

ON THE INFRARED AND ULTRAVIOLET
BEHAVIOR OF SCATTERING AMPLITUDES AND
WAVEFUNCTIONS

AARON JONATHAN HILLMAN

A DISSERTATION
PRESENTED TO THE FACULTY
OF PRINCETON UNIVERSITY
IN CANDIDACY FOR THE DEGREE
OF DOCTOR OF PHILOSOPHY

RECOMMENDED FOR ACCEPTANCE
BY THE DEPARTMENT OF
PHYSICS

ADVISER: PROFESSOR NIMA ARKANI-HAMED

SEPTEMBER 2023

© Copyright by Aaron Jonathan Hillman, 2023.

All rights reserved.

Abstract

We extend the on-shell program on three fronts: finite Λ , finite ϵ , and finite α' where these are the cosmological constant, the dimensional regularization parameter $\epsilon = \frac{4-D}{2}$, and a scale for putative ultraviolet completions of tree-level amplitudes with graviton exchange (string scale) respectively. In the first case we study conformally coupled scalars in the associated de Sitter background, whose wavefunction contributions are transcendental functions of uniform weight and are closely related to the density perturbations of direct interest for studying inflation. We derive a recursive rule for the symbol of such functions and study the implications for singularities, differential equations, and identify the symbol letters with graph associahedra. Most of this work appeared in [1]. In the second case we study the structure of UV and IR divergences for scalar Feynman integrals in the context of Feynman polytopes and tropical fans, providing a unified description for both and a prescription to compute the leading divergence. In the IR case we apply this to the study of infinite classes of diagrams appearing in the expansion of the four-point amplitude in $\mathcal{N} = 4$ SYM. This is based on work with Nima Arkani-Hamed and Sebastian Mizera [2]. In the case of UV divergences we introduce u -variables which realize an associated binary geometry and facilitate a novel subtraction scheme in dimensional regularization to compute the full Laurent expansion of the associated Feynman integral. On the third front, we consider unitarity constraints on a class of tree-level UV completions of amplitudes with graviton exchange. In particular we study completions of Higgs scattering and gauge boson scattering in the Standard Model and gauge boson scattering in diverse dimensions. We find surprising constraints connecting with intuitions from the swampland. This is based on work with Brad Bachu [3].

Acknowledgements

I must thank my adviser, Nima Arkani-Hamed, for imparting the essential wisdom behind practicing physics: it is Nature that is exquisite, not you. That cleverness may be a skill but it is not a virtue. And that these lessons in humility ought not cultivate a fear of breaking with convention, but the courage to surrender and listen to Nature; and then in perhaps a few bright moments, we may earn the privilege of articulating in inky flesh one of Nature's many wonders as it truly is.

I would also like to thank Isobel Ojalvo and Ruth Britto for agreeing to be on my thesis committee and Song He for agreeing to be an external reader on this thesis.

I am grateful to my many collaborators: Sebastian Mizera, Enrico Pajer, Daniel Baumann, Austin Joyce, Hayden Lee, Guilherme Pimentel, Alexander Atanasov, David Poland, Junchen Rong, and Ning Su. And I am grateful to Alex for all the many physics conversations over the years and the tremendously good and bad ideas and observations seated and unseated within them. I have also benefited greatly from the many conversations with Institute Postdocs and fellow graduate students: Hofie Hannesdottir, Giulio Salvatori, Lorenz Eberhardt, Yiming Chen, Pranay Gorantla, Wayne Zhao, and Brad Bachu among others.

I am grateful to the Princeton Physics Department and to the Institute for Advanced Study for welcoming me. I am indebted to the administrative staff at both places, especially: Catherine Brosowsky, Lisa Fleischer, Katherine Lamos, and Audrey Smerkanich.

Thank you to all who made DBar trivia possible, and especially our stalwart physics team members: Nicholas Haubrich and Erin Healy.

Finally, thank you to my loving family. To Darleny: you always understood. And above all, to my parents: for investing in ways said and unsaid, and in total, to my formal education through to its bittersweet end.

To my parents.

Contents

Abstract	iii
Acknowledgements	iv
1 Introduction	1
1.1 Wave Functions in Cosmology	3
1.2 UV and IR Divergences	4
1.3 Stringy Completions of Standard Model Amplitudes	7
I $\Lambda \neq 0$	9
2 de Sitter Symbology	10
2.1 Introduction	10
2.2 Background	13
2.2.1 Conformally Coupled Scalars in de Sitter	13
2.2.2 Polylogarithms and Symbols	16
2.3 Statement of the Recursive Rule	19
2.3.1 Time Integral Derivation	21
2.3.2 From the Cosmological Polytope	23
2.4 Some Examples	28
2.4.1 Two-Site Chain	28
2.4.2 Two-Site Loop	29

2.4.3	Three-Site Chain	31
2.5	Integrating the Symbols	32
2.5.1	Two-Site Chain	33
2.5.2	Two-Site Loop	34
2.5.3	Three-Site Chain	34
2.5.4	Four-Site and Beyond	36
2.6	Differential Equations from the Symbol	37
2.7	Singularities from Graph Associahedra	39
2.8	Conclusion and Outlook	46
II	$\epsilon \neq 0$	48
3	The Divergence Structure of Feynman Integrals	49
3.1	Symanzik Representation	51
3.1.1	Divergences	54
3.2	Newton Polytopes and Tropical Fans	55
3.3	Calculation of Leading Divergences	61
3.3.1	Systematics of the General Case	63
3.4	Feynman Polytopes	65
3.4.1	Summary of Feynman Polytopes	70
3.4.2	Divergences	71
3.4.3	UV Divergences	71
3.4.4	IR Divergences	76
3.4.5	Beyond Fishnets	83
3.5	Conclusions and Outlook	86
4	Subtraction Schemes for Feynman Integrals	88
4.1	Introduction	88

4.2	Euler Beta Function	90
4.3	u -variables for Feynman Integrals	92
4.3.1	Triangle in u -vars	93
4.4	UV u 's	98
4.5	Integrals in u -variables	107
4.6	Conclusions and Outlook	117

III $\alpha' \neq 0$ 119

5 Stringy Completions of Standard Model Amplitudes from the Bot- tom Up 120

5.1	Introduction	120
5.2	Unitarity and UV Completion	122
5.2.1	Review of Unitarity Constraints	123
5.2.2	Completions	125
5.3	Gauge Boson Scattering	136
5.3.1	Constraints in Four Spacetime Dimensions	137
5.3.2	Constraints in General Spacetime Dimensions	143
5.4	Standard Model Electroweak Sector	151
5.5	Conclusions Outlook	153

Chapter 1

Introduction

Our understanding of quantum field theory (QFT) is in flux. Starved of frequent encounter with experiment and spurred to pursue the still unanswered questions dropped on our plates at the end of the 20th century, recent decades have given occasion to study the structure of quantum field theory itself more closely. Different research programs have reacted to this state of affairs in different ways. One approach, motivated in part by the unfinished successes of the 60's S -matrix program (and reinvigorated with theoretical data and a different conceptual point of view), is to anchor ourselves to a scrupulous study of the observables and their intrinsic if often surprising properties. This is also motivated by thought experiments articulating that the only sharp observables in a theory of quantum gravity are boundary observables. But these a priori motivations are perhaps less compelling than the successes of this approach. Our current conception of quantum field theory attempts to hardwire, as much as is possible, Lorentz invariance, locality, and unitarity. These conditions are in tension and this tension leads to the great rigidity we find in the structure of theories consistent with these principles. But observables such as scattering amplitudes often exhibit further properties obfuscated by the attempt to manifest these physical principles at the outset. A classic and simple example which illustrates the

point are the surprising asymptotic scaling properties of tree-level gauge and gravity amplitudes at large complex momenta which allow BCFW recursion [4, 5]. The two complementary aspects of this quiet revolution in the study of scattering amplitudes are a commitment to finding the language which is most practical to calculate amplitudes in, and an openness to finding new representations for amplitudes in which we do not attempt to manifest locality and unitarity to the exclusion of all other nontrivial properties, but find a description in which all non-trivial properties have some natural, if not entirely manifest, significance. Though the latter is perhaps less responsible than the former, the two aspects of this approach are ultimately sides of the same coin, and we attempt to leverage both in the work contained herein. One might call the latter the principle of least magic¹. We do not want the apparent magic in our physics to be highly distributed throughout our description: we want it distilled, to the fewest essential, irreducible objects which all nontrivial statements descend from.

This thesis presents work advancing the on-shell program beyond statements about rational functions (tree-level amplitudes and loop integrands in QFT) on three fronts: wave functions in cosmology, calculation of ultraviolet (UV) and infrared (IR) divergences, and unitarity constraints on stringy amplitudes. We therefore label the three parts of this thesis by the three constant parameters which effect this deformation away from flat-space rational functions: the cosmological constant Λ , the deformation of the spacetime dimension ϵ , and the Regge slope α' .

¹Like the more correct “principle of stationary action”, we could elaborate on this and call it the principle of concentrated magic

1.1 Wave Functions in Cosmology

In the inflationary paradigm, density perturbations due to fluctuations in a scalar field ϕ the inflaton, whose vacuum energy sources an exponential expansion² with metric

$$ds^2 = -dt^2 + e^{2Ht} d\vec{x}^2 \quad (1.1)$$

where we have taken the spatial slices to be flat. This is a de Sitter metric and the Hubble constant H is proportional to the square root of the cosmological constant Λ sourced by the inflaton. There is a conformal symmetry associated with this metric which will act on boundary observables, and for perfect de Sitter the boundary correlation functions must respect this conformal symmetry. Indeed, the cosmic microwave background is consistent with this implied scale invariance up to deviations consistent with a mild tilt to the potential sourcing the exponential expansion.

As is well known, the temperature variation in the cosmic microwave background (CMB) is uniform to one part in ten thousand. This remarkable homogeneity is indeed a reason for invoking inflation: without inflation, incident radiation greater than one degree of separation on the celestial sphere should never have been in causal contact at the conventional big bang. But this approximate uniformity also means the interactions are very weakly coupled, and only provide a weakly coupled window on the early universe, with higher point functions heavily suppressed.

Despite this, calculations in cosmology are at a minimum of great theoretical importance. The most highfalutin motivation is that any conception of holography in this arena must have a notion of emergent time at a future spacelike boundary. This is a tremendous challenge and opportunity. But a more basic question begging consideration is how bulk unitarity is encoded in observables, even in perturbation theory.

²This expansion is approximately exponential. In reality it cannot be eternal, so the potential has a slight tilt, such that the appropriate number of rescalings of the metric (e-foldings) occurs, and then drops off. A good first approximation is exponential.

In flat space, we understand how to at least check unitarity of the S -matrix, but in cosmology the state of affairs is much less clear, and we are in fact motivated to consider generalizations of the notions familiar in flat space which do not amount to merely imposing $S^\dagger S = 1$. The most obvious of which is consistent factorization of the wave function at singularities.

We employ this approach in chapter 2, deriving a recursive rule for the symbol of perturbative wave function contributions for conformally coupled scalars in de Sitter space. This rule manifests the de Sitter analogue of consistent factorization at singularities and therefore unitarity. Though the symbol is constructed recursively from the discontinuities, we then observe what is implied for the differential equations of the wave function. Differential equations in kinematic space encode time evolution, and in this way we see the complementary physics of unitarity and time evolution encoded in this object.

Finally, we demonstrate that the set of multiplicatively independent symbol letters (and therefore singularities) of the symbol are described by a graph associahedron for a graph obtained in a simple way. The symbol letters will therefore obey the u -equations of the associated graph associahedron. We discuss what implications this might have for the symbol. This sets up an opportunity for a positive geometry to capture the complementary physics of unitarity and time evolution as they are encoded in the symbol of de Sitter wave functions, posing an exciting challenge for future work.

1.2 UV and IR Divergences

As on-shell thinking has progressed over the last few years, a lingering question has always been whether the new approaches could bring anything to bear on our understanding of the renormalization group, or indeed even accommodate the existing

one. This is perhaps the arena in which off-shell, fields-focused thinking in Euclidean space shines most. The work of Wilson, which was later further articulated beautifully by Polchinski [6], demonstrates both why there should exist a β function, the effect of which is to resum logarithms, and why the effects of irrelevant couplings can be ignored at low energies. The slogan for the essential insight here is that we should understand the effect of changing a UV cutoff scale-by-scale, not graph-by-graph in a perturbative expansion. These are great conceptual triumphs, but nonetheless they are primarily conceptual, and for that reason our understanding of UV divergences continues to exist in a very strange state of affairs.

When calculating a β function in practice, one does so graph-by-graph in a perturbative expansion. Often in the presentation of the subject confusing discussions of schemes and the assertion of renormalization scale independence of the amplitude serve to obfuscate the underlying remarkable facts about the perturbative expansion which are equivalent to the existence of a β -function. For example, the statement of the existence of a beta function at one-loop in dimensional regularization in $\lambda\phi^4$ theory is the following statement about the leading divergence at all loop order

$$(\mu^2)^\epsilon \mathcal{A}^{\text{leading}} = \lambda - (\mu^2)^{-\epsilon} \beta \frac{\lambda^2}{\epsilon} + (\mu^2)^{-2\epsilon} \beta^2 \frac{\lambda^3}{\epsilon^2} + \dots \quad (1.2)$$

that is, the leading divergences sum to

$$(\mu^2)^\epsilon \mathcal{A}^{\text{leading}} = \frac{\lambda}{1 + (\mu^2)^{-\epsilon} \beta \frac{\lambda}{\epsilon}} \quad (1.3)$$

This is the non-trivial fact “at leading log”. Once we have committed to dimensional regularization in $D = 4 - 2\epsilon$ dimensions with scale μ , there are no conditions we impose, only observations we make upon calculating. We should indeed observe all the nontrivial and remarkable predictions of Wilson by suitably organizing the perturbative expansion.

This is the opportunity for the on-shell approach in understanding UV divergences. It also appears not to be a luxury but a necessity for the study of the evermore complex IR divergences. These divergences, in contradistinction to UV divergences, are sensitive to the external kinematics. Moreover, there is nothing like the Wilsonian conceptual infrastructure to systematically guide our thinking as to what remarkable properties the perturbative expansion might exhibit. While there is an eclectic smorgasbord of physical intuition, the formulas we have are typically gleaned by heavy lifting in the perturbative expansion.

For this reason, we pursue in chapter 3 a systematic classification of the UV and IR divergence structure of Feynman integrals for a general kinematic regime, and a simple recipe to calculate the leading divergence of the associated Feynman integrals. The natural mathematical language for this work is Newton polytopes and tropical fans, which we motivate and describe.

We present an expression for the leading divergence of $M \times N$ fishnets in terms of multi-dimensional Catalan numbers, which is calculated via a novel expression for those Catalan numbers involving the hook-length formula. We then notice that in the case of power divergent graphs which are softened by a numerator in the $\mathcal{N} = 4$ SYM expansion, the fishnet calculation motivates a conjecture for calculating the leading divergence of the $\mathcal{N} = 4$ graphs which agrees with the four loop result.

In chapter 4 we move on to consider the systematic calculation of the Laurent expansion of a UV divergent Feynman integral with subdivergences. We are not aware of any general algorithm in the literature for the case of logarithmically divergent graphs with logarithmic subdivergences to extract the Laurent expansion in dimensional regularization. The method we introduce is completely canonical, and manifests the fact that the subtraction terms are products of lower loop integrals up to $\mathcal{O}(\epsilon)$ power deformations of the integrand. In this way, we can readily extract the Laurent expansion using only numerical integration. Furthermore, with the scheme maximally

leveraging the factorization properties of the Feynman integral, we hope it will constitute substantive progress toward understanding the renormalization group from an on-shell perspective.

1.3 Stringy Completions of Standard Model Amplitudes

In the final chapter, we consider unitarity constraints on UV complete Ansätze for tree-level amplitudes with graviton exchange. Graviton exchange produces unitarity violation at energies of order the Planck scale and requires UV completion. At its broadest level, this analysis attempts to make progress on constraining UV completions of the standard model given the tension that the only known UV completions of amplitudes come from string theory, but no vacuum resembling ours has been constructed in string theory from the top down. The way we choose to do this essentially just to use the Virasoro-Shapiro prefactor

$$\Gamma^{\text{str}} = -\frac{\Gamma(-\alpha's)\Gamma(-\alpha't)\Gamma(-\alpha'u)}{\Gamma(\alpha's)\Gamma(\alpha't)\Gamma(\alpha'u)} \quad (1.4)$$

to dress the field theory amplitudes. We motivate this Ansatz from the bottom up as well, and recent work has further demonstrated how rigid the constraints on such an Ansatz are [7]. Unitarity is then imposed as positivity of the partial wave expansion.

After presenting the unitarity constraints for gauge-boson scattering in the standard model, we move on to analyze gauge boson scattering in general dimensions, and find that both ten-dimensional heterotic string theories are at the boundary of the regions allowed by unitarity when the coupling and rank of the gauge-group are allowed to vary. We also find consistency with some swampland conjectures which were deduced

in the context of open strings. This indicates some universality which can hopefully be probed beyond the case of these restricted Ansätze .

Part I

$$\Lambda \neq 0$$

Chapter 2

de Sitter Symbology

2.1 Introduction

The flat-space S-matrix is tightly constrained by locality and unitarity. We understand how these constraints are manifest at tree-level and one-loop-level, though the constraints on the space of functions that might appear at two loops or more are still not known. Additionally, recent years have exposed the rich mathematical structures behind scattering amplitudes, such as the Grassmanian, Amplituhedron, Associahedron, and cluster algebras to name a few. These mathematical frameworks have furnished representations of scattering amplitudes as answers to elegant questions purely in kinematic space. In this context, Lorentz invariance and unitarity are derivative notions, emerging from the abstract conditions imposed in this space of kinematic data at infinity, such as positivity. The situation in cosmology is vastly underdeveloped by comparison. We have no significant constraints on the vacuum wave function beyond normalizability, and the stock of theoretical data is relatively meager. Moreover, the need for such a boundary description is more urgent, as our only windows on the early universe are spatial correlations in the late universe. We have no choice but to determine how causal evolution in the bulk has been encoded

in the future boundary, how time emerges in the kinematic space at infinity.

The study of cosmological polytopes, positive geometries in the kinematic space of spatial momenta at infinity, was initiated in [8], taking a necessary step in this direction. There, perturbative contributions to the vacuum wave function are the canonical forms of certain polytopes in the space of external and edge energies of Feynman graphs. In this setting, causal evolution and unitarity emerge as factorization properties of the underlying geometry [9]. A systematic cosmological bootstrap program exploiting the conformal symmetry of cosmological correlators in de Sitter was also initiated recently [10]. In this context, consistent time evolution in the bulk is encoded in differential equations obeyed by correlators in kinematic space. In both cases, the signatures of time-dependent physics emerge in a space with no explicit reference to time. The cosmological bootstrap also demonstrated the particular importance of correlators of conformally-coupled scalars as they can be used as seeds from which the correlators for externally massless fields and other mass and spin exchanges can be deduced by acting with differential operators. This further motivates a careful study of the conformally coupled case for reasons beyond its relative simplicity. We take up the task of studying such correlators in this work. Part of the story of the flurry of progress in understanding scattering amplitudes was the discovery of efficient recursion relations, such as BCFW recursion, which facilitate the production of a wealth of theoretical data to be studied. We take a modest step toward recursion relations for cosmological correlators, presenting a recursive rule for the symbol perturbative contributions to the vacuum wave function of conformally coupled scalars.

A somewhat heavy-handed formula for the symbol of perturbative contributions to the wave function was presented in [8] employing general methods which hold for general Aomoto polylogarithms built out of general pairs of polytopes. This picture, though general and providing a beautiful interpretation of the symbol as computed

via “walks” through the cosmological polytope, obfuscates the physics that is made manifest by the recursive rule. In particular, as is familiar from the S-matrix, the locations of singularities and factorization properties of the answer on singularities is how causality and unitarity are encoded in the answer. The situation is no different in cosmology, though we already encounter branch cuts at tree-level.

This chapter is organized as follows. In Section 2.2 we review the model under consideration and provide a brief review of symbol calculus. In Section 2.3 we state the recursive rule for the symbol of perturbative wave function contributions and provide a derivation directly in terms of time integrals as well as illustrating the interpretation in terms of the cosmological polytope. We illustrate the conceptual clarity and efficiency of the rule in Section 2.4 by computing the symbols for the two-site chain (Fig. 2.1), two-site loop, and the three-site chain. In Section 2.5 we fully integrate these examples reproducing the result of [11] at four-points and providing a novel result for the five-point tree-level contribution and the bubble integrand.

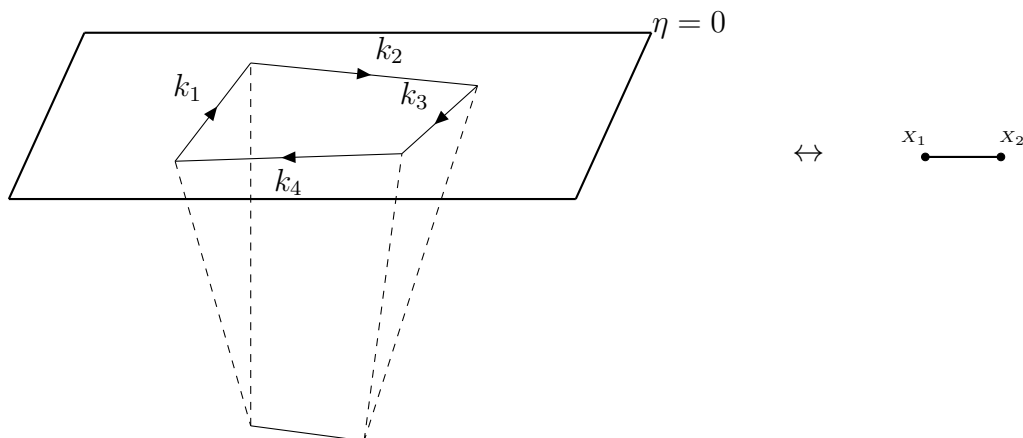


Figure 2.1: The dumbbell is a Feynman diagram where we have truncated the particles propagating to the boundary (dashed lines). The vertex energies are $X_1 = k_1 + k_2$ and $X_2 = k_3 + k_4$ and the solid edge is the propagation of an intermediate fluctuation.

2.2 Background

2.2.1 Conformally Coupled Scalars in de Sitter

We consider a conformally coupled scalar in a general FRW background as studied in [8, 9, 12]. The action is

$$S = \int d\eta d^d x \sqrt{-g} \left(\frac{1}{2} g^{\mu\nu} \partial_\mu \phi \partial_\nu \phi + \xi R \phi^2 + \sum_k \frac{\lambda_k}{k!} \phi^k \right) \quad (2.1)$$

Where

$$\xi = \frac{d-1}{4d} \quad ds^2 = a(\eta)^2 (d\eta^2 - dx^2) \quad (2.2)$$

and we use conformal time $\eta \in (-\infty, 0]$. This action is conformally equivalent to the following flat-space action

$$S[\phi] = \int d\eta d^d x \frac{1}{2} (\partial\phi)^2 + \sum_k \frac{\lambda_k(\eta)}{k!} \phi^k \quad (2.3)$$

where we now have time-dependent couplings given by

$$\lambda_k(\eta) = \lambda_k [a(\eta)]^{2 + \frac{d-1}{2}(2-k)} \quad (2.4)$$

We are interested in computing the vacuum wave-function given by

$$\Psi_{BD}[\Phi] = \int_{\phi(-\infty(1-i\epsilon))=0}^{\phi(0)=\Phi} \mathcal{D}\phi e^{iS[\phi]} \equiv \exp \left[i \sum_n \frac{1}{n!} \int \prod_i d^d k_i \Phi(k_i) \psi_n(\{k_i\}) \right] \quad (2.5)$$

Where the $i\epsilon$ prescription for η projects out the Bunch-Davies vacuum and the ψ_n are the wave function contributions that we are interested in computing in momentum space. These are n -point functions computed by the ordinary Feynman rules of the

theory. Here we mean spatial momentum space, as the times of interaction vertices are integrated over. Each ψ_n contains an overall momentum conserving δ -function from spatial translation invariance at the future boundary and we will be computing these “stripped” ψ_n ’s. We split the field into a classical momentum-space mode function and a fluctuation

$$\phi(\vec{p}, \eta) = \Phi(\vec{p})e^{iE_p\eta} + \delta\phi(\vec{p}, \eta) \quad (2.6)$$

We can expand the action and perform the path integral over fluctuations. We need the bulk-to-bulk propagator for fluctuations and the bulk to boundary propagator given by the classical solution

$$G_{\text{BB}}(\eta_v, \eta_{v'}) = \frac{1}{2E_e} \left(e^{-iE_e(\eta_v - \eta_{v'})} \theta(\eta_v - \eta_{v'}) + e^{-iE_e(\eta_{v'} - \eta_v)} \theta(\eta_{v'} - \eta_v) - e^{iE_e(\eta_v + \eta_{v'})} \right) \quad (2.7)$$

$$G_{\text{B}\partial\text{B}}(\eta) = e^{iE\eta} \quad (2.8)$$

This is the propagator for fluctuations is the Feynman propagator with an additional term ensuring that the fluctuations vanish at $\eta = 0$, consistent with the boundary condition in the path integral. The subscripts denote the vertices the propagator connects. We can use ordinary Feynman rules to compute correlators with final state vertices at $\eta = 0$. The result of this work applies exactly to the wave-function computed for couplings obeying

$$[a(\eta)]^{2+\frac{d-1}{2}(2-k)} = \frac{1}{\eta} \quad (2.9)$$

So that $\lambda_k(\eta) = \frac{\lambda_k}{\eta}$. We can pass to Fourier space most easily in this case, as

$$\lambda_k(\eta) = i \int_0^{\infty(1-i\epsilon)} e^{i\varepsilon\eta} d\varepsilon \quad (2.10)$$

Then the full wave function contributions take the form

$$F_{\mathcal{G}}(\{X_v\}, \{Y_e\}) = \int_{\mathcal{R}} \prod_{v \in \mathcal{V}} dx_v \psi_{\mathcal{G}}(\{x_v\}, \{Y_e\}) \quad (2.11)$$

where \mathcal{R} is the region defined by $x_v > X_v$ for each vertex, with X_v the sum external energies going out of a vertex. The $\psi_{\mathcal{G}}$ are the rational functions that can be computed using the tools in [8]. The sets in the argument of $\psi_{\mathcal{G}}$ are the shifted, unintegrated vertex energies in Fourier space and the Y 's are the fixed edge energies. For couplings and cosmologies consistent with Eq. 2.9, we know that, defining

$$F_{\mathcal{G}}(\{X_v\}, \{Y_e\}) = \left(\prod_{i=1}^n (2Y_i) \right) \tilde{F}_{\mathcal{G}}(\{X_v\}, \{Y_e\})$$

the function \tilde{F} will be a pure polylogarithm and so the \tilde{F} are what we study in this class of cosmologies and couplings. One concrete model in this class, $\lambda\phi^3$ theory in dS_4 , is of particular interest as it is a starting point relevant for understanding the structure of inflationary correlators [11, 10]. It is also worth noting at the onset that our results are formulated in the Lorentzian language of Ψ_{BD} but are just as good for Witten diagrams contributing to Z_{EAdS} , the partition function in Euclidean AdS. In particular, the recursive rule for contributions to the (partition)wave function as presented in this chapter can be ported to Euclidean AdS. This relation is discussed in greater detail in [13].

2.2.2 Polylogarithms and Symbols

Here we review the necessary aspects symbol calculus and multiple polylogarithms necessary for this work. A complete review of these topics and more can be found in [14]. The symbol map for polylogarithms has multiple definitions. Here we employ a definition considered in [15, 14]. If we have a complex-valued transcendental function $F^{(n)}$ of pure weight n which has a differential obeying the following

$$dF^{(n)} = \sum_i F_i^{(n-1)} d \log R_i \quad (2.12)$$

where the $F_i^{(n-1)}$ are transcendental functions of weight one lower and the R_i are rational functions, the symbol of the function is then defined recursively via

$$\mathcal{S}(F^{(n)}) = \sum_i \mathcal{S}(F_i^{(n-1)}) \otimes R_i \quad (2.13)$$

we see that the symbol inherits its linearity properties from the logarithm i.e.

$$a \otimes b - a \otimes c = a \otimes \frac{b}{c} \quad (2.14)$$

Moreover, the symbol of a function that is the product of two transcendental functions $F^{(n)}$ and $G^{(m)}$ of weights n and m respectively, has weight $n + m$ and its symbol is the “shuffle product” of the symbols of F and G , given by the sum of all ways of interleaving the symbol words of each, where this means interlacing the symbol entries between a pair of words while maintaining the internal order of each symbol. That is

$$a_1 \otimes \cdots \otimes a_n \sqcup a_{n+1} \otimes \cdots \otimes a_m = \sum_{\sigma \in \Sigma(n,m)} a_{\sigma(1)} \otimes \cdots \otimes a_{\sigma(m)}$$

where we introduced the notation \sqcup for the shuffle product and $\Sigma(n, m)$ is the set of all shuffles on a symbol word of length n and a word of length m , i.e. the subset of the S_{n+m} defined by

$$\Sigma(n, m) = \{ \sigma \in S_{n+m} \mid \sigma^{-1}(1) < \dots < \sigma^{-1}(n) \ \& \ \sigma^{-1}(n+1) < \dots < \sigma^{-1}(m) \}$$

As an example

$$\begin{aligned} a \otimes b \otimes c \sqcup e \otimes f &= a \otimes b \otimes c \otimes e \otimes f + a \otimes b \otimes e \otimes c \otimes f + a \otimes e \otimes b \otimes c \otimes f \\ &\quad e \otimes a \otimes b \otimes c \otimes f + e \otimes a \otimes b \otimes f \otimes c + e \otimes a \otimes f \otimes b \otimes c \\ &\quad e \otimes f \otimes a \otimes b \otimes c \end{aligned}$$

This is trivially extended to sums of words by linearity.

The classical polylogarithms have the following recursive definition as iterated integrals

$$\text{Li}_1(z) = -\log(1-z) \tag{2.15}$$

$$\text{Li}_n(z) = \int_0^z \text{Li}_{n-1}(t) d\log(t) \tag{2.16}$$

Therefore we see that the symbol of a classical polylogarithm is

$$\mathcal{S}(\text{Li}_n(z)) = -(1-z) \otimes z \cdots \otimes z \tag{2.17}$$

The classical polylogarithms each have a branch cut on the real axis from 1 to infinity. The discontinuity across this cut is readily obtained from the symbol. We see in the

case of classical polylogarithms that

$$\text{Disc}[\text{Li}_n(z)] = \int_1^z \text{Disc}[\text{Li}_{n-1}(t)] d \log(t) \quad (2.18)$$

Which from the discontinuity of the logarithm tells us that

$$\text{Disc}[\text{Li}_n(z)] = \frac{2\pi i}{n!} \log^n z \quad (2.19)$$

This is exactly $2\pi i$ times the function corresponding to the symbol we get from deleting the first entry of the symbol for the classical polylogarithm, the branch point being where the leading argument equals 0, in this case $z = 1$. This is a general phenomenon. Read from left to right, the symbol gives you the sequence of discontinuities of the function. Additionally, the differential indicates that read from right to left, the symbol gives you the differential equations obeyed by the function.

Finally we mention the integrability condition for symbols which is that for a symbol

$$S = \sum_{I=(i_1, \dots, i_n)} c_I a_{i_1} \otimes \cdots \otimes a_{i_n}$$

We must have

$$\sum_{I=(i_1, \dots, i_n)} c_I a_{i_1} \otimes \cdots \otimes a_{i_{p-1}} \otimes a_{i_{p+2}} \otimes a_{i_n} d \log a_{i_p} \wedge d \log a_{i_{p+1}} = 0$$

For every p , where we are replacing a symbol word by the dlog of adjacent entries wedged times the symbol with the entries deleted. This condition is non-trivial, as it is possible to write down symbols which do not obey the integrability condition and do not lift to a function. We also note here that the symbol only fixes the piece of the function of highest transcendental weight. It does not “see” the terms

of lower transcendental weight times transcendental numbers (still summing up to transcendental weight of the full function). For our purposes, these terms must be fixed by boundary conditions.

In order to illustrate the lucidity offered by the symbol, we will now prove the Abel dilogarithm identity.

$$\mathcal{S}\left(\text{Li}_2\left(\frac{wz}{(1-w)(1-z)}\right)\right) = -\frac{1-w-z}{(1-w)(1-z)} \otimes \frac{wz}{(1-w)(1-z)}$$

We can use the linearity properties of the symbol to separate this into

$$\begin{aligned} \mathcal{S}\left(\text{Li}_2\left(\frac{wz}{(1-w)(1-z)}\right)\right) &= -\frac{1-w-z}{1-z} \otimes \frac{w}{1-z} - \frac{1-z-w}{1-w} \otimes \frac{z}{1-w} - (1-w) \otimes w \\ &\quad - (1-z) \otimes z - (1-w) \otimes (1-z) - \otimes(1-z) \otimes (1-w) \end{aligned}$$

We clearly recognize these pieces as the symbol giving us the equality

$$\text{Li}_2\left(\frac{wz}{(1-w)(1-z)}\right) = \text{Li}_2\left(\frac{z}{1-w}\right) + \text{Li}_2\left(\frac{w}{1-z}\right) - \text{Li}_2(w) - \text{Li}_2(z) - \text{Li}_1(w)\text{Li}_1(z) + C$$

Where the last term allows the possibility of functions of lower transcendentality times transcendental numbers. At weight two we would only get constants, so it remains to fix the constant. Plugging in $w = 0$ we see that $C = 0$ and so the equality holds exactly where it is well defined i.e. $|w|, |z| < 1$.

2.3 Statement of the Recursive Rule

A graph \mathcal{G} corresponds to some contribution to the perturbative wave function, which is some function $F_{\mathcal{G}}$ of the final state energies. The discontinuities of this function occur when the positive sums of energies and cut edges corresponding to subgraphs of \mathcal{G} become negative. We denote such an energy corresponding to some subgraph \mathbf{g}

as $E_{\mathfrak{g}}^{tot} = \sum_{v \in \mathcal{V}} x_v + \sum_{e_c \in \mathcal{E}^c} y_{e_c}$ where e_c denotes an edge cut by circling the subgraph. We diagrammatically denote the discontinuity of $F_{\mathcal{G}}$ corresponding to the subgraph as

$$\text{Disc}_{\mathfrak{g}}[F_{\mathcal{G}}] = \text{Diagram} \quad (2.20)$$

With the dashed circle enclosing the subgraph corresponding to the discontinuity we are taking. Each complementary subgraph $\bar{\mathfrak{g}}_i$ has some set of vertices \mathcal{V}_i^c which has cut edges, the edges which connect to \mathfrak{g} , ending on them. We additionally denote this set of cut edges connecting to the complementary subgraph as \mathcal{E}_i^c . We will demonstrate that the discontinuity obeys the equality

$$\begin{aligned} \text{Disc}_{\mathfrak{g}}[F_{\mathcal{G}}(\{X_v\}, \{Y_e\})] &= \text{Disc}_{\mathfrak{g}}[F_{\mathfrak{g}}(\{X_v\}, \{X_{v_c} + Y_{e_c}\}, \{Y_e\})] \times \\ &\prod_i \sum_{\{\sigma_{e_c}\}} \left(\prod_{e_c} \sigma_{e_c} \right) F_{\bar{\mathfrak{g}}_i}(\{X_v\}, \{X_{v_c} + \sigma_{e_c} Y_{e_c}\}, \{Y_e\}) \end{aligned} \quad (2.21)$$

The set $\{X_{v_c}\}$ for any subgraph is the set of vertex energies corresponding to vertices with cut edges ending on them and the Y_{e_c} are the corresponding edge energies. The σ_{e_c} 's can take on values ± 1 and the sum is over all such combinations, with the product giving the a necessary sign depending on the number of negative σ 's. Because the discontinuities are produced by poles in the energy integrand, the sum of all the discontinuities vanishes by the residue theorem. This means we can say that the total energy discontinuity is minus the sum of all other discontinuities. This fact and the previous equation combine to produce a recursive rule for the symbol of any wave-function contribution. The discontinuity relation immediately translates to an equation for the symbol. Multiplication at the level of the integrated functions becomes a shuffle product at the level of the symbol and we tac on the appropriate

energy sum as the leading symbol entry corresponding to the discontinuity. That is,

$$\begin{aligned} \mathcal{S}(\tilde{F}_{\mathcal{G}}) &= \sum_{\mathfrak{g}} \frac{E_{\mathfrak{g}}^{tot}}{\left(\sum_v X_v\right)} \otimes \mathcal{S}(\text{Disc}_{\mathfrak{g}}[\tilde{F}_{\mathfrak{g}}(\{X_v\}, \{X_{v_c} + Y_{e_c}\}, \{Y_e\})]) \\ &\quad \prod_i \sqcup \sum_{\{\sigma_{e_c}\}} \left(\prod_{e_c} \sigma_{e_c} \right) \mathcal{S}(\tilde{F}_{\tilde{\mathfrak{g}}_i}(\{X_v\}, \{X_{v_c} + \sigma_{e_c} Y_{e_c}\}, \{Y_e\})) \end{aligned} \quad (2.22)$$

Already we see that the discontinuity structure is the avatar of unitarity, providing a consistency condition for the factorization of discontinuities of purely boundary data. Moreover, this rule reveals that the symbol is the object that can be systematically and uniquely built up from its most elementary building block, the three-particle symbol, X , and consistent factorization of the discontinuities. We illustrate simplicity and practicality of this recursive rule in Section 4.

2.3.1 Time Integral Derivation

Since the discontinuities are produced by poles in the integrand, it suffices to show that the energy integrand contributing to the perturbative wave function ψ_n factorizes appropriately on the residue of each of its poles. The poles of ψ correspond to the possible subgraphs of the diagram and are equal to the sum of the vertex energies plus the energies of edges running out of the subgraph. That is, we are interested in its residue as $E_{\mathfrak{g}}^{tot}$ goes to zero. This pole is generated by

$$\psi \sim \int_{-\infty}^0 d\eta_l e^{iE_{\mathfrak{g}}^{tot}\eta_l}$$

where η_l is the latest time within \mathfrak{g} for each contribution. As $E_{\mathfrak{g}}^{tot}$ goes to zero the integral is dominated by the region where $\eta_l \rightarrow -\infty$. As such, when taking the residue at this pole, the latest time (and therefore all times) contained in the subgraph are in the past of all other vertices, and so all θ functions in the propagators inconsistent

with this condition lose support in this region. We can express such a contribution generically as

$$\psi_{\mathfrak{g}} = \prod_{\bar{\mathfrak{g}}_i} \left(\int_{-\infty}^0 \prod_{v \in \mathcal{V}_i} d\eta_v e^{ix_v \eta_v} \prod_{e \in \mathcal{E}_i} G(\eta_{v_e}, \eta_{v'_e}, y_e) \right) \times \int_{-\infty}^0 \prod_{e_c \in \mathcal{E}_c} G_c(\eta_{e_c}^{out}, \eta_{e_c}^{in}, y_{e_c}) \left(\prod_{v \in \mathcal{V}_{\mathfrak{g}}} d\eta_v e^{ix_v \eta_v} \prod_{e \in \mathcal{E}_i} G(\eta_{v_e}, \eta_{v'_e}, y_e) \right) \quad (2.23)$$

The first piece is a product over all disconnected subgraphs, with all the vertices and edges internal to the graph contained in the parentheses. The right parenthetical contains the same but for the subgraph corresponding to the discontinuity we are taking. Between the two parentheticals we have the product over cut edges e_c leading out from the subgraph \mathfrak{g} . Because $\theta(\eta_{e_c}^{in} - \eta_{e_c}^{out})$ has no support and the other θ function is satisfied we replace the propagator with the “cut” propagator

$$G_c(\eta_{e_c}^{out}, \eta_{e_c}^{in}, y_{e_c}) = \frac{1}{2y_{e_c}} \left(e^{-iy_{e_c} \eta_{e_c}^{out}} - e^{iy_{e_c} \eta_{e_c}^{out}} \right) e^{iy_{e_c} \eta_{e_c}^{in}} \quad (2.24)$$

Where we omit the implicit theta function in the first term since it is already guaranteed when we are at the pole. It is worth pausing to note the exception to this which is possible with loops, where an edge leads out and back into the subgraph and so $\eta_{e_c}^{in} \equiv \eta_{e_c}^{out}$. In this case, neither θ function has support, and the cut propagator only has the last term. This is explicitly applied to computing the one-loop symbol in Section 3.

