Calculation,” *JHEP* **04** (2009) 040, [arXiv:0902.0477](#) [[hep-ph](#)].
- [22] M. D. Schwartz, K. Yan, and H. X. Zhu, “Collinear Factorization Violation and Effective Field Theory,” *Phys. Rev.* **D96** no. 5, (2017) 056005, [arXiv:1703.08572](#)

[hep-ph].

- [23] M. D. Schwartz, K. Yan, and H. X. Zhu, “Factorization Violation and Scale Invariance,” *Phys. Rev.* **D97** no. 9, (2018) 096017, [arXiv:1801.01138 \[hep-ph\]](#).
- [24] D. R. Yennie, S. C. Frautschi, and H. Suura, “The Infrared Divergence Phenomena and High-Energy Processes,” *Annals Phys.* **13** (1961) 379–452.
- [25] S. Weinberg, “Infrared Photons and Gravitons,” *Phys. Rev.* **140** (1965) B516–B524.
- [26] G. Grammer, Jr. and D. R. Yennie, “Improved Treatment for the Infrared Divergence Problem in Quantum Electrodynamics,” *Phys. Rev.* **D8** (1973) 4332–4344.
- [27] S. Weinberg, *The Quantum Theory of Fields. Vol. 1: Foundations*. Cambridge University Press, 2005.
- [28] T. Muta, *Foundations of Quantum Chromodynamics: An Introduction to Perturbative Methods in Gauge Theories, (3rd ed.)*, vol. 78 of *World Scientific Lecture Notes in Physics*. World Scientific, Hackensack, N.J., 2010.
- [29] C. Di’Lieto, S. Gendron, I. G. Halliday, and C. T. Sachrajda, “A Counter Example to the Bloch-Nordsieck Theorem in Nonabelian Gauge Theories,” *Nucl. Phys.* **B183** (1981) 223–250.
- [30] A. Andrasi, M. Day, R. Doria, J. Frenkel, and J. C. Taylor, “Soft Divergences in Perturbative QCD,” *Nucl. Phys.* **B182** (1981) 104–124.
- [31] C. E. Carneiro, M. Day, J. Frenkel, J. C. Taylor, and M. T. Thomaz, “Leading Noncancelling Infrared Divergences In Perturbative QCD,” *Nucl. Phys.* **B183** (1981) 445–470.
- [32] G. F. Sterman and S. Weinberg, “Jets from Quantum Chromodynamics,” *Phys. Rev. Lett.* **39** (1977) 1436.
- [33] T. Kinoshita, “Note on the Infrared Catastrophe,” *Progress of Theoretical Physics* **5** no. 6, (1950) 1045–1047.
- [34] N. Nakanishi, “General Theory of Infrared Divergence,” *Progress of Theoretical*

Physics **19** no. 2, (1958) 159–168.

- [35] T. Muta and C. A. Nelson, “Role of Quark-Gluon Degenerate States in Perturbative Quantum Chromodynamics,” *Phys. Rev. D* **25** (1982) 2222–2225.
- [36] A. Axelrod and C. A. Nelson, “Degenerate States and the Kinoshita-Lee-Nauenberg Theorem in Deep-Inelastic Scattering,” *Phys. Rev. D* **32** (1985) 2385–2395.
- [37] R. Akhoury, M. G. Sotiropoulos, and V. I. Zakharov, “The KLN Theorem and Soft Radiation in Gauge Theories: Abelian Case,” *Phys. Rev.* **D56** (1997) 377–387, [arXiv:hep-ph/9702270 \[hep-ph\]](#).
- [38] M. Lavelle and D. McMullan, “Collinearity, Convergence and Cancelling Infrared Divergences,” *JHEP* **03** (2006) 026, [arXiv:hep-ph/0511314 \[hep-ph\]](#).
- [39] R. Gastmans, J. Verwaest, and R. Meuldermans, “Dimensional Regularization in Massless QED,” *Nucl. Phys.* **B105** (1976) 454–460.
- [40] A. Khalil and W. A. Horowitz, “A Complete Diagrammatic Implementation of the Kinoshita-Lee-Nauenberg Theorem at Next-to-Leading Order,” [arXiv:1701.00763 \[hep-th\]](#).
- [41] T. Kinoshita and A. Sirlin, “Radiative Corrections to Fermi Interactions,” *Phys. Rev.* **113** (1959) 1652–1660.
- [42] T. Kinoshita and A. Ukawa, “New Approach to the Singularities of Feynman Amplitudes in the Zero Mass Limit,” *Phys. Rev.* **D13** (1976) 1573.
- [43] E. C. Poggio and H. R. Quinn, “The Infrared Behavior of Zero-Mass Green’s Functions and the Absence of Quark Confinement in Perturbation Theory,” *Phys. Rev.* **D14** (1976) 578.
- [44] G. F. Sterman, “Kinoshita’s Theorem in Yang-Mills Theories,” *Phys. Rev.* **D14** (1976) 2123–2125.
- [45] L. Landau, “On Analytic Properties of Vertex Parts in Quantum Field Theory,” *Nuclear Physics* **13** no. 1, (1959) 181 – 192.