We can now take the residue. It is known that the residue of the total energy pole gives the flat-space scattering amplitude [11, 10, 8], and this is the factor that remains from the subgraph corresponding to the residue. Focusing on the external

piece we are left with

$$\text{Res}_{E_g^{tot} \rightarrow 0} \psi_g^{ext} \propto \prod_{\bar{g}_i} \left(\int_{-\infty}^0 \prod_{v \in \mathcal{V}_i / \mathcal{V}_i^c} d\eta_v e^{ix_v \eta_v} \prod_{v \in \mathcal{V}_i^c} d\eta_v (e^{i(x_v - y_{e_c})\eta_v} - e^{i(x_v + y_{e_c})\eta_v}) \prod_{e \in \mathcal{E}_i} G(\eta_{v_e}, \eta_{v'_e}, y_e) \right) \quad (2.25)$$

where \mathcal{V}_i^c are the vertices in subgraph \bar{g}_i that have a cut edge ending on them. The above is merely the the sum over combinations of absorbing positive or negative edge energy for each vertex with a cut edge ending on it, with a product then taken over subgraphs. All the time-integrals between subgraphs are totally decoupled and so the full expression for the residue is

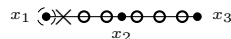
$$\text{Res}_{E_g^{tot} \rightarrow 0} \psi_g = 2\pi i \mathcal{A}[\bar{g}] \times \prod_{\bar{g}_i} \sum_{\{\sigma_{e_c}\}} \left(\prod \sigma_{e_c} \right) \psi_{\bar{g}_i}(\{x_v\}, \{x_{v_c} + \sigma_{e_c} y_{e_c}\}, \{y_e\}) \quad (2.26)$$

Where $\mathcal{A}[g]$ denotes the flat space scattering amplitude corresponding to the subgraph \bar{g} and we have distinguished between vertices v_c and v which do and do not have cut edges ending on them respectively. The σ_{e_c} 's are plus or minus one and their product gives the appropriate sign depending on whether there is an even or odd number of edges absorbing positive energy. Because the residue of the integrand is fully factorized, with no energy dependence mixing between subgraphs, the energy integral is just a product over the integrals for each subgraph and so this factorization trivially lifts to the level of the full wave function contribution, with the residue at the pole now corresponding to the discontinuity across a branch cut.

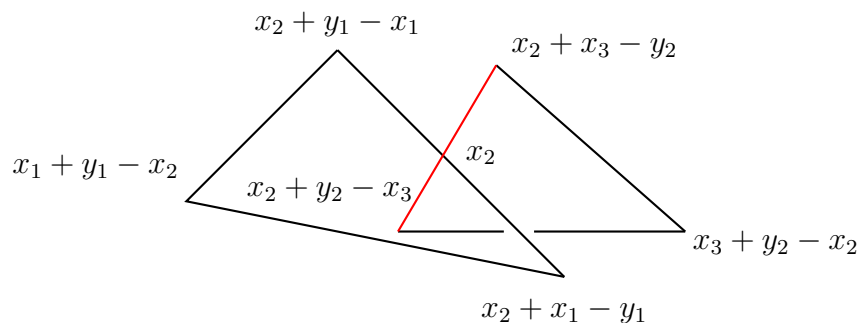
2.3.2 From the Cosmological Polytope

We can also understand the origin of the recursive rule from the perspective of the cosmological polytope, where the ingredients are factorization and triangulation. We will illustrate the rule utilizing two discontinuities of the three-site chain as the facets

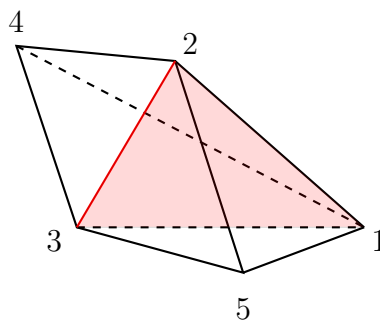
of corresponding polytope can be visualized in three dimensions. We use the formalism laid out in the latter half of [8]. The cosmological polytope corresponding to this graph is the “double-square pyramid” which has six vertices and lives in \mathbb{P}^4 . We first examine the discontinuity associated with $X_1 + Y_1$. The corresponding facet of the cosmological polytope has the five vertices indicated by the marking



We can visualize this facet in three-dimensional space as it's constructed out of intersecting triangles



The vertex opposite the red line at $x_1 + y_1 - x_2$ moves off in a fourth dimension and is the vertex not on this facet of the cosmological polytope. This red line also provides the natural triangulating hyperplane of the facet needed to compute the canonical form. The two tetrahedra that result from triangulating the pyramid give the canonical forms with x_2 absorbing positive and negative edge energies.



It is clear that the red shaded triangle comprising the intersection of a hyperplane with the pyramid is the polytope in \mathbb{P}^2 corresponding to the right triangle, which is the cosmological polytope corresponding to the Feynman diagram made of vertices x_2 and x_3 connected by edge y_2 . Therefore the canonical form on this two dimensional subspace is

$$\Omega = \frac{\langle X d^2 X \rangle \langle 123 \rangle^2}{\langle X 12 \rangle \langle X 23 \rangle \langle X 31 \rangle} \quad (2.27)$$

and its wave-function contribution is that of the two-site chain. In this triangulation, the canonical form for the pyramid is the difference of the canonical forms on two tetrahedra separated by the hyperplane. We see this is

$$\Omega = \frac{1}{\langle X 123 \rangle} \left(\frac{\langle X d^3 X \rangle \langle 1234 \rangle^3}{\langle X 234 \rangle \langle X 341 \rangle \langle X 412 \rangle} - \frac{\langle X d^3 X \rangle \langle 1325 \rangle^3}{\langle X 325 \rangle \langle X 251 \rangle \langle X 513 \rangle} \right) \quad (2.28)$$

Vertices 4 and 5 of the pyramid correspond to the two circles left of the two-site subgraph and are the only vertices lifted into the y_1 direction. We see that

$$\Omega = \frac{1}{2y_1} \left(\frac{\langle X d^3 X \rangle \langle 1234 \rangle^3}{(\langle X 23 \rangle - 2y_1)(\langle X 31 \rangle - y_1)(\langle X 12 \rangle - 2y_1)} - \frac{\langle X d^3 X \rangle \langle 1325 \rangle^3}{(\langle X 23 \rangle + 2y_1)(\langle X 12 \rangle + 2y_1)(\langle X 31 \rangle + 2y_1)} \right) \quad (2.29)$$

The piece corresponding to the wave function contribution is then

$$\text{Res}_{x_1+y_1 \rightarrow 0} \psi = \frac{1}{2y_1} \left(\frac{1}{(x_3 + y_2)(x_2 - y_1 + y_2)(x_3 + x_2 - y_1)} - \frac{1}{(x_3 + y_2)(x_2 + y_1 + y_2)(x_3 + x_2 + y_1)} \right) \quad (2.30)$$

consistent with the recursive rule. This is a general mechanism; the cosmological polytope corresponding to the subgraph is a triangulating hyperplane, separating polytopes whose canonical forms correspond to absorbing positive and negative edge energy.

The other aspect to understand from the polytope perspective is the factorization of disconnected subgraphs. This can be illustrated with the following discontinuity

$$x_1 \bullet \circ \circ \times \circ \bullet \times \circ \times \circ \bullet x_3$$

x_2

(2.31)

This facet of the cosmological polytope is a tetrahedron.

(2.32)

The red and blue lines are subspaces in \mathbb{P}^1 . We can call these projective subspaces X_1 and X_2 . The full tetrahedron in \mathbb{P}^3 is the Cartesian product of these subspaces. It can be understood as made from combining the subspaces with projective weights $\tau_1 X_1 + \tau_2 X_2$ where the τ 's are positive. Each subspace has one projective degree of freedom and the simplex of τ 's is the third projective degree of freedom which we can describe with one projective coordinate τ . The canonical form on the tetrahedron is then the three-form

$$\Omega_X = \Omega_{X_1} \wedge \Omega_\tau \wedge \Omega_{X_2}$$
(2.33)

We will show that this works explicitly. We can call the blue subspace X_1 and the red X_2 . The vertices on X_1 are $\{x_1 + y_1, x_1 - y_1\}$ which we call v_1 and v_2 . X_1 is

$$X_1 = \alpha_1 v_1 + \alpha_2 v_2$$
(2.34)

The canonical form is then

$$\Omega_{X_1} = d \log \left(\frac{\alpha_1}{\alpha_2} \right) \quad (2.35)$$

with poles at 0 and ∞ in the projective parameter $\alpha = \frac{\alpha_1}{\alpha_2}$. In homogenous coordinates we have

$$X_1 = \alpha_1 \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \alpha_2 \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} \quad (2.36)$$

and so we get the appropriate form

$$\Omega_{X_1} = d \log \left(\frac{x_1 - y_1}{x_1 + y_1} \right) \quad (2.37)$$

This is just the usual canonical form for an interval in \mathbb{P}^1 . The same holds for X_2 and we call the parameters β_1, β_2 . Now, with the combined space we'd simply have

$$X = \tau_1 v_1 + \tau_1 \alpha v_2 + \tau_2 v_3 + \tau_2 \beta v_4 \quad (2.38)$$

And we then say the canonical form is

$$\Omega_X = d \log \alpha \wedge d \log \beta \wedge d \log \left(\frac{\tau_1}{\tau_2} \right) \quad (2.39)$$

Now to extract the form in the homogenous coordinates we again proceed

$$X = \tau_1 \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} + \tau_1 \alpha \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix} + \tau_2 \begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \end{pmatrix} + \tau_2 \beta \begin{pmatrix} 0 \\ 0 \\ 1 \\ -1 \end{pmatrix} \quad (2.40)$$

and from these the form is given by

$$\Omega_X = d \log \left(\frac{x_1 - y_1}{x_1 + y_1} \right) \wedge d \log \left(\frac{x_3 - y_2}{x_3 + y_2} \right) \wedge d \log \left(\frac{x_1 + y_1}{x_3 + y_2} \right) \quad (2.41)$$

Which does give the correct residue for the wave-function contribution. If we added another subgraph, we would be working with the cartesian product of this original polytope and the polytope of the subgraph and the same would hold, and so the factorization can be seen inductively. It is worth commenting that different gauge-fixing choices for the projective weights would result in forms that do not in general have the same differentials. This is merely an artifact of having to make a choice, but the contribution to the perturbative wave function is the same no matter choice we make.

2.4 Some Examples

2.4.1 Two-Site Chain

The symbol for the two-site chain, corresponding to the four-point correlator, is readily acquired with the recursive rule. The graphical rule states

$$x_1 \text{---} \textcircled{\bullet} \text{---} x_2 = - \text{---} \bullet \text{---} x_2 - + \text{---} \bullet \text{---} x_2 \quad (2.42)$$

The same holds for the other vertex and so the symbol is immediately determined

$$\mathcal{S} = \frac{X_1 + Y}{X_1 + X_2} \otimes \frac{X_2 - Y}{X_2 + Y} + \frac{X_2 + Y}{X_1 + X_2} \otimes \frac{X_1 - Y}{X_1 + Y} \quad (2.43)$$

The symbol is naturally expressed in the following variables

$$u = \frac{X_1 - Y}{X_1 + X_2} \quad (2.44)$$

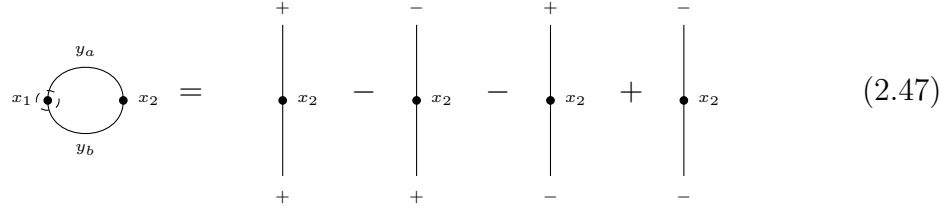
$$v = \frac{X_2 - Y}{X_1 + X_2} \quad (2.45)$$

and it takes the form

$$\mathcal{S} = -(1-u) \otimes \frac{u}{1-v} - (1-v) \otimes \frac{v}{1-u} \quad (2.46)$$

2.4.2 Two-Site Loop

Here, the graphical rule on the vertices says



$$\text{Diagram} = \text{Diagram}_1 - \text{Diagram}_2 - \text{Diagram}_3 + \text{Diagram}_4 \quad (2.47)$$

The same holds for the second vertex and so the vertices contribute

$$\frac{X_1 + Y_a + Y_b}{X_1 + X_2} \otimes [(X_2 + Y_a + Y_b) - (X_2 - Y_a + Y_b) - (X_2 + Y_a - Y_b) + (X_2 - Y_a - Y_b)]$$

$$\frac{X_2 + Y_a + Y_b}{X_1 + X_2} \otimes [(X_1 + Y_a + Y_b) - (X_1 - Y_a + Y_b) - (X_1 + Y_a - Y_b) + (X_1 - Y_a - Y_b)]$$

where we summed over all combinations of edge energy signs inside the brackets.

In order to get the contribution for the larger subgraphs, we use the fact that the discontinuity of the subgraph is minus the sum of the discontinuities of *its* subgraphs, with positive energy absorbed from cut edges. Graphically this means we have the

following

$$\begin{array}{c}
 \text{Loop with vertices } x_1, x_2 \text{ and internal lines } y_a, y_b \\
 = \\
 \begin{array}{cccc}
 - & + & - & + \\
 \text{---} x_1 & \text{---} x_1 & \text{---} x_2 & \text{---} x_2 \\
 + & + & + & +
 \end{array}
 \end{array} \quad (2.48)$$

The contributions for the two U-shaped subgraphs are therefore

$$\begin{aligned}
 & \frac{X_1 + X_2 + 2Y_b}{X_1 + X_2} \otimes \left[\frac{X_1 + Y_b - Y_a}{X_1 + Y_b + Y_a} + \frac{X_2 + Y_b - Y_a}{X_2 + Y_b + Y_a} \right] \\
 & \frac{X_1 + X_2 + 2Y_a}{X_1 + X_2} \otimes \left[\frac{X_1 + Y_a - Y_b}{X_1 + Y_b + Y_a} + \frac{X_2 + Y_a - Y_b}{X_2 + Y_b + Y_a} \right]
 \end{aligned}$$

If we add zero by both adding and subtracting each of $X_1 + Y_a + Y_b \otimes X_2 + Y_a + Y_b$ and $X_2 + Y_a + Y_b \otimes X_1 + Y_a + Y_b$, thereby contributing four new terms, we are motivated to introduce three sets of variables analogous to the two site

$$\begin{array}{lll}
 u_1 = \frac{X_1 + Y_a - Y_b}{X_1 + X_2 + 2Y_a} & u_2 = \frac{X_1 + Y_b - Y_a}{X_1 + X_2 + 2Y_b} & u_3 = \frac{X_1 - Y_a - Y_b}{X_1 + X_2} \\
 v_1 = \frac{X_2 + Y_a - Y_b}{X_1 + X_2 + 2Y_a} & v_2 = \frac{X_2 + Y_b - Y_a}{X_1 + X_2 + 2Y_b} & v_3 = \frac{X_2 - Y_a - Y_b}{X_1 + X_2}
 \end{array}$$

The symbol is neatly expressed as

$$\begin{aligned}
 \mathcal{S}_{1\text{-loop}} = & (1 - u_1) \otimes \frac{u_1}{1 - v_1} + (1 - v_1) \otimes \frac{v_1}{1 - u_1} \\
 & (1 - u_2) \otimes \frac{u_2}{1 - v_2} + (1 - v_2) \otimes \frac{v_2}{1 - u_2} \\
 & - (1 - u_3) \otimes \frac{u_3}{1 - v_3} - (1 - v_3) \otimes \frac{v_3}{1 - u_3}
 \end{aligned} \quad (2.49)$$

We recognize this as three two-site chains, a fact that can also be seen by appropriately rearranging the time-ordered Feynman diagrams.

2.4.3 Three-Site Chain

The previous two examples illustrate the intuitive diagrammatic organization of symbol contributions. The three-site chain is the first example which additionally demonstrates the efficiency of the recursive rule in computing the symbol. We begin by computing the contributions from vertices

$$\begin{array}{c} x_1 \text{---} \textcircled{\text{---}} \text{---} x_2 \text{---} x_3 \\ \text{---} x_2 \end{array} = \begin{array}{c} \text{---} x_2 \text{---} x_3 \\ \text{---} x_2 \end{array} - \begin{array}{c} \text{---} x_2 \text{---} x_3 \\ \text{---} x_2 \end{array} + \begin{array}{c} \text{---} x_2 \text{---} x_3 \\ \text{---} x_2 \end{array} \quad (2.50)$$

This is the difference of two-site chains where the absorbed edge energy is either positive or negative. Therefore the contributions for the external vertices are

$$\begin{aligned}
 & \frac{X_1 + Y_1}{X_1 + X_2 + X_3} \otimes \left[\frac{X_3 + Y_2}{X_3 + X_2 - Y_1} \otimes \frac{X_2 - Y_1 - Y_2}{X_2 - Y_1 + Y_2} + \frac{X_2 - Y_1 + Y_2}{X_3 + X_2 - Y_1} \otimes \frac{X_3 - Y_2}{X_3 + Y_2} \right. \\
 & \quad \left. - \frac{X_3 + Y_2}{X_3 + X_2 + Y_1} \otimes \frac{X_2 + Y_1 - Y_2}{X_2 + Y_1 + Y_2} - \frac{X_2 + Y_1 + Y_2}{X_3 + X_2 + Y_1} \otimes \frac{X_3 - Y_2}{X_3 + Y_2} \right] \\
 & \frac{X_3 + Y_2}{X_1 + X_2 + X_3} \otimes \left[\frac{X_1 + Y_1}{X_1 + X_2 - Y_2} \otimes \frac{X_2 - Y_2 - Y_1}{X_2 - Y_2 + Y_1} + \frac{X_2 - Y_2 + Y_1}{X_1 + X_2 - Y_1} \otimes \frac{X_1 - Y_1}{X_1 + Y_1} \right. \\
 & \quad \left. - \frac{X_1 + Y_1}{X_1 + X_2 + Y_2} \otimes \frac{X_2 + Y_2 - Y_1}{X_2 + Y_2 + Y_1} - \frac{X_2 + Y_2 + Y_1}{X_1 + X_2 + Y_2} \otimes \frac{X_1 - Y_1}{X_1 + Y_1} \right]
 \end{aligned}$$

Next we look at the middle vertex. We have

$$\begin{array}{c} x_1 \text{---} \textcircled{\text{---}} \text{---} x_3 \\ \text{---} x_2 \end{array} = \left(\begin{array}{c} x_1 \text{---} \text{---} \\ \text{---} x_1 \end{array} - \begin{array}{c} \text{---} x_1 \text{---} \text{---} \\ \text{---} x_1 \end{array} + \begin{array}{c} \text{---} x_1 \text{---} \text{---} \\ \text{---} x_1 \end{array} \right) \left(\begin{array}{c} \text{---} \text{---} x_2 \\ \text{---} x_2 \end{array} - \begin{array}{c} \text{---} \text{---} x_2 \\ \text{---} x_2 \end{array} + \begin{array}{c} \text{---} \text{---} x_2 \\ \text{---} x_2 \end{array} \right) \quad (2.51)$$

The factorization at the level of the answer is inherited from factorization at the level of the integrand. Factorization in the answer is manifest as a shuffle product in the symbol and so the contribution is

$$\frac{X_2 + Y_1 + Y_2}{X_1 + X_2 + X_3} \otimes \left[\frac{X_1 - Y_1}{X_1 + Y_1} \otimes \frac{X_3 - Y_2}{X_3 + Y_2} + \frac{X_3 - Y_2}{X_3 + Y_2} \otimes \frac{X_1 - Y_1}{X_1 + Y_1} \right] \quad (2.52)$$

Now we look at the larger subgraphs, where we have

$$\begin{aligned}
 \text{Diagram: } x_1 \text{ --- } x_2 \text{ --- } x_3 \text{ with a dashed loop on } x_1 \text{ and } x_2 &= - \left(\begin{array}{c} x_1 \text{ --- } \quad - \quad - \quad x_1 \text{ --- } + \quad + \quad - \quad \text{---} x_2 + \quad - \quad + \quad \text{---} x_2 + \end{array} \right) \\
 &\quad \times \left(\begin{array}{c} - \quad \text{---} x_3 - \quad + \quad \text{---} x_3 \end{array} \right) \\
 &\quad (2.53)
 \end{aligned}$$

The symbol contribution for the two subgraphs is therefore

$$\begin{aligned}
 &\frac{X_1 + X_2 + Y_2}{X_1 + X_2 + X_3} \otimes \left[\frac{X_1 - Y_1}{X_1 + Y_1} \otimes \frac{X_3 - Y_2}{X_3 + Y_2} + \frac{X_3 - Y_2}{X_3 + Y_2} \otimes \frac{X_1 - Y_1}{X_1 + Y_1} \right. \\
 &\quad \left. \frac{X_2 + Y_2 - Y_1}{X_2 + Y_2 + Y_1} \otimes \frac{X_3 - Y_2}{X_3 + Y_2} + \frac{X_3 - Y_2}{X_3 + Y_2} \otimes \frac{X_2 + Y_2 - Y_1}{X_2 + Y_2 + Y_1} \right] \\
 &\frac{X_3 + X_2 + Y_1}{X_1 + X_2 + X_3} \otimes \left[\frac{X_3 - Y_2}{X_3 + Y_2} \otimes \frac{X_1 - Y_1}{X_1 + Y_1} + \frac{X_1 - Y_1}{X_1 + Y_1} \otimes \frac{X_3 - Y_2}{X_3 + Y_2} \right. \\
 &\quad \left. \frac{X_2 + Y_1 - Y_2}{X_2 + Y_1 + Y_2} \otimes \frac{X_1 - Y_1}{X_1 + Y_1} + \frac{X_1 - Y_1}{X_1 + Y_1} \otimes \frac{X_2 + Y_1 - Y_2}{X_2 + Y_1 + Y_2} \right]
 \end{aligned}$$

We combine all these contributions to get the total symbol, which reproduces the 104 term expression in [8]. There does not appear to be a natural set variables which makes the symbol separate into pieces obviously corresponding to the symbols of classical polylogs as with the previous cases.

2.5 Integrating the Symbols

The symbol determines the polylogarithmic function up to functions of lower transcendental weight times transcendental numbers. By integrating the symbol to get the piece of highest transcendental weight, we can then impose boundary conditions to fix the other pieces. There is no general algorithm for associating a function with a symbol and beyond weight three we lose the guarantee that we only have to consider

classical polylogarithms. Nonetheless, we were able to compute the integrals for the symbols presented in the examples.

2.5.1 Two-Site Chain

In the u and v variables defined earlier the symbol for the two-site chain expands out to

$$\mathcal{S}_2 = -(1-u) \otimes u - (1-v) \otimes v - (1-u) \otimes (1-v) - (1-v) \otimes (1-u) \quad (2.54)$$

we immediately recognize these as corresponding to

$$\tilde{F} = \text{Li}_2(u) + \text{Li}_2(v) + \text{Li}_1(u)\text{Li}_1(v) + C \quad (2.55)$$

Where for n edges \tilde{F} is related to the full correlator via $F = \prod_{i=1}^n (2Y_i) \tilde{F}$. It remains to fix the constant. The correlator should vanish when the edge energy goes to zero. This subspace is defined by the constraint $u + v = 1$ and so we demand

$$\text{Li}_2(u) + \text{Li}_2(1-u) + \text{Li}_1(u)\text{Li}_1(1-u) + C = 0 \quad (2.56)$$

from which we deduce $C = -\frac{\pi^2}{6}$ due to the Euler identity, which holds for $u, v < 1$, clearly obeyed in the physical region. Therefore

$$\tilde{F}_2 = \text{Li}_2(u) + \text{Li}_2(v) + \text{Li}_1(u)\text{Li}_1(v) - \frac{\pi^2}{6} \quad (2.57)$$

which is the result found in [11] and again in [10]. It is amusing to note that this answer is actually naturally rearranged using the Abel identity into the following form

$$\tilde{F}_2 = \text{Li}_2\left(\frac{X_1 - Y}{X_1 + Y}\right) + \text{Li}_2\left(\frac{X_2 - Y}{X_2 + Y}\right) - \text{Li}_2\left(\frac{(X_1 - Y)(X_2 - Y)}{(X_1 + Y)(X_2 + Y)}\right) - \frac{\pi^2}{6} \quad (2.58)$$

2.5.2 Two-Site Loop

We recognized that the two-site loop was a sum of two-site chains from which we deduce

$$\begin{aligned}
\tilde{F}_{1\text{-loop}} = & \text{Li}_2(u_3) + \text{Li}_2(v_3) + \text{Li}_1(u_3)\text{Li}_1(v_3) \\
& -\text{Li}_2(u_2) + \text{Li}_2(v_2) + \text{Li}_1(u_2)\text{Li}_1(v_2) \\
& -\text{Li}_2(u_1) + \text{Li}_2(v_1) + \text{Li}_1(u_1)\text{Li}_1(v_1) \\
& +C
\end{aligned} \tag{2.59}$$

where it remains to fix the constant again. Setting either of the two edges to zero gives us an equation identical to the two-site chain simply with the opposite sign and so we have

$$\begin{aligned}
\tilde{F}_{1\text{-loop}} = & \text{Li}_2(u_3) + \text{Li}_2(v_3) + \text{Li}_1(u_3)\text{Li}_1(v_3) \\
& -\text{Li}_2(u_2) + \text{Li}_2(v_2) + \text{Li}_1(u_2)\text{Li}_1(v_2) \\
& -\text{Li}_2(u_1) + \text{Li}_2(v_1) + \text{Li}_1(u_1)\text{Li}_1(v_1) \\
& +\frac{\pi^2}{6}
\end{aligned} \tag{2.60}$$

we take up integration over spatial loop momenta in an upcoming work.

2.5.3 Three-Site Chain

The three-site chain has transcendental weight three. At weight three it is a theorem that you can integrate a symbol to a function expressed only in terms of classical polylogarithms. There is an approach to integration which consists of separating the pieces of the symbol based on their symmetry properties under exchange of slots. This approach was employed successfully in [15]. For example, at weight three the

only function which can produce a symbol with any antisymmetry in the last two slots is $\text{Li}_2(x) \log(y)$. We can therefore turn the rightmost tensor symbol into a \wedge and simplify in order to isolate this part of the symbol. We recognize that it corresponds to two-site chains times logs. When we have subtracted this piece, we can project onto antisymmetry in the first two slots to separate the $\text{Li}_3(z)$ piece, and then finally we are left with $\log^3(w)$'s. This algorithm still requires some guesswork in practice, but in this case we deduced the answer. We state the answer below

$$\begin{aligned}
\tilde{F}_3 = & \sum_{\{\sigma_1, \sigma_2\} \in \{-, +\}^2} \sigma_1 \sigma_2 \left[\text{Li}_3 \left(-\frac{u_{\sigma_1} v_{\sigma_2}}{1 - u_{\sigma_1} - v_{\sigma_2}} \right) + \text{Li}_3 \left(\frac{u_{\sigma_1}}{1 - v_{\sigma_2}} \right) + \text{Li}_3 \left(\frac{v_{\sigma_2}}{1 - u_{\sigma_1}} \right) \right. \\
& + \text{Li}_3 \left(\frac{1 - u_{\sigma_1} - v_{\sigma_2}}{1 - v_{\sigma_2}} \right) + \text{Li}_3 \left(\frac{1 - u_{\sigma_1} - v_{\sigma_2}}{1 - u_{\sigma_1}} \right) \\
& - \frac{1}{2} \text{Li}_1 \left(\frac{1 - u_{\sigma_1} - v_{\sigma_2}}{1 - v_{\sigma_2}} \right) \text{Li}_1^2 \left(\frac{u_{\sigma_1}}{1 - v_{\sigma_2}} \right) - \frac{1}{2} \text{Li}_1 \left(\frac{1 - u_{\sigma_1} - v_{\sigma_2}}{1 - u_{\sigma_1}} \right) \text{Li}_1^2 \left(\frac{v_{\sigma_2}}{1 - u_{\sigma_1}} \right) \\
& \left. + \frac{1}{6} \text{Li}_1^3 \left(\frac{u_{\sigma_1} + v_{\sigma_2}}{1 - u_{\sigma_1}} \right) + \frac{1}{6} \text{Li}_1^3 \left(\frac{u_{\sigma_1} + v_{\sigma_2}}{1 - v_{\sigma_2}} \right) - \frac{1}{6} \text{Li}_1^3 (u_{\sigma_1} + v_{\sigma_2}) \right] \\
& + \text{Li}_1 (1 - u_+) \left[\text{Li}_2 \left(\frac{1 - v_+ - u_+}{1 - v_- - u_+} \right) - \text{Li}_2 \left(\frac{v_- (1 - v_+ - u_+)}{v_+ (1 - v_- - u_+)} \right) \right. \\
& \left. - \text{Li}_2 \left(\frac{1 - v_+ - u_-}{1 - v_- - u_-} \right) + \text{Li}_2 \left(\frac{v_- (1 - v_+ - u_-)}{v_+ (1 - v_- - u_-)} \right) \right] \\
& + \text{Li}_1 (1 - v_+) \left[\text{Li}_2 \left(\frac{1 - v_+ - u_+}{1 - v_+ - u_-} \right) - \text{Li}_2 \left(\frac{u_- (1 - v_+ - u_+)}{u_+ (1 - v_+ - u_-)} \right) \right. \\
& \left. - \text{Li}_2 \left(\frac{1 - v_- - u_+}{1 - v_- - u_-} \right) + \text{Li}_2 \left(\frac{u_- (1 - v_- - u_+)}{u_+ (1 - v_- - u_-)} \right) \right] \\
& + \text{Li}_1 (1 - u_+) \text{Li}_1 (1 - v_+) \text{Li}_1 \left(\frac{(v_- - v_+) (u_- - u_+)}{(1 - u_+ - v_-) (1 - u_- - v_+)} \right)
\end{aligned} \tag{2.61}$$

Where we have used the variables

$$u_{\pm} = \frac{X_1 \pm Y_1}{X_1 + X_2 + X_3} \quad v_{\pm} = \frac{X_3 \pm Y_2}{X_1 + X_2 + X_3} \tag{2.62}$$

It remains to fix contributions not detected by the symbol. Setting $Y_1 = 0$ or $Y_2 = 0$ gives $u_+ = u_-$ or $v_+ = v_-$ respectively. Demanding that \tilde{F}_3 vanish in either of these

cases and additionally demanding that we recover the correct lower-point contributions on the appropriate discontinuities actually fixes any additional contribution to be zero, so the result is correct as stated.

2.5.4 Four-Site and Beyond

Already at weight four the algorithmic procedure for integrating the symbol by using its symmetry properties is contingent upon restricting to classical polylogarithms. For a generic integrable symbol, there is no guarantee beyond weight three that it can be produced by a function only consisting of classical polylogarithms. In [15] they are motivated to try this because of a conjecture, since proven, that a symbol at weight four obeying the following relation under permutation of slots

$$S_{abcd} - S_{bacd} - S_{abdc} + S_{badc} + S_{cdab} - S_{dcab} - S_{cdba} + S_{dcba} = 0 \quad (2.63)$$

can in fact be integrated to a function consisting only of classical polylogarithms. We can readily test this conjecture on the wave function at six points using the symbol rule. At six points (weight four) we now have two graph topologies that contribute at tree-level, the four-site chain and the four-site star. We input the symbols for each of these cases into the expression above and neither vanished. Therefore, it seems that at weight four we may lose the capacity to express even tree-level answers in terms of classical polylogarithms alone.

2.6 Differential Equations from the Symbol

Mass insertions Having fixed the symbol recursively by imposing factorization, we can recognize its differential properties. For one, the structure of the recursion makes manifest that the last entry in every word of the symbol will depend on only one vertex energy X_v . This makes it particularly simple to differentiate with respect to a single vertex energy, using the definition of the symbol (2.12). In particular, we observe that e.g. for the n -site chain we have

$$\partial_{X_n} F^{(n)}(X_1, \dots, X_n, Y_1, \dots, Y_{n-1}) = \frac{2Y_{n-1}}{X_n^2 - Y_{n-1}^2} \left[F^{(n-1)}(X_1, \dots, X_{n-1} + Y_{n-1}, Y_1, \dots, Y_{n-2}) - F^{(n-1)}(X_1, \dots, X_{n-1} + X_n, Y_1, \dots, Y_{n-2}) \right] \quad (2.64)$$

therefore, if we express the massive exchange graph as an infinite sum over mass



Figure 2.2: A contribution to the two-site chain now with six mass insertions. As in flat space, we are summing over an infinite number of mass insertions in order to produce the effects of a massive propagator. The wave-function does not factorize as in flat space, but we can at least derive a differential equation for the wave function.

insertions

$$\Psi_M(X_1, X_2, Y) = \sum_{N=0}^{\infty} (-M^2)^N F_M^N(X_1, X_2, Y) \quad (2.65)$$

Where

$$F_M^N(X_1, X_2, Y) = \int_0^{\infty} x_2 \dots x_{N-1} \psi^{(N)}(x_1 + X_1, x_2, \dots, x_N + X_2, Y, \dots, Y) \quad (2.66)$$

We see that the behavior of the two end-points is identical to the ordinary N -site chain. We have

$$\begin{aligned} \partial_{X_2} F_M^N(X_1, X_2, Y) = \frac{2Y_{n-1}}{X_n^2 - Y_{n-1}^2} \int_0^\infty x_2 \dots x_{N-1} \big[\psi^{(N-1)}(x_1 + X_1, x_2, \dots, x_{N-1} + Y, Y, \dots, Y) \\ - \psi^{(N-1)}(x_1 + X_1, x_2, \dots, x_{N-1} + X_2, Y, \dots, Y) \big] \end{aligned} \quad (2.67)$$

We now notice that in the second $\psi^{(N-1)}$ term, x_{N-1} and X_2 always come together. Moreover, the boundary terms in x_{N-1} of the integrand vanish since the behavior at infinity is good and the x_{N-1} guarantees vanishing at zero. This means, using integration by parts, that

$$\frac{1}{2Y} \partial_{X_2} (X_2^2 - Y^2) \partial_{X_2} \psi^N(X_1, X_2, Y) = \psi^{N-1}(X_1, X_2, Y) \quad (2.68)$$

which immediately implies, noting that on ψ^0 we just get the pure conformally coupled contact,

$$\left[\frac{1}{2Y} \partial_{X_2} (X_2^2 - Y^2) \partial_{X_2} + M^2 \right] \Psi(X_1, X_2, Y) = \frac{1}{(X_1 + X_2)} \quad (2.69)$$

This furnishes another derivation of the massive differential equation in de Sitter derived in [10]. There it was derived both from manipulations of the bulk time integral and is related to annihilation by the special conformal generator at the future boundary (though that in fact produces a homogenous differential equation). Here we have another derivation essentially in terms of boundary kinematic data.

All-n Conformally Coupled We can just as easily understand how to write an ODE in the external vertex energies on an n -site chain in the conformally coupled case. From (2.64) we see that only one of the two terms in (2.64) has X_n dependence.

Therefore, the recursive collapsing is clear. We see that

$$\partial_{X_n} \prod_{i=0}^{n-2} \left(\frac{2Y_{n-i-1}}{X_{n-i}^2 - Y_{n-i-1}^2} \partial_{X_n} \right) F^{(n)} = \frac{1}{X_1 + \cdots + X_n} \quad (2.70)$$

The generalization to m derivatives from one side and $n-m$ derivatives from the other side is also clear. For the three-site case, we will consider the case of differentiation with respect to the middle vertex in the next section, which will be clarified with the aid of a new representation of the symbol letters.

2.7 Singularities from Graph Associahedra

Stringy Canonical Forms First we will rapidly review stringy canonical forms [16], which for our purposes will merely be a crank we turn to generate symbol letters. For symbol letters to come from a stringy canonical form, certain conditions must be satisfied, which we will make clear. This is just to say that a random set of symbol letters will not necessarily come from a stringy canonical form. It is therefore a non-trivial fact that the set of multiplicatively independent symbol letters in our symbols are captured by a stringy canonical form and their associated binary geometries.

A stringy canonical form is an integral of the form:

$$I(\{X_i, c_j\}) = \int_{\mathbb{R}_+^n} d^n \log(x_i) x_i^{\alpha' X_i} \prod_{j=1}^m P_j(x_i)^{\alpha' c_j} \quad (2.71)$$

we can view the α' dependent (twisted) part of the integrand as regulating the logarithmically divergent $d^n \log(x_i)$ integral to give us something finite at finite α' . The name stringy canonical form comes from the fact that the leading behavior of this integral as $\alpha' \rightarrow 0$ (the leading term in the Laurent expansion) is the canonical form¹

¹The unique $d \log$ form with logarithmic singularities on the boundaries of the geometry.

of the Newton polytope associated to the integrand

$$\bigoplus_j c_j P_j - \vec{X} \quad (2.72)$$

where the Newton polytope of a polynomial is the convex hull of the exponent vectors e.g. $1 + x + y^2$ yields a triangle with vertices $\{(0, 0), (1, 0), (0, 2)\}$, and the sum means we add the hulls up linearly. The Newton polytope tells us about the convergence properties of the integrand. When some set of x_i become large or small, you are potentially probing a logarithmically divergent region. There are two basic points: the α' -dependent piece makes such limits of the integrand not single-valued on \mathbb{R}_+^n : the limits will depend on the relative rates at which x_i get small or big. The Newton polytope is the “blown up” space at the boundary of the integration domain on which all codimension-one limits are now single valued.

The punchline is that domain of convergence of the integral is cut out by complicated linear constraints in the X_i and c_j . So staring at the integrand, convergence looks complicated. Something very nice happens when $|F|$ the number of facets of the Newton polytope is equal to $m + n$. In this case, we can rewrite

$$I(\{X_i, c_j\}) = \int_{\mathbb{R}_+^n} d^n \log(x_i) u_k^{\alpha' p_k} \quad (2.73)$$

where the convergence condition for the integral is that $p_k > 0$. The u_k are rational functions of the x_i and P_j and the p_k are linear in the X_i and c_j . Moreover, very often the u_k have extremely nice properties such as u -equations, which are highly overconstrained non-linear equations between the variables. The symbol letters for our chains will indeed obey u -equations.

The crank that you turn to produce the u_k only works when $|F| = m + n$, because the u_k are calculated uniquely by projecting components along the rays generating a cone in the $m + n$ dimensional space. For uniqueness, this cone must be simplicial

in the $m + n$ dimensional space i.e. generated by $m + n$ rays. The upshot is that given this magical $|F| = m + n$ constraint is obeyed, there is a completely canonical algorithm to extract the u_k . We will not expound further here since it is very nicely covered in [16].