- [46] S. B. Libby and G. F. Sterman, “Jet and Lepton Pair Production in High-Energy Lepton-Hadron and Hadron-Hadron Scattering,” *Phys. Rev.* **D18** (1978) 3252.
- [47] G. F. Sterman, “Mass Divergences in Annihilation Processes. 1. Origin and Nature of Divergences in Cut Vacuum Polarization Diagrams,” *Phys. Rev.* **D17** (1978) 2773.
- [48] J. C. Collins and G. F. Sterman, “Soft Partons in QCD,” *Nucl. Phys.* **B185** (1981) 172–188.
- [49] I. Feige and M. D. Schwartz, “An On-Shell Approach to Factorization,” *Phys. Rev.* **D88** no. 6, (2013) 065021, [arXiv:1306.6341 \[hep-th\]](#).
- [50] I. Feige and M. D. Schwartz, “Hard-Soft-Collinear Factorization to All Orders,” *Phys. Rev.* **D90** no. 10, (2014) 105020, [arXiv:1403.6472 \[hep-ph\]](#).
- [51] R. J. Eden, P. V. Landshoff, D. I. Olive, and J. C. Polkinghorne, *The analytic S-matrix*. Cambridge Univ. Press, Cambridge, 1966.
- [52] J. Ware, R. Saotome, and R. Akhoury, “Construction of an Asymptotic S Matrix for Perturbative Quantum Gravity,” *JHEP* **10** (2013) 159, [arXiv:1308.6285 \[hep-th\]](#).
- [53] A. Manohar, B. Shotwell, C. Bauer, and S. Turczyk, “Non-Cancellation of Electroweak Logarithms in High-Energy Scattering,” *Phys. Lett.* **B740** (2015) 179–187, [arXiv:1409.1918 \[hep-ph\]](#).
- [54] G. ’t Hooft and M. J. G. Veltman, “DIAGRAMMAR,” *NATO Sci. Ser. B* **4** (1974) 177–322.
- [55] M. J. G. Veltman, “Diagrammatica: The Path to Feynman Rules,” *Cambridge Lect. Notes Phys.* **4** (1994) 1–284.
- [56] D. Zwanziger, “Reduction Formulas for Charged Particles and Coherent States in Quantum Electrodynamics,” *Phys. Rev.* **D7** (1973) 1082–1099.
- [57] H. F. Contopanagos and M. B. Einhorn, “Theory of the Asymptotic S matrix for Massless Particles,” *Phys. Rev.* **D45** (1992) 1291–1321.
- [58] V. Del Duca, L. Magnea, and G. F. Sterman, “Collinear Infrared Factorization and

- Asymptotic Evolution,” *Nucl. Phys.* **B324** (1989) 391–411.
- [59] D. A. Forde and A. Signer, “Infrared Finite Amplitudes for Massless Gauge Theories,” *Nucl. Phys.* **B684** (2004) 125–161, [arXiv:hep-ph/0311059](#) [[hep-ph](#)].
 - [60] G. Curci, M. Greco, and Y. Srivastava, “QCD Jets From Coherent States,” *Nucl. Phys.* **B159** (1979) 451–468.
 - [61] G. Giavarini and G. Marchesini, “IR Finite S Matrix in the QCD Coherent State Basis,” *Nucl. Phys.* **B296** (1988) 546–556.
 - [62] A. Strominger, “Lectures on the Infrared Structure of Gravity and Gauge Theory,” [arXiv:1703.05448](#) [[hep-th](#)].
 - [63] D. Kapec, M. Perry, A.-M. Raclariu, and A. Strominger, “Infrared Divergences in QED, Revisited,” *Phys. Rev.* **D96** no. 8, (2017) 085002, [arXiv:1705.04311](#) [[hep-th](#)].
 - [64] R. Haag, “On Quantum Field Theories,” *Kong. Dan. Vid. Sel. Mat. Fys. Med.* **29N12** (1955) 1–37.
 - [65] R. Haag, “Quantum Field Theories with Composite Particles and Asymptotic Conditions,” *Phys. Rev.* **112** (1958) 669–673.
 - [66] R. Haag, “The Framework of Quantum Field Theory,” *Nuovo Cim.* **14** no. 1suppl, (1959) 131–152.
 - [67] D. Ruelle, “On Asymptotic Condition in Quantum Field Theory,” *Helvetica Physica Acta* **35** no. 3, (1962) 147.
 - [68] H. Lehmann, K. Symanzik, and W. Zimmermann, “On the Formulation of Quantized Field Theories,” *Nuovo Cim.* **1** (1955) 205–225.
 - [69] J. Collins, “A New Approach to the LSZ Reduction Formula,” [arXiv:1904.10923](#) [[hep-ph](#)].
 - [70] C. Frye, H. Hannesdottir, N. Paul, M. D. Schwartz, and K. Yan, “Infrared Finiteness and Forward Scattering,” *Phys. Rev. D* **99** no. 5, (2019) 056015, [arXiv:1810.10022](#)

[hep-ph].

- [71] J. D. Dollard, “Quantum-Mechanical Scattering Theory for Short-Range and Coulomb Interactions,” *The Rocky Mountain Journal of Mathematics* (1971) 5–88.
- [72] P. A. Dirac, “Gauge-Invariant Formulation of Quantum Electrodynamics,” *Canadian Journal of Physics* **33** no. 11, (1955) 650–660.
- [73] T. W. B. Kibble, “Coherent Soft-Photon States and Infrared Divergences. i. Classical Currents,” *J. Math. Phys.* **9** no. 2, (1968) 315–324.
- [74] E. Bagan, M. Lavelle, and D. McMullan, “Charges From Dressed Matter: Construction,” *Annals Phys.* **282** (2000) 471–502, [arXiv:hep-ph/9909257](#) [hep-ph].
- [75] J. C. Collins, D. E. Soper, and G. F. Sterman, “Soft Gluons and Factorization,” *Nucl.Phys.* **B308** (1988) 833.
- [76] J. C. Collins, D. E. Soper, and G. F. Sterman, “Factorization of Hard Processes in QCD,” *Adv.Ser.Direct.High Energy Phys.* **5** (1988) 1–91, [arXiv:hep-ph/0409313](#) [hep-ph].
- [77] C. W. Bauer, S. Fleming, and M. E. Luke, “Summing Sudakov Logarithms in $B \rightarrow X_s \gamma$ in Effective Field Theory,” *Phys.Rev.* **D63** (2000) 014006, [arXiv:hep-ph/0005275](#) [hep-ph].
- [78] M. Beneke, A. Chapovsky, M. Diehl, and T. Feldmann, “Soft Collinear Effective Theory and Heavy to Light Currents Beyond Leading Power,” *Nucl.Phys.* **B643** (2002) 431–476, [arXiv:hep-ph/0206152](#) [hep-ph].
- [79] M. Beneke and T. Feldmann, “Multipole Expanded Soft Collinear Effective Theory With Non-Abelian Gauge Symmetry,” *Phys.Lett.* **B553** (2003) 267–276, [arXiv:hep-ph/0211358](#) [hep-ph].
- [80] C. W. Bauer, S. Fleming, D. Pirjol, I. Z. Rothstein, and I. W. Stewart, “Hard Scattering Factorization From Effective Field Theory,” *Phys.Rev.* **D66** (2002) 014017, [arXiv:hep-ph/0202088](#) [hep-ph].
- [81] C. W. Bauer, S. Fleming, D. Pirjol, and I. W. Stewart, “An Effective Field Theory for

- Collinear and Soft Gluons: Heavy to Light Decays,” *Phys.Rev.* **D63** (2001) 114020, [arXiv:hep-ph/0011336](#) [hep-ph].
- [82] T. Becher, A. Broggio, and A. Ferroglia, “Introduction to Soft-Collinear Effective Theory,” *Lect. Notes Phys.* **896** (2015) pp.1–206, [arXiv:1410.1892](#) [hep-ph].
 - [83] I. W. Stewart, “Lectures on the Soft-Collinear Effective Theory,” 2013.
 - [84] I. Z. Rothstein and I. W. Stewart, “An Effective Field Theory for Forward Scattering and Factorization Violation,” *JHEP* **08** (2016) 025, [arXiv:1601.04695](#) [hep-ph].
 - [85] A. V. Manohar, “Deep Inelastic Scattering as $x \rightarrow 1$ Using Soft Collinear Effective Theory,” *Phys.Rev.* **D68** (2003) 114019, [arXiv:hep-ph/0309176](#) [hep-ph].
 - [86] M. Beneke and V. A. Smirnov, “Asymptotic Expansion of Feynman Integrals Near Threshold,” *Nucl. Phys.* **B522** (1998) 321–344, [arXiv:hep-ph/9711391](#) [hep-ph].
 - [87] A. V. Manohar and I. W. Stewart, “The Zero-Bin and Mode Factorization in Quantum Field Theory,” *Phys.Rev.* **D76** (2007) 074002, [arXiv:hep-ph/0605001](#) [hep-ph].
 - [88] Z. Bern, L. J. Dixon, and V. A. Smirnov, “Iteration of Planar Amplitudes in Maximally Supersymmetric Yang-Mills Theory at Three Loops and Beyond,” *Phys. Rev.* **D72** (2005) 085001, [arXiv:hep-th/0505205](#) [hep-th].
 - [89] C. Anastasiou, Z. Bern, L. J. Dixon, and D. A. Kosower, “Planar Amplitudes in Maximally Supersymmetric Yang-Mills Theory,” *Phys. Rev. Lett.* **91** (2003) 251602, [arXiv:hep-th/0309040](#) [hep-th].
 - [90] V. Del Duca, C. Duhr, and V. A. Smirnov, “The Two-Loop Hexagon Wilson Loop in $N = 4$ SYM,” *JHEP* **05** (2010) 084, [arXiv:1003.1702](#) [hep-th].
 - [91] A. B. Goncharov, M. Spradlin, C. Vergu, and A. Volovich, “Classical Polylogarithms for Amplitudes and Wilson Loops,” *Phys.Rev.Lett.* **105** (2010) 151605, [arXiv:1006.5703](#) [hep-th].
 - [92] S. Caron-Huot, L. J. Dixon, A. McLeod, and M. von Hippel, “Bootstrapping a Five-Loop Amplitude Using Steinmann Relations,” *Phys. Rev. Lett.* **117** no. 24,

- (2016) 241601, [arXiv:1609.00669](#) [hep-th].
- [93] L. F. Alday, D. Gaiotto, and J. Maldacena, “Thermodynamic Bubble Ansatz,” *JHEP* **09** (2011) 032, [arXiv:0911.4708](#) [hep-th].
 - [94] J. Golden, A. J. McLeod, M. Spradlin, and A. Volovich, “The Sklyanin Bracket and Cluster Adjacency at All Multiplicity,” *JHEP* **03** (2019) 195, [arXiv:1902.11286](#) [hep-th].
 - [95] H. Hannesdottir and M. D. Schwartz, “ S -matrix for Massless Particles,” *Phys. Rev. D* **101** no. 10, (2020) 105001, [arXiv:1911.06821](#) [hep-th].
 - [96] J. A. Wheeler, “On the Mathematical Description of Light Nuclei by the Method of Resonating Group Structure,” *Phys. Rev.* **52** (1937) 1107–1122.
 - [97] W. Heisenberg, “Die „beobachtbaren Größen“ in der Theorie der Elementarteilchen,” *Zeits. f. Physik* **120** (1943) 513.
 - [98] R. P. Feynman, “Relativistic Cut-Off for Quantum Electrodynamics,” *Phys. Rev.* **74** (1948) 1430–1438.
 - [99] F. J. Dyson, “The S Matrix in Quantum Electrodynamics,” *Phys. Rev.* **75** (1949) 1736–1755.
 - [100] C. A. Nelson, “Origin of Cancellation of Infrared Divergences in Coherent State Approach: Forward Process $qq \rightarrow qq + \text{gluon}$,” *Nuclear Physics B* **181** no. 1, (1981) 141–156.
 - [101] S. Catani and L. Trentadue, “Resummation of the QCD Perturbative Series for Hard Processes,” *Nuclear Physics B* **327** no. 2, (1989) 323 – 352.
 - [102] M. Greco and G. Rossi, “A Note on the Infrared Divergence,” *Nuovo Cim.* **50** (1967) 168.
 - [103] R. J. Glauber, “Coherent and Incoherent States of the Radiation Field,” *Phys. Rev.* **131** no. 6, (1963) 2766.
 - [104] L. M. Frantz, “Compton scattering of an intense photon beam,” *Phys. Rev.* **139**

no. 5B, (1965) B1326.