Variables for de Sitter Symbols We seek a positive parameterization of our symbol letters, so that we can feed them into a stringy canonical form. We see that this is in fact possible by construction. We now describe an algorithm which lets us do this. Given a Feynman graph, we define a new doubled graph, where we mark

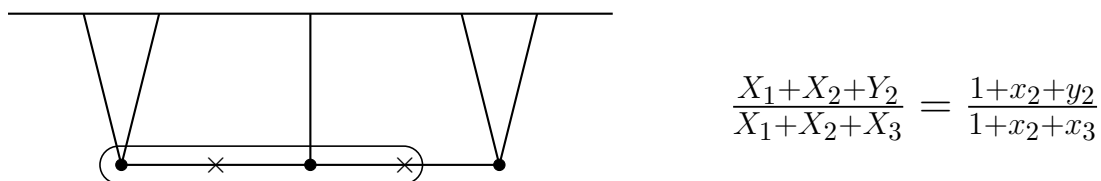


Figure 2.3: The doubled graph has another vertex on every edge. The \times have positive variables y_i associated to them and the \cdot have x_i . The equation to the right of the graph is the mapping of the symbol letters into this positive parameterization for the particular tubing. When the tubing does not contain the boundary \times , the associated Y_e comes with a minus sign in the numerator.

an \times on all of the edges. The point of doing is so that we can identify tubings on this graph with different signs of absorbing the edge energy that the \times is one. It is natural to do this since both choices of sign appear in the symbol. The rule we will introduce is that for each tubing on the doubled graph, the numerator will be the sum of all vertex energies for the black vertices enclosed, and then $\pm Y_j$ for all boundary \times , with the sign being $-$ for \times on the boundary but external to the tubing and $+$ for a \times which is on the boundary but internal to the tubing. An example is provided in fig. 2.3. It is worth noting here now that there is no canonical about this sign convention, and of course the convention changes by simply changing the sign of all edge energies.

Next we want to positively parameterize these letters. This means we want to identify

each of these ratios with some ratio of polynomials in positive variables. We do this by first identifying each vertex with an x_i and each \times with a y_i and use the parameterization

$$\frac{\sum_v X_v - \sum(-)Y_e}{X_1 + \dots X_n} = \frac{\sum_v x_v + \sum(+y_e)}{1 + x_2 + \dots x_n} \quad (2.74)$$

where $x_1 = 1$. This means for any tubing on the doubled graph, we add the x_v for all enclosed black vertices, and for every boundary \times which is *inside* the tubing, we add a y_e . This gives us a polynomial for each tubing, plus the denominator polynomial. We can then write the stringy canonical form

$$I = \int d^{V-1} \log(x_v) (1 + x_2 \dots x_n)^{p_G} \prod_v x_v^{p_v} \prod_T P_T(\{x_v\}, \{y_e\})^{p_T} \quad (2.75)$$

So what is going to appear is clearly all sums of x_v and y_e for tubings on the doubled graph. It is known that this is the stringy canonical form for the graph associahedron associated to graph G [17]. Therefore, when we turn the stringy canonical forms crank, we will get u -variables which furnish binary realizations of the associated graph associahedron. We now illustrate this in some examples.

It is worth commenting that we do not use the positivity of the x_v and y_e per se, as their essential content is dictated merely by their exponents not their coefficients. That being said, it would be interesting to consider in this parameterization what the positive region $x_v > 0$ and $y_e > 0$ corresponds to in kinematic space. The singularities

Two-Site Chain We have the following parameterization of the variables in the two-site chain

$$\frac{X_1 + Y}{X_1 + X_2} = \frac{1}{1 + y + x_2} \quad \frac{X_2 + Y}{X_1 + X_2} = \frac{x_2}{1 + y + x_2} \quad (2.76)$$

This means the polynomials appearing are $\{1 + y, x_2 + y, 1 + y + x_2\}$ the Minkowski sum of which is a pentagon. This means we satisfy the condition via $5 = 3 + 2$ and we can ask for u -variables. Once a choice of parameterization is made, there are no further choices. In this case, we get the u -variables

$$\begin{aligned} u_{1,3} &= \frac{X_1 + Y}{X_1 + X_2} & u_{2,5} &= \frac{X_2 + Y}{X_1 + X_2} \\ u_{1,4} &= \frac{X_1 - Y}{X_1 + Y} & u_{2,4} &= \frac{X_2 - Y}{X_2 + Y} \\ u_{3,5} &= \frac{2Y(X_1 + X_2)}{(X_1 + Y)(X_2 + Y)} \end{aligned} \quad (2.77)$$

The subscripts identify these u -variables with chords on a pentagon, the numbers denoting the vertices the associated chord connects. This pentagon is the A_2 pentagon, the two-dimensional associahedron. This squares with the graph associahedron for the marked graph indeed being A_2 . The u -equations can be stated as

$$1 - u_{i,j} = \prod_{k,l \text{ cross } i,j} u_{k,l} \quad (2.78)$$

where the product on the right-hand side is over chords k, l which cross i, j . For example, we see that

$$1 - u_{1,3} = u_{2,4}u_{2,5} \quad (2.79)$$

Note that because in the positive region $x_v > 0$ and $y_e > 0$ the u 's vary from 0 to 1, the above equations imply compatibility conditions for the u -variables. When $u_{i,j}$ goes to zero, all the u 's associated with crossing chords must equal one.

Finally we write the symbol in these variables

$$S = u_{1,3} \otimes u_{2,4} + u_{2,5} \otimes u_{1,4} \quad (2.80)$$

with a first-entry condition only allowing $u_{1,3}$ and $u_{2,4}$ (imposing absence of folded), the symbol seems to be produced by summing over *incompatible* pairs. Sadly, such a simple rule does not seem to hold in the three-site case, though there is perhaps a generalization accounting for the discrepancy.

Three-Site Chain The three-site case proceeds in the same way, now with a five vertex doubled graph producing the A_4 associahedron which has 14 facets. Two of the letters are associated with the edges going to zero and will not appear in the symbol.

The outermost X_v energy differentials are always clear, but differentials with respect to middle vertex energies are not. But the symbol in this case does seem to give something canonical. There are three canonical rational functions dependent solely on X_2 and the edge energies which are furnished by symbol letters or their products

$$\left\{ \frac{X_2 + Y_1 - Y_2}{X_2 + Y_2 + Y_1}, \quad \frac{X_2 - Y_1 + Y_2}{X_2 + Y_2 + Y_1}, \quad \frac{(X_2 - Y_1 - Y_2)(X_2 + Y_1 + Y_2)}{(X_2 - Y_1 + Y_2)(X_2 + Y_1 - Y_2)} \right\} \quad (2.81)$$

and produce the differential in X_2 :

$$\begin{aligned} d_{X_2}\psi = & - \left(\frac{X_3 + Y_2}{X_1 + X_2 + X_3} \otimes \frac{X_1 + X_2 - Y_2}{X_1 + X_2 + Y_2} + \frac{X_1 + X_2 + Y_2}{X_1 + X_2 + X_3} \otimes \frac{X_3 - Y_2}{X_3 + Y_2} \right) d \log \left(\frac{X_2 + Y_1 - Y_2}{X_2 + Y_2 + Y_1} \right) \\ & - (X_1 \leftrightarrow X_3, Y_1 \leftrightarrow Y_2) \\ & \left(\frac{X_3 + Y_2}{X_1 + X_2 + X_3} \otimes \frac{X_1 + Y_1}{X_1 + X_2 - Y_2} + (X_1 \leftrightarrow X_3, Y_1 \leftrightarrow Y_2) \right) d \log \left(\frac{(X_2 - Y_1 - Y_2)(X_2 + Y_1 + Y_2)}{(X_2 - Y_1 + Y_2)(X_2 + Y_1 - Y_2)} \right) \end{aligned} \quad (2.82)$$

the all of these $d \log$'s have two-site chains as coefficients. In particular, if we denote that two-site chain with vertex energies X_L, X_R and exchange energy Y as

$\psi_2(X_L, X_R, Y)$ then we see that we have the differential

$$\begin{aligned}
d_{X_2}\psi &= -\psi_2(X_1 + X_2, X_3, Y_2)d\log\left(\frac{X_2 + Y_1 - Y_2}{X_2 + Y_2 + Y_1}\right) - (X_1 \leftrightarrow X_3, Y_1 \leftrightarrow Y_2) \\
&- \psi_2(X_1 + (X_2 - Y_2)/2, X_3 + (X_2 - Y_1)/2, (Y_1 + Y_2 - X_2)/2)d\log\left(\frac{(X_2 - Y_1 - Y_2)(X_2 + Y_1 + Y_2)}{(X_2 - Y_1 + Y_2)(X_2 + Y_1 - Y_2)}\right)
\end{aligned} \tag{2.83}$$

and so the total differential of the three-site chain is a sum over two-site chains with shifted energies times $d\log$'s of a single vertex energy. This grouping was far from manifest at the level of the symbol produced by the recursion. In particular, only ratios of two factors linear in edges and vertex energies ever appeared. Recognizing this decomposition into three canonical differentials was made possible by expressing the symbol in terms of these canonical letters. Indeed, this can even motivate a guess for the differential with respect to any middle vertex and furnish a recursive rule for constructing the symbol from differentials. This would furnish a more time-evolution centric construction, dovetailing beautifully with the unitarity and factorization centric picture presented earlier.

Comments on Compatibility and Uniqueness As stated earlier our choice of parameterization is not unique. In particular we could have chosen the opposite identification of $(\pm Y)$ with including or excluding y_e in the positive parameterization. When we choose the other parameterization, we just flip the sign of Y to find

$$\begin{aligned}
u_{1,3} &= \frac{X_1 - Y}{X_1 + X_2} & u_{2,5} &= \frac{X_2 - Y}{X_1 + X_2} \\
u_{1,4} &= \frac{X_1 + Y}{X_1 - Y} & u_{2,4} &= \frac{X_2 + Y}{X_2 - Y} \\
u_{3,5} &= \frac{-2Y(X_1 + X_2)}{(X_1 + Y)(X_2 + Y)}
\end{aligned} \tag{2.84}$$

Now the symbol is

$$S = (u_{1,3}u_{1,4}) \otimes u_{2,4} + (u_{2,5}u_{2,4}) \otimes u_{1,4} \quad (2.85)$$

We see that the $u_{1,4}$ and $u_{2,4}$ partial energy singularities are compatible, but an incompatible cross term still lingers, making this convention seem definitively worse. As it stands, there is no obvious choice, and it would be nice to understand what is the invariant content or at least if one manifests some nice properties.

2.8 Conclusion and Outlook

In this section we have derived a simple recursive rule for computing the symbols of perturbative contributions to the vacuum wave function. The rule applies exactly to a case of particular interest, $\lambda\phi^3$ in dS_4 . This case is interesting for reasons both concrete and speculative. It was demonstrated [10] that the four-point function of conformally coupled scalars exchanging a general massive scalar is sufficient to systematically generate four-point functions with massless external legs exchanging fields of general spin. Our case considered exchange of a conformally coupled scalar, from which one can obtain correlators of massless external legs exchanging conformally coupled fields with spin, which includes gauge fields and partially massless fields of spin 2. It would be interesting to explore generalizing the cosmological bootstrap tools to the five-point function computed here. The results presented here also apply to the computation of Witten diagrams in Euclidean AdS. It would be interesting to explore whether symbol recursion has a nice boundary interpretation in this setting, as the rule is manifest in momentum-space, an as yet unfamiliar setting for 3D Euclidean CFT. This specific case is also interesting due to its connection to flat-space physics. A differential operator acting on the total energy discontinuity of a given wave-function contribution produces the flat space scattering amplitude. It is known that a very rich geometric and algebraic structure describes $\lambda\phi^3$ amplitudes in flat

space and is connected to the geometry of the open string moduli space. It would be surprising if this structure only existed on a codimension one subspace of kinematic space.

More adventurously, this work opens the door to new tools and perspectives on probing the complementary notions of unitarity and time-evolution in de Sitter space and in particular how these are encoded in analytic properties of the de Sitter wavefunction. The symbol, literally read from left to right and from right to left captures these two properties respectively. The recursive rule makes the factorization structure clear, but the differential side is far from completely understood. The added structure of graph associahedra encoding the symbol letters motivate the de Sitter wave function as an exciting playground for probing how these two notions, always in tension, are encoded in a perhaps more abstract but canonical setting. We leave such exploration to future work.

Part II

$$\epsilon \neq 0$$

Chapter 3

The Divergence Structure of Feynman Integrals

Generically, the amplitudes in the four-dimensional quantum field theories we care about are divergent. The most familiar and well understood kind of divergence which arises is a UV divergence. In a momentum space representation of the amplitude, these divergences come from high momenta running in loops of the associated Feynman graphs. In the Feynman/Schwinger/world-line parameter representation of the integral, these divergences come from subgraphs of the Feynman graph shrinking and the associated edge parameters α_e going to zero. These high-momentum, short-distance divergences of scattering amplitudes can be studied once in the context of a “regulator”, which is simply a crude guess for the behavior of the short-distance physics which renders the amplitudes finite and allows one to begin calculating. In brief, the remarkable fact about UV divergences in renormalizable theories is that irrespective of the regulator, these divergences resum in a controlled fashion, allowing one to make invariant predictions in terms of the parameters of the renormalizable theory. This resummation is typically described in the context of the β -functions of the theory, which dictate how one should evolve the parameters of their theory to

effect the aforementioned resummation.

These non-trivial facts about observables are indirectly justified through the work of Wilson beautifully articulated later by Polchinski [6], where the arguments are made in the context of path integrals. The work argues using very much off-shell and Euclidean reasoning both for the existence of a Wilsonian β -function which effects the resummation capturing the change of physics with scale, and for the fact that values of irrelevant couplings at higher scales indeed become irrelevant as one evolves to a low scale. This work is a seminal contribution to the understanding of renormalization as it extinguishes two myths: that renormalization is a mere trick, and that the laws of Nature somehow care that we ought to be able to make all predictions using a finite menu of interactions in a quantum field theory. Despite the indisputable conceptual and technical successes relevant to renormalization in quantum field theory, the state of affairs is still somewhat disturbing. The fact of the matter is that despite the conceptual triumph of off-shell Wilsonian thinking, amplitudes and β -functions are typically computed in dimensional regularization. Such a divorce between the conceptual and calculational leaves something to be desired.

One could argue that a quantitatively and conceptually unified description of UV divergences and renormalization ought not be high on the priority list; we have the right picture of what is going on and can calculate essentially all we need to. Even if that is granted, an even greater motivation for obtaining a clear understanding of renormalization in the perturbative expansion is that we have no alternative for the case of IR divergences. There simply is no Wilsonian conceptual guiding light. What we need are formulas and observations about the structure of the perturbative expansion. To that end, in this chapter we consider the divergence structure of scalar Feynman integrals. We are able to fully characterize the ultraviolet (UV) and infrared (IR) divergence structure of a scalar Feynman graph in a very general kinematic regime. We emphasize the role of the tropical fan as the fundamental object characterizing

the divergence structure and demonstrate how the dual Newton polytope contains all of this information. We highlight the two essential differences in the characterization of UV versus IR divergences: the latter have a condition for divergence of subgraphs sensitive to external kinematics, and this sensitivity also carries through to the compatibility of subgraphs, which in the IR case does not obey the typical Zimmerman forest rules as in the UV.

After characterizing the divergences, we leverage our simple recipe for the calculation of leading divergences to illustrate the compatibility of different triangulations of the fan for UV divergences, one of which is the Zimmerman forest triangulation. We also calculate the leading divergence of $M \times N$ fishnet graphs and present a conjecture for how this can be extended to apply to all “rung-rule” graphs (which appear in the $\mathcal{N} = 4$ SYM perturbative expansion), and therefore may reduce the observance of the exponentiation of the leading divergence in $\mathcal{N} = 4$ SYM to a combinatorial question about such graphs. The work in this section sets us up to study all subleading divergences in chapter 4.

3.1 Symanzik Representation

Our starting place for the study of scalar Feynman integrals is the Symanzik representation. Beginning with a Feynman integral associated with some graph G with L loops and E edges in D spacetime dimensions.

$$\mathcal{A}_G = \int \prod_e \frac{d^D k_e}{\pi^{LD/2}} \frac{1}{p_e^2 + m_e^2} \quad (3.1)$$

By representing each propagating with a worldline time α_e via

$$\frac{1}{p_e^2 + m_e^2} = \int_0^\infty e^{-\alpha_e(p_e^2 + m_e^2)} \quad (3.2)$$

and performing the Gaussian integrals, we arrive at the Symanzik representation of the Feynman integral

$$\mathcal{A}_G = \int_0^\infty \frac{d^E \alpha_e}{\mathcal{U}^{D/2}} e^{-\mathcal{F}/\mathcal{U}} \quad (3.3)$$

where we have introduced two polynomials in the parameters α_e , the Symanzik polynomials. They are defined as

$$\mathcal{U} = \sum_{T^1} \prod_{e \notin T^1} \alpha_e \quad (3.4)$$

$$\mathcal{F} = \sum_{T^2} p_{T^2}^2 \prod_{e \notin T^2} \alpha_e + \left(\sum_e m_e^2 \alpha_e \right) \mathcal{U} \quad (3.5)$$

Where T^1 and T^2 denote spanning one-trees and spanning two-trees in the graph G , respectively. A spanning one-tree is a single connected tree which contains every vertex in the graph G . A spanning two-tree is a pair of disjoint, connected trees which together contain every vertex in G . The momentum p_{T^2} denotes the external momentum flowing between these two disjoint trees.

It is worth noting some of the obvious features of these polynomials, for example that \mathcal{U} and \mathcal{F} are homogeneous of degrees L and $L + 1$, respectively. Because of these homogeneity properties, we can consider a coordinate transformation of the form

$$\alpha_e \rightarrow \lambda \alpha_e \quad \alpha_{e'} \rightarrow \lambda \quad (3.6)$$

for some chosen $\alpha_{e'}$ and $e \neq e'$. We can then perform the integral of equation (3.3) over λ in order to arrive at

$$\mathcal{A}_G = \Gamma(d_G) \int \frac{d^E \alpha_e}{\text{GL}(1)} \frac{1}{\mathcal{U}^{D/2}} \left(\frac{\mathcal{U}}{\mathcal{F}} \right)^{d_G} \quad (3.7)$$

where the modding by $\text{GL}(1)$ denotes a choice of $\alpha_{e'}$ which we set equal to one and integrate the remaining variables from zero to infinity. We have also introduced the

degree of divergence of a Feynman graph d_G defined as

$$d_G = E - LD/2 \tag{3.8}$$

Strictly speaking, in order for this integral representation over λ to converge, we need $d_G > 0$ which is to say that we have a superficially convergent integral. Nonetheless, since this integration in λ is merely that of a Gamma function, the integral is defined for $d_G \leq 0$ in dimensional regularization with $D = 4 - 2\epsilon$ and furthermore via analytic continuation for e.g. power law divergences which are not regulated by an arbitrarily small deviation of the spacetime dimension. The integral representation in equation (3.7) will be our primary object of study.

Kinematic Region Before moving on to the classification of IR and UV divergences of scalar Feynman integrals, we need to state which kinematic region we will be working in. Previous work on this subject (e.g. [18]) has placed strict genericity constraints on the external momenta and internal masses, which in fact disallow the kinematics that furnish soft/collinear divergences. This work is particularly interested in the study of IR divergences, and therefore certainly relaxes the constraints on kinematics well beyond this, to include massless internal and external propagators. Nonetheless, there are graphs for which our results do not apply. The examples which have escaped the classification provided herein are particular Lorentzian configurations. The regime in which we operate is one in which all monomials in the \mathcal{F} polynomial are present which would be present if all kinematic invariants associated with spanning two-trees are nonnegative. This is the case for planar graphs, for instance.

3.1.1 Divergences

Our objective is to fully characterize the UV and IR divergence structure of scalar Feynman integrals. More precisely, by UV and IR divergences we mean divergences coming from the boundary of the integration region, not singularities in the bulk of the integration region. Singularities in the bulk of the integration region typically place codimension-1 or higher constraints on the external kinematics and are called Landau singularities. These are associated with thresholds for particle production or more generically anomalous thresholds, but they are beyond the scope of this work. We will always assume our kinematics are away from such thresholds so our integrand is single valued in the bulk of the integration region.

Recalling the representation of the propagator in equation (3.2), we see that asymptotically high-momentum, UV processes correspond to small α_e while IR processes correspond to large α_e . It is worth commenting that once we commit to a particular GL(1)-fixed form of (3.7), large or small α_e is defined with respect to the $\alpha_{e'}$ which was fixed to one.

Because we are interested in singularities coming from the boundary of the integration region, we must fully characterize the asymptotic behavior of the integrand in all directions of the integration region. This is a question we can ask for a general integral of a product of polynomials to general powers, but which will have a particularly nice answer in the case of Feynman integrals. Before articulating this simple solution in the case of Feynman integrals, we now describe the more general mathematics that Feynman integrals will naturally call home.

3.2 Newton Polytopes and Tropical Fans

Our starting point to motivate tropical fans and Newton polytopes will be the general toy integral

$$I = \int_{\mathbb{R}_+^n} d^n \log(X) X^a P(X)^{-c} \quad (3.9)$$

where we have introduced the notation

$$X^a = \prod_i X_i^{a_i} \quad (3.10)$$

and $P(X)$ is a polynomial in the X_i which does not produce singularities in the bulk of the integration region

$$P(X) = \sum_v c_v X^v \quad (3.11)$$

our Feynman integrals in the Symanzik representation clearly fall into this class. If we are interested in studying the asymptotic behavior of the integrand, it is natural to hope to linearize the problem by considering the change of variables $X_i = e^{-t_i}$

$$I = \int_{\mathbb{R}^n} d^n t e^{-t \cdot a} P(e^{-t})^{-c} \quad (3.12)$$

And we can ask along a particular ray $t \sim \lambda t$ for $\lambda > 0$, what is the leading behavior of the integrand. This is clearly

$$e^{c \text{Min}_v[t \cdot v] - t \cdot a} (1 + \dots) \quad (3.13)$$

where the (\dots) represents subleading terms and $\text{Min}_v[t \cdot v]$ denotes taking the minimum of $t \cdot v$ over all monomials in P . Note that given that choice of definition $X_i = e^{-t_i}$ the monomial(s) chosen by $\text{Min}_v[t \cdot v]$ is(are) minimally damped for large $|t|$. This clearly selects the leading term along this direction, and the integral will be convergent along

this direction so long as the exponent is less than zero. This exponent is sufficiently important that we will name it

$$\text{Trop}[X^a P(X)^{-c}](t) = c \text{Min}_v[t \cdot v] - t \cdot a \quad (3.14)$$

This is the tropicalization of the integrand. In the future we will often abuse notation and it should be clear from context what is being tropicalized and the relevant logarithmic t_i variables which the tropicalization is a function of. The global negativity of Trop guarantees convergence of the integral. Therefore, we would like to understand in the simplest terms the global structure of the Trop function on our integration domain.

The Min condition means that Trop is not linear on our integration domain, but is piecewise linear. The function therefore divides the space into top-dimensional domains of linearity: n -dimensional subspaces of the integration domain in which Trop is linear, which together tile the full integration domain. At this point we could stop and return to our integrals of the form (3.7). We can merely use the direct definition of the tropicalization of the integrand in equation (3.14), plonk down the many inequalities implied by the Min condition, and obtain the space they cut out. This space, on which Trop is piecewise linear, together with the value of Trop in the domains of linearity, fully characterizes the divergence structure of the integrand. This would check for global negativity by laboriously checking in each cone. We will find that such a brutal approach is in fact far too computationally complex, though for a random integrand, something roughly like it would be necessary. But Feynman polytopes are not random integrands. We will see that the best thing we could hope to find is a clean characterization of the primitive single rays which together bound and positively generate the cones. By checking the asymptotic behavior on these cones alone, we can check convergence. This is what we now describe.

It is not difficult to see that the domains of linearity are associated with particular v 's. That is, if we choose a particular v , we might ask: what is the set of t 's for which

$$\text{Trop}_v = t \cdot v - t \cdot a \quad (3.15)$$

For a given v this space can be less than n -dimensional or even empty, but it is easy to see that the space is indeed a cone, as the Min condition is respected under positive linear combinations, and so we denote it as

$$C_v := \{t \cdot v \leq t \cdot v'\} \quad (3.16)$$

Some subset of v 's in the polynomial P furnish the top-dimensional cones which together tile the space. As these cones tile the space, there are lower-dimensional boundaries between top dimensional cones, the lowest dimensional of which are merely rays. Any given cone is bounded by such *extremal rays*. Indeed, the specification of all the extremal rays in addition to the compatibility conditions which group them into cones, completely specifies the *tropical fan*.

Tropical Fan and Dual Newton Polytope Thus far, we have been describing the properties of the space in which t lives, motivated by understanding the asymptotic behavior of the integrand at large $|t|$. But it is natural to discuss the tropical fan which we want to describe in the context of the dual space in which positive linear combinations of the v 's live. This space is the translation by a of the Newton polytope of the polynomial P , which is the convex hull of the exponent vectors of P

$$-a \oplus c\text{Newt}[P] = -a \oplus c \left\{ \frac{\sum \lambda_v v}{\sum \lambda_v} \mid \lambda_v \geq 0 \right\} \quad (3.17)$$

and this space is clearly dual to the t space. The space is a convex hull of points and in fact defines a polytope. A top dimensional cone C_v in t space corresponds to a vertex v of this polytope. The extremal rays r in t space correspond to directions along which many C_v 's are compatible. These are therefore facets of the Newton polytope, and must furnish a facet inequality. This is readily obtained by evaluating Trop on this ray r

$$\text{Trop}[r] = \text{Min}_v[r \cdot v] - r \cdot a \quad (3.18)$$

which in particular implies the inequality for the Newton polytope

$$r \cdot p \geq \text{Trop}[r] \quad (3.19)$$

for all p in $a \oplus c\text{Newt}[P]$. Note that the constant in the inequality above is the constant characterizing the convergence behavior of the integrand along the ray r . Therefore, knowledge of the facet presentation of the Newton polytope of P , in particular the facet inequalities and compatibilities, fully characterizes the divergence structure of our integral I .

Minkowski Sums The integrands we will be interested in for Feynman integrals will not consist of a single polynomial but two polynomials, \mathcal{U} and \mathcal{F} raised to generic powers. We therefore need to suitably generalize the notions discussed above. The generalizations proceed in the obvious way. Given a set of Newton polytopes $\text{Newt}[P_i]$, the Minkowski sum

$$\bigoplus_i c_i \text{Newt}[P_i] \quad (3.20)$$

is simply given by the set of points obtained by taking a point in each $\text{Newt}[P_i]$ added componentwise with the weighting c_i . This Newton polytope is the same as that

for the associated product of polynomials raised to those powers. For example,

$$\text{Newt}[P_1^{c_1} P_2^{c_2}] = c_1 P_1 \oplus c_2 P_2 \quad (3.21)$$

This only strictly makes sense for $c_i > 0$, but we will not be concerned, strictly speaking, with the precise convex hull. What we are concerned about is the combinatorics of the tropical fan: what its extremal rays are, what the compatibilities are, and then the value of Trop on those rays. These clearly only depend on the monomials in the polynomials, not the signs in the Minkowski sum. Said differently, we could imagine the exponents are positive, but then notice that everything of value which we calculate is obviously smooth in these powers.

Example: In order to illustrate how the myriad pieces fit together, we consider the example

$$I_{\text{tri}} = \int_0^\infty d \log(x_1) d \log(x_2) x_1^a x_2^b (1 + x_1 + x_2)^{-a-b-c} \quad (3.22)$$

which evaluates to

$$I_{\text{tri}} = \frac{\Gamma(a)\Gamma(b)\Gamma(c)}{\Gamma(a+b+c)} \quad (3.23)$$

It is simple to directly write down the Newton polytope associated to the integrand multiplying the $d \log$ measure, and this is depicted in figure 3.1. It is also trivial to directly write down the tropicalization of the integrand

$$\text{Trop} = (a+b+c)\text{Min}[0, t_1, t_2] - at_1 - bt_2 \quad (3.24)$$

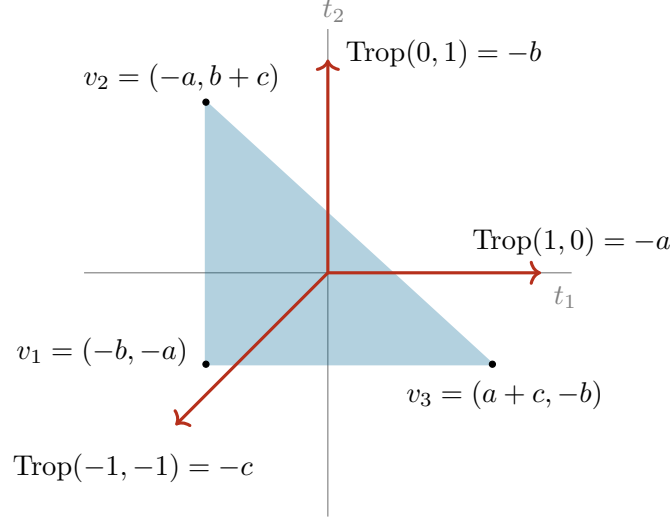


Figure 3.1: The Newton polytope (shaded blue) associated with the integrand in (3.22). The value of the tropicalization of the integrand along the inward pointing facet normal defines the constant in the inequality for that facet. These inequalities are in (3.25). The integral converges when the value of the tropicalization is everywhere negative and therefore this Newton polytope properly contains the origin.

It is readily confirmed that the facet presentation of the polytope in figure (3.1) is given by

$$\begin{cases} -p_1 - p_2 & \geq -c \\ p_1 & \geq -a \\ p_2 & \geq -b \end{cases} \quad (3.25)$$

Where $p = (p_1, p_2)$ are the coordinates for a point in the Newton polytope. For each inequality, the coefficients of the p_i define a dual object, the inward pointing normal to the facet. These are depicted in red in figure (3.1). The non-trivial fact, trivial to check in this instance, is that these facet normals are precisely the extremal rays of the tropical fan defined by the tropicalization of the integrand (3.24), and we see that evaluating Trop on these inward pointing normals indeed yields the constants in the inequalities of (3.25). In this way, we see that if we have the facet presentation of the Newton polytope associated with the integrand of (3.7), then we have the complete characterization of the divergence structure of the integral in its simplest form. We

finally briefly comment on the notion of a Minkowski sum. Given a newton polytope P_1 , which is described as a convex hull of vertices (3.17), we can Minkowski sum it with another polytope P_2 . This is done in the obvious way, by simply taking the hull generated by adding any point in P_1 to any point in P_2 . The generalization to arbitrarily many polytopes is clear. Instead of taking the Minkowski sum, one can also simply compute the Newton polytope of the product of the underlying polynomials directly. This will obviously yield the same result. Later on, we will be interested in the Newton polytope of \mathcal{UF} , as it contains all of the information we are interested in regarding the divergence structure of our integral.

3.3 Calculation of Leading Divergences

Characterizing the divergence structure of a Feynman integral is valuable, but we can be forgiven for having greater ambitions. In particular, we can at least hope to calculate the leading logarithmic divergence of a graph, i.e. the leading divergences where Trop is regulated by a parameter arbitrarily close to zero. Before stating how this is systematically done in general, we can continue with the example integral (3.22) to get a sense of how a systematic means of extracting leading divergences from complicated integrals can come about.

We will first seek to probe the divergence coming from $c \rightarrow 0$ with a and b generic and ensuring convergent behavior otherwise. This tells us that the divergence comes from when x_1 and x_2 are getting large (note that the ray in figure 3.1 indicates $(-t_1, -t_2) = (-1, -1)$ therefore x_1 and x_2 are growing). We can consider the limiting behavior of the integrand in this limit

$$I_{\text{tri}}^c = \int d\log(x_1)d\log(x_2)x_1^a x_2^b (x_1 + x_2)^{-a-b-c} \quad (3.26)$$

where we have been deliberately ambiguous about the integration region. Note that for $c = 0$, this integrand develops a rescaling symmetry under $x_1 \rightarrow \lambda x_1$ and $x_2 \rightarrow \lambda x_2$

$$x_1 \rightarrow \lambda x_1 \quad x_2 \rightarrow \lambda x_2 \quad (3.27)$$

and the integrand is homogenous under this rescaling for finite c . This is a generic phenomenon; taking the limiting behavior of an integrand along a tropical ray, the integrand develops a homogeneity along that direction, as we have only kept monomials with leading weight under the associated rescaling.

We can make the implications of this more transparent with a coordinate transformation. Consider the change of variables in (t_1, t_2) space

$$\begin{pmatrix} t_1 \\ t_2 \end{pmatrix} = w \begin{pmatrix} 1 \\ 1 \end{pmatrix} + \bar{w} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \quad (3.28)$$

Now we revisit the question of the integration region. We are interested in calculating the leading divergence as $c \rightarrow 0$ and the limit of the integrand we have taken is clearly only justified for $w > 0$. And so we have

$$I_{\text{tri}}^c = |\alpha - \beta| \int_0^\infty dw e^{-cw} \int d\bar{w} e^{(a\alpha+b\beta)\bar{w}} (e^{\alpha\bar{w}} + e^{\beta\bar{w}})^{-a-b} \quad (3.29)$$

Note that we have set $c = 0$ in the \bar{w} part of the integrand. This is because the \bar{w} integral is now regular around $c = 0$ and any c -dependence would be subleading in our small c limit. We see that the result of (3.29) is

$$I_{\text{tri}}^c = \frac{1}{c} \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)} \quad (3.30)$$

which is of course the correct leading term of the Laurent expansion of (3.23) around $c = 0$. We used generic α and β to illustrate that so long as the determinant of the

jacobian did not vanish, we could choose whatever direction we like, as the integration is dominated by large w and the integrations decouple no matter what. As such, it is ideal to make a convenient choice which makes transparent some convenient features of (3.29). But crucial here is that the ω and $\bar{\omega}$ integration decoupled. This is a generic aspect of our prescription as any entangled aspect of the integration region is sent to zero at large ω and will again contribute subleading terms in our expansion, which we are not calculating here.

Also notice that the w integral merely produces a factor of $\frac{1}{\text{Trop}}$ on the ray in question. If we chose $\alpha = 0$ and $\beta = 1$ then the coordinate transformation has unit jacobian and is essentially gauge-fixing the λ rescaling symmetry by setting $x_1 \rightarrow 1$ to yield the \bar{w} integral. We can write this as

$$I_{\text{tri}}^c = \frac{1}{\text{Trop}} \int_0^\infty \frac{d \log(x_1) d \log(x_2)}{\text{GL}(1)} x_1^a x_2^b (x_1 + x_2)^{-a-b} \quad (3.31)$$

which suggests the generalization. The pointed, half-space integration along the divergent rays will always produce inverse products of the value of Trop along those rays. This is responsible for the poles in the regulator, the “divergent” part of the integral. What will remain is a simplified integral, which is the limiting behavior of the integrand in this divergent cone under consideration, with the enhanced GL(1) scaling symmetries of the integrand fixed, yielding a finite, lower dimensional integral.

3.3.1 Systematics of the General Case

We can now state in generality how to compute the leading divergence. We consider an integral

$$I = \int_{\mathbb{R}^n} d^n \lambda R(e^\lambda) \quad (3.32)$$

where we have used the logarithmic variables λ . We can derive the formula for the case of a single top-dimensional divergent cone as the full answer is obtained by summing over all top dimensional cones. In an n -dimensional space, and m -dimensional divergent cone C_F is generated by a set of $m < n$ extremal rays $\{w^{(i)}\}$ where the index labels the ray, not its components. We assume that we are working in a triangulation of the cone if it is not simplicial. The F denotes that this cone is dual to some $n - m$ dimensional face F of the dual Newton polytope. We can consider a coordinate transformation of the t variables

$$t_i = \lambda_j w^{(j)} + \bar{\lambda}_k \bar{w}^{(k)} \quad (3.33)$$

where the repeated indices are summed over and we introduce $n - m$ vectors $\bar{w}^{(k)}$. These vectors are not canonically determined, they are merely a choice, though we will see that some choices are simpler than others. First we note that the jacobian for the change to $\lambda, \bar{\lambda}$ variables is

$$J = |w^{(1)} \dots w^{(m)} \bar{w}^{(1)} \dots \bar{w}^{(n-m)}| \quad (3.34)$$

Because we are considering the leading behavior associated with the m -dimensional cone C_F , the limiting behavior of the integrand becomes homogenous under λ rescalings and the integration in the λ variables is pointed i.e. $\lambda > 0$. In particular the integral behaves as

$$I \rightarrow |w^{(1)} \dots w^{(m)} \bar{w}^{(1)} \dots \bar{w}^{(n-m)}| \int_{\mathbb{R}_+^m} d^m \lambda e^{-\lambda_i \text{Trop}[w^{(i)}]} \int_{\mathbb{R}^{n-m}} d^{n-m} \bar{\lambda} \tilde{R}(e^{\bar{\lambda}_j}) \quad (3.35)$$

Where \tilde{R} means we only keep the monomials of R which are leading in this limit. These are monomials whose vertices lie on all the facets $w^{(i)}$. And so we see that this

contribution is

$$I_{C_F} = \frac{|w^{(1)} \dots w^{(m)} \bar{w}^{(1)} \dots \bar{w}^{(n-m)}|}{\prod_i \text{Trop}[w^{(i)}]} \int_{\mathbb{R}^{n-m}} \frac{d^n t}{\text{GL}(1)^m} \tilde{R}(e^t) \quad (3.36)$$

where we introduce the short hand notation for the measure in the remaining finite integral to indicate that we take the limiting behavior of the integrand, and gauge fix the m new rescaling symmetries of this limiting integrand. The jacobian remains out front as a particular choice of $\text{GL}(1)^m$ might carry numerical factors. In the case of Feynman integrals, it is most convenient to both define the of scale rays $\omega^{(i)}$ such that we get the constants in the inequalities defined in the subsequent section and then a gauge-fixing such that the jacobian is one.

This is all the general technology we will need for the calculation of leading divergences of Feynman integrals. The integrand is composed of the two Symanzik polynomials \mathcal{U} and \mathcal{F} . We will identify the leading divergences as coming from certain cones, generated by their associated rays. We can then apply (3.36). The remaining work is therefore the clear articulation of the facet presentation of the Newton polytope of \mathcal{UF} . This what we turn our attention to now.

3.4 Feynman Polytopes

In order to state the facet presentation of the Newton polytope relevant to our study of Feynman integrals, we will need to introduce some operations on graphs. The first notion we introduce is modding a graph G by a subgraph γ to produce G/γ the graph wherein all connected components of γ have been shrunk to single vertices. The momenta that now flow into the resultant vertices are added linearly. This is depicted in an example in figure 3.2. In the modded graph G_B/γ_{14} the resultant vertex now has external momentum equal to the sum of the three massive external momenta. By momentum conservation, this equals minus the remaining massless

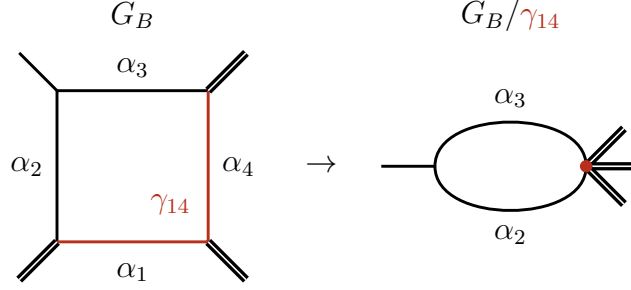


Figure 3.2: A subgraph γ_{14} such that $\mathcal{F}_{G_B/\gamma_{14}} = 0$.

external momentum and therefore squares to zero, making G_B/γ_{14} *scaleless* i.e.

$$\mathcal{F}_{G_B/\gamma_{14}} = 0 \quad (3.37)$$

scaleless G/γ 's will be the stars of our study of IR divergences.

Another graphical notion of value is that of one-vertex irreducibility (1VI). If a graph cannot be disconnected by removal of a single vertex, it is said to be (1VI). With these notions we are now equipped to state the facet presentation of the Newton polytope and therefore the tropical fan. The facet presentation of \mathbf{U}_G is

$$\mathbf{U}_G: \quad a_G = L_G \quad \text{and} \quad a_\gamma \geq L_\gamma \quad \text{for all } \gamma\text{'s}. \quad (3.38)$$

Analogously, \mathbf{F}_G is cut out by

$$\mathbf{F}_G: \quad a_G = L_G + 1 \quad \text{and} \quad \begin{cases} a_\gamma \geq L_\gamma & \text{if } \mathcal{F}_{G/\gamma} \neq 0, \\ a_\gamma \geq L_\gamma + 1 & \text{if } \mathcal{F}_{G/\gamma} = 0, \end{cases} \quad (3.39)$$

for all γ 's.

And the inequalities cutting out the polytope \mathbf{S}_G , the primary object of our interest, are obtained simply by adding the inequalities for \mathbf{U}_G and \mathbf{F}_G .