- [105] A. Ilderton and D. Seipt, “Backreaction on background fields: A coherent state approach,” *Phys. Rev. D* **97** no. 1, (2018) 016007, [arXiv:1709.10085 \[hep-th\]](#).
- [106] T. W. B. Kibble, “Coherent Soft-Photon States and Infrared Divergences. iii. Asymptotic States and Reduction Formulas,” *Phys. Rev.* **174** (1968) 1882–1901.
- [107] T. W. B. Kibble, “Coherent Soft-Photon States and Infrared Divergences. iv. The Scattering Operator,” *Phys. Rev.* **175** (1968) 1624–1640.
- [108] S. Catani, M. Ciafaloni, and G. Marchesini, “Noncancelling Infrared Divergences in QCD Coherent State,” *Nucl. Phys.* **B264** (1986) 588–620.
- [109] R. Gonzo, T. Mc Loughlin, D. Medrano, and A. Spiering, “Asymptotic Charges and Coherent States in QCD,” [arXiv:1906.11763 \[hep-th\]](#).
- [110] D. R. Butler and C. A. Nelson, “Nonabelian Structure of Yang-Mills Theory and Infrared Finite Asymptotic States,” *Phys. Rev.* **D18** (1978) 1196.
- [111] C. A. Nelson, “Avoidance of Counter Example to Nonabelian Bloch-Nordsieck Conjecture by Using Coherent State Approach,” *Nucl. Phys.* **B186** (1981) 187–204.
- [112] J. G. M. Gatheral, “Exponentiation of Eikonal Cross-sections in Nonabelian Gauge Theories,” *Phys. Lett.* **133B** (1983) 90–94.
- [113] J. Frenkel and J. C. Taylor, “Nonabelian Eikonal Exponentiation,” *Nucl. Phys.* **B246** (1984) 231–245.
- [114] G. Curci and M. Greco, “Mass Singularities and Coherent States in Gauge Theories,” *Physics Letters B* **79** no. 4-5, (1978) 406–410.
- [115] F. Havemann, “Collinear divergences and asymptotic states,” *Zeuthen Report* (1985) PHE–85–14.
- [116] D. A. Kosower, “All Order Collinear Behavior in Gauge Theories,” *Nucl. Phys.* **B552** (1999) 319–336, [arXiv:hep-ph/9901201 \[hep-ph\]](#).

- [117] S. Catani and M. Ciafaloni, “Gauge Covariance of QCD Coherent States,” *Nuclear Physics B* **289** (1987) 535–556.
- [118] D. Carney, L. Chaurette, D. Neuenfeld, and G. W. Semenoff, “Infrared Quantum Information,” *Phys. Rev. Lett.* **119** no. 18, (2017) 180502, [arXiv:1706.03782 \[hep-th\]](#).
- [119] D. Carney, L. Chaurette, D. Neuenfeld, and G. Semenoff, “On the Need for Soft Dressing,” *JHEP* **09** (2018) 121, [arXiv:1803.02370 \[hep-th\]](#).
- [120] C. Gómez, R. Letschka, and S. Zell, “The Scales of the Infrared,” *JHEP* **09** (2018) 115, [arXiv:1807.07079 \[hep-th\]](#).
- [121] Y.-T. Chien, M. D. Schwartz, D. Simmons-Duffin, and I. W. Stewart, “Jet Physics from Static Charges in AdS,” *Phys. Rev.* **D85** (2012) 045010, [arXiv:1109.6010 \[hep-th\]](#).
- [122] G. Korchemsky and A. Radyushkin, “Renormalization of the Wilson Loops Beyond the Leading Order,” *Nucl. Phys.* **B283** (1987) 342–364.
- [123] E. Gardi, J. M. Smillie, and C. D. White, “The Non-Abelian Exponentiation Theorem for Multiple Wilson Lines,” *JHEP* **06** (2013) 088, [arXiv:1304.7040 \[hep-ph\]](#).
- [124] M. D. Schwartz, *Quantum Field Theory and the Standard Model*. Cambridge University Press, 2014.
- [125] C. W. Bauer and M. D. Schwartz, “Event Generation from Effective Field Theory,” *Phys.Rev.* **D76** (2007) 074004, [arXiv:hep-ph/0607296 \[hep-ph\]](#).
- [126] C. W. Bauer and M. D. Schwartz, “Improving Jet Distributions With Effective Field Theory,” *Phys.Rev.Lett.* **97** (2006) 142001, [arXiv:hep-ph/0604065 \[hep-ph\]](#).
- [127] T. Becher and M. Neubert, “On the Structure of Infrared Singularities of Gauge-Theory Amplitudes,” *JHEP* **06** (2009) 081, [arXiv:0903.1126 \[hep-ph\]](#). [Erratum: *JHEP*11,024(2013)].
- [128] E. Gardi and L. Magnea, “Factorization Constraints for Soft Anomalous Dimensions in QCD Scattering Amplitudes,” *JHEP* **03** (2009) 079, [arXiv:0901.1091 \[hep-ph\]](#).

- [129] W. Magnus, “On the Exponential Solution of Differential Equations for a Linear Operator,” *Commun. Pure Appl. Math.* **7** (1954) 649–673.
- [130] G. F. Sterman, *An Introduction to Quantum Field Theory*. Cambridge University Press, 1993.
- [131] I. Feige, M. D. Schwartz, and K. Yan, “Removing Phase-Space Restrictions in Factorized Cross Sections,” *Phys. Rev.* **D91** (2015) 094027, [arXiv:1502.05411 \[hep-ph\]](#).
- [132] A. Hornig, C. Lee, and G. Ovanesyan, “Infrared Safety in Factorized Hard Scattering Cross-Sections,” *Phys. Lett.* **B677** (2009) 272–277, [arXiv:0901.1897 \[hep-ph\]](#).
- [133] M. D. Schwartz, “Resummation and NLO Matching of Event Shapes With Effective Field Theory,” *Phys. Rev.* **D77** (2008) 014026, [arXiv:0709.2709 \[hep-ph\]](#).
- [134] E. Laenen, K. J. Larsen, and R. Rietkerk, “Imaginary Parts and Discontinuities of Wilson Line Correlators,” *Phys. Rev. Lett.* **114** no. 18, (2015) 181602, [arXiv:1410.5681 \[hep-th\]](#).
- [135] E. Laenen, K. J. Larsen, and R. Rietkerk, “Position-Space Cuts for Wilson Line Correlators,” *JHEP* **07** (2015) 083, [arXiv:1505.02555 \[hep-th\]](#).
- [136] E. Farhi, “Quantum Chromodynamics Test for Jets,” *Phys. Rev. Lett.* **39** (1977) 1587–1588.
- [137] T. Becher and M. D. Schwartz, “A Precise Determination of α_s From LEP Thrust Data Using Effective Field Theory,” *JHEP* **07** (2008) 034, [arXiv:0803.0342 \[hep-ph\]](#).
- [138] J. Golden and A. J. Mcleod, “Cluster Algebras and the Subalgebra Constructibility of the Seven-Particle Remainder Function,” *JHEP* **01** (2019) 017, [arXiv:1810.12181 \[hep-th\]](#).
- [139] S. Catani, “The Singular Behaviour of QCD Amplitudes at Two-Loop Order,” *Physics Letters B* **427** no. 1-2, (1998) 161–171, [arXiv:9802439 \[hep-ph\]](#).
- [140] G. F. Sterman and M. E. Tejeda-Yeomans, “Multiloop Amplitudes and

- Resummation,” *Phys. Lett.* **B552** no. 1, (2003) 48–56, [arXiv:hep-ph/0210130](#) [[hep-ph](#)].
- [141] T. Becher and M. Neubert, “Infrared Singularities of Scattering Amplitudes in Perturbative QCD,” *Phys. Rev. Lett.* **102** (2009) 162001, [arXiv:0901.0722](#) [[hep-ph](#)]. [Erratum: *Phys. Rev. Lett.* 111, no. 19, 199905 (2013)].
 - [142] A. Broggio, C. Gnendiger, A. Signer, D. Stöckinger, and A. Visconti, “SCET Approach to Regularization-Scheme Dependence of QCD Amplitudes,” *JHEP* **2016** no. 1, (2016) 78, [arXiv:1506.05301](#) [[hep-ph](#)].
 - [143] Z. Bern, L. J. Dixon, D. C. Dunbar, and D. A. Kosower, “One Loop n Point Gauge Theory Amplitudes, Unitarity and Collinear Limits,” *Nucl. Phys.* **B425** (1994) 217–260, [arXiv:hep-ph/9403226](#) [[hep-ph](#)].
 - [144] L. J. Dixon, M. von Hippel, and A. J. McLeod, “The Four-Loop Six-Gluon NMHV Ratio Function,” *JHEP* **01** (2016) 053, [arXiv:1509.08127](#) [[hep-th](#)].
 - [145] O. Steinmann, *About the Relationship Between the Wightman Functions and the Retarded Commutators*. PhD thesis, ETH Zurich, 1960.
 - [146] S. Caron-Huot, L. J. Dixon, F. Dulat, M. von Hippel, A. J. McLeod, and G. Papathanasiou, “Six-Gluon Amplitudes in Planar $\mathcal{N}=4$ Super-Yang-Mills Theory at Six and Seven Loops,” *JHEP* **08** (2019) 016, [arXiv:1903.10890](#) [[hep-th](#)].
 - [147] **Particle Data Group** Collaboration, M. Tanabashi *et al.*, “Review of Particle Physics,” *Phys. Rev. D* **98** no. 3, (2018) 030001.
 - [148] K.-T. Chen, “Iterated Path Integrals,” *Bull. Amer. Math. Soc.* **83** no. 5, (1977) 831–879.
 - [149] A. B. Goncharov, “Geometry of Configurations, Polylogarithms, and Motivic Cohomology,” *Adv. Math.* **114** no. 2, (1995) 197–318.
 - [150] A. B. Goncharov, “Multiple Polylogarithms, Cyclotomy and Modular Complexes,” *Math Res. Letters* **5** (1998) 497–516, [arXiv:1105.2076](#) [[math.AG](#)].
 - [151] A. B. Goncharov, “Galois Symmetries of Fundamental Groupoids and

Noncommutative Geometry,” *Duke Math. J.* **128** (2005) 209, [arXiv:math/0208144](#) [[math.AG](#)].

- [152] M. Deneufchâtel, G. H. E. Duchamp, V. H. N. Minh, and A. I. Solomon, “Independence of Hyperlogarithms Over Function Fields via Algebraic Combinatorics,” [arXiv:1101.4497](#) [[math.CO](#)].
- [153] F. Brown, “On the Decomposition of Motivic Multiple Zeta Values,” [arXiv:1102.1310](#) [[math.NT](#)].
- [154] C. Duhr, H. Gangl, and J. R. Rhodes, “From Polygons and Symbols to Polylogarithmic Functions,” *JHEP* **10** (2012) 075, [arXiv:1110.0458](#) [[math-ph](#)].
- [155] F. C. S. Brown and A. Levin, “Multiple Elliptic Polylogarithms,” [arXiv:1110.6917](#) [[math.NT](#)].
- [156] J. Broedel, C. Duhr, F. Dulat, and L. Tancredi, “Elliptic Polylogarithms and Iterated Integrals on Elliptic Curves. Part I: General Formalism,” *JHEP* **05** (2018) 093, [arXiv:1712.07089](#) [[hep-th](#)].
- [157] J. Broedel, C. Duhr, F. Dulat, B. Penante, and L. Tancredi, “Elliptic Symbol Calculus: from Elliptic Polylogarithms to Iterated Integrals of Eisenstein Series,” *JHEP* **08** (2018) 014, [arXiv:1803.10256](#) [[hep-th](#)].
- [158] S. Bloch, M. Kerr, and P. Vanhove, “Local Mirror Symmetry and the Sunset Feynman Integral,” *Adv. Theor. Math. Phys.* **21** (2017) 1373–1453, [arXiv:1601.08181](#) [[hep-th](#)].
- [159] J. L. Bourjaily, Y.-H. He, A. J. McLeod, M. Von Hippel, and M. Wilhelm, “Traintracks Through Calabi-Yaus: Amplitudes Beyond Elliptic Polylogarithms,” *Phys. Rev. Lett.* **121** no. 7, (2018) 071603, [arXiv:1805.09326](#) [[hep-th](#)].
- [160] J. L. Bourjaily, A. J. McLeod, M. von Hippel, and M. Wilhelm, “A (Bounded) Bestiary of Feynman Integral Calabi-Yau Geometries,” *Phys. Rev. Lett.* **122** no. 3, (2019) 031601, [arXiv:1810.07689](#) [[hep-th](#)].
- [161] C. Bogner, S. Müller-Stach, and S. Weinzierl, “The Unequal Mass Sunrise Integral Expressed through Iterated Integrals on $\overline{\mathcal{M}}_{1,3}$,” *Nucl. Phys. B* **954** (2020) 114991, [arXiv:1907.01251](#) [[hep-th](#)].