Facets and Factorization The inequalities as stated above are not minimal. That is, some γ 's will not correspond to codimension-one faces, i.e. facets. We now describe the conditions for facets in \mathbf{S}_G . For the $\mathcal{F}_{G/\gamma} \neq 0$ case, facets correspond to 1VI subgraphs. In the $\mathcal{F}_{G/\gamma} = 0$ case, we find that the facet condition agrees with the motivic condition introduced by Brown [18], which is that for Γ to be a subgraph, all subgraphs $\gamma \subset \Gamma$ with which also satisfy $\mathcal{F}_{G/\gamma} = 0$ must have $L_\gamma < L_\Gamma$ ¹. In the $\mathcal{F}_{G/\gamma} \neq 0$ case, the polynomials and the geometry factorize as well

$$\begin{aligned}\mathcal{U}_G &\rightarrow \mathcal{U}_\gamma \times \mathcal{U}_{G/\gamma} \\ \mathcal{F}_G &\rightarrow \mathcal{U}_\gamma \times \mathcal{F}_{G/\gamma} \\ \mathbf{S}_G &\rightarrow \mathbf{U}_\gamma \times \mathbf{S}_{G/\gamma}\end{aligned}\tag{3.40}$$

In the $\mathcal{F}_{G/\gamma} = 0$, there is no such factorization. This one of the first key differences to be observed between IR and UV divergences at the level of Feynman polytopes. Amusingly, previous works have stated essentially the same inequalities as (3.39)[19], but for Feynman integrals in a very restricted kinematic regime, one which precludes the study of IR divergences. Moreover, these works did not state a compatibility condition, likely because they were only considering UV divergences and it suffices to employ Zimmerman's forest rules, which are stronger than the compatibility condition for \mathbf{S}_G and indeed constitutes a specific triangulation of the fan. But this does not apply for the case of infrared divergences. It is worth stressing that nothing here is of any value without understanding compatibility. Without compatibility one cannot even state what the degree of the leading divergence will be, let alone calculate its coefficient. In order to have a complete *ab initio* description of the tropical fan, compatibility is therefore crucial and we describe it in the next section. First, we

¹Brown's work, which proved this condition, was in a restricted set of kinematics which do not apply to the IR divergences of interest here. We do not have a proof here, but observe it in our extended kinematic regime.

can now employ (3.19) in order to obtain the value of the tropicalization of our the integrand (3.3) along any ray associated with a subgraph γ .

In order to perform this simple evaluation, it is convenient to present the integrand of (3.3) in the form of (3.9). That is,

$$\frac{d^E \log(\alpha_e)}{\text{GL}(1)} \left[\Pi_\alpha^{-\frac{E}{L}} \mathcal{U} \right]^{d_G - D/2} \left[\Pi_\alpha^{-\frac{E+1}{L}} \mathcal{F} \right]^{-d_G} \quad (3.41)$$

where we introduced the notation

$$\Pi_\alpha = \prod_e \alpha_e \quad (3.42)$$

it is a trivial exercise to check that this is a mere rewriting of (3.3). It is also readily verified that sprinkling these powers of Π_α serves to shift the Newton polytopes to live on a common hyperplane $a_G = 0$. In this case the shifted inequalities become for \mathbf{U}_G

$$a_\gamma \geq \frac{L_\gamma E_G - L_G E_\gamma}{E_G} \quad (3.43)$$

and for \mathbf{F}_G :

$$a_\gamma \geq \begin{cases} \frac{L_\gamma E_G - (L_G + 1) E_\gamma}{E_G} & \text{if } \mathcal{F}_{G/\gamma} \neq 0, \\ \frac{(L_\gamma + 1) E_G - (L_G + 1) E_\gamma}{E_G} & \text{if } \mathcal{F}_{G/\gamma} = 0, \end{cases} \quad (3.44)$$

Finally, leveraging (3.19) in the two cases we find

$$\text{Trop}(\gamma) = \begin{cases} -d_\gamma & \text{if } \mathcal{F}_{G/\gamma} \neq 0, \\ d_{G/\gamma} & \text{if } \mathcal{F}_{G/\gamma} = 0 \end{cases} \quad (3.45)$$

a beautiful result which places IR and UV divergences on parallel footing. At the level of single rays, this characterizes the divergence structure. But for complicated graphs, a mere characterization of single directions is essentially useless. The leading divergences are those higher dimensional cones in the integration region generated by the maximal number of compatible divergent rays. Again, for this reason is it crucial to understand the compatibility condition between rays.

Compatibility Building upon the statement of inequalities and evaluation of Trop along any ray in the fan, we finalize our complete understanding of the fan by stating the compatibility criterion. We introduce the function z_γ which is simply the constant in the inequality as defined above. For \mathbf{S}_G these are

$$z_\gamma = \begin{cases} 2L_\gamma & \text{if } \mathcal{F}_{G/\gamma} \neq 0 \\ 2L_\gamma + 1 & \text{if } \mathcal{F}_{G/\gamma} = 0 \end{cases} \quad (3.46)$$

In terms of the function z_γ for the polytope \mathbf{S}_G we claim that the compatibility condition is

$$z_{\gamma_1} + z_{\gamma_2} \geq z_{\gamma_1 \cap \gamma_2} + z_{\gamma_1 \cup \gamma_2} \quad (3.47)$$

This compatibility condition is in tension with supermodularity, which would instead have a \leq in the expression above. Polytopes cut out by inequalities of the form above where z_γ obeys supermodularity are proven biconditionally to be generalized permutahedra. In particular, graphs with totally generic kinematics are generalized permutahedra, in which case the compatibility condition above must reduce to an equality. More interestingly, graphs in which two compatible subgraphs γ_1 and γ_2 realize (3.47) as a *strict* inequality demonstrate that \mathbf{S}_G is not a generalized permutahedron. Such graphs are plentiful and are associated with the presence of soft-collinear divergences.

Generalized permutahedra aside, the inequalities and compatibility condition constitute a complete characterization of the divergence structure of scalar Feynman integrals. Crucially, this characterization is articulated in a form which is immediately useful for the calculation of such divergences. This is what we will turn our attention to in chapter 4. Before doing so, it is valuable to review a variety of statements that can already be made about the nature of UV and IR divergences without calculating any amplitudes at all.

3.4.1 Summary of Feynman Polytopes

Here we succinctly summarize the relevant facts about the tropical fan associated with the integrand of (3.3). The relevant data defining the tropical fan are contained in table 3.4.1. The information therein as well as the compatibility condition in (3.47)

	$\mathcal{F}_{G/\gamma} \neq 0$	$\mathcal{F}_{G/\gamma} = 0$
extremal rays	γ 1VI	γ motic
Trop	$-d_\gamma$	$d_{G/\gamma}$
z_γ	$2L_\gamma$	$2L_\gamma + 1$

Table 3.1: Table of relevant data characterizing the tropical fan associated to a Feynman integral. In particular the specific coefficient 2's in the last line come from choosing $\mathbf{U} \oplus \mathbf{F}$, which provides equivalent data. This presentation puts UV and IR divergences on as a symmetric footing as possible. And distills the essential differences, which are reducible to the difference in z_γ .

completely define the divergence structure of a Feynman integral. In particular, leveraging knowledge of the $\text{Trop} = 0$ extremal rays as well as compatibility, one can determine the highest dimensional divergent cone in the fan. This dimension is the leading pole in ϵ in the Laurent expansion of the integral.

It is also worth noting that everything is derivable from the difference in z_γ . The

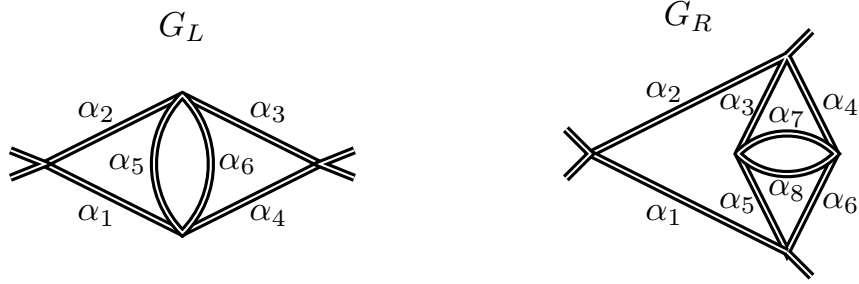


Figure 3.3: Two UV-divergent Feynman diagrams G_L and G_R considered in (3.48) and (3.50). These are each classic examples of overlapping divergences. The ugliest manner in which this can rear its head is that the divergent subspace for G_R is not simplicial and requires triangulation (see Fig. 3.4).

inequalities are of the same form and the compatibility condition is of the same form in all cases. The only difference is the value of z_γ , a difference in value sensitive to external kinematics. In the setting of Newton polytopes, we can therefore distill the difference between UV and IR divergences to the presence of a (+1) in this simple function on the graph.

3.4.2 Divergences

Equation (3.45) manifests some historically more non-trivial results, and can be leveraged to make further statements. Immediately clear is Weinberg's theorem that a superficially convergent ($d_G > 0$) graph with no subdivergences ($d_\gamma > 0$ for all γ) is UV finite. And there is of course the analogous statement for IR divergences: a graph with $d_{G/\gamma} < 0$ for all γ such that $\mathcal{F}_{G/\gamma} = 0$ is IR finite.

3.4.3 UV Divergences

We now provide some illustrative examples. As a UV example, we consider the diagram G_L in Fig. 3.3 (left). It contributes a leading logarithmic divergence in $D = 4 - 2\varepsilon$ and therefore has a two-dimensional divergent $\text{Trop} = 0$ space (recall that one divergence comes from the overall $\Gamma(d_{G_L}) = \frac{1}{3\varepsilon} + \dots$). In order to calculate the leading divergence, we first need to identify the divergent rays spanning the $\text{Trop} = 0$

space. These correspond to the superficially divergent subgraphs, of which there are three: γ_{1256} , γ_{3456} , γ_{56} where the subscripts denote the edges of the subgraph. Next, we need to know the compatibility of these divergent rays in \mathbf{S}_{G_L} . This is given by the compatibility criterion (3.47) and demonstrates that γ_{56} is compatible with each of the other two subgraphs, but γ_{1256} and γ_{3456} are not compatible with each other. This means that we have two two-dimensional cones and they share a ray. We must compute the contribution to the leading divergence from each two-dimensional cone separately and add them.

We may now leverage (3.36) to obtain

$$I_{G_L} = \Gamma(d_{G_L}) \frac{1}{d_{\gamma_{56}}} \left(\frac{1}{d_{\gamma_{1256}}} + \frac{1}{d_{\gamma_{3456}}} \right) + \dots, \quad (3.48)$$

where the finite integral is accounted for as it is merely a product of bubble integrals which equal one. Since $d_{\gamma_{1256}} = d_{\gamma_{3456}} = 2\varepsilon$ and $d_{\gamma_{56}} = \varepsilon$, the leading UV divergence contributes

$$I_{G_L} = \frac{1}{3\varepsilon^3} + \dots \quad (3.49)$$

The structure of the rays in this example captures the familiar RG intuition of sequentially shrinking loops. One simplicial cone is generated by rays r_{1256} and r_{56} and captures the shrinking of the two corresponding subgraphs, one nested in the other. Any ray in this cone corresponds to a direction of scaling integration variables where the bubble γ_{56} shrinks at least as fast as the collective γ_{1256} , capturing the physics of shrinking successive UV subgraphs to effective vertices.

Leading divergences and the RG. In general, the maximal divergent cones need not be simplicial. A given cone's contribution then need not be interpreted in the familiar nested shrinking picture consistent with our familiar RG intuitions, or in

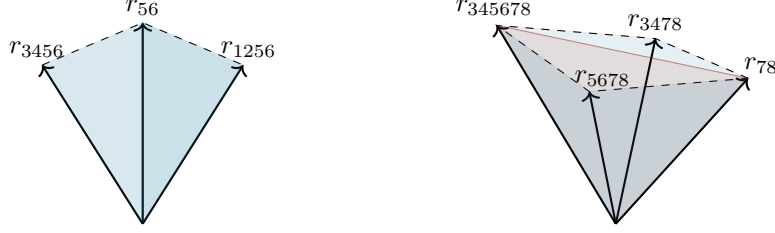


Figure 3.4: Divergent rays of the diagrams from Fig. 3.3. Left: Two two-dimensional cones which share a single ray r_{56} , reflecting the fact that there is an overlapping divergence. Right: Triangulation of the three-dimensional cone into two simplicial cones separated by the red plane. In this triangulation each simplex has the interpretation of nested shrinking subgraphs with ordering implied by the rays.

particular the rules for summing over divergences prescribed by BPHZ. Nonetheless, there does of course exist such a BPHZ triangulation. Consider the diagram G_R from Fig. 3.3 (right), whose rays are depicted in Fig. 3.4 (right) and grouped into simplicial cells by a triangulation. This triangulation is the BPHZ triangulation, as it groups the rays into simplices equivalent to imposing the additional BPHZ compatibility criterion, which is being either nested or disjoint subgraphs. In particular, r_{3478} and r_{5678} are no longer in the same simplicial cell in this triangulation. This BPHZ triangulation realizes an intuitive ordering of scaling rate. That is, any ray in the simplex generated by $r_{345678}, r_{3478}, r_{78}$ has components γ_{78} shrink the fastest, followed by γ_{34} and then γ_{56} . It is interesting that this nested shrinking is one natural way of organizing the calculation of the leading the divergence, though not uniquely so. Later on, when confronting subleading divergences, we will find that the BPHZ triangulation emerges as canonical for deeper reasons than this. Using (3.13), we can verify that the two triangulations produce the same leading contribution:

$$\begin{aligned} & \frac{1}{d_{\gamma_{345678}} d_{\gamma_{78}}} \left(\frac{1}{d_{\gamma_{3478}}} + \frac{1}{d_{\gamma_{5678}}} \right) \\ &= \frac{1}{d_{\gamma_{3478}} d_{\gamma_{5678}}} \left(\frac{1}{d_{\gamma_{78}}} + \frac{1}{d_{\gamma_{345678}}} \right) = \frac{1}{3\varepsilon^3}, \end{aligned} \tag{3.50}$$

and hence the leading contribution in this case is

$$I_{G_R} = \Gamma(d_{G_R}) \frac{1}{3\varepsilon^3} + \dots = \frac{1}{12\varepsilon^4} + \dots \quad (3.51)$$

Similarly, the phenomenon of overlapping divergences for the diagram G_L is illustrated in Fig. 3.4 (left), where the ray r_{56} is shared between two two-dimensional cones, but presents no confusion in this language. We see that when the study of the UV divergence structure of Feynman integrals is framed in the language of the tropical fan, not only is the issue of overlapping divergences in the worst case merely a question of triangulating a non-simplicial fan, the calculation of the leading divergence is itself made totally transparent. When calculating subleading divergences, we will see that the tropical fan and a canonical refinement of it make the solution of the overlapping divergences problem at all subleading orders a triviality.

Example and Feynman Polytopes as Blowups Another perspective one can take on the function of the tropical fan and the dual Newton polytope is that it provides a “blowup” of the boundary of the integration region on which the integrand is now single-valued. To understand the significance of this, we will first illustrate the problem of single-valuedness on the boundary of the integration region. In order to do this, we consider a low-dimensional ultraviolet example, the sunrise graph in $D = 4 - 2\varepsilon$ dimensions with masses $m_e = 1$ and no external kinematics depicted in fig. 3.5.

The Symanzik polynomials in this case are

$$\mathcal{U} = \alpha_1\alpha_2 + \alpha_1\alpha_3 + \alpha_2\alpha_3 \quad \mathcal{F} = (\alpha_1 + \alpha_2 + \alpha_3)\mathcal{U} \quad (3.52)$$

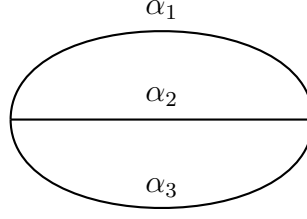


Figure 3.5: Sunrise graph with equal nonzero masses and no external kinematics.

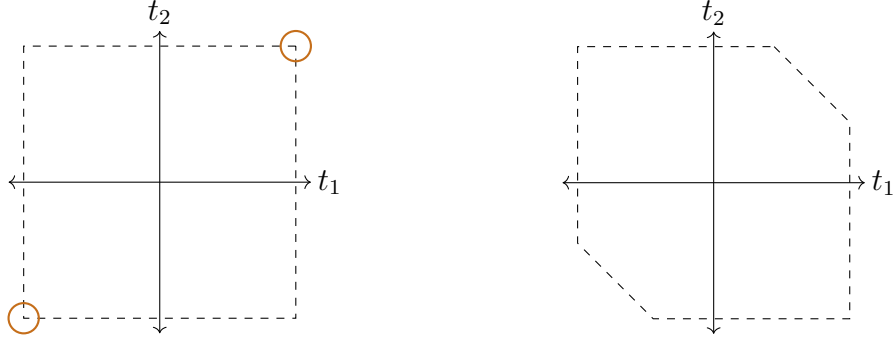


Figure 3.6: The integration region in logarithmic Symanzik variables (left) is simply a square. The integrand does not have single valued limits as it approaches the circled corners, as shown in (3.55). These corners must be blown up (right). The necessary blow-up is captured by the Newton polytope and its dual normal fan.

and the integral with $\alpha_3 \rightarrow 1$

$$I_G = \int d\alpha_1 d\alpha_2 \frac{\mathcal{F}^{-1+2\epsilon}}{\mathcal{U}^{2-3\epsilon}} \quad (3.53)$$

It suffices to set $\epsilon = 0$ for this analysis. Now we consider two different limits to $(\alpha_1, \alpha_2) = (0, 0)$. One way of probing this issue is to observe that the iterated residue in α_1 and α_2 do not commute, and so there should be a face between these two boundaries. We can also probe it parameterizing different one-parameter limits to the same point. Consider the change of variables $(\alpha_1, \alpha_2) = (\rho, \rho^k \alpha)$ for positive integer k in which the integrand is

$$\Omega = d\log(\rho) d\log(\alpha) \frac{\rho^{k+1} \alpha (1 + \rho + \rho^k \alpha)}{(\rho^{k+1} \alpha + \rho + \rho^k \alpha)^2} \quad (3.54)$$

with α generic, the $\rho \rightarrow 0$ limit probes the point of interest. We see that

$$\text{Res}_{\rho \rightarrow 0} \Omega = \begin{cases} \frac{d\alpha}{(1+\alpha)^2} & k = 1 \\ 0 & k > 1 \end{cases} \quad (3.55)$$

This reveals that our variables α_e do not furnish a global description of the boundary of the integration region on which Ω is single valued. The integer k defines different trajectories approaching the $(0,0)$ and the behavior of the integrand varies with k . We discover that there is a new distinguished direction in this neighborhood, on which the behavior of the integrand is not a mere composition of the small α_1 and small α_2 behavior. This new ray has an associated dual facet, which is a blow-up of the corner as depicted in fig. 3.6. The integrand has a residue on this facet, which is the $k = 1$ result of (3.55).

Mechanically doing blowups by hand would be a prohibitively laborious task. By obtaining a simple description of the tropical fan, we automatically perform the blowup of the boundary. Such an efficient characterization of this blowup, but in globally well-defined variables, is what will be the bedrock of the renormalization prescription discussed later on in chapter 4.

3.4.4 IR Divergences

In which dimensions do IR divergences exist? This discussion allows us to bound the dimensions in which IR divergences appear. In particular, we can ask: what is the highest dimension in which a marginal IR divergence may exist? For $n > 3$, we will maximize $E - LD/2$ when we have a bubble filled with trivalent vertices. In this case, with $2v$ trivalent vertices, we have

$$d_{G/\gamma} = \frac{(4 - D) + v(6 - D)}{2}. \quad (3.56)$$

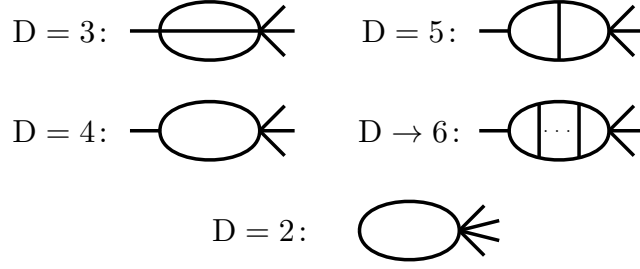


Figure 3.7: Leading G/γ graphs which are scaleless and marginal in the relevant dimensions. We write $D \rightarrow 6$ as marginality is only attained in the limit of an infinite number of rungs.

We see that as $v \rightarrow \infty$ then $D \rightarrow 6$ becomes marginal, therefore one will not find IR divergences in six or higher dimensions. Since adding numerators can only soften IR divergences, these bounds apply to a general theory. Fig. 3.7 catalogues the leading examples of marginal scaleless graphs in the relevant dimensions.

It is worth noting that the existence of marginal IR divergences in a given dimension implies the existence of power IR divergences in lower dimensions. In particular, scalar theories such as ϕ^3 exhibit power IR divergences in $D = 4$, giving an additional *raison d'être* for numerators in gauge theories: softening IR divergences to logarithmic.

Fishnet diagrams. For scalar integrals in $D = 4$, it is natural to study the generalization to *fishnet* diagrams, i.e., planar arrays of $M \times N$ squares, see Fig. 3.9 for $M = N = 4$ with two massless corners. Calculation of the divergent contribution from a graph is performed by identifying the rays generating divergent cones and applying (3.36). In the case of IR-divergent fishnets, the combinatorics of these rays is rich, but admits a simple description. In space time, we associated the infrared divergences with corners of the fishnet, where internal edges meet external momentum. In the α_e parameters, these can be associated to shrinking subgraphs as described hitherto. The shrinking subgraphs which furnish an infrared divergence emanate from the diagonally opposed corner to the one we familiarly describe as having the collinear

momenta. The nested compatible subgraphs emanating from a corner describe a cascade of nested infrared divergences. The nesting of edges implies an ordering, just as in the case of UV divergences and the RG. In α_e parameters, this ordering is manifest as which edges are “scaling hardest” to zero, or the complementary set of edges to infinity, becoming soft as described in 3.9. We now describe these combinatorics more precisely.

It suffices to study the case with two diagonally-opposed massless corners, as the combinatorics of IR rays associated with neighboring corners is trivial, i.e., the geometry is a product. This means that for any assignment of the external masses, we can treat the diagonals independently and multiply the result. This can be seen by considering the compatibility condition. Two infrared subgraphs γ_1 and γ_2 associated with neighboring corners will have

$$z_{\gamma_1} = 2L_{\gamma_1} + 1 \quad z_{\gamma_2} = 2L_{\gamma_2} + 1 \quad (3.57)$$

We check the compatibility condition (3.47) which reduces to checking

$$2L_{\gamma_1} + 2L_{\gamma_2} + 1 \geq 2L_{\gamma_1 \cap \gamma_2} + 2L_{\gamma_1 \cup \gamma_2} \quad (3.58)$$

where in simplifying we used the fact that their union will be an infrared subgraph, but their intersection cannot be, as it will not touch three external edges (notice that this crucially depends on the subgraphs being associated with neighboring corners and not diagonally opposed corners). We can simplify the above as

$$L_{\gamma_1} + L_{\gamma_2} + 1 > L_{\gamma_1 \cap \gamma_2} + L_{\gamma_1 \cup \gamma_2} \quad (3.59)$$

The concern for failing to meet the compatibility condition is that the union γ_1 and γ_2 will make a new loop. Because the subgraphs emanate from neighboring corners, we

see that this is not possible. Therefore all infrared subgraphs from neighboring corners are compatible, and we simply furnish a triangulation for the highly non-simplicial subspaces associated with rays emanating from diagonally opposed corners, and any pair of simplicial cells from each (of dimension $M \times N$) will furnish a simplicial cell of dimension $2MN$. And in fact, we can simply compute the dual cone volumes independently, and multiply the result.

A single leading-dimensional divergent cone is in correspondence with a staircase walk (dashed lines) separating the fishnet into two Young tableau-shaped sets of edges (blue and red) as in Fig. 3.9. This defines a cone because it defines a maximal set of mutually compatible subgraphs on either side of the dashed line. In the growing picture, the compatible rays γ are given by the complements of all the Young *subdiagrams* based at the top-left (blue) or bottom-right corner (red), including the dashed bordering edges. For example, the blue shading with the dashed lines is the γ corresponding to the tableaux given by the red shading.

In the shrinking picture, we have the complementary graphs. Said differently, each staircase walk defines two subgraphs, which are the sets of all edges on either side of the dashed line (including the dashed line).

Since such cones are in general not simplicial, in order to leverage (3.36) one needs a triangulation. For concreteness, consider the red subdiagram in Fig. 3.9. All sub-tableaux of the red subgraph are mutually compatible, but this number of tableaux (nine) exceeds the dimension of the subspace in which the cones live (six). One triangulation is given by cells corresponding to the *standard* tableaux, see Fig. 3.9. The numbers in the boxes indicate an ordering. Scaling of the bottom and right edges of every red box labeled with integer i is given by $\rho_i \rho_{i+1} \cdots \rho_6$, where each ρ_i parameterizes scaling along a ray generating the simplicial cell. This is clearly simplicial since it furnishes a grouping of six sub-tableaux. Each cell clearly contributes the volume $\frac{1}{6!}$, so the non-trivial combinatorics is in counting the cells of this triangulation i.e.

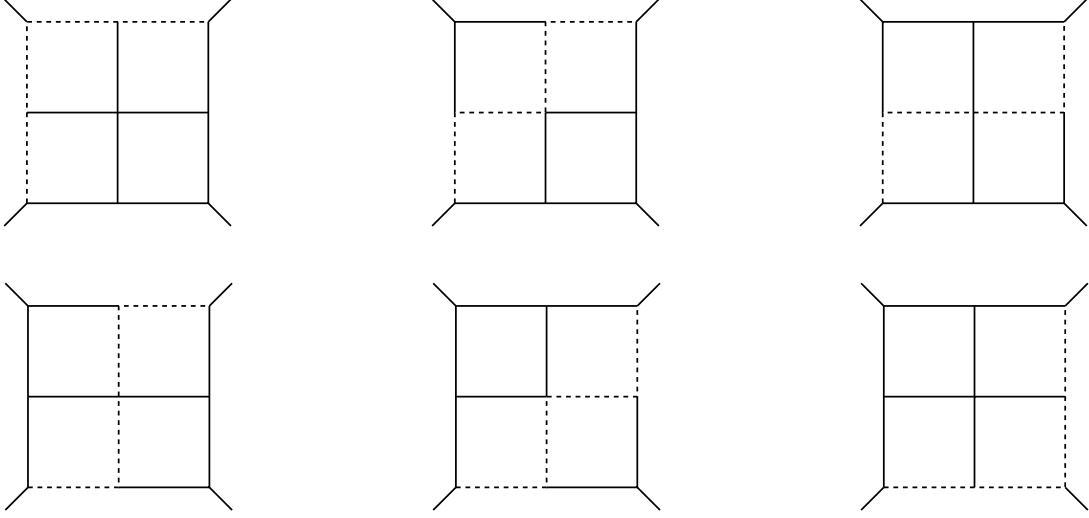


Figure 3.8: All leading staircase walks corresponding to leading contributions from one diagonal pairing.

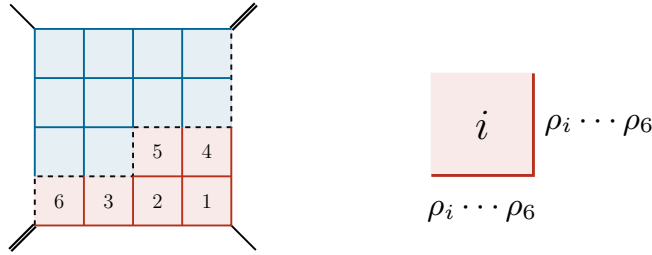


Figure 3.9: Left: Half of a top-dimensional contribution to a four-point fishnet diagram with opposite massless corners. The IR subgraphs (red and blue plus respective sub-tableaux) are either of the colored subgraphs plus the dashed lines and all sub-tableaux contained in (subtableaux of squares plus the bordering edges). The complementary set of edges are the edges becoming soft. We additionally label for the red portion one of the standard tableaux corresponding to a cell in the triangulation. Right: Scaling parameters $\rho_i \rightarrow \infty$ (the growing picture) of the lower-right edges for a square labeled by integer i in the standard tableaux.

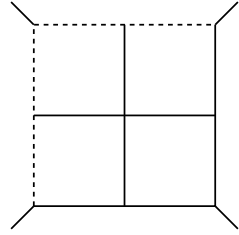
counting the number of standard tableaux. This is given by the hook-length formula. This counting can be done separately for the tableaux on either side of the dashed staircase walk, and the results multiplied.

We list all lower-left to upper-right staircase walks for the two-by-two fishnet in Fig. 3.8. There are three different shapes which will give three different contributions to the leading divergence. The first is for the walk which goes all along the outer edges of the graph. There is only one relevant side of the staircase walk, and so we simply

need to count the number of standard two-by-two tableaux using the hook-length formula. This is

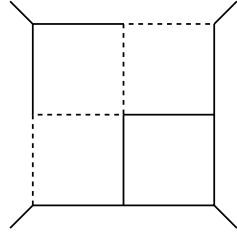
$$f^{(2,2)} = 2 \quad (3.60)$$

where the exponent of f denotes the shape of the tableaux (list of number of boxes in each row). The contribution from this walk is



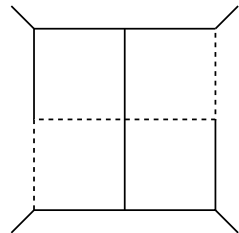
$$= \frac{2}{4!\epsilon^4} \quad (3.61)$$

The remaining two types of staircase walk produce non-trivial terms on either side of the dashed line and are



$$= \frac{f^{(2,1)} f^{(1)}}{3!1!\epsilon^4} \quad (3.62)$$

and



$$= \frac{f^{(1,1)} f^{(1,1)}}{2!2!\epsilon^4} \quad (3.63)$$

where the hook length formula yields $f^{(2,1)} = 2$ and $f^{(1,1)} = 1$. We can sum all contributions to yield

$$2 \times \left(\frac{1}{12} + \frac{1}{3} + \frac{1}{4} \right) \times \frac{1}{\epsilon^4} = \frac{4}{3\epsilon^4} \quad (3.64)$$

For the fully massless two-by-two fishnet, we multiply with the identical result for the upper left to lower right staircase walks to get

$$\begin{array}{|c|c|} \hline & \\ \hline & \\ \hline \end{array} = \frac{16}{9\epsilon^4} + \dots \quad (3.65)$$

We can generalize the result to arbitrary $M \times N$ fishnets. Again, for the fully massless fishnet, we focus on the calculation associated with lower left to upper right staircase walks and square it. The staircase walks define two tableaux that the fishnet is divided into. The shapes of these tableaux are complementary in the sense that when one tableaux has shape $(\lambda_1, \dots, \lambda_M)$ then the other has shape $(\bar{\lambda}_M, \dots, \bar{\lambda}_1)$ where $\bar{\lambda}_i = N - \lambda_i$. We claim that summing over the volume contributions from all the cells yields

$$\frac{2^{MN}}{(MN)!} C_{M,N} = \sum_{\lambda_i > \lambda_{i+1}} \frac{f^{(\lambda_1, \dots, \lambda_M)} f^{(\bar{\lambda}_M, \dots, \bar{\lambda}_1)}}{(\lambda_1 + \dots + \lambda_M)! (\bar{\lambda}_1 + \dots + \bar{\lambda}_M)!}, \quad (3.66)$$

and as before, $f^{(\lambda_1, \dots, \lambda_M)}$ counts the number of the standard Tableaux with shape $\lambda = (\lambda_1, \dots, \lambda_M)$. The numbers $C_{M,N}$ are the multi-dimensional Catalan numbers [20] with standard definition:

$$C_{M,N} = (MN)! \frac{\text{sf}(M-1) \text{sf}(N-1)}{\text{sf}(M+N-1)} \quad (3.67)$$

with the superfactorial $\text{sf}(N) = 0!1!2! \dots N!$. The result is squared to produce the leading divergence of a fully massless $M \times N$ fishnet. We call this a claim since while the right hand side of (3.66) is deducible from compatibility and (3.45), equaling the left-hand side is a claim. This in fact provides an interesting expression for the $C_{M,N}$'s which we are not aware of in the math literature.

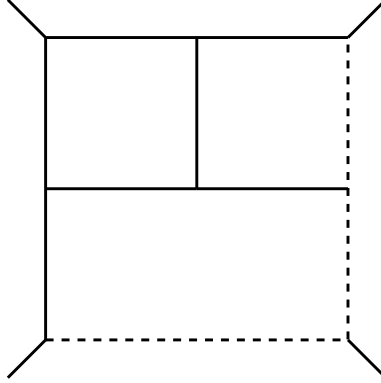


Figure 3.10: Staircase walk and indeed subgraph γ (γ is simply the dashed edges) which corresponds to a power divergence in the associated scalar graph (simply note that $d_{G/\gamma} = 7 - 3 * 2 = 1$ in $D = 4$). In the perturbative expansion of the amplitude for $\mathcal{N} = 4$ SYM this graph comes with a numerator that tames this logarithmic divergence.

3.4.5 Beyond Fishnets

We can also consider “rung-rule” graphs [21, 22] which also contribute for example to the leading infrared divergence of the four-point amplitude in $\mathcal{N} = 4$ SYM. The first such example of a graph is the tennis court, presented in Fig. 3.10. The actual Feynman integral crucially has a numerator, not merely the scalar propagators. We can recognize the need for such an infrared-divergence-softening numerator by considering the dashed subgraph in Fig. 3.10. The result graph has $d_{G/\gamma} = 1$ and therefore produces a power divergence in the infrared. The numerator softens this divergence to be logarithmic. While we have checked that accounting for the numerator, our analysis reproduces the correct answer for the leading divergence, perhaps more interestingly we can follow our nose ignoring this detailed accounting of the numerator and motivate a general method to extract the leading divergence from these rung-rule graphs.

Notice that in $D = 4 - 2\epsilon$ dimensions, we actually find that $d_{G/\gamma} = 1 + 3\epsilon$ for the above. When the numerator is accounted for, all compatibilities of the scalar analysis are respected, the tropical limit of the integrand along already logarithmically diver-

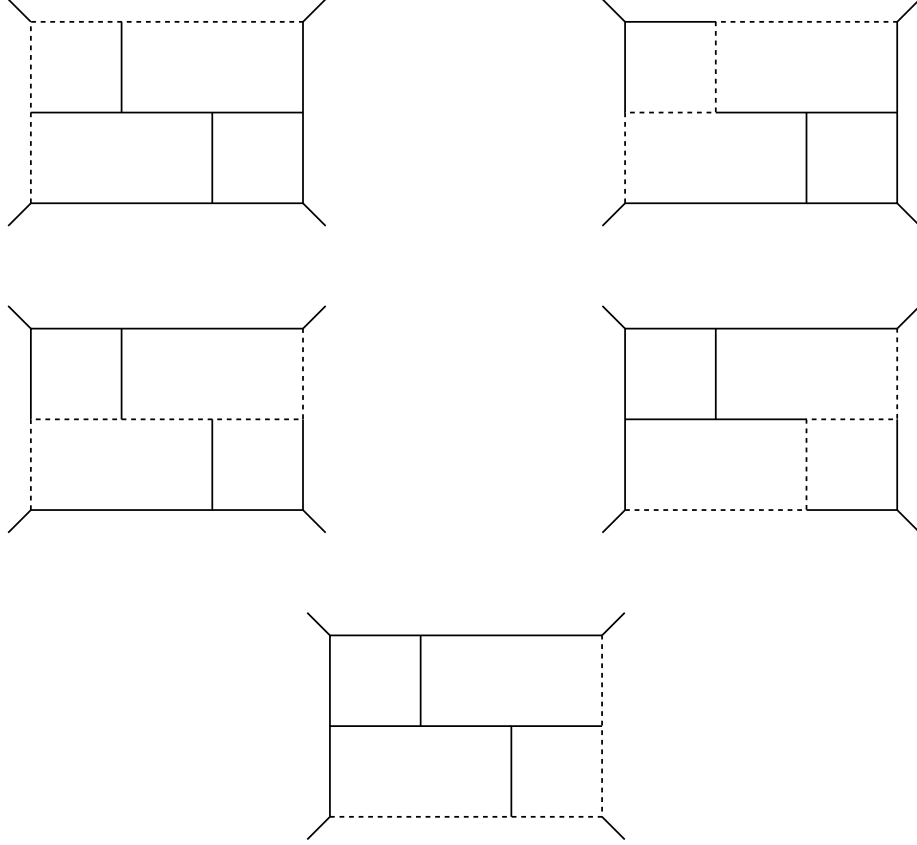


Figure 3.11: All staircase walks in the conjectured expansion. Note that the opposite diagonal pairing will have a different set of walks as the symmetry is broken. Following the procedure as with fishnets and simply throwing out the power divergent constant, we recover the correct leading divergence for this Mondrian graph.

gent rays is respected, and merely 3ϵ is recovered as the tropical limit along the limit shrinking the dashed subgraph. For this reason, we are motivated to simply carry on as in the fishnet case, and when we encounter a power divergent subgraph, we throw out the power divergent constant and keep the ϵ -dependent piece as the value to be used in our formulas.

Applying this method in the case of the tennis court correctly reproduces the leading expression $\frac{16}{9\epsilon^6}$ for the leading divergence. We will now illustrate this calculation in a more complicated, four-loop rung-rule graph which appears in the expansion of the the four-point amplitude in $\mathcal{N} = 4$ SYM. The graph and the relevant staircase walks *for one* diagonal direction are presented in Fig. 3.11. With the prescription to throw

out the power divergent piece of $d_{G/\gamma}$ along a ray, we can simply proceed exactly as we did with the fishnets in order to compute the contribution to the leading divergence. Recall that in the shrinking picture, all staircase walks also define all relevant subgraphs and therefore divergent rays. This is by inclusion of the dashed line and then all edges to one side of the dashed line. This is no different in the case of Fig. 3.11. And again recall that the ϵ dependent piece of the value of Trop along these divergent rays is $L_{G/\gamma}\epsilon$. We can compute, top-to-bottom and left-to-right, the contributions from these diagonal staircase walks: $\{\frac{1}{4!\epsilon^4}, \frac{1}{3!\epsilon^4}, \frac{1}{2!\times 2!\epsilon^4}, \frac{1}{3!\epsilon^4}, \frac{1}{4!\epsilon^4}\}$. Notice that the numerators of these factors are all one. This is because the subgraphs in Fig. 3.11 are unambiguously nested. Therefore the cones are simplicial and there is no need for counting of cells in a triangulation. This is not the case for the staircase walks for the other diagonal. It is easy to see that the staircase walks along the other diagonal (of which there are six) do not have a purely nested structure: there are cones requiring a triangulation into two simplicial cells. This is readily deduced by simply drawing the subgraphs and applying the compatibility condition. In this case, we get $\{\frac{2}{4!\epsilon^4}, \frac{2}{3!\epsilon^4}, \frac{1}{2!\times 2!\epsilon^4}, \frac{1}{2!\times 2!\epsilon^4}, \frac{2}{3!\epsilon^4}, \frac{2}{4!\epsilon^4}\}$. Summing up the contributions we find

$$\begin{array}{c} \diagup \quad \diagdown \\ \hline \begin{array}{|c|c|} \hline \square & \square \\ \hline \square & \square \\ \hline \end{array} \\ \hline \diagdown \quad \diagup \end{array} = \frac{2}{3\epsilon^4} \times \frac{4}{3\epsilon^4} + \dots \quad (3.68)$$

which is the correct leading term in ϵ [22]. It is possible that the leading divergence of the full amplitude comes from rung-rule graphs (this is observed to be the case through four loops). This would reduce the question of demonstrating the exponentiation of the leading divergence of the four point amplitude in $\mathcal{N} = 4$ SYM to a combinatorial problem enumerating rung-rule graphs and leveraging a formula like (3.66) generalized to rung-rule graphs, as we just did in this example.