- [162] S. Coleman and R. Norton, “Singularities in the Physical Region,” *Il Nuovo Cimento (1955-1965)* **38** no. 1, (1965) 438–442.
- [163] J. Collins, “A New and Complete Proof of the Landau Condition for Pinch Singularities of Feynman Graphs and Other Integrals,” [arXiv:2007.04085](#) [[hep-ph](#)].
- [164] G. ’t Hooft and M. J. G. Veltman, “Combinatorics of Gauge Fields,” *Nucl. Phys.* **B50** (1972) 318–353.
- [165] S. Abreu, R. Britto, C. Duhr, and E. Gardi, “From Multiple Unitarity Cuts to the Coproduct of Feynman Integrals,” *JHEP* **10** (2014) 125, [arXiv:1401.3546](#) [[hep-th](#)].
- [166] F. Pham, “Singularités des processus de diffusion multiple,” *Annales de l’I.H.P. Physique théorique* **6** no. 2, (1967) 89–204.
- [167] J. Leray, “Le Calcul Différentiel et Intégral sur une Variété Analytique Complexe. (Problème de Cauchy. III.),” *Bulletin de la Société Mathématique de France* **87** (1959) 81–180.
- [168] S. Abreu, R. Britto, C. Duhr, and E. Gardi, “Cuts from Residues: the One-Loop Case,” *JHEP* **06** (2017) 114, [arXiv:1702.03163](#) [[hep-th](#)].
- [169] S. Bloch and D. Kreimer, “Feynman Amplitudes and Landau Singularities for 1-Loop Graphs,” *Commun. Num. Theor. Phys.* **4** (2010) 709–753, [arXiv:1007.0338](#) [[hep-th](#)].
- [170] S. Bloch and D. Kreimer, “Cutkosky Rules and Outer Space,” [arXiv:1512.01705](#) [[hep-th](#)].
- [171] H. Lehmann, K. Symanzik, and W. Zimmermann, “On the Formulation of Quantized Field Theories. II,” *Nuovo Cim.* **6** (1957) 319–333.
- [172] O. Steinmann, “Wightman-Funktionen und Retardierte Kommutatoren. II,” *Helvetica Physica Acta* **33** (1960) 347–362.
- [173] D. Ruelle, “Connection between Wightman Functions and Green Functions in p -Space,” *Il Nuovo Cimento* **19** (1961) 356–376.

- [174] H. Araki and N. Burgoyne, “Properties of the Momentum Space Analytic Function,” *Il Nuovo Cimento (1955-1965)* **18** no. 2, (1960) 342–346.
- [175] H. Araki, “Generalized Retarded Functions and Analytic Function in Momentum Space in Quantum Field Theory,” *Journal of Mathematical Physics* **2** no. 2, (1961) 163–177.
- [176] M. Lassalle, “Analyticity Properties Implied by the Many-Particle Structure of the n Point Function in General Quantum Field Theory. 1. Convolution of n Point Functions Associated with a Graph,” *Commun. Math. Phys.* **36** (1974) 185.
- [177] K. E. Cahill and H. P. Stapp, “Optical Theorems and Steinmann Relations,” *Annals Phys.* **90** (1975) 438.
- [178] L. J. Dixon, J. M. Drummond, and J. M. Henn, “Bootstrapping the Three-Loop Hexagon,” *JHEP* **11** (2011) 023, [arXiv:1108.4461 \[hep-th\]](#).
- [179] L. J. Dixon, J. M. Drummond, M. von Hippel, and J. Pennington, “Hexagon Functions and the Three-Loop Remainder Function,” *JHEP* **12** (2013) 049, [arXiv:1308.2276 \[hep-th\]](#).
- [180] L. J. Dixon, J. M. Drummond, C. Duhr, and J. Pennington, “The Four-Loop Remainder Function and Multi-Regge Behavior at NNLLA in Planar $\mathcal{N}=4$ super-Yang-Mills Theory,” *JHEP* **06** (2014) 116, [arXiv:1402.3300 \[hep-th\]](#).
- [181] J. M. Drummond, G. Papathanasiou, and M. Spradlin, “A Symbol of Uniqueness: The Cluster Bootstrap for the 3-Loop MHV Heptagon,” *JHEP* **03** (2015) 072, [arXiv:1412.3763 \[hep-th\]](#).
- [182] L. J. Dixon and M. von Hippel, “Bootstrapping an NMHV Amplitude through Three Loops,” *JHEP* **10** (2014) 065, [arXiv:1408.1505 \[hep-th\]](#).
- [183] L. J. Dixon, M. von Hippel, A. J. McLeod, and J. Trnka, “Multi-Loop Positivity of the Planar $\mathcal{N}=4$ SYM Six-Point Amplitude,” *JHEP* **02** (2017) 112, [arXiv:1611.08325 \[hep-th\]](#).
- [184] L. J. Dixon, J. Drummond, T. Harrington, A. J. McLeod, G. Papathanasiou, and M. Spradlin, “Heptagons from the Steinmann Cluster Bootstrap,” *JHEP* **02** (2017) 137, [arXiv:1612.08976 \[hep-th\]](#).

- [185] J. Drummond, J. Foster, Ö. Gürdoğan, and G. Papathanasiou, “Cluster Adjacency and the Four-Loop NMHV Heptagon,” *JHEP* **03** (2019) 087, [arXiv:1812.04640 \[hep-th\]](#).
- [186] S. Caron-Huot, L. J. Dixon, M. von Hippel, A. J. McLeod, and G. Papathanasiou, “The Double Pentaladder Integral to All Orders,” *JHEP* **07** (2018) 170, [arXiv:1806.01361 \[hep-th\]](#).
- [187] B. Basso, A. Sever, and P. Vieira, “Spacetime and Flux Tube S-Matrices at Finite Coupling for $\mathcal{N}=4$ Supersymmetric Yang-Mills Theory,” *Phys. Rev. Lett.* **111** no. 9, (2013) 091602, [arXiv:1303.1396 \[hep-th\]](#).
- [188] B. Basso, A. Sever, and P. Vieira, “Space-Time S-matrix and Flux Tube S-matrix II. Extracting and Matching Data,” *JHEP* **01** (2014) 008, [arXiv:1306.2058 \[hep-th\]](#).
- [189] A. Belitsky, “Nonsinglet Pentagons and NMHV Amplitudes,” *Nucl. Phys. B* **896** (2015) 493–554, [arXiv:1407.2853 \[hep-th\]](#).
- [190] B. Basso, A. Sever, and P. Vieira, “Hexagonal Wilson Loops in Planar $\mathcal{N}=4$ SYM Theory at Finite Coupling,” *J. Phys. A* **49** no. 41, (2016) 41LT01, [arXiv:1508.03045 \[hep-th\]](#).
- [191] Y. Li and H. X. Zhu, “Bootstrapping Rapidity Anomalous Dimensions for Transverse-Momentum Resummation,” *Phys. Rev. Lett.* **118** no. 2, (2017) 022004, [arXiv:1604.01404 \[hep-ph\]](#).
- [192] Ø. Almeland, C. Duhr, E. Gardi, A. McLeod, and C. D. White, “Bootstrapping the QCD Soft Anomalous Dimension,” *JHEP* **09** (2017) 073, [arXiv:1706.10162 \[hep-ph\]](#).
- [193] B. Basso and L. J. Dixon, “Gluing Ladder Feynman Diagrams into Fishnets,” *Phys. Rev. Lett.* **119** no. 7, (2017) 071601, [arXiv:1705.03545 \[hep-th\]](#).
- [194] J. M. Drummond, J. M. Henn, and J. Trnka, “New Differential Equations for On-Shell Loop Integrals,” *JHEP* **04** (2011) 083, [arXiv:1010.3679 \[hep-th\]](#).
- [195] V. Del Duca, L. J. Dixon, J. M. Drummond, C. Duhr, J. M. Henn, and V. A. Smirnov, “The One-Loop Six-Dimensional Hexagon Integral with Three Massive Corners,” *Phys. Rev. D* **84** (2011) 045017, [arXiv:1105.2011 \[hep-th\]](#).

- [196] J. L. Bourjaily, A. J. McLeod, M. von Hippel, and M. Wilhelm, “Rationalizing Loop Integration,” *JHEP* **08** (2018) 184, [arXiv:1805.10281 \[hep-th\]](#).
- [197] J. Henn, E. Herrmann, and J. Parra-Martinez, “Bootstrapping Two-Loop Feynman Integrals for Planar $\mathcal{N} = 4$ sYM,” *JHEP* **10** (2018) 059, [arXiv:1806.06072 \[hep-th\]](#).
- [198] A. J. McLeod, H. Munch, G. Papathanasiou, and M. von Hippel, “A Novel Algorithm for Nested Summation and Hypergeometric Expansions,” [arXiv:2005.05612 \[hep-th\]](#).
- [199] J. Bartels, L. Lipatov, and A. Sabio Vera, “BFKL Pomeron, Reggeized Gluons and Bern-Dixon-Smirnov Amplitudes,” *Phys. Rev. D* **80** (2009) 045002, [arXiv:0802.2065 \[hep-th\]](#).
- [200] H. Hannesdottir and M. D. Schwartz, “A Finite S-matrix,” [arXiv:1906.03271 \[hep-th\]](#).
- [201] A. Beilinson and P. Deligne, “Interprétation Motivique de la Conjecture de Zagier reliant Polylogarithmes et Régulateurs,” in *Motives (Seattle, WA, 1991)*, vol. 55 of *Proc. Sympos. Pure Math.*, pp. 97–121. Amer. Math. Soc., Providence, RI, 1994.
- [202] J. Zhao, *Multiple Zeta Functions, Multiple Polylogarithms and their Special Values*, vol. 12 of *Series on Number Theory and its Applications*. World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 2016.
- [203] V. A. Smirnov, “Evaluating Feynman Integrals,” *Springer Tracts Mod. Phys.* **211** (2004) 1–244.
- [204] E. Panzer, *Feynman Integrals and Hyperlogarithms*. PhD thesis, Humboldt U., 2015. [arXiv:1506.07243 \[math-ph\]](#).
- [205] G. Duplancic and B. Nizic, “Dimensionally Regulated One Loop Box Scalar Integrals with Massless Internal Lines,” *Eur. Phys. J. C* **20** (2001) 357–370, [arXiv:hep-ph/0006249](#).
- [206] H. P. Stapp, “Inclusive cross sections are discontinuities,” *Phys. Rev. D* **3** (1971) 3177–3184.