3.5 Conclusions and Outlook

In this chapter we have characterized the UV and IR divergence structure of Feynman integrals for a broad range of kinematics. We did this in the Symanzik representation of the Feynman integral and in this setting, the tropical fan is the star of the show. By providing a maximally concise description of the tropical fan, we not only characterize the divergences completely and in irreducible terms, but isolate the single kernel which distinguishes UV from IR divergences: the function

$$z_\gamma = \begin{cases} 2L_\gamma & \mathcal{F}_{G/\gamma} \neq 0 \\ 2L_\gamma + 1 & \mathcal{F}_{G/\gamma} = 0 \end{cases} \quad (3.69)$$

using z_γ , UV and IR divergences are described in a completely unified fashion. We then see the distinct value

$$\text{Trop} = \begin{cases} -d_\gamma & \mathcal{F}_{G/\gamma} \neq 0 \\ d_{G/\gamma} & \mathcal{F}_{G/\gamma} = 0 \end{cases} \quad (3.70)$$

and even more crucially the distinct manifestation of compatibility of graphs emerge solely from the difference in z_γ . We then leveraged this understanding to provide a simple prescription for the calculation of the leading divergence of a Feynman integral. This was most non-trivially applied to make all-loop statements about $M \times N$ fishnets and yield well-motivated speculations about leading divergence between the fishnet case, for more general rung-rule graphs in the $\mathcal{N} = 4$ SYM expansion.

This all opens some exciting opportunities for future work. In particular, demonstrating the exponentiation of the leading divergence in $\mathcal{N} = 4$ SYM is a sharp calculation target. Beyond this, we would of course like to extend the kinematic regime in which these results apply. We already know from examples that the facet inequalities will no longer have entries 0 and 1 alone, but at least 2 as well. Nonetheless, a com-

parably principled characterization in more general Minkowski regimes may still be possible, and will be relevant for the study of divergences in theories of interest, or even recovering classic results such as those in QED. We will address such questions in future work.

Chapter 4

Subtraction Schemes for Feynman Integrals

4.1 Introduction

In the previous chapter we characterized the tropical fan associated to a scalar Feynman integral. This fully characterized the divergence structure of the Feynman integral and facilitated the calculation of leading divergences. This is of course not the end of the story. In order to understand UV and IR divergences beyond leading order, we need to extract the subleading divergences of a Feynman integral. In dimensional regularization, what we want is a means by which to systematically extract the Laurent expansion of a Feynman integral around $\epsilon = 0$, an oft underappreciated fact. This is in fact a problem with no solution for the generic case. To be clear, such a solution means that for generic external kinematics, with potentially multiple scales, a means by which to prescribe a set of finite integrals paired with inverse powers of ϵ which reproduce the Laurent expansion of the full integral. For UV divergences, the BPHZ method furnishes a prescription to obtain subtraction terms which will render a Feynman integral finite by expanding in masses and momenta about

a subtraction point. The main triumph of the BPHZ method is to understand the issues regarding overlapping divergences, in particular how to consistently perform subtractions associated to divergent subgraphs which overlap. But as a practical tool this method leaves much to be desired. We often want our amplitudes calculated in dimensional regularization, and the momentum subtraction method does not facilitate that. Moreover, since the subtraction terms are generated in practice by taking derivatives of the integrand, they proliferate many complicated terms, failing to manifest the simplification and factorization associated with shrinking subgraphs, which we observed for leading divergences in the previous section. Furthermore, this scheme attempts to obviate a regulator entirely and simply speak directly in terms of subtracted integrals, making it more difficult to state sharply the remarkable analytic phenomena that equate to the existence of Wilsonian beta function.

A marriage of these ideas in the form of single-scale subtraction for dimensionally regulated integrals was elegantly employed to calculate high-loop renormalization group functions in ϕ^4 theory [23], but again is not a recipe for the Laurent expansion of amplitudes generically, as it utilizes particular facts about single-scale “ p -integrals” (which are not even always Feynman integrals). All interesting statements about the renormalization group are statements about the regulated (UV completed) amplitude itself. What we would like is simply a means by which, only with the capacity to compute finite integrals i.e. without access to analytic continuation, extract the Laurent expansion in ϵ of our Feynman integrals. Anything else is simply a non-trivial observations about the coefficients in the Laurent expansion and infinite sums of diagrams.

The desire to have a prescription for such a basic problem is motivation enough, but perhaps the main motivation for adopting this point of view is as a warm-up case for the study of IR divergences. In the case of IR divergences, we do not have an analogue of Wilsonian thinking to motivate formulas which we can extract from more

abstract reasoning and then simply verify in perturbation theory when calculating the things we actually care about. Even analogous attempts at reproducing the success of Wilson for the infrared ultimately lead back to the shaky footing of statements about Feynman integrals relying on analytic continuation. We need to go straight for the formulas, and in the case of infrared divergences, this likely means observing them as nontrivial facts about how divergences in the perturbative expansion are organized. Toward this aim, we first review the basic problem of extracting a Laurent expansion only using finite integrals, without analytic continuation. We then illustrate an elegant subtraction scheme which is inspired by the work of Brown and Dupont on α' expansion of string integrals employing dihedral coordinates[24]. Afterwards, we go back to UV divergences and prescribe a canonical subtraction scheme leveraging u -variables which furnish a binary realization of the Feynman polytope to the study of UV divergences. This scheme is completely algorithmic, relatively economical, and completely canonical. Moreover, it manifests and leverages the factorization properties of the integrand and lifts them as much as possible to the integral. We illustrate the scheme and a two and three loop example and compare to verify against previous results.

4.2 Euler Beta Function

We can motivate the basic problem simply by considering the Euler beta function

$$\frac{\Gamma(\alpha' X_1)\Gamma(\alpha' X_2)}{\Gamma(\alpha' X_1 + \alpha' X_2)} = \int_0^1 \frac{dx}{x} x^{\alpha' X_1} (1+x)^{-\alpha' X_1 - \alpha' X_2} \quad (4.1)$$

We would like to extract the α' expansion of this integral, but without recourse to analytic continuation. We will not be able to avail ourselves of analytic continuation later on, and we will be rewarded for learning how to do things properly on simple

examples. The issue is that this integral generates poles in α' . Clearly when we set $\alpha' = 0$ at the level of the integrand, we have logarithmic divergences as $x \rightarrow 0$ and as $x \rightarrow \infty$. Recall that our leading divergence prescription from before would furnish integrals

$$\left(\int_0^1 \frac{dx}{x} x^{\alpha' X_1} \right) + \left(\int_1^\infty \frac{dx}{x} x^{-\alpha' X_2} \right) = \frac{1}{\alpha' X_1} + \frac{1}{\alpha' X_2} \quad (4.2)$$

which is of course the correct leading term in the α' expansion. The most naive thing to attempt would be to subtract the limiting behavior of the integrand. But notice that the limiting behavior of the integrand in one direction is not integrable in the other direction. What we need is a way to subtract something which is locally equal to this limiting behavior, but is globally well defined. We can find variables which do this, and indeed they also provide a natural way to write the integrand. Consider

$$u_1 = \frac{x}{1+x} \quad u_2 = \frac{1}{1+x} \quad (4.3)$$

with measure $\omega = d \log(x)$ we can write our integral as

$$\int \omega \ u_1^{\alpha' X_1} u_2^{\alpha' X_2} \quad (4.4)$$

the thing to notice is that locally in the vicinity of the singularities, these variables possess the correct leading behavior of the integrand, but they are bounded between 0 and 1 on the whole integration domain. They therefore let us extend the integration region. In particular consider

$$\int_1^\infty \frac{dx}{x} x^{-\alpha' X_2} \rightarrow \int_0^\infty d \log(u_2) u_2^{\alpha' X_2} \quad (4.5)$$

at large x , the relevant region, these behave identically, but the latter is a globally well-defined term which can be subtracted against our original integrand. Moreover,

it in fact equals the leading behavior:

$$\int_0^\infty d\log(u_2) u_2^{\alpha' X_2} = \frac{1}{\alpha' X_2} \quad (4.6)$$

Therefore we can write

$$\frac{\Gamma(\alpha' X_1) \Gamma(\alpha' X_2)}{\Gamma(\alpha' X_1 + \alpha' X_2)} = \frac{1}{\alpha' X_1} + \frac{1}{\alpha' X_2} + \int \left[\omega u_1^{\alpha' X_1} u_2^{\alpha' X_2} - d\log(u_1) u_1^{\alpha' X_1} - d\log(u_2) u_2^{\alpha' X_2} \right] \quad (4.7)$$

the subtracted integrand can now be expanded order-by-order in α' and integrated numerically or otherwise. This is a convergent representation of the α' expansion of the Euler beta function . The method we have employed here has been proven for open string amplitudes [24], of which the above is one example. We will be able to abstract the lesson here to the case of Feynman integrals.

4.3 u -variables for Feynman Integrals

In this section we will describe a systematic procedure for renormalizing an integral which is a product of polynomials raised to powers. This requires that the integral have only logarithmic singularities and the expansion produced is a Laurent expansion in a single external parameter around zero. The basis for this renormalization procedure is laid out in [24], where it is formulated specifically for the case of the open and closed string moduli spaces. The language employed is one tied to that specific problem, but the lessons of the renormalization scheme are naturally abstracted to a general setting, and we will articulate them in our own language of tropical geometry and generalizations of the “dihedral coordinates” on moduli space. As usual, before stating any general rules, we will revisit our triangle example and illustrate the procedure explicitly.

4.3.1 Triangle in u -vars

Recall the triangle integral (3.22) which we now describe with variables

$$I = \int_0^\infty \frac{dx_1 dx_2}{x_1 x_2} x_1^{\alpha' X_1} x_2^{\alpha' X_2} (1 + x_1 + x_2)^{-\alpha' X_1 - \alpha' X_2 - \alpha' X_3} \quad (4.8)$$

if we introduce the three variables

$$u_1 = \frac{x_1}{1 + x_1 + x_2} \quad u_2 = \frac{x_2}{1 + x_1 + x_2} \quad u_3 = \frac{1}{1 + x_1 + x_2} \quad (4.9)$$

we can write the integral as

$$I = \int_0^\infty \frac{dx_1 dx_2}{x_1 x_2} u_1^{\alpha' X_1} u_2^{\alpha' X_2} u_3^{\alpha' X_3} \quad (4.10)$$

Naively, this is a content-free rewriting of our integrand. And if this were merely done with random ratios that reproduce the integrand, it indeed would be; but these are not random. The first thing to notice is that we have written the integrand in terms of as many u 's as there are rays of the fan. That is because each of these u 's is associated to such a ray. In particular, a single u_i shrinking to zero corresponds to a divergence of the integrand regulated by the corresponding $\alpha' X_i$. What's more, the individual u_i themselves of course have a tropical fan associated to their tropicalization. The rays of these fans associated to each u_i are just the subsets of the rays associated to our integrand. In this case, since there is only one non-trivial polynomial: $(1 + x_1 + x_2)$, the fan associated to a given u_i is simply the fan associated to the full integrand. We can evaluate the tropicalizations of the u_i on these rays in order to probe their asymptotic behavior. We find the relation

$$\text{Trop}[u_i](r_j) = \delta_{ij} \quad (4.11)$$

so that r_i is in fact the unique ray along which $u_i \rightarrow 0$. Therefore, each u_i manifests a single divergence, uniquely associating it with a ray. Moreover, this property of each u_i is maintained upon simplification along the support of another u_j going to zero. This will be essentially later on for constructing subtraction terms associated with multiple u 's going to zero. Though the technical improvement in manifesting the structure of divergences is only mild in this case, the ability to write the integrand this way is highly non-trivial and advantageous in general. It is in fact, more than we can hope for in the case of Feynman integrals, but we will see that the essential structure relevant for renormalization does not evaporate as we move to the cases of physical interest.

The non-trivial aspect of the prescription for renormalization has nothing to do with the existence of integral representation such as (4.10) which manifests the divergence structure, but simply with the existence of the u 's. That these u 's exist means (4.11) in addition to criteria which we will confront later on.

Modified Residue The operation we need to introduce in order to prescribe subtraction terms is the modified residue. First we define it for a single degeneration, $u_1 \rightarrow 0$. First split the integrand Ω into

$$\Omega = u_1^{\alpha' X_1} d \log(u_1) \wedge \Omega^{u_1} \quad (4.12)$$

where Ω^{u_1} has full projective invariance under rescalings along the ray associated to u_1 (in this case the x_1 shrinking ray) in the limit $u_1 \rightarrow 0$. The modified residue of the integrand Ω associated with this degeneration is

$$R_{u_1} \Omega = \Omega^{u_1} \Big|_{u_1 \rightarrow 0} \quad (4.13)$$

where the evaluation means evaluating Ω^{u_1} in the one parameter limit taking $u_1 \rightarrow 0$ (if chosen as a variable, this parameter can be u_1 itself). The subtraction term associated with r_1 is then

$$\Omega_1 = u_1^{\alpha' X_1} d \log(u_1) \wedge R_{u_1} \Omega \quad (4.14)$$

As written, this is a top-form (two-form) which serves as one of the subtraction terms to be subtracted from Ω . But as it is a subtraction, we also need to calculate its α' expansion. Like our familiar $GL(1)^p$ simplified integrals we obtained studying leading divergences, here we can also “integrate out” the divergent direction. This is clear if we explicitly write our integral in u_1 and x_2 variables. The $d \log$ measure is

$$d \log(x_1) d \log(x_2) = (1 - u_1)^{-1} d \log(u_1) \wedge d \log(x_2) \quad (4.15)$$

which simply maps to $d \log(x_2)$ after the modified residue. So we have

$$\Omega_1 = u^{\alpha' X_1} d \log(u_1) \wedge \left[\left(\frac{x_2}{1 + x_2} \right)^{\alpha' X_2} \left(\frac{1}{1 + x_2} \right)^{\alpha' X_3} d \log(x_2) \right] \quad (4.16)$$

where again, the modified residue in the rightmost factor sets $x_1 \rightarrow 0$. In these variables, u_1 is integrated from zero to one and x_2 from zero to infinity. The u_1 integral is trivial and we therefore get

$$S_1 = \frac{1}{\alpha' X_1} \int_0^\infty d \log(x_2) \left(\frac{x_2}{1 + x_2} \right)^{\alpha' X_2} \left(\frac{1}{1 + x_2} \right)^{\alpha' X_3} \quad (4.17)$$

Which is a pole multiplying the Euler beta function, whose renormalization must also be performed (and we did perform above). Let’s take stock of what this prescription has achieved. From a mechanical point of view, u_1 is useful as it is a globally well behaved rational function which manifests the divergence of the integrand as $x_1 \rightarrow 0$

and only that divergence. We can therefore use it not just to perform a local change of variables in the neighborhood of x_1 , but to define a simplified integrand which agrees with Ω along this degeneration. This has allowed us to prescribe a top-form which is a subtraction term, as well as integrate out the divergent direction to produce a one-lower-dimensional integral equal to our subtraction term. This integral will itself be renormalized, but the process clearly terminates.

This addresses the construction of the subtraction terms in codimension one, but that only constitutes half of the nontriviality. The higher codimension subtractions, associated with cones generated by compatible rays, are what ultimately place stringent constraints on the kinds of rational functions which can constitute suitable u -variables. We will now construct the Ω_{12} subtraction term for our triangle integral and state the full renormalization of the integrand.

On a double degeneration we generalize (4.12) and (4.13) in the obvious way, writing

$$\Omega = u_1^{\alpha' X_1} d\log(u_1) \wedge u_2^{\alpha' X_2} d\log(u_2) \wedge \Omega^{u_1, u_2} \quad (4.18)$$

where of course now Ω^{u_1, u_2} is just a function (0-form). The generalization of the modified residue

$$R_{u_1, u_2} \Omega = \Omega^{u_1, u_2} \Big|_{u_1, u_2 \rightarrow 0} \quad (4.19)$$

is not the only simplification we make. We also simplify each $u_i^{\alpha' X_i} d\log(u_i)$ on the support of the other u_i going to zero. This furnishes

$$\Omega_{12} = \left(u_1^{\alpha' X_1} d\log(u_1) \Big|_{u_2 \rightarrow 0} \right) \wedge \left(u_2^{\alpha' X_2} d\log(u_2) \Big|_{u_1 \rightarrow 0} \right) \wedge R_{u_1, u_2} \Omega \quad (4.20)$$

It is easy to see in this case that $R_{u_1, u_2} \Omega = 1$ and so our subtraction term, integrated in u_1 and u_2 each running from 0 to 1, trivially yields $\frac{1}{\alpha'^2 X_1 X_2}$.

Combinatorics of Subtraction In order to render our integrand Ω finite we need to subtract each S_i corresponding to divergences coming from a single ray. But each S_i contains two subdivergences corresponding to further degenerating along one of the two compatible rays. Therefore when we subtract all S_i , we have double counted the mutual degenerations. This is why we need each S_{ij} , which manifests the double degeneration, and contributes with opposite sign to compensate for the double degeneration. Therefore, our integrand is

$$\Omega^R = \Omega - \Omega_1 - \Omega_2 - \Omega_3 + \Omega_{12} + \Omega_{23} + \Omega_{13} \quad (4.21)$$

But of course, the sum of the subtraction terms is not zero. We need to calculate it, and we need to do so in a way that is again a collection of integrals that can be expanded at integrand level. We simply apply the same procedure from above to our S_1 in (4.17). And note that there is really no new work to do at all; we already have these subtraction terms at our disposal, as the result of integrating Ω_{12} over u_1 (and relatedly Ω_{13}) which we can denote as $\Omega_{12}^{u_1}$ and $\Omega_{13}^{u_1}$. Therefore, we have

$$\Omega_1^{u_1, R} = \Omega_1^{u_1} - \Omega_{12}^{u_1} - \Omega_{13}^{u_1} \quad (4.22)$$

So we see that at the level of the full integral, we have

$$I^R = \sum_I \int \Omega^{u_I, R} \quad (4.23)$$

where u_I denotes a set I of compatible rays defining which $u_i \rightarrow 0$. In each case this defines an $|I|$ lower-dimensional form, which is then renormalized in the way described above, denoted by R . The same expression (4.23) applies in general and in particular for the Feynman integrals which we will consider below.

4.4 UV u 's

The existence of u -variables in the context of open string moduli spaces and more generally binary geometries [16, 17] requires at minimum a simplicial fan. This is because each u is associated with a facet of the associated Newton polytope (or dually, an extremal ray in the fan) and a maximal degeneration should correspond to a unique set of u 's going to zero. A simplicial fan is one in which all d -dimensional cones are generated by d extremal rays. A non-simplicial fan can be non-uniquely triangulated such that the fan is now comprised of simplicial cells. As was discussed previously, this is what must be done in order to compute a leading divergence for a generic Feynman integral, as their fans are generically non-simplicial. Therefore, naively we cannot hope to find anything like u -variables in the conventional sense associated with a Feynman integral; for a non-simplicial fan it is simply a nonsense hope. But we can hope for something similar: u -variables associated with a fan which is equivalent to a triangulation of our fan. That is, a fan with identical rays, but additional compatibility conditions between rays such that all cones become simplicial. In the dual Newton polytope, this could be realized by the same facet inequalities but with slightly different constants.

Beyond a simplicial fan, the most general machinery available to produce u -variables comes from stringy canonical forms i.e. integrals of products of polynomials raised to regulated powers (each polynomial with its own power called a summand), in which the requirement is that

$$\#\text{variables} + \#\text{summands} = \#\text{rays of fan} \quad (4.24)$$

where the fan is the dual fan associated to the Minkowski sum of the Newton polytopes of the summands. This is a highly non-trivial constraint. Nonetheless, in the case of Feynman integrals with sufficiently generic kinematics i.e. those which remain

generalized permutahedra (see discussion around (3.47)), we can claim to meet it for a fan equivalent to ours but with additional compatibilities. For such cases we claim to have an efficient algorithm for the determination of u -variables associated with all facets of the polytope \mathbf{S}_G , which is not that in [16]. These u -variables therefore come with a triangulation of the fan dual to \mathbf{S}_G , which turns out to agree with Zimmerman's forest formula applied to all subgraphs.

Before moving on to explain everything in detail, we will state the formulas that constitute the final results without any explanation. The following sections will build toward this formula and its simple algorithmic implementation. First we simply write down a generalization of a Feynman integral depending upon some integers p_γ

$$I_G(\{p_\gamma\}) = \int \omega_G \left(\prod_\gamma u_\gamma^{p_\gamma \epsilon} \right) \frac{\prod \alpha_e}{\mathcal{U}^{D/2}} \left(\frac{\mathcal{U}}{\mathcal{F}} \right)^{d_G} \quad (4.25)$$

where we have deformed an ordinary Feynman integral associated with graph G by product of rational functions of α_e 's called u_γ raised to powers $p_\gamma \epsilon$. This generalized Feynman integral therefore limits to a Feynman integral when the $p_\gamma = 0$. Therefore, the Laurent expansion of this integral for generic p_γ contains that of a Feynman integral of interest as a special case. The first formula is for the top-dimensional (at integrand level) subtraction term associated with a divergent face F (or forest of subgraphs F) generated by rays associated to subgraphs Γ_i

$$S_F = \int (u_{\Gamma_i/F})^{d_{\Gamma_i}} d \log(u_{\Gamma_i/F}) \left[\omega_{G/F} \frac{\prod^{G/F} U_{e/F}}{\mathcal{U}_{G/F}^{D/2}} \left(\frac{\mathcal{F}_{G/F}}{\mathcal{U}_{G/F}} \right)^{d_G} u_{\gamma_j/F}^{p'_j \epsilon} \right] \times \left[\omega_{\Gamma_i/F} \frac{\prod^{\Gamma_i/F} U_{e/F}}{\mathcal{U}_{\Gamma_i/F}^{D/2}} u_{\gamma_j/F}^{p'_j \epsilon} \right] \quad (4.26)$$

which in turn is equal to

$$S_F = \frac{1}{d_{\Gamma_i}} I_{G/F}(\{p'_\gamma\}_{G/F}) \prod_i I_{\gamma_i/F}(\{p'_\gamma\}_{\gamma_i/F}) \quad (4.27)$$

Products of lower loop graphs of the same type. The p'_γ denote in general new integer shifted powers on the lower integrals. It is worth emphasizing just why this pair of equations is a triumph, before explaining precisely where they come from and how they are used. We will articulate below that the variables u_{Γ_i} are completely canonical and dictated by the Feynman polytope. There is no guesswork in obtaining them and the algorithm for obtaining them is simple and efficient given knowledge of the facets, which we have articulated in the previous chapter. These two equations therefore allow us to furnish a subtraction term S_F associated with a forest F of subgraphs which manifests the factorization properties of the integrand and measure. Most importantly this factorization manifests what S_F is equal to: products of integrals calculated at lower loop order. Therefore, with the minimal work we could hope for, we can with any method capable of evaluating *finite* integrals, systematically extract the Laurent expansion in ϵ of our Feynman integral, and do so in a way reflecting the factorization properties of these integrals. This is a subtraction scheme sharing all the desirable properties of that for string amplitudes. Moreover, because the subtraction terms are demonstrably equal to products of lower-loop Feynman integrals generalized by powers $p_\gamma \epsilon$, and they manifest a recursive procedure for producing the Laurent expansion, they give a hope for understanding the remarkable phenomenon of resummation of divergences intrinsically in an “on-shell” fashion.

Constructing the u ’s The unexplained objects in the expression above are the u_γ ’s (and the U_e ’s which are products of u_γ ’s). These are completely canonical rational functions of the underlying α_e ’s. When simply listed, they do not obviously

furnish any progress toward evaluating the Feynman integrals. Indeed, there are more u_γ 's than there are edges of the Feynman graph (there is one for every facet of \mathbf{S}_G), so their introduction naively constitutes an increase in complexity. But these are of course the wrong measures of complexity. As we will find, the purpose in life of the u_γ is to be the canonical variables which blow up singularities associated with subgraphs γ and are globally well defined. Beyond this, the u_γ 's enjoy many remarkable properties which are necessary to furnish (4.26) and (4.27). Here we note that strictly speaking, these u_γ 's are the correct u_γ 's to associated to a Feynman integral when it is a generalized permutahedron. Nonetheless, so long as we are only interested in subdivergences coming from subgraphs in a specific class, these u_γ 's can be used for renormalization beyond the generalized permutahedron case (i.e. when the divergences are the same as those for the same graph but with generic kinematics, as is the case for UV divergences).

In this section, in order to construct the u_γ 's, we will consider scalar graphs with completely generic kinematics, where the structure of the feynman polytope is

$$\mathbf{S}_G = \Delta_{E-1} \oplus \mathbf{U}_G \tag{4.28}$$

the polytope \mathbf{S}_G in this case is guaranteed to be a generalized permutahedron. It is a theorem due to Ardila, Benedetti, and Doker that a Generalized permutahedron is always representable as a signed Minkowski sum of simplices [25]. The notion of a Minkowski difference is subtle and not completely canonical. Here it is simply taken to mean that $P - Q = R$ if $P = Q \oplus R$. Nonetheless, we will not use this theorem to prove the legitimacy of our construction, as our construction in fact takes a leap beyond what was proven, but we comment as it has nonetheless served as valuable inspiration.

We claim that in one triangulation, the combinatorial structure of the fan of \mathbf{S}_G is

identical to that of

$$\text{Newt} \left[\alpha_G \prod_{\text{facet } \gamma} \alpha_\gamma \right] \quad (4.29)$$

this is powerful because it means that a triangulation of the fan we care about is identical to that associated with an integrand with as many summands as facets. Therefore, there are u -variables associated with this fan. We can leverage the technology of stringy canonical forms to compute them. We can feed these summands into the stringy u -variables crank to compute the u -variables. This process becomes very computationally expensive, but application in a variety of examples motivates an ansatz, which can be used to fix the u -variables much more efficiently applying the constraints. We conjecture that the correct u -variables are of the form

$$u_\gamma = \prod_{\Gamma \supseteq \gamma} \left(\frac{\alpha_\Gamma}{\alpha_G} \right)^{p_\Gamma} \quad (4.30)$$

where the p_Γ need to be determined. But this immediately computable recursively using (4.11). The above implies that

$$\text{Trop}[u_\gamma] = \sum_{\Gamma \supseteq \gamma} p_\Gamma \text{Trop}[\alpha_\Gamma] \quad (4.31)$$

note that

$$\text{Trop}[\alpha_\Gamma](r_{\gamma'}) = \begin{cases} 1 & \Gamma \subseteq \gamma' \\ 0 & \text{else} \end{cases} \quad (4.32)$$

which allows to proceed starting with γ and building up the values of p_Γ . In particular, clearly $p_\gamma = 1$ in (4.30). Then we proceed to subgraphs $\Gamma \supset \gamma$ such that $L_\Gamma = L_\gamma + 1$ and since none can be a subset of any other, we clearly have $p_\Gamma = -1$ for such Γ . The steps are graded by loop number and at step i we then have

1. Consider all graphs $\Gamma \supset \gamma$ with loop number $L_\Gamma + i$.

2. For each Γ we have $p_\Gamma = \sum_{\Gamma \supset \Gamma' \supseteq \gamma} p_{\Gamma'}$

this unambiguously and systematically defines u_γ . Since we only care about performing subtractions associated to the divergent part of the fan, we could consider only using the α_Γ from these subgraphs. We leave such explorations to future work.

Properties of the u 's These variables obey other remarkable properties which are valuable for performing renormalization. One property is the inversion formula for the α_e 's in terms of the u_γ 's, which is

$$\frac{\alpha_e}{\alpha_G} = \prod_{\gamma} u_\gamma^{\gamma|e} \quad (4.33)$$

where $\gamma|e$ simple denotes whether e is an edge in γ or not i.e.

$$\gamma|e = \begin{cases} 1 & e \in \gamma \\ 0 & e \notin \gamma \end{cases} \quad (4.34)$$

Additionally, the u 's provide a binary realization of the geometry which manifests the factorization properties. If we denote the u variable for subgraph γ with respect to total graph Γ as $u_\gamma|_\Gamma$ then we can state the following fact about the u_γ 's as a given $u_{\gamma_i} \rightarrow 0$:

$$u_{\gamma_j} \rightarrow \begin{cases} u_{\gamma_j}|_{\gamma_i} & \gamma_j \subset \gamma_i \\ u_{\gamma_j}|_{G/\gamma_i} & \gamma_j \text{ disjoint to } \gamma_i \\ 1 & \gamma_j \text{ incompatible with } \gamma_i \end{cases} \quad (4.35)$$

so the u_γ 's simplify in a way manifesting the factorization properties of the integral and its associated geometry on the support of a given degeneration. Moreover, they realize the geometry as a binary geometry, where all u_γ 's are bounded between zero and one and all faces are probed by u_γ 's going to zero and one.

Because the u_γ 's describe a simplicial fan, there are only as many simultaneously compatible u_γ 's as the dimension of the space. Any set of compatible u_γ 's would also be a good set of integration variables. As a set of compatible u_γ 's vary independently between 0 and 1, they cover the integration domain, and would therefore be integrated over the hypercube $[0, 1]^{E-1}$. The final property we could hope to extract are u -equations, which are necessary to describe the problem intrinsically in terms of the u -variables if we so choose. In specific cases the u -equations can be found, but at present we do not have a general prescription. We now move on to describe an example.

Parachute We now workout the u -variables for the parachute graph fig 4.1, in which all edges are massive. The set of facets is readily identified as $\{\gamma_1, \gamma_2, \gamma_3, \gamma_4, \gamma_{123}, \gamma_{124}, \gamma_{34}\}$. The easiest u_γ 's to compute are those of $\{\gamma_{123}, \gamma_{124}, \gamma_{34}\}$ as they have no supergraphs which are facets. Therefore, they are

$$u_{123} = \frac{\alpha_{123}}{\alpha_G} \quad u_{124} = \frac{\alpha_{124}}{\alpha_G} \quad u_{34} = \frac{\alpha_{34}}{\alpha_G} \quad (4.36)$$

Then we can proceed with the individual edges. Given that the algorithm above is graded by loop order and this is merely a two-loop graph, the process terminates in one step. We simply divide α_e by the one-loop graphs containing that edge. And so we find

$$u_{1,2} = \frac{\alpha_{1,2}\alpha_G}{\alpha_{123}\alpha_{124}} \quad \alpha_{3,4} = \frac{\alpha_{3,4}\alpha_G}{\alpha_{34}\alpha_{12(3,4)}} \quad (4.37)$$

In table 4.1 we present the limits of the u_γ 's for the parachute as any one u_γ is taken to zero. This demonstrates that the u_γ 's computed for the parachute indeed obey the rules (4.35). This is highly non-trivial and essential to manifesting the factorization structure of (4.26) and (4.27). Naively, it could have been that we constructed u_γ 's which obey the tropical condition before degeneration, and then found that on a

	u_1	u_2	u_3	u_4	u_{34}	u_{123}	u_{124}
u_1	$\frac{\alpha_1 \alpha_G}{\alpha_{123} \alpha_{124}}$	$\frac{\alpha_2 \alpha_{234}}{\alpha_{23} \alpha_{24}}$	$\frac{\alpha_3 \alpha_{234}}{\alpha_{23} \alpha_{34}}$	$\frac{\alpha_4 \alpha_{234}}{\alpha_{24} \alpha_{34}}$	$\frac{\alpha_{34}}{\alpha_{234}}$	$\frac{\alpha_{23}}{\alpha_{234}}$	$\frac{\alpha_{24}}{\alpha_{234}}$
u_2	$\frac{\alpha_1 \alpha_{134}}{\alpha_{13} \alpha_{14}}$	$\frac{\alpha_2 \alpha_G}{\alpha_{123} \alpha_{124}}$	$\frac{\alpha_3 \alpha_{134}}{\alpha_{13} \alpha_{34}}$	$\frac{\alpha_4 \alpha_{134}}{\alpha_{14} \alpha_{34}}$	$\frac{\alpha_{34}}{\alpha_{134}}$	$\frac{\alpha_{13}}{\alpha_{134}}$	$\frac{\alpha_{14}}{\alpha_{134}}$
u_3	$\frac{\alpha_1}{\alpha_{12}}$	$\frac{\alpha_2}{\alpha_{12}}$	$\frac{\alpha_3 \alpha_G}{\alpha_{34} \alpha_{123}}$	1	$\frac{\alpha_4}{\alpha_{124}}$	$\frac{\alpha_{12}}{\alpha_{124}}$	1
u_4	$\frac{\alpha_1}{\alpha_{12}}$	$\frac{\alpha_2}{\alpha_{12}}$	1	$\frac{\alpha_4 \alpha_G}{\alpha_{34} \alpha_{124}}$	$\frac{\alpha_3}{\alpha_{123}}$	1	$\frac{\alpha_{12}}{\alpha_{123}}$
u_{34}	$\frac{\alpha_1}{\alpha_{12}}$	$\frac{\alpha_2}{\alpha_{12}}$	$\frac{\alpha_3}{\alpha_{34}}$	$\frac{\alpha_4}{\alpha_{34}}$	$\frac{\alpha_{34}}{\alpha_G}$	1	1
u_{123}	$\frac{\alpha_1}{\alpha_{123}}$	$\frac{\alpha_2}{\alpha_{123}}$	$\frac{\alpha_3}{\alpha_{123}}$	1	1	$\frac{\alpha_{123}}{\alpha_G}$	1
u_{124}	$\frac{\alpha_1}{\alpha_{124}}$	$\frac{\alpha_2}{\alpha_{124}}$	1	$\frac{\alpha_4}{\alpha_{124}}$	1	1	$\frac{\alpha_{124}}{\alpha_G}$

Table 4.1: The limit of the u_γ variable in a given row i as the u_γ in column j goes to zero. Notice that the simplifications reflect the compatibility described in (4.35). These u_γ therefore furnish a factorizing and binary realization of the geometry.

degeneration associated with a particular γ , that the tropical condition was no longer obeyed, let alone that the set of u_γ 's became something meaningful as defined on the resulting graphs G/γ and γ . It is indeed possible to construct such impostor u_γ 's. The u_γ 's in table 4.1 limit to exactly the u_γ 's that would have been defined on the product graph $G/\gamma \times \gamma$, and the u_γ 's associated with facets incompatible with the degeneration in question, limit to one.

We can also write u -equations in this case

$$1 - u_{34} = u_{123} u_{124} (1 - u_3 u_4 u_{34}^2)$$

$$1 - u_{124} = u_{123} u_3 u_{34}$$

$$1 - u_{123} = u_{124} u_4 u_{34}$$

$$1 - u_3 = u_4 u_{124}^2 (1 - u_3 u_4 u_{34}^2)$$

$$1 - u_4 = u_3 u_{123}^2 (1 - u_3 u_4 u_{34}^2)$$

there are no relevant equations for u_1 and u_2 as they are compatible with everyone. One could use the equations above which are naively overdetermined to eliminate $u_{124}, u_{123}, u_3, u_4$ in favor of u_{34} and express the Feynman integral in terms of u_1, u_2 and u_{34} . We leave general explorations of u -equations, and expressing Feynman integrals in terms of them, to future work.

This gets us much of the way to understanding (4.26) and (4.27). In particular, we can make clear the notation $u_{\gamma/F}$. A subtraction term will be associated with a set F of compatible $\{\Gamma_i\}$ which are going to zero, and $u_{\gamma/F}$ denotes the simplification of a given u_γ in this degeneration limit. In particular, for Γ_i which are in F , they are simplified on support of all other Γ_j in F going to zero. For γ *not* in F , we simplify on support of all members of F going to zero. As we know from the above, the u_γ 's simply become what they are as defined on the product graph

$$G_F = (G/\Gamma_1 \dots \Gamma_F) \times \Gamma_1 \dots \times \Gamma_F \quad (4.38)$$

It is worth noting that the fact that we can take the degeneration limit in any order is crucially a reflection of the compatibility of the facets. Different orders of such limits dropping monomials are not necessarily equivalent for random sets of facets. Before moving on, we provide a few more miscellaneous comments in some special cases.

Finite Type Cluster Algebras Though we cannot give a general prescription for u -equations, there are instances where the u -variables associated to a Feynman graph are those of a finite type cluster algebra, and obey perfect u -equations. One such example is the sunrise integral. We state the u -variables here for convenience

$$u_i = \frac{\alpha_i \alpha_G}{\alpha_{ii+1} \alpha_{ii-1}} \quad u_{ij} = \frac{\alpha_{ij}}{\alpha_G} \quad (4.39)$$

where u_i is associated to the single edge i and ij is associated to the bubble subgraph made out of edges i and j . Moreover, $i + 1$ and $i - 1$ should be understood cyclically over $\{1, 2, 3\}$. The u -equations are given by

$$1 - u_i = u_{i-1}u_{i+1}u_{i-1i+1}^2 \quad (4.40)$$

$$1 - u_{ii+1} = u_{i-1}u_{ii-1}u_{i-1i+1} \quad (4.41)$$

which are the equations associated to the B_2 cluster algebra, which is also the two-dimensional permutahedron. Additionally, certain kinematic arrangements of the triangle furnish an A_2 associahedron. Nonetheless, the relevance of finite-type cluster algebras is merely an amusement. Of greater interest, of course, would be u -equations for Feynman integrals in general.

Going Bananas The multi-edge generalizations of the sunrise are the so-called “bananas”, some number of edges running between two vertices. The Feynman polytopes of these graphs are not simple, and there has therefore been no reason to suspect they should be associated with anything canonical, despite the fact they exhibit full permutation symmetry. We claim that the simple polytope which furnishes the u -variables associated the banana graphs, using the procedure described hitherto, is the permutahedron. This follows from the claims contained here. In particular, given the simplices whose Minkowski sum we claim shares the same normal fan as a triangulation of the our normal fan, the work of [17] implies that this is indeed the permutahedron.

4.5 Integrals in u -variables

Using the properties of the u -variables outlined above we can revisit our integrals (3.3) and express them in terms of u ’s. In order to do so we will denote $\frac{\alpha_e}{\alpha_G} = U_e$

which equals the product in (4.33). We restate our original integral of Symanzik polynomials for convenience this time with the $d^{E-1} \log$ measure ω_G :

$$I_G = \int_0^\infty \omega_G \prod_e \alpha_e \frac{1}{\mathcal{U}^{D/2}} \left(\frac{\mathcal{U}}{\mathcal{F}} \right)^{d_G} \quad (4.42)$$

where we can write this projective measure covariantly as

$$\omega_G = \sum_e (-1)^e d \log(\alpha_1) \dots \widehat{\log(\alpha_e)} \dots d \log(\alpha_E) \quad (4.43)$$

where $\widehat{\log(\alpha_e)}$ is omitted in each term in the sum. This is equivalent to writing for some choice of α_{e_\star} :

$$\omega_G = d \log(\alpha_1/\alpha_{e_\star}) \dots d \log(\alpha_E/\alpha_{e_\star}) \quad (4.44)$$

Note that the factor multiplying the $d \log$ measure is invariant under a uniform rescaling in all α_e , in particular by α_G^{-1} . Therefore, we can utilize the inversion formula (4.33) to write

$$I_G = \int \omega_G \prod_e U_e \frac{1}{\mathcal{U}^{D/2}(U)} \left(\frac{\mathcal{U}(U)}{\mathcal{F}(U)} \right)^{d_G} \quad (4.45)$$

where U_e is simply the associated product of u_γ 's furnished by the inversion. It is also valuable at this point to introduce the generalized integral

$$I_G(\{p_\gamma\}) = \int \omega_G \prod_e U_e \frac{\prod u_\gamma^{p_\gamma \epsilon}}{\mathcal{U}^{D/2}(U)} \left(\frac{\mathcal{U}(U)}{\mathcal{F}(U)} \right)^{d_G} \quad (4.46)$$

this clearly reproduces the Feynman integral of interest when the p_γ are all set to zero. Integrals in this class are what subtraction terms will generically be equal to. As it stands, this constitutes a redundant rewriting of the integral in (4.42), but it will prove to be an eminently useful rewriting, manifesting the factorization properties of the measure as well as what multiplies it, and making the renormalization transparent.

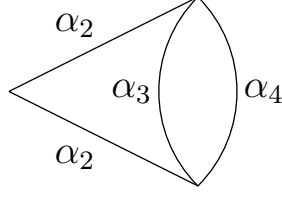


Figure 4.1: Two loop contribution to the expansion of the four-point amplitude in ϕ^4 . This particular graph is computed with $m_e^2 = 1$ and no external kinematics in (4.61).