- [207] N. I. Usyukina and A. I. Davydychev, “Exact Results for Three and Four Point Ladder Diagrams with an Arbitrary Number of Rungs,” *Phys. Lett. B* **305** (1993) 136–143.
- [208] F. Chavez and C. Duhr, “Three-Mass Triangle Integrals and Single-Valued Polylogarithms,” *JHEP* **11** (2012) 114, [arXiv:1209.2722 \[hep-ph\]](#).
- [209] D. J. Broadhurst, “Summation of an Infinite Series of Ladder Diagrams,” *Phys. Lett. B* **307** (1993) 132–139.
- [210] J. Drummond, G. Korchemsky, and E. Sokatchev, “Conformal Properties of Four-Gluon Planar Amplitudes and Wilson Loops,” *Nucl. Phys. B* **795** (2008) 385–408, [arXiv:0707.0243 \[hep-th\]](#).
- [211] S. Abreu, R. Britto, C. Duhr, and E. Gardi, “Algebraic Structure of Cut Feynman Integrals and the Diagrammatic Coaction,” *Phys. Rev. Lett.* **119** no. 5, (2017) 051601, [arXiv:1703.05064 \[hep-th\]](#).
- [212] S. Abreu, R. Britto, C. Duhr, and E. Gardi, “Diagrammatic Hopf Algebra of Cut Feynman Integrals: the One-Loop Case,” *JHEP* **12** (2017) 090, [arXiv:1704.07931 \[hep-th\]](#).
- [213] S. Abreu, R. Britto, C. Duhr, E. Gardi, and J. Matthew, “Diagrammatic Coaction of Two-Loop Feynman Integrals,” in *14th International Symposium on Radiative Corrections: Application of Quantum Field Theory to Phenomenology*. 2019. [arXiv:1912.06561 \[hep-th\]](#).
- [214] S. Caron-Huot, L. J. Dixon, F. Dulat, M. Von Hippel, A. J. McLeod, and G. Papathanasiou, “The Cosmic Galois Group and Extended Steinmann Relations for Planar $\mathcal{N} = 4$ SYM Amplitudes,” *JHEP* **09** (2019) 061, [arXiv:1906.07116 \[hep-th\]](#).
- [215] N. Arkani-Hamed, P. Benincasa, and A. Postnikov, “Cosmological Polytopes and the Wavefunction of the Universe,” [arXiv:1709.02813 \[hep-th\]](#).
- [216] N. Arkani-Hamed and P. Benincasa, “On the Emergence of Lorentz Invariance and Unitarity from the Scattering Facet of Cosmological Polytopes,” [arXiv:1811.01125 \[hep-th\]](#).

- [217] A. M. Polyakov, “Gauge Fields as Rings of Glue,” *Nucl. Phys.* **B164** (1980) 171–188.
- [218] I. Y. Arefeva, “Quantum Contour Field Equations,” *Phys. Lett.* **B93** (1980) 347–353.
- [219] V. S. Dotsenko and S. N. Vergeles, “Renormalizability of Phase Factors in the Nonabelian Gauge Theory,” *Nucl. Phys.* **B169** (1980) 527.
- [220] R. A. Brandt, F. Neri, and M.-a. Sato, “Renormalization of Loop Functions for All Loops,” *Phys. Rev.* **D24** (1981) 879.
- [221] G. P. Korchemsky and A. V. Radyushkin, “Loop Space Formalism and Renormalization Group for the Infrared Asymptotics of QCD,” *Phys. Lett.* **B171** (1986) 459–467.
- [222] G. Korchemsky and A. Radyushkin, “Infrared Asymptotics of Perturbative QCD: Renormalization Properties of the Wilson Loops In Higher Orders of Perturbation Theory,” *Sov. J. Nucl. Phys.* **44** (1986) 877.
- [223] C. Duhr, “Hopf Algebras, Coproducts and Symbols: an Application to Higgs Boson Amplitudes,” *JHEP* **08** (2012) 043, [arXiv:1203.0454 \[hep-ph\]](#).
- [224] S. Caron-Huot, L. J. Dixon, J. M. Drummond, F. Dulat, J. Foster, Ö. Gürdoğan, M. von Hippel, A. J. McLeod, and G. Papathanasiou, “The Steinmann Cluster Bootstrap for $\mathcal{N}=4$ Super Yang-Mills Amplitudes,” in *19th Hellenic School and Workshops on Elementary Particle Physics and Gravity*. 2020. [arXiv:2005.06735 \[hep-th\]](#).
- [225] J. Drummond, J. Foster, and O. Gürdoğan, “Cluster Adjacency Properties of Scattering Amplitudes in $N = 4$ Supersymmetric Yang-Mills Theory,” *Phys. Rev. Lett.* **120** no. 16, (2018) 161601, [arXiv:1710.10953 \[hep-th\]](#).
- [226] A. B. Goncharov, “Multiple Polylogarithms and Mixed Tate Motives,” [arXiv:math/0103059 \[math.AG\]](#).
- [227] F. Brown, “Mixed Tate Motives Over \mathbb{Z} ,” *Ann. of Math. (2)* **175** no. 2, (2012) 949–976, [arXiv:1102.1312 \[math.AG\]](#).
- [228] S. Caron-Huot, “Superconformal Symmetry and Two-Loop Amplitudes in Planar

- $\mathcal{N}=4$ Super Yang-Mills,” *JHEP* **1112** (2011) 066, [arXiv:1105.5606 \[hep-th\]](#).
- [229] C. Duhr, “Mathematical Aspects of Scattering Amplitudes,” in *Theoretical Advanced Study Institute in Elementary Particle Physics: Journeys Through the Precision Frontier: Amplitudes for Colliders*, pp. 419–476. 2015. [arXiv:1411.7538 \[hep-ph\]](#).
- [230] F. C. Brown, “Polylogarithmes Multiples Uniformes en Une Variable,” *Comptes Rendus Mathematique* **338** no. 7, (2004) 527 – 532.
- [231] L. J. Dixon, C. Duhr, and J. Pennington, “Single-Valued Harmonic Polylogarithms and the Multi-Regge Limit,” *JHEP* **10** (2012) 074, [arXiv:1207.0186 \[hep-th\]](#).
- [232] V. Del Duca, S. Druc, J. Drummond, C. Duhr, F. Dulat, R. Marzucca, G. Papathanasiou, and B. Verbeek, “Multi-Regge Kinematics and the Moduli Space of Riemann Spheres with Marked Points,” *JHEP* **08** (2016) 152, [arXiv:1606.08807 \[hep-th\]](#).
- [233] Ø. Almeland, C. Duhr, and E. Gardi, “Three-Loop Corrections to the Soft Anomalous Dimension in Multileg Scattering,” *Phys. Rev. Lett.* **117** no. 17, (2016) 172002, [arXiv:1507.00047 \[hep-ph\]](#).
- [234] O. Schlotterer and S. Stieberger, “Motivic Multiple Zeta Values and Superstring Amplitudes,” *J. Phys. A* **46** (2013) 475401, [arXiv:1205.1516 \[hep-th\]](#).
- [235] O. Schnetz, “Graphical Functions and Single-Valued Multiple Polylogarithms,” *Commun. Num. Theor. Phys.* **08** (2014) 589–675, [arXiv:1302.6445 \[math.NT\]](#).