Before moving on to the generalities, we work through a simple two-loop example to illustrate the details.

Parachute 1 We will consider all masses $m_e^2 = 1$ and no external kinematics. In this case

$$\mathcal{U} = (\alpha_1 + \alpha_2)(\alpha_3 + \alpha_4) + \alpha_3\alpha_4 \quad (4.47)$$

and

$$\mathcal{F} = (\alpha_1 + \alpha_2 + \alpha_3 + \alpha_4)\mathcal{U} \quad (4.48)$$

The u -variables in this case are

$$\begin{aligned} u_{1,2} &= \frac{\alpha_{1,2}\alpha_G}{\alpha_{123}\alpha_{124}} & \alpha_{3,4} &= \frac{\alpha_{3,4}\alpha_G}{\alpha_{34}\alpha_{12(3,4)}} \\ u_{12(3,4)} &= \frac{\alpha_{12(3,4)}}{\alpha_G} & \alpha_{34} &= \frac{\alpha_{34}}{\alpha_G} \end{aligned} \quad (4.49)$$

and the inversion expressions yield

$$U_{1,2} = u_{1,2}u_{123}u_{124} \quad U_{3,4} = u_3u_{34}u_{12(3,4)} \quad (4.50)$$

In which case, up to rescaling by α_G we need to know

$$\mathcal{U}(U) = u_{123}u_{124}u_{34}[(u_1 + u_2)(u_3u_{123} + u_4u_{124}) + u_3u_4u_{123}u_{124}u_{34}^2] \quad (4.51)$$

and our integral is, choosing U_1 as our gauge-fixing parameter

$$I_G = \int d\log(u_2/u_1) d\log(u_3 u_{34}/u_1 u_{124}) d\log(u_4 u_{34}/u_1 u_{123}) \frac{\prod U_e}{\mathcal{U}(U)^{2-\epsilon}} \quad (4.52)$$

We want to construct a subtraction term and know what we subtracted. These are our two objectives. One nice feature of the form of (4.51) is it manifests the leading behavior in the $u_{34} \rightarrow 0$ limit, which is

$$\mathcal{U}(U) = u_{123} u_{124} u_{34} [(u_1 + u_2)(u_3 u_{123} + u_4 u_{124})] \quad (4.53)$$

we will later make substitutions corresponding to the simplifications of (4.35). Next we look at the $d\log$ measure in this limit. We want to extract a $d\log(u_{34})$ in accordance with the renormalization prescription. In doing so, we implicitly choose another α_e which is effectively set to 1 as we trade for $d\log(u_{34})$. In this case let's choose u_4 and we have

$$d\log(u_2/u_1) d\log(u_3 u_{123}/u_4 u_{124}) d\log(u_{34}) \quad (4.54)$$

We can now put things together and apply (4.35) in which case

$$u_{1,2} = \frac{\alpha_{1,2}}{\alpha_{12}} \quad u_{3,4} = \frac{\alpha_{3,4}}{\alpha_{34}} \quad (4.55)$$

$$u_{12(3,4)} = 1 \quad (4.56)$$

note that on the support of this simplification we find $u_1 + u_2 = u_3 + u_4 = 1$ and therefore

$$\mathcal{U}(U) = 1 \quad (4.57)$$

in this limit. We then find

$$S_{34} = \int u_{34}^\epsilon d\log(u_{34}) d\log(u_2/(1-u_2)) d\log(u_3/(1-u_3)) u_2(1-u_2) u_3(1-u_3) \quad (4.58)$$

in the α_e 's this expression is

$$S_{34} = \int d\alpha_2 d\alpha_3 d\alpha_4 \left(\frac{\alpha_3 + \alpha_4}{1 + \alpha_2 + \alpha_3 + \alpha_4} \right)^\epsilon \frac{1}{(1 + \alpha_2 + \alpha_3 + \alpha_4)(\alpha_1 + \alpha_2)(\alpha_3 + \alpha_4)^2} \quad (4.59)$$

and constitutes a good subtraction term which renders the integrand expandable order-by-order in ϵ . But we also notice that integrating (4.58) readily yields

$$S_{34} = \frac{1}{\epsilon} \quad (4.60)$$

a delightfully compact expression to have subtracted. Ultimately this lets us write

$$I_G = \frac{1}{\epsilon} + \int d\alpha_2 d\alpha_3 d\alpha_4 \left[\frac{(1 + \alpha_2 + \alpha_3 + \alpha_4)^{-2\epsilon}}{[(\alpha_1 + \alpha_2)(\alpha_3 + \alpha_4) + \alpha_3 \alpha_4]^{2-\epsilon}} - \left(\frac{\alpha_3 + \alpha_4}{1 + \alpha_2 + \alpha_3 + \alpha_4} \right)^\epsilon \frac{1}{(1 + \alpha_2 + \alpha_3 + \alpha_4)(\alpha_1 + \alpha_2)(\alpha_3 + \alpha_4)^2} \right] \quad (4.61)$$

a convergent representation of I_G .

Parachute 2 In order to illustrate in a more general setting and for ease of comparison to results calculated previously [23], we will calculate the parachute again for the case $m_e^2 = 0$ and a single scale such that

$$\mathcal{F} = s\alpha_1\alpha_2(\alpha_3 + \alpha_4) \quad (4.62)$$

since there is a single scale and the s -dependence can be restored by dimensional analysis, we now set $s = 1$. From the point of view of the algorithm, nothing changes in the subtractions here, just the explicit form of the terms due to the difference in

the \mathcal{F} polynomial.

$$I_G = \frac{\Gamma(1-2\epsilon)^2}{\epsilon\Gamma(2-4\epsilon)} + \int d\alpha_2 d\alpha_3 d\alpha_4 \left[\frac{\left(\frac{(\alpha_1+\alpha_2)(\alpha_3+\alpha_4)+\alpha_3\alpha_4}{\alpha_1\alpha_2(\alpha_3+\alpha_4)} \right)^{2\epsilon}}{[(\alpha_1+\alpha_2)(\alpha_3+\alpha_4)+\alpha_3\alpha_4]^{2-\epsilon}} - \left(\frac{\alpha_3+\alpha_4}{1+\alpha_2+\alpha_3+\alpha_4} \right)^\epsilon \frac{\left(\frac{(\alpha_1+\alpha_2)^2}{\alpha_1\alpha_2} \right)^{2\epsilon}}{(1+\alpha_2+\alpha_3+\alpha_4)(\alpha_1+\alpha_2)(\alpha_3+\alpha_4)^2} \right] \quad (4.63)$$

It is worth noting that the subtraction term is clearly $1/\epsilon$ times the bubble in $D = 4 - 4\epsilon$ dimensions. We can expand the subtracted integrand and integrate using `hyperInt` to get the expansion for the contribution to the amplitude

$$\Gamma(2\epsilon) \left[\frac{1}{\epsilon} + 5 + (19 - 6\zeta(2))\epsilon + \dots \right] \quad (4.64)$$

which we have verified is in agreement with [23].

General subtraction terms: Before describing the full subtraction scheme we define a new deformation of the integral in (4.45):

$$I_G(\{p_\gamma\}) = \int_0^\infty \omega_G \left(\prod_\gamma u_\gamma^{p_\gamma \epsilon} \right) \left(\prod_e U_e \right) \frac{1}{\mathcal{U}^{D/2}(U)} \left(\frac{\mathcal{U}(U)}{\mathcal{F}(U)} \right)^{d_G} \quad (4.65)$$

and we call the measure

$$\omega_G = d^{E-1} \log(U_e/U_{e_*}) \quad (4.66)$$

If we compute these integrals for finite p_γ , we clearly recover our Feynman integral of interest when $p_\gamma \rightarrow 0$. We will be able to recognize our subtraction terms as products of integrals of this type. It is amusing to note that for generic p_γ , we are dressing the stringy canonical integral associated with our geometry by an additional projectively invariant function which respects the factorization properties of our geometry. In this sense, Feynman integrals are natural deformations of stringy canonical integrals

evaluated at specific p_γ . We now consider the general case of a subtraction term associated with a set $\{u_\gamma\}_F$ of multiple $u_\gamma \rightarrow 0$. Our starting point is the representation of the integral as (4.45). If we leverage the binary property of the u -variables, it is manifest that we get the factorization

$$\prod_{e \in G} U_e \rightarrow \prod_{e \in G/\gamma}^{G/\gamma} U_e \prod_{e \in \gamma} U_e \quad (4.67)$$

where it is understood that in the U_e on the right-hand side, the underlying u 's are simplified and can be intrinsically defined with respect to their associated graphs G/γ and γ now. Moreover, the factorization properties of the \mathcal{U} and \mathcal{F} guarantee the factorization of the the rest of the integrand up to factors which we now describe. Note that the \mathcal{U} polynomial has a term proportional to $u_\gamma^{L_\gamma}$ for all u 's. Therefore, if we consider the factorization

$$\mathcal{U}_G \rightarrow \mathcal{U}_\gamma \times \mathcal{U}_{G/\gamma} \quad (4.68)$$

Lastly, we need to understand how the measure factorizes in this limit. Again leveraging the binary property of the u_γ 's, the factorization of the measure into that on G/γ and on γ is manifest. We want to extract

$$d \log(u_{\gamma_1}) \dots d \log(u_{\gamma_F}) \quad (4.69)$$

from the measure and then have a remaining form on the γ_i and $G/\gamma_1 \dots \gamma_F$ spaces of u 's. It is clear how to do this. We can first consider the γ 's in the forest which have no subgraphs themselves.

$$\omega_G = \sum_e (-1)^e d \log(U_1) \dots d \widehat{\log(U_e)} \dots d \log(U_E) \quad (4.70)$$

where the hat means we omit that term in the sum. This can of course be written by choosing a particular e_* and writing

$$\omega_G = d \log(U_1/U_{e_*}) \dots d \log(U_E/U_{e_*}) \quad (4.71)$$

And in the limiting case we simply have

$$\omega_G \rightarrow d \log(u_{\Gamma_1}) \dots d \log(u_{\Gamma_F}) \prod_{\Gamma_i/F} \omega_{\Gamma_i/F} \quad (4.72)$$

where

$$\omega_{\Gamma_i/F} = \prod_e^{\Gamma_i/F} \wedge d \log(U_e/U_{(e_*)_i}) \quad (4.73)$$

Putting everything together this means we again get

$$S_F = \int (u_{\Gamma_i/F})^{d_{\Gamma_i}} d \log(u_{\Gamma_i/F}) \left[\omega_{G/F} \frac{\prod_e^{G/F} U_{e/F}}{\mathcal{U}_{G/F}^{D/2}} \left(\frac{\mathcal{F}_{G/F}}{\mathcal{U}_{G/F}} \right)^{d_G} u_{\gamma_j/F}^{p_{\gamma_j} \epsilon + \delta p_{\gamma_j} \epsilon} \right] \times \left[\omega_{\Gamma_i/F} \frac{\prod_e^{\Gamma_i/F} U_{e/F}}{\mathcal{U}_{\Gamma_i/F}^{D/2}} u_{\gamma_j/F}^{p_{\gamma_j} \epsilon + \delta p_{\gamma_j} \epsilon} \right] \quad (4.74)$$

Then because the u_{γ_i} are integrated between 0 and 1, it is completely trivial that this produces the inverse powers of d_{Γ_i} .

$$S_F = \frac{1}{d_{\Gamma_i}} I_{G/F}(\{p'_\gamma\}_{G/F}) \prod_i I_{\gamma_i/F}(\{p'_\gamma\}_{\gamma_i/F}) \quad (4.75)$$

It is worth reiterating what we have done to emphasize what is nontrivial. The factorization properties of the \mathcal{U} and \mathcal{F} polynomials were always present of course, and the u_γ 's cannot add or detract from these in any way.

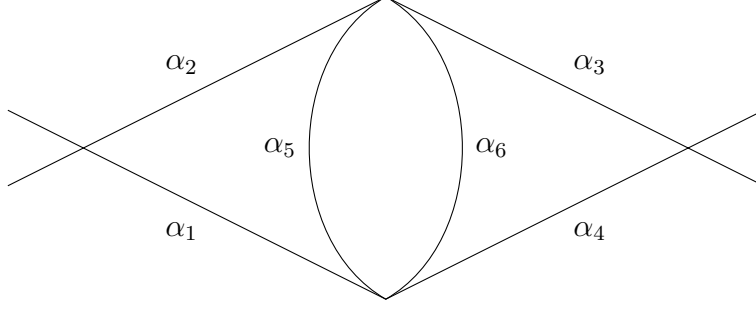


Figure 4.2: Three loop single scale graph in ϕ^4 with $m_e^2 = 0$

Eye Graph As a final example we calculate the Laurent expansion of the three-loop graph in Fig. 4.2 again with $m_e^2 = 0$ and a single scale s which we set to 1. The Symanzik polynomials in this case are

$$\mathcal{U} = (\alpha_1 + \alpha_2)(\alpha_3 + \alpha_4)(\alpha_5 + \alpha_6) + \alpha_5\alpha_6(\alpha_1 + \alpha_2 + \alpha_3 + \alpha_4) \quad (4.76)$$

$$\mathcal{F} = \alpha_1\alpha_2[(\alpha_3 + \alpha_4)(\alpha_5 + \alpha_6) + \alpha_5\alpha_6] + \alpha_3\alpha_4[(\alpha_1 + \alpha_2)(\alpha_5 + \alpha_6) + \alpha_5\alpha_6] + \alpha_5\alpha_6(\alpha_1\alpha_3 + \alpha_2\alpha_4) \quad (4.77)$$

We understand from our facet and compatibility rules that there are three divergent rays associated with subgraphs $\gamma_{1256}, \gamma_{3456}, \gamma_{56}$ which have degrees of divergence $2\epsilon, 2\epsilon, \epsilon$ respectively. These rays generate two, two-dimensional cones. One cone is generated by the rays of γ_{56} and γ_{1256} and the other by the rays of γ_{56} and γ_{3456} . Therefore, merely at the combinatorial level, the subtracted integrand will be

$$\Omega' = \Omega - \Omega_{1256} - \Omega_{3456} - \Omega_{56} + \Omega_{56,1256} + \Omega_{56,3456} \quad (4.78)$$

where these integrands are computed via the prescription above. We will not state the explicit integrand as it is unambiguously furnished by the prescription. We will though, state which products of lower dimensional integrals these subtraction terms

are equal to. We have that

$$\int \Omega_{1256} = \frac{\Gamma(1-3\epsilon)^2}{\Gamma(2-6\epsilon)} I_G^{\text{par}} \quad (4.79)$$

where I_G^{par} is the Feynman integral for the parachute in $D = 4 - 2\epsilon$ with $m_e^2 = 1$ and no external kinematics, an expansion we already computed previously. Note that this subtraction term indeed reflects the expected factorization structure into a product of a bubble G/γ_{1256} (though with a shifted value of ϵ) and the parachute γ_{1256} .

Combining everything, this means that the associated subtraction term, including the inverse power of ϵ , is equal to

$$S_{1256} = \frac{1}{2\epsilon} \frac{\Gamma(1-3\epsilon)^2}{\Gamma(2-6\epsilon)} I_G^{\text{par}} \quad (4.80)$$

which is of course equal to S_{3456} . The term $\Omega_{56,1256}$ is even simpler, leaving merely that residual bubble in $D = 4 - 6\epsilon$

$$S_{56,1256} = \frac{1}{2\epsilon} \frac{1}{\epsilon} \frac{\Gamma(1-3\epsilon)^2}{\Gamma(2-6\epsilon)} \quad (4.81)$$

and this is of course equal to $\Omega_{56,3456}$. The only term we have computed already is

$$\int \Omega_{56} = \int \omega_{1234} u_{12}^{\epsilon/2} u_{34}^{\epsilon/2} \frac{\left(\frac{(\alpha_1 + \alpha_2)^3 (\alpha_3 + \alpha_4)^3}{(\alpha_1 \alpha_2 (\alpha_3 + \alpha_4) + \alpha_3 \alpha_4 (\alpha_1 + \alpha_2))^2} \right)^{3\epsilon/2}}{(\alpha_1 + \alpha_2)^2 (\alpha_3 + \alpha_4)^2} \quad (4.82)$$

this is simply an integral representation for a product of two bubbles combined in a single projective integral, but for the fact that we have the $p_{12}, p_{34} = 1/2$ exponents. We simply perform the associated subtractions on this integrand to compute its Laurent expansion. It would be interesting to explore whether mild deformations of the subtraction scheme allow one to map in the simplest way ever subtraction term to a product of lower integrals calculated previously in the expansion of the amplitude.

This reduces the problem of finding the Laurent expansion to the calculation of finite integrals. We verified via numerical integration that our results are in agreement with those found in [23, 26]

$$\Gamma(3\epsilon)I_G = \frac{\Gamma(1-\epsilon)^4\Gamma(1+3\epsilon)}{\Gamma(1-4\epsilon)3\epsilon^3} (1 + 7\epsilon + 31\epsilon^2 + (103 + 36\zeta(3))\epsilon^3 + \dots) \quad (4.83)$$

Generalities The aforementioned subtraction scheme is easily implemented in general, reducing the problem of extracting the Laurent expansion to the intrinsically complex problem of performing finite integrals which expand in transcendental functions. Beyond its utility in evaluating Feynman integrals, and hopefully spreading evaluation of multi-loop UV divergent Feynman integrals to the masses, this form of the expansion manifesting factorization to lower loop Feynman integrals might help manifest the remarkable properties that the existence of renormalization group equations imply about the amplitude. In this way, this subtraction scheme constitutes a step toward understanding renormalization group equations from a more on-shell perspective.

4.6 Conclusions and Outlook

In this section we have addressed a not-well-appreciated problem: extracting the Laurent expansion in ϵ of a generic Feynman integral with UV divergences. We did this by discovering u -variables associated to such a Feynman integral, which realize the Feynman polytope as a binary geometry. The renormalization prescription is motivated by and closely resembles that of [24] for string amplitudes. The prescription is completely canonical and algorithmic. This also opens directions for further inquiry. In examples we have seen that the u -variables obey a kind of u -equations which are not perfect but have slightly more structure than a generic polynomial consistent with a binary geometry. It would be interesting to specify how to obtain the u -equations

from the graph itself. This would then facilitate an entirely intrinsic description of the Feynman integral in terms of u -variables associated to subgraphs, which would be a new fascinating representation of Feynman integrals. Along this line of thinking, it would also be interesting to see what u -variables might tell us about combining graphs. There is of course the direction of extending these results to IR divergences. We will be presenting closely related progress on this front in a future work.

Part III

$$\alpha' \neq 0$$

Chapter 5

Stringy Completions of Standard Model Amplitudes from the Bottom Up

5.1 Introduction

Already at tree level one can diagnose the need for the UV completion of amplitudes with graviton exchange. Such amplitudes grow with center-of-mass energy and need to be unitarized below the Planck scale. This problem is of course not unique to gravitational amplitudes; scattering of longitudinally polarized W bosons exhibits the same growth with energy, a violation of unitarity which we now understand to be remedied by a weakly coupled Higgs. What is more surprising is that the force we have known the longest has proven the hardest to UV complete.

This state of affairs can be understood in the context of unitarity constraints on massless scattering amplitudes. From this point of view, the graviton mediating a universally attractive force is understood as a consequence of consistent factorization of massless scattering amplitudes with a helicity two particle. By dimensional

analysis, the helicity demands gravity has an irrelevant coupling and an associated growth in energy for amplitudes exchanging gravitons. So, consistency of massless scattering both make sense of gravity's privileged position as our oldest force on the books and implies it is the most immediately problematic long-range force at high energies.

But why is it harder to UV complete than W scattering? *Because* it is a long-range force. High energy growth of the amplitude is not the only way to violate unitarity; it can be violated at low energies as well, by not having a positively expandable imaginary part. Contact interactions produce neither singularities nor imaginary parts for tree level amplitudes and therefore do not directly impose additional *a priori* constraints on the amplitude. This affords a certain freedom in engineering UV-completions by resolving a contact interaction into particle production. This is not the case for amplitudes with massless exchanges. The same consistency conditions which imply gravity's universal attraction are positivity constraints which must be respected by any putative UV-completion. This proves to be a surprisingly stringent constraint.

But we also know gravity need not merely be a fly in our unitarity ointment; in the context of string theory, gravity is *required* for unitarity. And by studying consistency conditions on amplitudes from the right point of view, we can see in an analogous way that gravity is not an obstruction, but rather a facilitator of certain means of unitarization. That is what we find in particular for the class of Ansätze considered in this work. And by considering stringy Ansätze motivated by amplitudes with graviton exchange, we can initiate a program of building UV completions of the standard model (SM) from the bottom up.

The outline of this work is as follows. In Sec. 5.2 we review unitarity constraints on amplitudes and motivate a certain class of stringy Ansätze for $2 \rightarrow 2$ scattering which unitarize field theory amplitudes. These amplitudes resemble closed string

amplitudes, but are strictly studied from the point of view of unitarity constraints on $2 \rightarrow 2$ scattering, not any explicit realization in some string background. In Sec. 5.3, with the goal of studying standard model and beyond the standard model (BSM) amplitudes, we study $SO(N)$ and $SU(N)$ gauge boson amplitudes. While in this neighborhood, we also probe unitarity constraints in general spacetime dimensions, finding evidence for example that the maximum allowed N for $SO(N)$ is 32 as realized by the $SO(32)$ heterotic string. Finally, in Sec. 5.4 we study Higgs scattering and comment on the compatibility of our amplitudes with the SM.

5.2 Unitarity and UV Completion

Before moving on to specific Ansätze, it is useful to review the constraints of Lorentz invariance and unitarity on the amplitude and some well-known instances of tree-level UV-completion in the absence of gravity. This will illustrate the fundamental challenge with even tree-level UV-completion in the presence of gravity and motivate the stringy form factor which will dress all the amplitudes considered herein. We always consider amplitudes with massless external states and employ spinor helicity with mandelstams

$$s = 2p_1 \cdot p_2 \qquad t = 2p_2 \cdot p_3 \qquad u = 2p_1 \cdot p_3 \qquad (5.1)$$

and $s + t + u = 0$. For on-shell kinematics the spinor-brackets obey $\langle ij \rangle = \pm [ij]^*$ with the sign depending whether the states are ingoing or outgoing. When imposing positive expandability on a basis of orthogonal polynomials, the argument of the polynomials is $\cos \theta$ where

$$t = -\frac{s}{2}(1 - \cos \theta) \qquad u = -\frac{s}{2}(1 + \cos \theta) \qquad (5.2)$$

5.2.1 Review of Unitarity Constraints

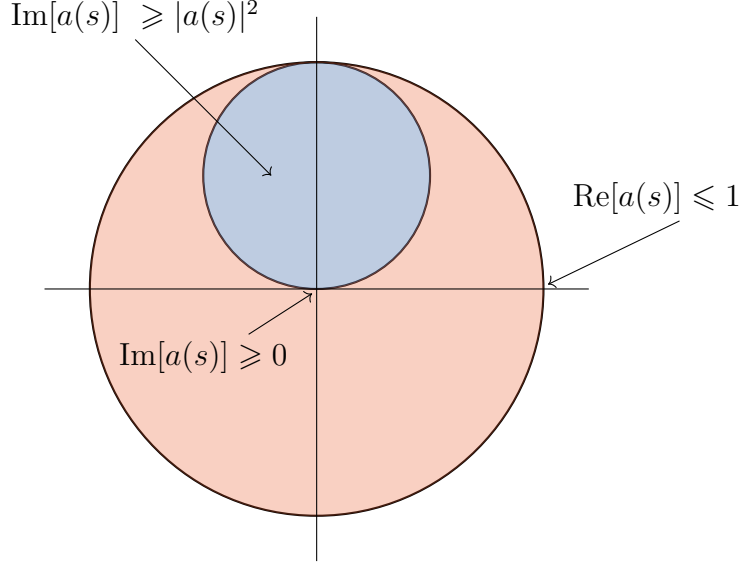


Figure 5.1: Above we depict in the complex $a(s)$ the bounds relevant for our analysis. The non-perturbative bound on $a(s)$ is shaded in blue. At weak coupling, one can diagnose unitarity violation at large s using the weaker, pink shaded condition. We then impose positivity of the imaginary part which is the weak-coupling portion of the blue shading (where $|a(s)|^2$ is parametrically suppressed.)

Lorentz invariance allows us to expand the elastic two-to-two amplitude for massless particles as

$$\mathcal{A}(1_{r_1}^{h_1}, 2_{r_2}^{h_2}, 3_{r_3}^{h_3}, 4_{r_4}^{h_4}) = 16\pi \sum_J (2J+1) a_{J,R}^{\{h_i\}}(s) \mathbb{G}_{D,J}^{\{h_i\}}(\cos \theta) \mathbb{P}_R(\{r_i\}) \quad (5.3)$$

The $\mathbb{G}_{D,J}^{\{h_i\}}$ are the relevant basis of orthogonal (in general spinning) polynomials for the scattering process in question, corresponding to spin- J exchange in D space-time dimensions with external helicities $\{h_i\}$. We will state the particular polynomial basis for each process we consider. The $\mathbb{P}_R(\{r_i\})$ are projectors in the internal symmetry space for representations r_i exchanging representation R (in the s -channel here). The weights $a_{J,R}^{\{h_i\}}(s)$ in front of these two sets of orthogonal projectors are partial wave coefficients, which depend on Mandelstam s and the constants in the amplitude.

Unitarity then imposes the bound on partial waves

$$\text{Im}[a(s)] \geq |a(s)|^2 \quad (5.4)$$

where the constraint holds for each $a_{J,R}^{\{h_i\}}(s)$ individually and we suppress the labels on $a(s)$.¹ Noting that this equation is identical to

$$(\text{Im}[a(s)] - 1/2)^2 + \text{Re}[a(s)]^2 \leq \frac{1}{4} \quad (5.5)$$

the unitarity bound clearly forces each $a(s)$ to lie in the Argand circle depicted in figure 5.1. Two limits of this non-perturbative constraint will be useful for our weak coupling analysis, each simplifying either the left or right-hand side of equation (5.4). These limits are

$$\begin{cases} \text{Im}[a(s)] \geq 0 & \text{positive expandability at singular loci} \\ |a(s)| \leq 1 & \text{boundedness at large } s \end{cases} \quad (5.6)$$

The tension between these two conditions and causality imposed via Regge boundedness in the complex s plane, places stringent constraints on the form of weakly-coupled gravitational UV completions. We will be considering processes with graviton exchange, which exhibit unitarity violation at large s and need to be UV completed below the Planck scale. Rather than give up on weak coupling, we will posit an Ansatz for a weakly coupled completion which softens the high-energy behavior. This will be done by the introduction of new heavy states and positive expandability at the associated singularities must be checked. This constrains the low-energy data in the context of the specific Ansatz to which we have committed. In this work we will

¹The S -matrix is required to be a positive operator in general, therefore in general this constraint is a statement about positivity of a matrix, but we will be studying amplitudes for which each individual $a(s)$ obeys (5.4).

manifest all necessary conditions on the amplitude other than positive expandability of residues, leaving it as the non-trivial condition which must be checked.

In addition to boundedness at fixed angle, positive expandability of the imaginary part, and Regge boundedness, we impose additional reasonable constraints on the spectrum. In particular we require that there are a fixed number of spins at a given mass i.e. that the Chew-Frautschi plot has no spikes. We summarize the full set of constraints below

- Boundedness at fixed angle and high energy via

$$|a(s)| \leq 1 \tag{5.7}$$

- Regge boundedness for complex s :

$$\lim_{|s| \rightarrow \infty} \mathcal{A}/s^2 \rightarrow 0 \quad \text{fixed } t < 0 \tag{5.8}$$

- Fixed number of spins at a given mass by imposing that any residue in s is a polynomial in t .
- Positive expandability of the imaginary part, the only condition which we will not manifest.

5.2.2 Completions

Before moving on to constructing gravitational completions, it is useful to recapitulate a well-known example of a non-gravitational tree level amplitude requiring UV completion: the non-linear sigma model. This will help frame the unique challenge and opportunities afforded in studying gravitational completions.

Non-linear Sigma Model: We consider the amplitude for four Goldstones

$$\mathcal{A}_{\text{NLSM}} = -\frac{1}{f^2} \left[(s+t) (\text{Tr}(T^a T^b T^c T^d) + \text{Tr}(T^d T^c T^b T^a)) + (s+u) (\text{Tr}(T^b T^a T^d T^c) + \text{Tr}(T^c T^d T^a T^b)) \right. \\ \left. + (t+u) (\text{Tr}(T^a T^c T^b T^d) + \text{Tr}(T^d T^b T^c T^a)) \right] \quad (5.9)$$

With the T^a standard generators of the defining representation of $SU(N)$. We have scalar partial wave for any exchanged representation

$$a_0(s) \sim \frac{s}{f^2} \quad (5.10)$$

which means the amplitude needs to be unitarized before $s \sim f^2$. Already in this case we can illustrate how the tensions between boundedness at fixed-angle and Regge boundedness in the complex plane don't allow the most naive softenings of the amplitude. For example, we could try the exponential form factor

$$\hat{\mathcal{A}} = e^{-\frac{s^2+t^2+u^2}{M^4}} \mathcal{A}_{\text{NLSM}} \quad (5.11)$$

which is exponentially soft at fixed angle, but diverges exponentially for imaginary s at fixed t . If we settle on more humble expectations for softening, the most obvious way to soften the amplitude is to divide it by something i.e. introduce a massive exchange, which should in particular have positive mass-squared. Therefore the simplest Ansatz is

$$\hat{\mathcal{A}}_{\text{NLSM}} = -\frac{1}{f^2} \left[\frac{s}{1 - \frac{s}{M^2}} (\mathbb{P}_1^s + \mathbb{P}_2^s) + \frac{t}{1 - \frac{t}{M^2}} (\mathbb{P}_1^t + \mathbb{P}_2^t) + \frac{u}{1 - \frac{u}{M^2}} (\mathbb{P}_1^u + \mathbb{P}_2^u) \right] \quad (5.12)$$

where we decomposed the usual flavor-ordered form into the form manifesting exchanges in respective channels, with

$$\begin{cases} \mathbb{P}_1^s = \frac{2}{N} \delta_{ab} \delta_{cd} \\ \mathbb{P}_2^s = d_{abe} d_{cde} \end{cases} \quad (5.13)$$

which correspond to the singlet and symmetric adjoint irreducible flavor exchanges. Now, the partial waves are clearly bounded so long as $M^2 < f^2$ i.e. the UV completion scale is below the UV cutoff f^2 . But we are not done, the amplitude now has a singularity at M^2 in each channel and the residue must be positive. This additionally gives the condition $f^2 > 0$. The low-energy amplitude alone was agnostic about the sign of f^2 , but once committed to a certain UV completion, it is possible to constrain low-energy data by imposing positive expandability of residues. We will find in the case of gravitational amplitudes that constructing a UV completion is more difficult, the upshot being that the constraints on low energy data are far more interesting.

Graviton Exchange: Our objective is to study UV completions in the presence of graviton exchange, so let's see if we can apply the lessons from the non-linear sigma model to the case of four identical scalars exchanging gravitons. The amplitude is

$$\mathcal{A}_{\text{grav}} = -\frac{1}{M_P^2} \left(\frac{tu}{s} + \frac{su}{t} + \frac{st}{u} \right) \quad (5.14)$$

Where M_P is the reduced Planck mass $M_P^{-2} = 8\pi G$. Again we see that the scalar partial wave grows with energy

$$a_0(s) \sim \frac{s}{M_P^2} \quad (5.15)$$

As with the NSLM amplitude, we do not want to give up on perturbativity; we will pursue a tree level UV completion. It was easy to engineer a unitarization of the NLSM amplitude, what could be the difficulty with gravity? Particle production at

low energies. More explicitly, unlike the NLSM amplitude, the graviton exchanges already impose a sign condition on the coupling squared. This fixed sign in the low-energy amplitude is something our putative completion will have to respect, and it has surprisingly drastic consequences. The issue has everything to do with these low energy poles, and can already be seen in the context of ϕ^3 amplitudes. Suppose we wanted to UV-improve the behavior of the ϕ^3 amplitude; we can try to mimic the strategy employed in the NLSM. Focusing on a single channel

$$\frac{g^2}{s} \rightarrow \frac{g^2}{s} \times \frac{1}{\prod_{i=1}^n \left(1 - \frac{s}{M_{s,i}^2}\right)} \quad (5.16)$$

where we have introduced n new massive poles to soften the high-energy behavior to $\frac{1}{s^{n+1}}$. As with the initial $s = 0$ pole, we require the residue on each of these massive poles to be positive. But notice that because there is no pole at infinity, a residue theorem guarantees that some of the residues must have the wrong sign. We have arrived at the need for an infinite number of poles to have a hope of softening this amplitude in a way consistent with unitarity. We can make an Ansatz for a factor with an infinite number of poles in each channel which will dress the amplitude:

$$\mathcal{D}(s, t, u) = \frac{N(s, t, u)}{\prod_i (s - M_{s,i}^2) \prod_j (t - M_{t,j}^2) \prod_k (u - M_{u,k}^2)} \quad (5.17)$$

In order to further constraint this putative dressing factor, we will impose the well-motivated condition that there is a finite number of spins at given mass-level. Analytically, this means that the residue at each massive pole must be a polynomial in the remaining Mandelstam invariant, in particular the numerator must cancel the poles in the other channel upon taking such a residue. The denominator is a product

of polynomials so it's natural to have the same Ansatz for the numerator

$$N(s, t, u) = \prod_i (s + r_{s,i}) \prod_j (t + r_{t,j}) \prod_k (u + r_{u,k}) \quad (5.18)$$

The residue is then

$$\text{Res}_{s \rightarrow M_{s,i}^2} \mathcal{D} \propto \frac{\prod_j (t + r_{t,j}) \prod_k (t + M_{s,i}^2 - r_{u,k})}{\prod_j (t - M_{t,j}^2) \prod_k (t + M_{u,k}^2 + M_{s,i}^2)} \quad (5.19)$$

requiring cancellation of the remaining poles in all channels yields the condition:

$$M_{s,i}^2 + M_{t,j}^2 \in \{r_{u,k}\} \quad (5.20)$$

plus its cyclic rotations. We will proceed with the most obvious way of solving this constraint, which is to make all six of these putative sets one mass scale times the positive integers. We will call this mass scale M_s in which case we find

$$\mathcal{D}(s, t, u) = \frac{\prod_{i=1}^{\infty} (s + M_s^2 i) \prod_{j=1}^{\infty} (t + M_s^2 j) \prod_{k=1}^{\infty} (u + M_s^2 k)}{\prod_{i=1}^{\infty} (s - M_s^2 i) \prod_{j=1}^{\infty} (t - M_s^2 j) \prod_{k=1}^{\infty} (u - M_s^2 k)} \quad (5.21)$$

Which we recognize as a famous function

$$\Gamma^{\text{str}} = - \frac{\Gamma(-\alpha' s) \Gamma(-\alpha' t) \Gamma(-\alpha' u)}{\Gamma(\alpha' s) \Gamma(\alpha' t) \Gamma(\alpha' u)} \quad (5.22)$$

with $\alpha' = \frac{1}{M_s^2}$. Though we did not impose these conditions, it is readily verified from the form in (5.22) that the amplitude satisfies Regge boundedness and boundedness at fixed angle and high energy. We summarize the properties of this amplitude below:

- Exponentially soft at high-energy, fixed angle

◦ Regge limit:

$$\lim_{|s| \rightarrow \infty} \Gamma^{\text{str}} = s^{\alpha' t} \quad \text{fixed } t < 0 \quad (5.23)$$

◦ A residue in s at level n i.e. mass $M^2 = nM_s^2$ is a polynomial of degree $2n$ in t and is in particular

$$\text{Res}_{s \rightarrow n/\alpha'} \Gamma^{\text{str}} = \frac{1}{n!(n-1)!} \left(\prod_{i=1}^{n-1} (i + \alpha' t)^2 \right) t(n + \alpha' t) \quad (5.24)$$

If we dress our graviton exchange amplitude (5.14) with Γ^{str} this is in fact the four dilaton amplitude in type IIB. This quasi-derivation of the Virasoro-Shapiro amplitude is far from new; Virasoro remarked on it in his original paper after all [27], but it is surprising. Though we by no means claim to be proving that this is the unique way of unitarizing graviton exchange, it is remarkable that with very little input and at each stage making the simplest of moves, we arrive at a form factor with Regge boundedness and *exponential* softness at fixed angle. These two conditions are extremely non-trivial to engineer, but we were able to build an amplitude with both properties by seeking a UV completion in the presence of gravity. We will find that the arrow goes both ways. Gravity does not merely necessitate the discovery of this somewhat sophisticated completion, it is in fact necessary for the use of such a completion.

Partial Wave Unitarity: With an Ansatz in hand for gravitational amplitudes, we can now impose positive expandability of the residues. The amplitude is by construction positive on the massless poles, so we only need to check the positive expandability on the new poles in the completion. These come from Γ^{str} and the residue of Γ^{str} at general mass level n was stated above in (5.24).

Now that we have Γ^{str} and all of its wonderful properties at our disposal, one might hope that with this unitarizing hammer, any field theory amplitude looks like

a nail. We can test it on something simple such as ϕ^4 theory, in which case we merely check the partial wave expansion of Γ^{str} . At the first mass-level one can readily verify that

$$\text{Res}_{s \rightarrow 1/\alpha'} \Gamma^{\text{str}} \propto \left(P_2(\cos \theta) - P_0(\cos \theta) \right) \quad (5.25)$$

so we find a negative residue. We could try coupling it to gravity; the amplitude is just augmenting the four-dilaton amplitude in type IIB by a contact interaction:

$$\mathcal{A}^\lambda = \Gamma^{\text{str}} \left[-\frac{1}{M_P^2} \left(\frac{tu}{s} + \frac{su}{t} + \frac{st}{u} \right) + \lambda \right] \quad (5.26)$$

We have

$$\text{Res}_{s \rightarrow 1/\alpha'} \mathcal{A}^\lambda \propto \frac{1}{70} P_4(\cos \theta) + \left(\frac{2}{7} + \frac{\alpha' M_P^2 \lambda}{6} \right) P_2(\cos \theta) + \left(\frac{7}{10} - \frac{\alpha' M_P^2 \lambda}{6} \right) P_0(\cos \theta) \quad (5.27)$$

And so we find the two-sided bound on λ :

$$\frac{-12}{7} \leq \lambda \frac{M_P^2}{M_s^2} \leq \frac{21}{5} \quad (5.28)$$

So we can have a quartic coupling so long as it does not overwhelm the graviton exchange piece of the amplitude. In this sense, we are able to UV-improve $\lambda\phi^4$ theory at high energies so long as we couple it to gravity. In the context of these Ansätze, gravity is not an obstruction to unitary UV softening, but a necessity. One could imagine that perhaps this was merely the absence of long-distance physics at low energies that was the problem, not gravity in particular. We can test this on pure gauge boson scattering. Checking unitarity in this case requires an additional layer of calculation: 6- j coefficients.

6- j Coefficients: Since we are working at tree level, taking the imaginary part merely amounts to taking the residue (5.19). Moreover, in practice we find that the

strongest constraints come at low mass levels, stabilizing at the latest by mass level three for the amplitudes considered in this analysis. The most non-trivial aspect of testing positive expandability is in re-expanding the color structures of (5.41), which as such are not expressed in the basis of s -channel projectors. The full set of projectors in any one of the three channels, full set meaning the projectors for all representations that can be exchanged in this channel, are an orthogonal basis of projectors depending on the external indices. Therefore, each projector in the t and u channels can be uniquely expanded in terms of the projectors in the s -channel:

$$\mathbb{P}_R^t = \sum_{R'} C_{R,R'}^{t,s} \mathbb{P}_{R'}^s \quad (5.29)$$

where we have suppressed the external indices which these projectors are functions of. This is merely the crossing the equation, and it is solved trivially by contraction given that the basis of projectors corresponding to the exchange of irreps R are orthogonal. That is

$$C_{R,R'}^{t,s} = \frac{\mathbb{P}_R^t \cdot \mathbb{P}_{R'}^s}{\mathbb{P}_{R'}^s \cdot \mathbb{P}_{R'}^s} \quad (5.30)$$

where the dot denotes contraction with the relevant invariants on the external indices which have been suppressed. This coefficient $C_{R,R'}^{t,s}$ is a 6- j symbol, depending on the four external representations in addition to the two exchanged representations R and R' . These must be calculated for the group and representations in question. Once this is done, the projection onto irreducible exchanged states is straightforward. A useful resource for the determination of these projectors is [28].

We can consider an example to illustrate how dependence on group theory factors can emerge in the unitarity constraints. For example, we can consider dressing our scalar amplitude with some global symmetry charged under representation R

$$\mathcal{A}^{abcd} = \Gamma^{\text{str}} \left(\delta^{ab} \delta^{cd} \frac{tu}{s} + \delta^{bc} \delta^{ad} \frac{su}{t} + \delta^{ac} \delta^{bd} \frac{st}{u} \right) \quad (5.31)$$

in preparation to expand onto exchange of a singlet in the s -channel we, need to expand the kronecker deltas in the t and u in terms of those in the s -channel. Using (5.29) this is a simple task contracting indices

$$C_{\cdot,\cdot}^{t,s} = \frac{1}{d_R} \quad (5.32)$$

Therefore we have

$$\mathcal{A}^{abcd} = \delta^{ab}\delta^{cd}\Gamma^{\text{str}} \left(\frac{tu}{s} + \frac{1}{d_R} \frac{su}{t} + \frac{1}{d_R} \frac{st}{u} \right) \quad (5.33)$$

where we now do the ordinary kinematic unitarity analysis on the function multiplying this s -channel projector. The spin-0 exchange at the first mass-level produces the bound

$$d_R \leq 7 \quad (5.34)$$

Gauge Bosons We can return to analyzing our four gauge boson amplitude

$$\mathcal{A} = \frac{g_{YM}^2}{3} \langle 12 \rangle^2 [34]^2 \Gamma^{\text{str}} \left(\frac{\mathbb{P}_{\text{Adj}}^s - \mathbb{P}_{\text{Adj}}^t}{st} + \frac{\mathbb{P}_{\text{Adj}}^t - \mathbb{P}_{\text{Adj}}^u}{tu} + \frac{\mathbb{P}_{\text{Adj}}^u - \mathbb{P}_{\text{Adj}}^s}{su} \right) \quad (5.35)$$

where we have the adjoint projectors

$$\mathbb{P}_{\text{Adj}}^s = f^{abe} f^{edc} \quad (5.36)$$

and similarly for the t and u channels. We can focus on the explicit case of $SO(N)$. Positive expandability is already violated at level one for the most subleading Regge trajectory in the largest representation exchanged between the adjoints. The ex-

changed state and associated coefficient in the partial wave expansion are:

$$\begin{array}{|c|c|} \hline & \\ \hline & \\ \hline \end{array} \quad : \quad -g_{YM}^2$$

$m^2 = 1, J = 0$

(5.37)

Where the left is the Young tableaux for the exchanged representation of $SO(N)$, the mass-squared in units of M_s^2 , and the spin. The coefficient is identical for the analogous representation of $SU(N)$. Perhaps the most salient feature of this violation is the indication that the actual heterotic string amplitude becomes unitarity violating once gravity becomes *too weak* relative to the gauge interactions; we cannot have the gauge interactions all on their own. Yet again we find that unitarity of these amplitudes requires graviton exchange. String theory of course necessitates gravity, so unitarity violation in the complete absence of gravity, is perhaps not so surprising from that point of view. But what is interesting about this analysis is that we can move around in the space of parameters untethered to any choice of background or class of compactifications and make what appear to be robust observations about constraints on this class of amplitudes with graviton exchange. For instance, extending the above to our full Ansätze will reveal that the generalization of (5.37) is in tension, but compatible, with the weak gravity conjecture (WGC).

Summary of Gravitational Ansätze: Now that we have motivated both the prefactor Γ^{str} and that it must be used for completions along with graviton exchange, we can discuss more completely the rules bounding our Ansätze. The basic game is clear: multiply a massless field theory amplitude by the stringy form-factor Γ^{str} and check positive expandability. Now we discuss the restrictions we impose on the massless field theory amplitudes.

In motivating Γ^{str} we already took as assumptions that the residues of the amplitude in s were polynomials in t . Additionally, we impose Regge boundedness of the amplitude, which in the case of our gravitational amplitudes means bounded by s^2 . Seeing as Γ^{str} goes as $s^{\alpha' t}$ in the Regge limit, the field theory amplitude which this factor multiplies can scale at most as s^2 in the Regge limit, which the graviton exchange piece indeed will. This bounds the dimension of interactions coming from operators that we may put in by hand. As for denominators, one can consider the heterotic-type deformation

$$\frac{1}{s} \rightarrow \frac{1}{s(1 + \alpha' s)} \quad (5.38)$$

It is crucial that this potential tachyon pole has vanishing residue

$$\text{Res}_{s \rightarrow -1/\alpha'} \frac{1}{s(1 + \alpha' s)} \Gamma^{\text{str}} = 0 \quad (5.39)$$

Indeed any product of $(n + \alpha' s)$ for distinct integers n can be added to the denominator and will obey the vanishing residue condition above. But they will violate our condition for a finite number of spins at a given mass level. Recall from (5.19) that at the first mass level we have

$$\text{Res}_{s \rightarrow 1/\alpha'} \Gamma^{\text{str}} \propto t(1 + \alpha' t) = u(1 + \alpha' u) \quad (5.40)$$

which means, considering the t and u channel terms in our amplitude, that we can have our massless poles and the $(1 + \alpha' t)$ or $(1 + \alpha' u)$ factors as well, but no more. Additional factors would fail to cancel upon taking the residue at the first mass level and introduce an infinite number of spins. For simplicity, we only consider these poles on the graviton exchange part of the amplitude.

Therefore, we can compute a massless scattering amplitude, the form of which is itself fixed by consistent factorization into three-particle amplitudes on the massless poles.

This amplitude can have contact terms put in by hand up to scaling as s^2 in the Regge limit. The amplitude is then dressed by Γ^{str} which ensures that the amplitude satisfies all necessary criteria, except potentially positive expandability on the new massive poles. Checking this condition produces constraints on the low energy data.

5.3 Gauge Boson Scattering

First we take up massless gauge boson scattering. At tree-level, the non-trivial $2 \rightarrow 2$ gauge boson amplitudes are those of W 's and gluons. There is a $2 \rightarrow 2$ amplitude for B 's (gauge boson of $U(1)_Y$), but it is purely mediated by graviton exchange and therefore not further constrained by unitarity. In addition to these, in many GUT models we have gauge bosons of $SO(N)$ or $SU(N)$, and we will carry out the analysis for general N in both cases. The massless scattering amplitude is

$$\mathcal{A}(1^{-a}, 2^{-b}, 3^{+c}, 4^{+d}) = \langle 12 \rangle^2 [34]^2 \left[\frac{1}{M_P^2} \left(\frac{\mathbb{P}_1^s}{s} + \frac{\mathbb{P}_1^t}{t} + \frac{\mathbb{P}_1^u}{u} \right) + \frac{g_{\text{YM}}^2}{3} \left(\frac{\mathbb{P}_{\text{Adj}}^s - \mathbb{P}_{\text{Adj}}^t}{st} + \frac{\mathbb{P}_{\text{Adj}}^t - \mathbb{P}_{\text{Adj}}^u}{tu} + \frac{\mathbb{P}_{\text{Adj}}^u - \mathbb{P}_{\text{Adj}}^s}{su} \right) \right] \quad (5.41)$$

we write the color structures in this way as they are the color structures to be expanded in on a factorization channel. In particular, these are

$$\begin{cases} \mathbb{P}_1^s & \delta^{ab} \delta^{cd} \\ \mathbb{P}_{\text{Adj}}^s & f^{abe} f^{edc} \end{cases} \quad (5.42)$$

Note that in the Regge limit, this amplitude already scales as s^2 . Any additional contribution consistent with Regge behavior would require new poles, which we do not have at low energies. Our Ansatz is then

$$\mathcal{A}^{\text{UV}}(1^{-a}, 2^{-b}, 3^{+c}, 4^{+d}) = \Gamma^{\text{str}} \mathcal{A}(1^{-a}, 2^{-b}, 3^{+c}, 4^{+d}) \quad (5.43)$$

or $\mathcal{A}_{\text{Het}}^{\text{UV}}$ which deforms the graviton poles to look heterotic

$$\mathcal{A}_{\text{Het}}^{\text{UV}}(1^{-a}, 2^{-b}, 3^{+c}, 4^{+d}) = \langle 12 \rangle^2 [34]^2 \left[\frac{1}{M_P^2} \left(\frac{\mathbb{P}_1^s}{s(1 + \alpha' s)} + \frac{\mathbb{P}_1^t}{t(1 + \alpha' t)} + \frac{\mathbb{P}_1^u}{u(1 + \alpha' u)} \right) + \frac{g_{\text{YM}}^2}{3} \left(\frac{\mathbb{P}_{\text{Adj}}^s - \mathbb{P}_{\text{Adj}}^t}{st} + \frac{\mathbb{P}_{\text{Adj}}^t - \mathbb{P}_{\text{Adj}}^u}{tu} + \frac{\mathbb{P}_{\text{Adj}}^u - \mathbb{P}_{\text{Adj}}^s}{su} \right) \right] \quad (5.44)$$

We introduce g_s via

$$g_s^2 = \frac{M_s^2}{M_P^2} \quad (5.45)$$

which is the dimensionless strength characterizing how far below the Planck scale the UV completion scale M_s is i.e. how weakly coupled a completion of gravity it is. The strongest constraints come from the pole in the s -channel with the helicity configuration in (5.41). Furthermore, we note that this Ansatz is perfectly good in any number of spacetime dimensions, and we will also consider the necessary but insufficient condition of positive expandability on the *scalar* D -dimensional Gegenbauer polynomials.

5.3.1 Constraints in Four Spacetime Dimensions

In four dimensions, the polynomials are the spinning Gegenbauer's

$$\mathbb{G}_{D=4,J}^{\{h_i\}}(\cos \theta) = d_{h_{12}, h_{34}}^J(\cos \theta) \quad (5.46)$$

where $h_{ij} = h_i - h_j$. In the s -channel for (5.41) these polynomials actually correspond to the scalar Legendre's, and we find the strongest constraints.

SO(N): For $SO(N)$ the adjoints exchange six representations. Positivity of g_{YM}^2 is already a consistency condition imposed at massless level. As can be seen in the left plot of figure 5.2, both the heterotic and non-heterotic cases have an N -independent upper bound on the ratio $\frac{g_{YM}^2}{g_s^2}$ and an upper bounding curve depending on both the

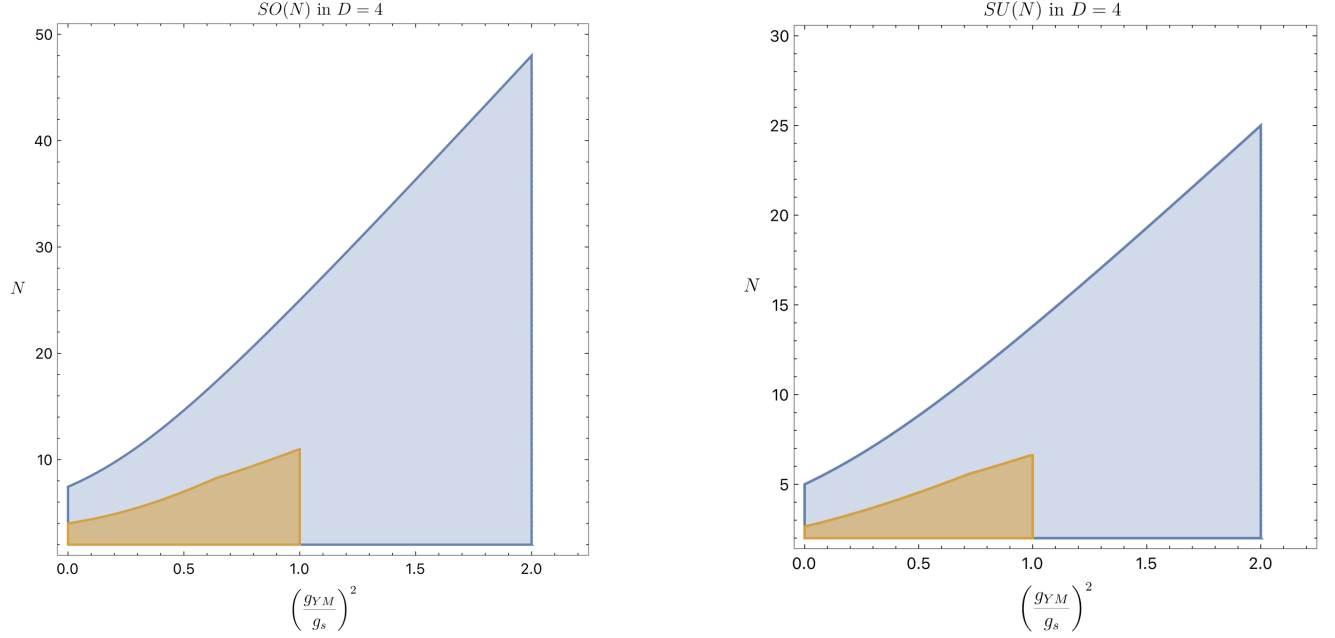


Figure 5.2: Allowed space of N and g_{YM}/g_s in four spacetime dimensions, for $SO(N)$ (left) and $SU(N)$ (right). In both cases, the larger allowed region (blue) has heterotic poles and contains the region without heterotic poles (orange). In the heterotic case, the maximum rank is 24 for both groups and occurs at the coupling $g_{YM}^2 = 2g_s^2$, the value fixed in the heterotic string.

coupling and N . In the non-heterotic case this upper curve is only piecewise smooth, the result of two smooth constraints. In particular, we find for the non-heterotic case, the constraints and the associated mass-level, spin, and $SO(N)$ representation are the

following:

$$\left\{ \begin{array}{ll} \begin{array}{c} \begin{array}{|c|c|} \hline & \\ \hline & \\ \hline \end{array} \\ m^2 = 1, J = 0 \end{array} & g_{YM}^2 \leq g_s^2 \\ \\ \bullet & 1 + \left(\frac{g_{YM}}{g_s} \right)^2 (N - 2) - \frac{N(N-1)}{12} \geq 0 \\ m^2 = 1, J = 0 & \\ \\ \bullet & 4 + 2 \left(\frac{g_{YM}}{g_s} \right)^2 (N - 2) - \frac{N(N-1)}{5} \geq 0 \\ m^2 = 2, J = 0 & \end{array} \right. \quad (5.47)$$

And in the heterotic case simply the two constraints:

$$\left\{ \begin{array}{ll} \begin{array}{c} \begin{array}{|c|c|} \hline & \\ \hline & \\ \hline \end{array} \\ m^2 = 1, J = 0 \end{array} & g_{YM}^2 \leq 2g_s^2 \\ \\ \bullet & 2 + \left(\frac{g_{YM}}{g_s} \right)^2 (N - 2) - \frac{N(N-1)}{24} \geq 0 \\ m^2 = 1, J = 0 & \end{array} \right. \quad (5.48)$$

Where the masses m^2 are measured in units of M_s^2 . The two-by-two tableaux is the largest representation exchanged between the gauge bosons, having dimension $\frac{N(N+1)(N+2)(N-3)}{12}$, and is the exchanged state enforcing the anti-weak-gravity bound.

We note that all constraints come from the most subleading Regge trajectory.

SU(N): For $SU(N)$ the adjoints exchange seven representations, but the situation is analogous, with analogous representations imposing similar bounds. We find for

the non-heterotic case:

$$\left\{ \begin{array}{ll} \begin{array}{c} N_c - 1 \\ \left\{ \begin{array}{c} \begin{array}{|c|c|c|c|} \hline & & & \\ \hline \vdots & \vdots & & \\ \hline & & & \end{array} \end{array} \right. & g_{YM}^2 \leq g_s^2 \\ m^2 = 1, J = 0 & \\ \\ \bullet & 5 + 4 \left(\frac{g_{YM}}{g_s} \right)^2 N - \frac{1+2N^2}{3} \geq 0 \\ m^2 = 1, J = 0 & \\ \\ \bullet & 11 + 5 \left(\frac{g_{YM}}{g_s} \right)^2 N - N^2 \geq 0 \\ m^2 = 2, J = 0 & \end{array} \right. \quad (5.49)$$

and in the heterotic case:

$$\left\{ \begin{array}{ll} \begin{array}{c} N_c - 1 \\ \left\{ \begin{array}{c} \begin{array}{|c|c|c|c|} \hline & & & \\ \hline \vdots & \vdots & & \\ \hline & & & \end{array} \end{array} \right. & g_{YM}^2 \leq 2g_s^2 \\ m^2 = 1, J = 0 & \\ \\ \bullet & \left(\frac{g_{YM}}{g_s} \right)^2 N - \frac{1}{12}(N^2 - 25) \geq 0 \\ m^2 = 1, J = 0 & \end{array} \right. \quad (5.50)$$

We see that for $SU(N)$ again the largest representation enforces an identical bound on the coupling to the $SO(N)$ case, this time the dimension of the representation is $\frac{N^2(N+3)(N-1)}{4}$. These bounds are not the same function of rank r for each group, but the bound on the rank is identical at the maximum value of g_{YM}^2 in the heterotic case, which is $2g_s^2 = \frac{2M_s^2}{M_P^2}$. This is the condition relating the Yang-Mills coupling, the

string scale, and the Planck scale in the heterotic string, a condition which survives compactification as both ten-dimensional couplings get the same volume dilution [29].

Generalization It is possible to derive the second constraint in the heterotic case, that for spin-zero singlet exchange at mass-level one, for a general group in terms of the dimension of the adjoint d_A and the adjoint casimir C_A . This is because when projecting onto the s -channel singlet exchange, the relevant 6- j coefficients are

$$C_{\bullet,\bullet}^{s,t} = \frac{1}{d_A} \quad (5.51)$$

And using the fact that

$$f^{ars} f^{rsb} = C_A \delta^{ab} \quad (5.52)$$

we also find

$$C_{A,\bullet}^{s,t} = \frac{C_A}{d_A} \quad (5.53)$$

with A again denoting the adjoint representation. We do not derive the generalization of the bound $g_{YM}^2 \leq 2g_s^2$ which is observed to be identical in the $SO(N)$ and $SU(N)$ cases. We can unify the statement of the bounds at least for these two groups though, and one might expect that the bound holds in general

$$\left\{ \begin{array}{l} \frac{1}{2} g_{YM}^2 M_P^2 \leq M_s^2 \\ 2 + g_{YM}^2 \frac{M_P^2}{M_s^2} C_A - \frac{d_A}{12} \geq 0 \end{array} \right. \quad (5.54)$$

Phenomenology Here we comment on the implications for four-dimensional phenomenology. First we note the obvious, which is that only the combination of constraints from singlets of the gauge group, massless adjoints, and the largest exchanged representation both for $SO(N)$ and $SU(N)$ combine to produce a bounded region.

The largest representation furnishes the bound relating the UV completion scale, Planck scale, and the coupling:

$$g_{YM}^2 M_P^2 \leq 2M_s^2 \quad (5.55)$$

or with no factor of two on the right-hand side when we do not have heterotic denominators. This is an anti-weak-gravity type of bound: consistent with weak-gravity, but pointing in the opposite direction. But for sufficiently large ranks of the gauge-group, we cannot afford for the gauge-coupling to be too-weak with a low string scale, either. In particular, no global symmetry is allowed for these massless adjoint vectors for $N > 7$ for $SO(N)$ and $N > 5$ for $SU(N)$.

We can discuss these constraints in the context of W bosons and gluons. For the case of heterotic poles, the constraint is the same for both W 's and gluons and is (5.55). For couplings around the GUT scale, we have $g_2^2, g_3^2 \sim \frac{1}{2}$, in which case the putative string scale is bound by

$$\frac{M_P}{2} \lesssim M_s \quad (5.56)$$

meaning the lowest putative string scale is around 10^{18} GeV. For the case without heterotic poles we see that the constraint on M_s^2 changes by a factor of two. Though there is also a lower bound for $SU(3)$ in the non-heterotic case, the bound is not interesting unless the Yang-Mills coupling is sufficiently weak. Otherwise, the constraint $g_s < 1$ justifying the perturbative analysis is much stronger anyway. Nonetheless, for sufficiently weak g_{YM}^2 (of order a tenth) we see that we get a two-sided bound on M_s . It is also amusing to note that in the case of non-heterotic poles, $SU(5)$ and $SO(10)$ gauge bosons are essentially pegged to have $g_{YM}^2 = 2g_s^2$.

5.3.2 Constraints in General Spacetime Dimensions

In general dimensions, the full constraints require D -dimensional spinning polynomials. The Yang-Mills amplitude in D spacetime dimensions is of the form

$$\mathcal{A} = \mathcal{F}^4 \mathcal{A}^{\text{scalar}} \quad (5.57)$$

where $\mathcal{A}^{\text{scalar}}$ has no dependence on polarization vectors and \mathcal{F}^4 is the famous polynomial permutation-invariant in field strengths which sits in front of the field theory Yang-Mills amplitude. In [30], positivity of \mathcal{F}^4 alone was remarkably shown to contain the critical dimension constraint $D \leq 10$ and information about the spectrum of 11-dimensional supergravity. This is the only condition for positive expandability of \mathcal{F}^4 . For our analysis, what is relevant is that so long as we satisfy the critical dimension constraint $D \leq 10$, positive expandability of $\mathcal{A}^{\text{scalar}}$ on the scalar D -dimensional Gegenbauer polynomials then implies positive expandability of the full amplitude's residues, as we merely have a product of two positively expandable functions, which in turn must be positively expandable. So for general dimensions equal to or below ten, we will consider the sufficient though not strictly necessary condition of positive expandability on the *scalar* D -dimensional Gegenbauer polynomials. This means that the true constraints could in principle be weaker, but merely expanding on the scalar polynomial basis already provides interesting constraints. So we have²

$$\mathbb{G}_{D,J}^{\{h_i\}}(\cos \theta) = G_J^{(D)}(\cos \theta) \quad (5.58)$$

It is worth commenting that in four spacetime dimensions, where we did the full spinning analysis, the strongest constraints for residues in the s -channel come from the process with $h_1 = h_2 = -1$ and $h_3 = h_4 = +1$ for which the spinning polynomials

²These would be denoted in the math literature as $C_J^{(\frac{D-3}{2})}(\cos \theta)$

reduce to the Legendre polynomials. This might hint at the constraints on $\mathcal{A}^{\text{scalar}}$ constituting the full set of constraints, but we will leave such analysis to future work.

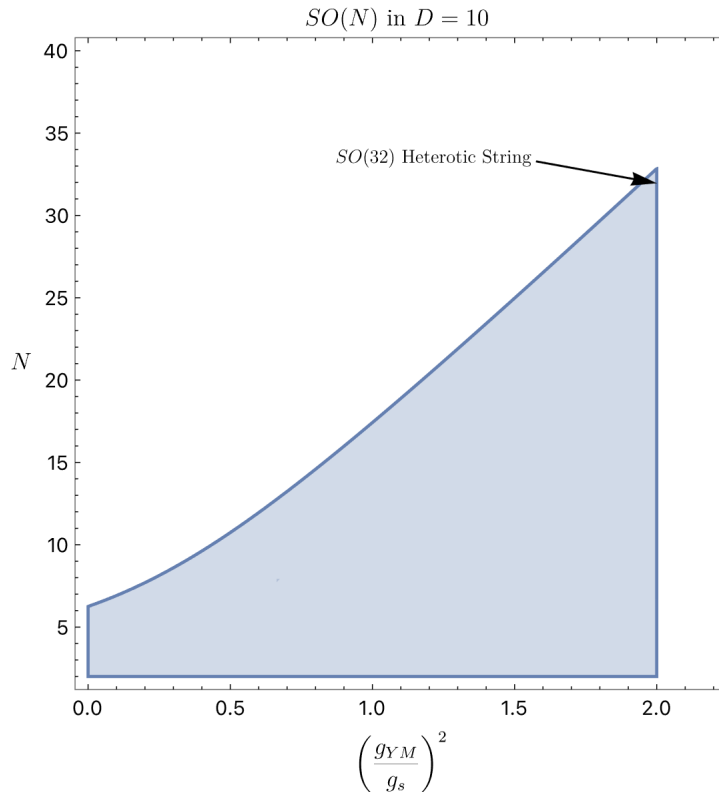


Figure 5.3: Imposing positive expandability of the heterotic form of (5.41) on scalar Gegenbauer's in $D = 10$, we find the allowed region shaded in blue. The value of coupling and N fixed in $SO(32)$ heterotic string theory is indicated by the arrow, $g_{YM}^2 = 2g_s^2$ and $N = 32$.

Heterotic $SO(N)$ in Ten Dimensions Given that our Ansatz is essentially the heterotic string amplitude (the gauge group and gauge-coupling are not fixed), it is natural to check the constraints of the previous section but in ten dimensions, which can be found in figure 5.3. Similar exchanged states produce the analogous curves in the $D = 10$ case, the only difference being that the strongest upper bounding curve occurs at mass level three rather than one. In particular, the maximum allowed value of the coupling ratio $\frac{g_{YM}^2}{g_s^2}$ is 2, enforced by the same exchanged state as in four dimensions, and this meets the upper bounding curve from mass-level three singlet

exchange at the corner of the allowed region with the maximum allowed

$$N \leq 32 + \frac{16}{19} \quad (5.59)$$

which of course means the only theory allowed by this corner of the allowed region is the $SO(32)$ heterotic string.

Dimension vs. Rank By varying the dimension, we can consider the generalization of (5.54) to general spacetime dimension D

$$\left\{ \begin{array}{l} g_{YM}^2 \leq \frac{16\pi G}{\alpha'} \\ 2 + g_{YM}^2 \frac{\alpha'}{8\pi G} C_A - \frac{D-2}{8(D-1)} d_A \geq 0 \quad M^2 = 1 \\ 2 + g_{YM}^2 \frac{\alpha'}{8\pi G} C_A - \frac{D-2}{6(D+1)} d_A \geq 0 \quad M^2 = 2 \\ \frac{528-6D(D-6)}{224+D(D-18)} + g_{YM}^2 \frac{\alpha'}{8\pi G} C_A - \frac{3(D-2)(D(D-14)+192)}{16(D+3)(D(D-18)+224)} d_A \geq 0 \quad M^2 = 3 \end{array} \right. \quad (5.60)$$

These are all constraints that are relevant across dimension $4 \leq D \leq 10$. Notice that again, in general dimensions we observe that the maximum allowed rank of the gauge group always occurs at $g_{YM}^2 = 2g_s^2$, the maximum allowed value of the coupling and also the value in the $SO(32)$ heterotic string. This bound comes from the same gauge group representations as in $D = 4$ and the dimension-dependent bounds always come from the spin-zero singlet exchange. We can then track how the maximum allowed rank varies with the spacetime dimension, which is presented in figure 5.4. We look at bounds between four and ten dimensions, over which three distinct smooth bounding curves comprise the full upper bounding curve. As stated

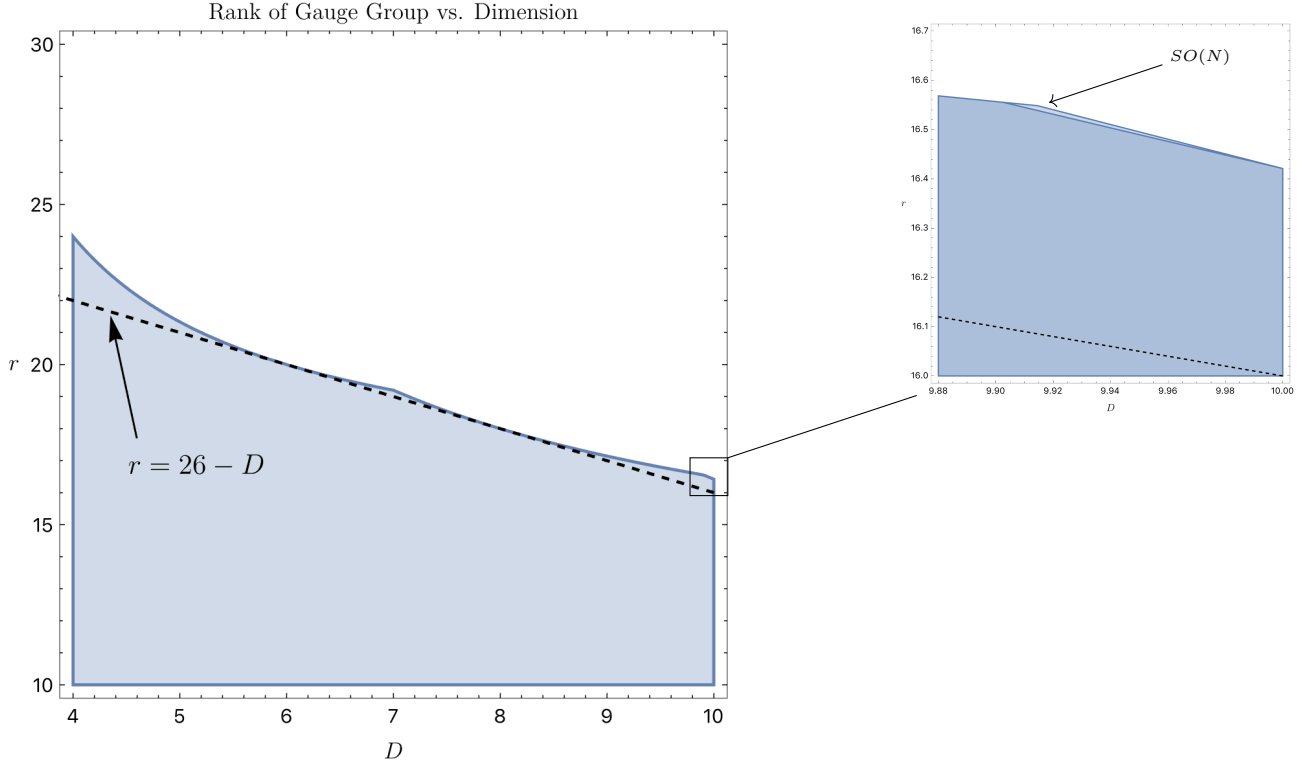


Figure 5.4: Imposing positive expandability of the heterotic form of (5.41) on scalar D -dimensional Gegenbauer's with $g_{YM}^2 = 2g_s^2$, we find the allowed region in rank r vs. dimension D shaded in blue. There are three-piecewise segments, the two leftmost are identical for $SO(N)$ and $SU(N)$. The zoomed portion shows that in the vicinity of $D = 10$, the third curve is not the same curve in r and D for $SO(N)$ and $SU(N)$ but agrees in $D = 10$, the constraints agreeing in all physical dimensions.

before, the bounding curve always corresponds to a spin-zero singlet of the gauge group, the only quantum number varying being the mass. For dimensions $4 \leq D \leq 7$ the constraint comes from the first mass level. In higher dimensions the strongest constraint comes from the second mass level until in the vicinity of $D = 10$, at which a yet stronger constraint comes in from the third mass level. This is seen in the zoomed portion of figure 5.4. In $D = 10$ this constraint from mass level three is necessary to disallow $N = 33$ in the case of $SO(N)$. The dashed line on figure 5.4 represents the swampland conjecture $r < 26 - D$, derived assuming BPS completeness with 16 supercharges [31]. We emphasize that the first two of these upper bounding curves is identical for $SO(N)$ and $SU(N)$, the agreement being not at all manifest even at

the level of the analytic expressions in r and D which are required to be positive; the expressions merely share a common factor enforcing positivity. For example, at the first mass level the expressions are

$$\begin{cases} (2+r) \left(2 - r \frac{D-2}{8(D-1)}\right) & SO(N) \\ (2r-1) \left(2 - r \frac{D-2}{8(D-1)}\right) & SU(N) \end{cases} \quad (5.61)$$

Even more dramatically, the zoomed portion reveals the constraint at the third mass level is not even the same curve for $SU(N)$ and $SO(N)$, but the two curves are such that they meet at $D = 10$ constraining the rank to be $r \leq 16 + \frac{8}{19}$, so the rank constraint in any integer dimension is identical. This non-trivial conspiracy of the Gegenbauer polynomials and recoupling coefficients suggests a universality to these constraints when centered on the correct data which could perhaps be exploited in greater generality.

Exceptional and Product Groups Though we do not carry out the analysis to verify the $g_{YM}^2 \leq \frac{16\pi G}{\alpha'}$ is the bound for other gauge groups, if we assume this bound, we can apply the remaining constraints and verify that all exceptional groups are allowed. It is worth also dwelling on what is non-trivial about the constraint of the form

$$1 + C_A - x d_A \geq 0 \quad (5.62)$$

this is the constraint on the gauge group when the purported coupling bound is saturated. What is crucial here is that the constant and the coefficient of C_A are the same. This is true for the expressions in (5.60) and only for the final constraint in $D = 10$. One way in which this constraint is non-trivial, which we have partially already appreciated, is that on all classical groups it factorizes as a function of rank.

We have already seen this for $SO(N)$ and for $SU(N)$, and finally including $Sp(2N)$ we find

$$\left\{ \begin{array}{ll} (2+r)(1-rx) & SO(N) \\ (2r-1)(1-rx) & SU(N) \\ (2r+1)(1-rx) & Sp(2N) \end{array} \right. \quad (5.63)$$

This also lets us readily analyze product groups. For simplicity we can consider $G \times G$ and we scatter adjoints of each simple factor of the gauge group. There are therefore three distinct $2 \rightarrow 2$ amplitudes which must be considered for positivity. If we call the pair of states in one factor G of the gauge group state A and a pair of gauge bosons from the other simple factor B then we need to consider the two-by-two matrix of amplitudes

$$S_{2 \times 2} = \begin{pmatrix} S^{A \rightarrow A} & S^{A \rightarrow B} \\ S^{B \rightarrow A} & S^{B \rightarrow B} \end{pmatrix} \quad (5.64)$$

and the unitarity condition is then

$$S_{2 \times 2}^\dagger S_{2 \times 2} \prec 1 \quad (5.65)$$

which then implies that the following matrix of partial wave coefficients is positive semi-definite

$$\begin{pmatrix} a_{J,R}^{A \rightarrow A} & a_{J,R}^{A \rightarrow B} \\ a_{J,R}^{B \rightarrow A} & a_{J,R}^{B \rightarrow B} \end{pmatrix} \quad (5.66)$$

the amplitudes for the diagonals are simply those considered above. The amplitudes for the off-diagonal term are even simpler, as only graviton exchange in a single

channel is present

$$\mathcal{A}^{A \rightarrow B} = 8\pi G \times \Gamma^{\text{str}} \frac{\delta_A^{ab} \delta_B^{cd}}{s(1 + \alpha' s)} \quad (5.67)$$

Therefore, the only exchanged representations R with non-trivial off-diagonal dependence are singlets. In particular, to consider general groups we can be motivated as in the previous section to assume that the condition $g_{YM}^2 \leq \frac{16\pi G}{\alpha'}$ holds in general. If we then consider the scalar exchanges at spin zero, we find the following matrix must be positive semi-definite

$$\begin{pmatrix} 1 + C_A - f(D)d_A & -f(D)d_A \\ -f(D)d_A & 1 + C_A - f(D)d_A \end{pmatrix} \quad (5.68)$$

where $f(D)$ is simply the dimension dependent coefficient from earlier

$$f(D) = \begin{cases} \frac{D-2}{16(D-1)} & M^2 = 1 \\ \frac{D-2}{12(D+1)} & M^2 = 2 \\ \frac{3(D-2)(D(D-14)+192)}{32(D+3)(D(D-18)+224)} & M^2 = 3 \end{cases} \quad (5.69)$$

imposing positivity of the eigenvalues of this matrix we find the condition

$$1 + C_A - 2f(D)d_A \geq 0 \quad (5.70)$$

From (5.63) we see that on the classical groups this simply cuts the rank constraint in half. In particular, in $D = 10$ where the bound on the rank had been $16 + \frac{8}{19}$, for the product it is $8 + \frac{4}{19}$ for each identical factor. As such, $SO(16) \times SO(16)$ is the maximal rank product of special orthogonal groups allowed by unitarity.

It is also amusing to apply the $D = 10$ form of (5.70) to the exceptional groups. In

this case, as the constraint is

$$1 + C_A - \frac{19d_A}{156} \geq 0 \quad (5.71)$$

which is violated by E_6 , E_7 , but indeed satisfied E_8 . The groups $G_2 \times G_2$ and $F_4 \times F_4$ are also allowed. We also note that when the groups are not taken to be identical it is readily verified that of all products made of E_6 , E_7 and E_8 only $E_8 \times E_8$ is consistent with unitarity.

Swampland Conjectures In the context of these Ansätze we are able to make contact with some swampland conjectures, in particular weak gravity and consequences of completeness, such as the conjecture about the maximum allowed rank of the gauge group. In the context of gravitational completions, we also see the mechanism for generating something like completeness. Placing the field theory amplitude in front of a common stringy form-factor means that the u and t channel projectors for the gauge-group must be re-expanded in terms of the s -channel projectors when we take a residue in s . This is done via the group theory crossing equation which is solved via 6- j symbols. In this way, we see that scattering some specified representations builds up the need for other representations in their tensor product. This kind of mechanism for completeness has everything to do with gravity, as one can note that with the non-gravitational NLSM completion (5.12) the poles merely produce the singlet and adjoint exchanges. Even for a stringy completion of the NLSM amplitude via Lovelace-Shapiro, we still only generate either anti-symmetric adjoint exchange or symmetric adjoint and singlet exchange. But in the context of our gravitational amplitudes, we find every possible representation that can be exchanged between the external states is indeed exchanged. This provides a mechanism to bootstrap the completeness hypothesis in the context of these amplitudes, one which crucially relies on the presence of gravity.

5.4 Standard Model Electroweak Sector

Now we direct our attention to the electroweak sector. We already studied Ansätze for the scattering of $SU(N)$ gauge bosons, and found a constraint on the relation between the gauge-coupling, the UV completion scale, and the Planck scale. The only constraint coming from scattering $SU(N)$ gauge bosons in the heterotic case with $N \leq 5$ was

$$g_{YM}^2 M_P^2 \leq 2M_s^2 \quad (5.72)$$

Then the lowest string scale we can obtain is

$$M_s^2 \sim 10^{18} \text{ GeV}^2 \quad (5.73)$$

and the bound pushes M_s^2 up by a factor of two in the non-heterotic case. We can further probe the electroweak sector by studying the scattering of four Higgs's. The amplitude in this case is

$$\begin{aligned} \mathcal{A}(1, \bar{2}, 3, \bar{4}) = \Gamma^{\text{str}} \Bigg(& -\frac{1}{M_P^2} \left(\frac{tu}{s} \mathbb{P}_{1,1}^s + \frac{su}{t} \mathbb{P}_{1,1}^t \right) + \frac{t-u}{2s} \left(\frac{g_1^2}{4} \mathbb{P}_{1,1}^s + g_2^2 \mathbb{P}_{\text{Adj},1}^s \right) \\ & + \frac{s-u}{2t} \left(\frac{g_1^2}{4} \mathbb{P}_{1,1}^t + g_2^2 \mathbb{P}_{\text{Adj},1}^t \right) + 2\lambda(\mathbb{P}_{1,1}^s + \mathbb{P}_{1,1}^t) \Bigg) \quad (5.74) \end{aligned}$$

and similarly for the configuration with massless t and u channel poles only. The subscript labels on the projector denote the exchanged representation of the corresponding factor of $SU(2) \times U(1)_Y$ with 1 denoting singlet exchange (with zero charge in the $U(1)_Y$ case). The projectors are normalized such that these are the conventional normalizations for the gauge-couplings and Higgs quartic coupling in the standard model. If we fix the gauge-couplings such that $\alpha_S^{-1} = \alpha_W^{-1} = 25$ we can produce a plot relating the Higgs quartic coupling and the putative string scale in Planck units. The kink in the allowed region minimizing M_s is just outside of the bounds of the running

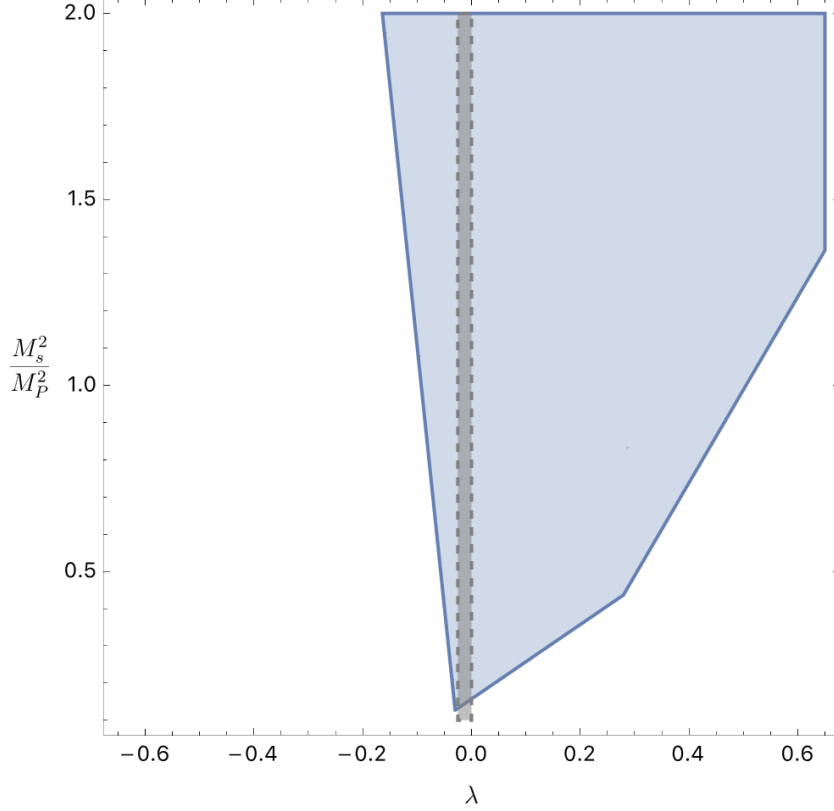


Figure 5.5: Plot of allowed region (blue) for Higgs quartic coupling λ versus the string scale squared M_s^2 in reduced Planck units. The dashed lines bounding the shaded green band are roughly 3σ bounds for the SM running of λ with uncertainty coming from the uncertainty in the mass of the top quark (taken from [32]). The bounds come from imposing unitarity of (5.74). With heterotic gravitational denominators, λ is strictly positive.

of λ predicted by the standard model [32]. The running predicted has λ become negative at $\sim 10^{10}$ GeV and asymptote to within the gray shaded region of figure 5.5 below the Planck scale. The lowest scale allowed by this bound is slightly below that allowed by the four gauge boson scattering, constituting a weaker constraint.

It is crucial to note though, that we find this bound on the Higgs quartic coupling in the absence of heterotic gravitational denominators. If instead we have the heterotic denominators, then we find that λ is strictly positive, with the explicit lower bound varying with the gauge couplings. It is also worth noting that in analogy with these poles being associated with gauge bosons' and graviton's non-minimal couplings

to the dilaton via $F^2\phi$ and $R^2\phi$ in the heterotic string, we would expect a coupling $(H^\dagger H)\phi$ for these amplitudes. In the context of many BSM models, the Higgs quartic running is modified by new states below $\sim 10^{10}\text{GeV}$, and could be consistent with the small positive λ 's allowed by this heterotic analogue of (5.74). But this is at least naively in tension with the only new states coming in at this putative string scale, a scale which gauge-boson scattering required be only slightly below the Planck scale for α 's $\sim \frac{1}{25}$. This basic tension in the $2 \rightarrow 2$ Higgs amplitude already poses an interesting puzzle in this particular bottom up approach to building stringy completions of the standard model.

5.5 Conclusions Outlook

In this work we considered a class of stringy Ansätze for $2 \rightarrow 2$ amplitudes and constrained them via unitarity. The stringy nature of these amplitudes flipped gravity's role from an obstruction to UV softening to a necessary condition for UV softening consistent with positive expandability of the imaginary part, with constraints bounding the strength of other interactions relative to gravity from above. We found that gauge boson scattering in particular furnished stringent constraints, especially in higher dimensions, where we found that the $SO(32)$ heterotic string is barely consistent with perturbative unitarity. Perhaps most compelling is the agreement between the suggested constraints on rank versus spacetime dimension for both $SO(N)$ and $SU(N)$ gauge boson scattering not only with each other but with the swampland conjecture $r < 26 - D$ in greater than four dimensions. These results indicate that perturbative unitarity constraints can become particularly potent in the context of gravity. There are a variety of further questions to pursue.

These amplitudes morally resolved the particles into closed strings. It would be interesting to pursue the open string analogue of this analysis which would correspond to

braneworld scenarios. In this case the leading unitarizing interactions are the gauge-boson exchanges, not gravity, and a genus 0 analysis will not furnish constraints on gauge or global symmetry groups, though one can still expect constraints on couplings. In order to get gravity in the low-energy limit, which makes symmetry constraints more likely, we will need control over Ansätze at genus 1. It is worth reviewing here why we do not obtain interesting constraints at genus 0, in the absence of gravity. We could consider completions of the form

$$\mathcal{A} = \langle 12 \rangle^2 [34]^2 \frac{\Gamma(-\alpha' s) \Gamma(-\alpha' t)}{\Gamma(1 - \alpha' s - \alpha' t)} \text{Tr}(T^a T^b T^c T^d) + \text{other orderings} \quad (5.75)$$

this could even be done to produce a different completion of the NLSM a la Lovelace-Shapiro. The reason we do not obtain interesting constraints for this class of Ansätze is simple. The residue at mass level n is

$$\text{Res}_{s \rightarrow n/\alpha'} \mathcal{A} = \frac{\langle 12 \rangle^2 [34]^2}{n!} \left(\prod_{i=1}^{n-1} (i + t) \right) \quad (5.76)$$

$$\times [\text{Tr}(T^a T^b T^c T^d) + (-1)^{n-1} \text{Tr}(T^a T^b T^d T^c)] \quad (5.77)$$

And so the color structures, appearing in the factor in brackets, are simply those appearing at tree-level, proportional to the s -channel projectors and trivially positive.

Another interesting extension of this work would be considering deformations like the Coon amplitude or the more general deformations considered in [33]. Again here, deformations of closed string amplitudes will provide more readily amenable to this analysis as they will provide generalizations of the above constraints already at genus 0.

Finally, in order to continue to probe the consistency of the amplitudes considered herein, two obvious avenues are to develop the analysis both for higher points and for

massive external legs. The analysis above not only obtains bounds from requiring the exchange of positive norm states, but tells you the quantum numbers of such states. One can iteratively build up amplitudes with these massive states as external legs and impose further consistency.

Acknowledgments We would like thank Wayne Zhao for initial collaboration. We would also like to thank Sebastian Mizera, Lorenz Eberhardt, and Yu-tin Huang for useful discussions. And we would like to especially thank Nima Arkani-Hamed for many valuable discussions and comments on the draft.

Bibliography

- [1] A. Hillman, “Symbol Recursion for the dS Wave Function,” [arXiv:1912.09450 \[hep-th\]](#).
- [2] N. Arkani-Hamed, A. Hillman, and S. Mizera, “Feynman polytopes and the tropical geometry of UV and IR divergences,” *Phys. Rev. D* **105** no. 12, (2022) 125013, [arXiv:2202.12296 \[hep-th\]](#).
- [3] B. Bachu and A. Hillman, “Stringy Completions of the Standard Model from the Bottom Up,” [arXiv:2212.03871 \[hep-th\]](#).
- [4] R. Britto, F. Cachazo, and B. Feng, “New recursion relations for tree amplitudes of gluons,” *Nucl. Phys. B* **715** (2005) 499–522, [arXiv:hep-th/0412308](#).
- [5] R. Britto, F. Cachazo, B. Feng, and E. Witten, “Direct proof of tree-level recursion relation in Yang-Mills theory,” *Phys. Rev. Lett.* **94** (2005) 181602, [arXiv:hep-th/0501052](#).
- [6] J. Polchinski, “Renormalization and Effective Lagrangians,” *Nucl. Phys. B* **231** (1984) 269–295.
- [7] C. Cheung and G. N. Remmen, “Stringy Dynamics from an Amplitudes Bootstrap,” [arXiv:2302.12263 \[hep-th\]](#).
- [8] N. Arkani-Hamed, P. Benincasa, and A. Postnikov, “Cosmological Polytopes and the Wavefunction of the Universe,” [arXiv:1709.02813 \[hep-th\]](#).
- [9] N. Arkani-Hamed and P. Benincasa, “On the Emergence of Lorentz Invariance and Unitarity from the Scattering Facet of Cosmological Polytopes,” [arXiv:1811.01125 \[hep-th\]](#).
- [10] N. Arkani-Hamed, D. Baumann, H. Lee, and G. L. Pimentel, “The Cosmological Bootstrap: Inflationary Correlators from Symmetries and Singularities,” *JHEP* **04** (2020) 105, [arXiv:1811.00024 \[hep-th\]](#).
- [11] N. Arkani-Hamed and J. Maldacena, “Cosmological Collider Physics,” [arXiv:1503.08043 \[hep-th\]](#).

- [12] P. Benincasa, “From the flat-space S-matrix to the Wavefunction of the Universe,” [arXiv:1811.02515 \[hep-th\]](#).
- [13] J. M. Maldacena, “Non-Gaussian features of primordial fluctuations in single field inflationary models,” *JHEP* **05** (2003) 013, [arXiv:astro-ph/0210603](#).
- [14] C. Duhr, H. Gangl, and J. R. Rhodes, “From polygons and symbols to polylogarithmic functions,” *JHEP* **10** (2012) 075, [arXiv:1110.0458 \[math-ph\]](#).
- [15] 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 \[hep-th\]](#).
- [16] N. Arkani-Hamed, S. He, and T. Lam, “Stringy canonical forms,” *JHEP* **02** (2021) 069, [arXiv:1912.08707 \[hep-th\]](#).
- [17] S. He, Z. Li, P. Raman, and C. Zhang, “Stringy canonical forms and binary geometries from associahedra, cyclohedra and generalized permutohedra,” *JHEP* **10** (2020) 054, [arXiv:2005.07395 \[hep-th\]](#).
- [18] F. Brown, “Feynman amplitudes, coaction principle, and cosmic Galois group,” *Commun. Num. Theor. Phys.* **11** (2017) 453–556, [arXiv:1512.06409 \[math-ph\]](#).
- [19] K. Schultka, *Microlocal analyticity of Feynman integrals*. PhD thesis, Humboldt U., Berlin, 2019.
- [20] OEIS Foundation Inc., “Multidimensional Catalan numbers, Entry A060854 in The On-Line Encyclopedia of Integer Sequences,” [\[https://oeis.org/A060854\]](https://oeis.org/A060854) (2022) .
- [21] 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. D* **72** (2005) 085001, [arXiv:hep-th/0505205](#).
- [22] Z. Bern, M. Czakon, L. J. Dixon, D. A. Kosower, and V. A. Smirnov, “The Four-Loop Planar Amplitude and Cusp Anomalous Dimension in Maximally Supersymmetric Yang-Mills Theory,” *Phys. Rev. D* **75** (2007) 085010, [arXiv:hep-th/0610248](#).
- [23] M. Kompaniets and E. Panzer, “Renormalization group functions of ϕ^4 theory in the MS-scheme to six loops,” *PoS LL2016* (2016) 038, [arXiv:1606.09210 \[hep-th\]](#).
- [24] F. Brown and C. Dupont, “Single-valued integration and superstring amplitudes in genus zero,” *Commun. Math. Phys.* **382** no. 2, (2021) 815–874, [arXiv:1910.01107 \[math.NT\]](#).

- [25] F. Ardila, C. Benedetti, and J. Doker, “Matroid polytopes and their volumes,” 2011.
- [26] K. G. Chetyrkin, A. L. Kataev, and F. V. Tkachov, “New Approach to Evaluation of Multiloop Feynman Integrals: The Gegenbauer Polynomial x Space Technique,” *Nucl. Phys. B* **174** (1980) 345–377.
- [27] M. A. Virasoro, “Alternative constructions of crossing-symmetric amplitudes with regge behavior,” *Phys. Rev.* **177** (1969) 2309–2311.
- [28] P. Cvitanovic, *Group theory: Birdtracks, Lie’s and exceptional groups*. 2008.
- [29] D. J. Gross, J. A. Harvey, E. J. Martinec, and R. Rohm, “Heterotic String Theory. 2. The Interacting Heterotic String,” *Nucl. Phys. B* **267** (1986) 75–124.
- [30] N. Arkani-Hamed, L. Eberhardt, Y.-t. Huang, and S. Mizera, “On unitarity of tree-level string amplitudes,” *JHEP* **02** (2022) 197, [arXiv:2201.11575 \[hep-th\]](#).
- [31] H.-C. Kim, H.-C. Tarazi, and C. Vafa, “Four-dimensional $\mathcal{N} = 4$ SYM theory and the swampland,” *Phys. Rev. D* **102** no. 2, (2020) 026003, [arXiv:1912.06144 \[hep-th\]](#).
- [32] G. Degrandi, S. Di Vita, J. Elias-Miro, J. R. Espinosa, G. F. Giudice, G. Isidori, and A. Strumia, “Higgs mass and vacuum stability in the Standard Model at NNLO,” *JHEP* **08** (2012) 098, [arXiv:1205.6497 \[hep-ph\]](#).
- [33] C. Cheung and G. N. Remmen, “Veneziano Variations: How Unique are String Amplitudes?,” [arXiv:2210.12163 \[hep-th\]](#).
- [34] L. Adams, C. Bogner, and S. Weinzierl, “The two-loop sunrise integral around four space-time dimensions and generalisations of the Clausen and Glaisher functions towards the elliptic case,” *J. Math. Phys.* **56** no. 7, (2015) 072303, [arXiv:1504.03255 \[hep-ph\]](#).
- [35] Z. Bern, L. J. Dixon, and D. A. Kosower, “Dimensionally regulated pentagon integrals,” *Nucl. Phys. B* **412** (1994) 751–816, [arXiv:hep-ph/9306240](#).
- [36] N. Arkani-Hamed, T.-C. Huang, and Y.-T. Huang, “The EFT-Hedron,” *JHEP* **05** (2021) 259, [arXiv:2012.15849 \[hep-th\]](#).
- [37] S. Caron-Huot, Y.-Z. Li, J. Parra-Martinez, and D. Simmons-Duffin, “Graviton partial waves and causality in higher dimensions,” [arXiv:2205.01495 \[hep-th\]](#).
- [38] S. Caron-Huot, Y.-Z. Li, J. Parra-Martinez, and D. Simmons-Duffin, “Causality constraints on corrections to Einstein gravity,” [arXiv:2201.06602 \[hep-th\]](#).

- [39] Z. Bern, D. Kosmopoulos, and A. Zhiboedov, “Gravitational effective field theory islands, low-spin dominance, and the four-graviton amplitude,” *J. Phys. A* **54** no. 34, (2021) 344002, [arXiv:2103.12728 \[hep-th\]](#).
- [40] A. L. Guerrieri, J. Penedones, and P. Vieira, “S-matrix bootstrap for effective field theories: massless pions,” *JHEP* **06** (2021) 088, [arXiv:2011.02802 \[hep-th\]](#).
- [41] A. Guerrieri, J. Penedones, and P. Vieira, “Where Is String Theory in the Space of Scattering Amplitudes?,” *Phys. Rev. Lett.* **127** no. 8, (2021) 081601, [arXiv:2102.02847 \[hep-th\]](#).
- [42] J. Albert and L. Rastelli, “Bootstrapping pions at large N,” *JHEP* **08** (2022) 151, [arXiv:2203.11950 \[hep-th\]](#).
- [43] M. Soldate, “Partial Wave Unitarity and Closed String Amplitudes,” *Phys. Lett. B* **186** (1987) 321–327.
- [44] N. Arkani-Hamed, L. Motl, A. Nicolis, and C. Vafa, “The String landscape, black holes and gravity as the weakest force,” *JHEP* **06** (2007) 060, [arXiv:hep-th/0601001](#).
- [45] A. Adams, N. Arkani-Hamed, S. Dubovsky, A. Nicolis, and R. Rattazzi, “Causality, analyticity and an IR obstruction to UV completion,” *JHEP* **10** (2006) 014, [arXiv:hep-th/0602178](#).
- [46] M. van Beest, J. Calderón-Infante, D. Mirfendereski, and I. Valenzuela, “Lectures on the Swampland Program in String Compactifications,” *Phys. Rept.* **989** (2022) 1–50, [arXiv:2102.01111 \[hep-th\]](#).
- [47] B. W. Lee, C. Quigg, and H. B. Thacker, “The Strength of Weak Interactions at Very High-Energies and the Higgs Boson Mass,” *Phys. Rev. Lett.* **38** (1977) 883–885.
- [48] B. W. Lee, C. Quigg, and H. B. Thacker, “Weak Interactions at Very High-Energies: The Role of the Higgs Boson Mass,” *Phys. Rev. D* **16** (1977) 1519.
- [49] S. Caron-Huot, D. Mazac, L. Rastelli, and D. Simmons-Duffin, “Sharp boundaries for the swampland,” *JHEP* **07** (2021) 110, [arXiv:2102.08951 \[hep-th\]](#).
- [50] S. Caron-Huot and V. Van Duong, “Extremal Effective Field Theories,” *JHEP* **05** (2021) 280, [arXiv:2011.02957 \[hep-th\]](#).
- [51] S. Caron-Huot, D. Mazac, L. Rastelli, and D. Simmons-Duffin, “AdS bulk locality from sharp CFT bounds,” *JHEP* **11** (2021) 164, [arXiv:2106.10274 \[hep-th\]](#).

- [52] N. Arkani-Hamed, Y.-t. Huang, J.-Y. Liu, and G. N. Remmen, “Causality, unitarity, and the weak gravity conjecture,” *JHEP* **03** (2022) 083, [arXiv:2109.13937 \[hep-th\]](#).
- [53] C. Cheung and G. N. Remmen, “Infrared Consistency and the Weak Gravity Conjecture,” *JHEP* **12** (2014) 087, [arXiv:1407.7865 \[hep-th\]](#).
- [54] N. Arkani-Hamed, T.-C. Huang, and Y.-t. Huang, “Scattering amplitudes for all masses and spins,” *JHEP* **11** (2021) 070, [arXiv:1709.04891 \[hep-th\]](#).
- [55] M. C. Bergère and J. B. Zuber, “Renormalization of Feynman amplitudes and parametric integral representation,” *Communications in Mathematical Physics* **35** no. 2, (Jun, 1974) 113–140. <https://doi.org/10.1007/BF01646611>.
- [56] K. Schultka, “Toric geometry and regularization of Feynman integrals,” [arXiv:1806.01086 \[math-ph\]](#).
- [57] M. Borinsky, “Tropical Monte Carlo quadrature for Feynman integrals,” [arXiv:2008.12310 \[math-ph\]](#).
- [58] E. Panzer, “Hepp’s bound for Feynman graphs and matroids,” [arXiv:1908.09820 \[math-ph\]](#).
- [59] A. Pak and A. Smirnov, “Geometric approach to asymptotic expansion of Feynman integrals,” *Eur. Phys. J. C* **71** (2011) 1626, [arXiv:1011.4863 \[hep-ph\]](#).
- [60] T. Y. Semenova, A. V. Smirnov, and V. A. Smirnov, “On the status of expansion by regions,” *Eur. Phys. J. C* **79** no. 2, (2019) 136, [arXiv:1809.04325 \[hep-th\]](#).
- [61] A. V. Smirnov, “FIESTA4: Optimized Feynman integral calculations with GPU support,” *Comput. Phys. Commun.* **204** (2016) 189–199, [arXiv:1511.03614 \[hep-ph\]](#).
- [62] B. Jantzen, A. V. Smirnov, and V. A. Smirnov, “Expansion by regions: revealing potential and Glauber regions automatically,” *Eur. Phys. J. C* **72** (2012) 2139, [arXiv:1206.0546 \[hep-ph\]](#).
- [63] B. Ananthanarayan, A. Pal, S. Ramanan, and R. Sarkar, “Unveiling Regions in multi-scale Feynman Integrals using Singularities and Power Geometry,” *Eur. Phys. J. C* **79** no. 1, (2019) 57, [arXiv:1810.06270 \[hep-ph\]](#).
- [64] T. Kaneko and T. Ueda, “A Geometric method of sector decomposition,” *Comput. Phys. Commun.* **181** (2010) 1352–1361, [arXiv:0908.2897 \[hep-ph\]](#).
- [65] T. Kaneko and T. Ueda, “Sector Decomposition Via Computational Geometry,” *PoS ACAT2010* (2010) 082, [arXiv:1004.5490 \[hep-ph\]](#).

- [66] F. Brown and D. Kreimer, “Angles, Scales and Parametric Renormalization,” *Lett. Math. Phys.* **103** (2013) 933–1007, [arXiv:1112.1180 \[hep-th\]](#).
- [67] O. I. Zavialov, *Renormalized Quantum Field Theory*, vol. 21. Springer Science & Business Media, 2012.
- [68] V. A. Smirnov, *Renormalization and Asymptotic Expansions*. Birkhäuser, 1991.
- [69] O. Amini, S. J. Bloch, J. I. Burgos Gil, and J. Fresán, “Feynman amplitudes and limits of heights,” *Izvestiya: Mathematics* **80** no. 5, (Oct, 2016) 813–848. <http://dx.doi.org/10.1070/IM8492>.
- [70] P. Tourkine, “Tropical Amplitudes,” *Annales Henri Poincaré* **18** no. 6, (2017) 2199–2249, [arXiv:1309.3551 \[hep-th\]](#).
- [71] P. Dai and W. Siegel, “Worldline Green Functions for Arbitrary Feynman Diagrams,” *Nucl. Phys. B* **770** (2007) 107–122, [arXiv:hep-th/0608062](#).
- [72] K. Roland and H.-T. Sato, “Multiloop worldline Green functions from string theory,” *Nucl. Phys. B* **480** (1996) 99–124, [arXiv:hep-th/9604152](#).
- [73] M. G. Schmidt and C. Schubert, “Worldline Green functions for multiloop diagrams,” *Phys. Lett. B* **331** (1994) 69–76, [arXiv:hep-th/9403158](#).
- [74] B. Bollobás, *Modern graph theory*, vol. 184. Springer Science & Business Media, 2013.
- [75] S. Bloch, H. Esnault, and D. Kreimer, “On Motives Associated to Graph Polynomials,” *Commun. Math. Phys.* **267** no. 1, (Oct, 2006) 181–225. <https://doi.org/10.1007/s00220-006-0040-2>.
- [76] F. Cachazo, N. Early, A. Guevara, and S. Mizera, “Scattering Equations: From Projective Spaces to Tropical Grassmannians,” *JHEP* **06** (2019) 039, [arXiv:1903.08904 \[hep-th\]](#).
- [77] N. Arkani-Hamed, S. He, T. Lam, and H. Thomas, “Binary Geometries, Generalized Particles and Strings, and Cluster Algebras,” [arXiv:1912.11764 \[hep-th\]](#).
- [78] N. Arkani-Hamed, T. Lam, and M. Spradlin, “Positive configuration space,” *Commun. Math. Phys.* **384** no. 2, (2021) 909–954, [arXiv:2003.03904 \[math.CO\]](#).
- [79] N. Arkani-Hamed, S. He, and T. Lam, “Cluster configuration spaces of finite type,” [arXiv:2005.11419 \[math.AG\]](#).
- [80] J. Drummond, J. Foster, O. Gürdoğan, and C. Kalousios, “Tropical fans, scattering equations and amplitudes,” [arXiv:2002.04624 \[hep-th\]](#).

- [81] J. Drummond, J. Foster, O. Gürdogan, and C. Kalousios, “Algebraic singularities of scattering amplitudes from tropical geometry,” *JHEP* **04** (2021) 002, [arXiv:1912.08217 \[hep-th\]](#).
- [82] D. Chicherin, J. M. Henn, and G. Papathanasiou, “Cluster algebras for Feynman integrals,” *Phys. Rev. Lett.* **126** no. 9, (2021) 091603, [arXiv:2012.12285 \[hep-th\]](#).
- [83] T. Binoth and G. Heinrich, “An automatized algorithm to compute infrared divergent multiloop integrals,” *Nucl. Phys. B* **585** (2000) 741–759, [arXiv:hep-ph/0004013](#).
- [84] T. Binoth and G. Heinrich, “Numerical evaluation of multiloop integrals by sector decomposition,” *Nucl. Phys. B* **680** (2004) 375–388, [arXiv:hep-ph/0305234](#).
- [85] C. Bogner and S. Weinzierl, “Resolution of singularities for multi-loop integrals,” *Comput. Phys. Commun.* **178** (2008) 596–610, [arXiv:0709.4092 \[hep-ph\]](#).
- [86] S. Borowka, G. Heinrich, S. P. Jones, M. Kerner, J. Schlenk, and T. Zirke, “SecDec-3.0: numerical evaluation of multi-scale integrals beyond one loop,” *Comput. Phys. Commun.* **196** (2015) 470–491, [arXiv:1502.06595 \[hep-ph\]](#).
- [87] S. Mizera and S. Telen, “Landau Discriminants,” [arXiv:2109.08036 \[math-ph\]](#).
- [88] L. Nilsson and M. Passare, “Mellin transforms of multivariate rational functions,” *Journal of Geometric Analysis* **23** no. 1, (Jan, 2013) 24–46, [arXiv:1010.5060 \[math.CV\]](#), <https://doi.org/10.1007/s12220-011-9235-7>.
- [89] C. Berkesch, J. Forsgård, and M. Passare, “Euler–Mellin integrals and A-hypergeometric functions,” [arXiv:1103.6273 \[math.CV\]](#).
- [90] S. Mizera, “Crossing symmetry in the planar limit,” *Phys. Rev. D* **104** no. 4, (2021) 045003, [arXiv:2104.12776 \[hep-th\]](#).
- [91] I. Gelfand, M. Kapranov, and A. Zelevinsky, *Discriminants, Resultants, and Multidimensional Determinants*. Modern Birkhäuser Classics. Birkhäuser Boston, 2009.
- [92] S. He, Z. Li, and Q. Yang, “Truncated cluster algebras and Feynman integrals with algebraic letters,” [arXiv:2106.09314 \[hep-th\]](#).
- [93] F. A. Berends, A. I. Davydychev, and N. I. Ussyukina, “Threshold and pseudothreshold values of the sunset diagram,” *Phys. Lett. B* **426** (1998) 95–104, [arXiv:hep-ph/9712209](#).

- [94] A. Yellespur Srikant, “Spherical Contours, IR Divergences and the geometry of Feynman parameter integrands at one loop,” *JHEP* **07** (2020) 236, [arXiv:1907.05429 \[hep-th\]](#).
- [95] J. M. Drummond, G. P. Korchemsky, and E. Sokatchev, “Conformal properties of four-gluon planar amplitudes and Wilson loops,” *Nucl. Phys. B* **795** (2008) 385–408, [arXiv:0707.0243 \[hep-th\]](#).
- [96] J. M. Henn, T. Peraro, M. Stahlhofen, and P. Wasser, “Matter dependence of the four-loop cusp anomalous dimension,” *Phys. Rev. Lett.* **122** no. 20, (2019) 201602, [arXiv:1901.03693 \[hep-ph\]](#).
- [97] S. Brannetti, M. Melo, and F. Viviani, “On the tropical Torelli map,” *Advances in Mathematics* **226** no. 3, (Feb, 2011) 2546–2586. <http://dx.doi.org/10.1016/j.aim.2010.09.011>.
- [98] L. Caporaso, “Geometry of tropical moduli spaces and linkage of graphs,” *Journal of Combinatorial Theory, Series A* **119** no. 3, (Apr, 2012) 579–598. <http://dx.doi.org/10.1016/j.jcta.2011.11.011>.
- [99] L. Landau, “On analytic properties of vertex parts in quantum field theory,” *Nucl. Phys.* **13** no. 1, (1960) 181–192.
- [100] J. Tevelev, “Compactifications of subvarieties of tori,” *American Journal of Mathematics* **129** no. 4, (2007) 1087–1104, [arXiv:math/0412329 \[math.AG\]](#). <https://people.math.umass.edu/~tevelev/trop60.pdf>.
- [101] P. Cvitanović and T. Kinoshita, “New approach to the separation of ultraviolet and infrared divergences of Feynman-parametric integrals,” *Phys. Rev. D* **10** (Dec, 1974) 3991–4006. <https://link.aps.org/doi/10.1103/PhysRevD.10.3991>.
- [102] T. Appelquist, “Parametric integral representations of renormalized Feynman amplitudes,” *Annals of Physics* **54** no. 1, (1969) 27–61. <https://www.sciencedirect.com/science/article/pii/0003491669903339>.
- [103] N. Nakanishi, *Graph Theory and Feynman Integrals*. Mathematics and its applications: a series of monographs and texts. Gordon and Breach, 1971.
- [104] C. Chandler, “Some physical region mass shell properties of renormalized Feynman integrals,” *Commun. Math. Phys.* **19** (1970) 169–188.
- [105] D. Maclagan and B. Sturmfels, *Introduction to Tropical Geometry*. Graduate Studies in Mathematics. American Mathematical Society, 2015. <https://books.google.com/books?id=zFsoCAAQBAJ>.
- [106] N. Arkani-Hamed, J. L. Bourjaily, F. Cachazo, A. B. Goncharov, A. Postnikov, and J. Trnka, *Grassmannian Geometry of Scattering Amplitudes*. Cambridge University Press, 4, 2016. [arXiv:1212.5605 \[hep-th\]](#).

- [107] H. Elvang and Y.-t. Huang, *Scattering Amplitudes in Gauge Theory and Gravity*. Cambridge University Press, 4, 2015.
- [108] E. Speer, *Generalized Feynman Amplitudes*. Annals of Mathematics Studies. Princeton University Press, 1969.
- [109] N. Arkani-Hamed, J. L. Bourjaily, F. Cachazo, and J. Trnka, “Local Integrals for Planar Scattering Amplitudes,” *JHEP* **06** (2012) 125, [arXiv:1012.6032 \[hep-th\]](#).
- [110] N. Arkani-Hamed, A. Hodges, and J. Trnka, “Positive Amplitudes In The Amplituhedron,” *JHEP* **08** (2015) 030, [arXiv:1412.8478 \[hep-th\]](#).
- [111] C. Itzykson and J. Zuber, *Quantum Field Theory*. International Series In Pure and Applied Physics. McGraw-Hill, New York, 1980. Sec. 8.2.
- [112] N. Arkani-Hamed, Y. Bai, and T. Lam, “Positive Geometries and Canonical Forms,” *JHEP* **11** (2017) 039, [arXiv:1703.04541 \[hep-th\]](#).
- [113] F. Tellander and M. Helmer, “Cohen-Macaulay Property of Feynman Integrals,” [arXiv:2108.01410 \[hep-th\]](#).
- [114] N. Arkani-Hamed, A. Hillman, and S. Mizera, “in preparation,”.
- [115] N. N. Bogoliubow and O. S. Parasiuk, “Über die Multiplikation der Kausalfunktionen in der Quantentheorie der Felder,” *Acta Mathematica* **97** no. none, (1957) 227 – 266. <https://doi.org/10.1007/BF02392399>.
- [116] K. Hepp, “Proof of the Bogoliubov-Parasiuk theorem on renormalization,” *Communications in Mathematical Physics* **2** no. 1, (Dec, 1966) 301–326. <https://doi.org/10.1007/BF01773358>.
- [117] W. Zimmermann, “Convergence of Bogoliubov’s method of renormalization in momentum space,” *Communications in Mathematical Physics* **15** no. 3, (Sep, 1969) 208–234. <https://doi.org/10.1007/BF01645676>.
- [118] V. A. Smirnov, *Analytic tools for Feynman integrals*, vol. 250. 2012.
- [119] 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)].
- [120] J. C. Collins, D. E. Soper, and G. F. Sterman, “Factorization of Hard Processes in QCD,” *Adv. Ser. Direct. High Energy Phys.* **5** (1989) 1–91, [arXiv:hep-ph/0409313](#).
- [121] T. Becher, A. Broggio, and A. Ferroglia, *Introduction to Soft-Collinear Effective Theory*, vol. 896. Springer, 2015. [arXiv:1410.1892 \[hep-ph\]](#).

- [122] E. Olivucci and P. Vieira, “Stampedes I: Fishnet OPE and Octagon Bootstrap with Nonzero Bridges,” [arXiv:2111.12131 \[hep-th\]](#).
- [123] B. Basso, L. J. Dixon, D. A. Kosower, A. Krajenbrink, and D.-l. Zhong, “Fishnet four-point integrals: integrable representations and thermodynamic limits,” *JHEP* **07** (2021) 168, [arXiv:2105.10514 \[hep-th\]](#).
- [124] F. Brown, “Multiple zeta values and periods: from moduli spaces to Feynman integrals,” *Contemp. Math* (2011) 27–52.
- [125] F. Brown, “Motivic periods and the cosmic Galois group (IHES, May 2015).”.
- [126] S. Weinberg, “High-Energy Behavior in Quantum Field Theory,” *Phys. Rev.* **118** (May, 1960) 838–849.
<https://link.aps.org/doi/10.1103/PhysRev.118.838>.
- [127] A. Connes and D. Kreimer, “Hopf algebras, renormalization and noncommutative geometry,” *Commun. Math. Phys.* **199** (1998) 203–242, [arXiv:hep-th/9808042](#).
- [128] A. Connes and D. Kreimer, “Renormalization in quantum field theory and the Riemann-Hilbert problem. 1. The Hopf algebra structure of graphs and the main theorem,” *Commun. Math. Phys.* **210** (2000) 249–273, [arXiv:hep-th/9912092](#).
- [129] A. Connes and D. Kreimer, “Renormalization in quantum field theory and the Riemann-Hilbert problem. 2. The beta function, diffeomorphisms and the renormalization group,” *Commun. Math. Phys.* **216** (2001) 215–241, [arXiv:hep-th/0003188](#).
- [130] F. Cachazo and N. Early, “Planar Kinematics: Cyclic Fixed Points, Mirror Superpotential, k-Dimensional Catalan Numbers, and Root Polytopes,” [arXiv:2010.09708 \[math.CO\]](#).
- [131] Iain W. Stewart, “Lectures on the Soft-Collinear Effective Theory,” *EFT Course 8.851, SCET Lecture Notes, Massachusetts Institute of Technology* (2013) .
- [132] J. K. Schlenk, *Techniques for higher order corrections and their application to LHC phenomenology*. PhD thesis, Munich, Tech. U., 8, 2016.
- [133] G. Heinrich, S. Jahn, S. P. Jones, M. Kerner, F. Langer, V. Magerya, A. Pöldaru, J. Schlenk, and E. Villa, “Expansion by regions with pySecDec,” *Comput. Phys. Commun.* **273** (2022) 108267, [arXiv:2108.10807 \[hep-ph\]](#).
- [134] E. Gardi and L. Magnea, “Factorization constraints for soft anomalous dimensions in QCD scattering amplitudes,” *JHEP* **03** (2009) 079, [arXiv:0901.1091 \[hep-ph\]](#).

- [135] N. Agarwal, L. Magnea, C. Signorile-Signorile, and A. Tripathi, “The Infrared Structure of Perturbative Gauge Theories,” [arXiv:2112.07099 \[hep-ph\]](#).
- [136] A. H. Mueller, “On the Asymptotic Behavior of the Sudakov Form-factor,” *Phys. Rev. D* **20** (1979) 2037.
- [137] J. C. Collins, “Algorithm to Compute Corrections to the Sudakov Form-factor,” *Phys. Rev. D* **22** (1980) 1478.
- [138] A. Sen, “Asymptotic Behavior of the Sudakov Form-Factor in QCD,” *Phys. Rev. D* **24** (1981) 3281.
- [139] O. Gürdoğan and V. Kazakov, “New Integrable 4D Quantum Field Theories from Strongly Deformed Planar $\mathcal{N} = 4$ Supersymmetric Yang-Mills Theory,” *Phys. Rev. Lett.* **117** no. 20, (2016) 201602, [arXiv:1512.06704 \[hep-th\]](#). [Addendum: *Phys.Rev.Lett.* 117, 259903 (2016)].
- [140] D. Chicherin, V. Kazakov, F. Loebbert, D. Müller, and D.-l. Zhong, “Yangian Symmetry for Fishnet Feynman Graphs,” *Phys. Rev. D* **96** no. 12, (2017) 121901, [arXiv:1708.00007 \[hep-th\]](#).
- [141] F. Loebbert and J. Miczajka, “Massive Fishnets,” *JHEP* **12** (2020) 197, [arXiv:2008.11739 \[hep-th\]](#).
- [142] B. Ananthanarayan, S. Banik, S. Friot, and S. Ghosh, “Multiple Series Representations of N-fold Mellin-Barnes Integrals,” *Phys. Rev. Lett.* **127** no. 15, (2021) 151601, [arXiv:2012.15108 \[hep-th\]](#).
- [143] Y. Ma, “A Forest Formula to Subtract Infrared Singularities in Amplitudes for Wide-angle Scattering,” *JHEP* **05** (2020) 012, [arXiv:1910.11304 \[hep-ph\]](#).
- [144] L. de la Cruz, “Feynman integrals as A-hypergeometric functions,” *JHEP* **12** (2019) 123, [arXiv:1907.00507 \[math-ph\]](#).
- [145] R. P. Klausen, “Hypergeometric Series Representations of Feynman Integrals by GKZ Hypergeometric Systems,” *JHEP* **04** (2020) 121, [arXiv:1910.08651 \[hep-th\]](#).
- [146] T.-F. Feng, C.-H. Chang, J.-B. Chen, and H.-B. Zhang, “GKZ-hypergeometric systems for Feynman integrals,” *Nucl. Phys. B* **953** (2020) 114952, [arXiv:1912.01726 \[hep-th\]](#).
- [147] See Supplemental Material at [URL] for further examples of computing leading UV and IR divergences, as well as an extension to more general integrals. It includes Refs. [34, 35].
- [148] G. P. Korchemsky, “Asymptotics of the Altarelli-Parisi-Lipatov Evolution Kernels of Parton Distributions,” *Mod. Phys. Lett. A* **4** (1989) 1257–1276.

- [149] G. F. Sterman and M. E. Tejeda-Yeomans, “Multiloop amplitudes and resummation,” *Phys. Lett. B* **552** (2003) 48–56, [arXiv:hep-ph/0210130](#).
- [150] P. Mastrolia and S. Mizera, “Feynman Integrals and Intersection Theory,” *JHEP* **02** (2019) 139, [arXiv:1810.03818 \[hep-th\]](#).
- [151] S. Abreu, R. Britto, C. Duhr, E. Gardi, and J. Matthew, “From positive geometries to a coaction on hypergeometric functions,” *JHEP* **02** (2020) 122, [arXiv:1910.08358 \[hep-th\]](#).
- [152] R. Beekveldt, M. Borinsky, and F. Herzog, “The Hopf algebra structure of the R^* -operation,” *JHEP* **07** (2020) 061, [arXiv:2003.04301 \[hep-th\]](#).
- [153] J. H. Lowenstein and W. Zimmermann, “The Power Counting Theorem for Feynman Integrals with Massless Propagators,” *Commun. Math. Phys.* **44** (1975) 73–86.
- [154] P. Benincasa, “Cosmological Polytopes and the Wavefunction of the Universe for Light States,” [arXiv:1909.02517 \[hep-th\]](#).
- [155] C. Duhr and F. Dulat, “PolyLogTools — polylogs for the masses,” *JHEP* **08** (2019) 135, [arXiv:1904.07279 \[hep-th\]](#).
- [156] E. Panzer, “Algorithms for the symbolic integration of hyperlogarithms with applications to Feynman integrals,” *Comput. Phys. Commun.* **188** (2015) 148–166, [arXiv:1403.3385 \[hep-th\]](#).

