

All loop scattering as a counting problem

N. Arkani-Hamed,^a H. Frost^b, G. Salvatori^c, P-G. Plamondon^d and H. Thomas^e

^a*School of Natural Sciences, Institute for Advanced Study,
Princeton, NJ, 08540, U.S.A.*

^b*Mathematical Institute,
Andrew Wiles Building, Woodstock Rd, Oxford, U.K.*

^c*Max-Planck-Institut für Physik,
Werner-Heisenberg-Institut, D-80805 München, Germany*

^d*Laboratoire de Mathématiques de Versailles, UVSQ, CNRS, Université Paris-Saclay, IUF,
Paris, France*

^e*LaCIM, Département de Mathématiques, Université du Québec à Montréal,
Montréal, QC, Canada*

E-mail: arkani@ias.edu, frost@maths.ox.ac.uk, giulios@mpp.mpg.de,
pierre-guy.plamondon@uvsq.fr, thomas.hugh_r@uqam.ca

ABSTRACT: This is the first in a series of papers presenting a new understanding of scattering amplitudes based on fundamentally combinatorial ideas in the kinematic space of the scattering data. We study the simplest theory of colored scalar particles with cubic interactions, at all loop orders and to all orders in the topological 't Hooft expansion. We find novel integral formulas for the amplitudes of this theory, with no trace of the conventional sum over Feynman diagrams, but instead determined by a beautifully simple counting problem attached to any order of the topological expansion. These results represent a significant step forward in the decade-long quest to formulate the fundamental physics of the real world in a radically new language, where the rules of spacetime and quantum mechanics, as reflected in the principles of locality and unitarity, are seen to emerge from deeper mathematical structures.

KEYWORDS: $1/N$ Expansion, Scattering Amplitudes

ARXIV EPRINT: [2309.15913](https://arxiv.org/abs/2309.15913)

Contents

1	Introduction and summary	2
1.1	Kinematic space	7
1.2	The first miracle: discovering Feynman diagrams	9
1.3	An infinity of diagrams and the spectre of gravity	11
1.4	The amplitudes	12
1.5	The second miracle: the counting problem	14
1.6	Back to the amplitude!	18
2	The partial amplitude expansion	18
3	Momenta and curves	21
3.1	Mountainscapes	21
3.2	Intersections	22
3.3	Momentum assignments	23
3.4	Spirals	25
4	The Feynman fan	26
4.1	Example: tree level at 5-points	27
4.2	The fan	28
4.3	The mapping class group	28
4.4	Example: the non-planar 1-loop propagator	30
4.5	The delta plane	30
4.6	Example: the planar 1-loop propagator	32
5	A counting problem for curves	33
5.1	Curve matrices	33
5.2	Headlight functions	35
5.3	Example: tree level at 5-points	36
5.4	Example: the non-planar 1-loop propagator	37
5.5	Spirals	38
5.6	Example: the planar 1-loop propagator	39
5.7	Example: the genus one 2-loop vacuum	39
6	Integrand curve integrals	41
6.1	Example: the tree level 5-point amplitude	42
6.2	Example: the planar 1-loop propagator	42
6.3	Example: the planar 1-loop 3-point amplitude	43
6.4	Note on factorization	44
7	Amplitude curve integrals	45
7.1	Example: the planar 1-loop propagator	47

7.2	Example: the non-planar 1-loop propagator	48
7.3	Example: the non-planar 3-point amplitude	49
7.4	Example: genus-one 2-loop amplitudes	49
8	Modding out by the mapping class group	50
8.1	Warm up	50
8.2	A tropical Mirzakhani kernel	51
8.3	Example: the non-planar 1-loop propagator	52
8.4	General tropical Mirzakhani kernels	53
8.5	The general iterative method	54
8.6	Example: the genus one 2-loop vacuum amplitude	55
9	Exemplitudes	56
9.1	The non-planar 1-loop 3-point amplitude	56
9.2	The genus one 2-loop vacuum amplitude	58
9.3	The planar 2-loop tadpole	60
9.4	The planar 3-loop vacuum amplitude	62
10	A first look at recursion	64
11	Outlook	66
A	Deriving the curve integral formula	68
B	Factorization in detail	69
B.1	MCG invariant curve	69
B.2	MCG non-invariant curve	70
C	The surface Symanzik polynomials	71
C.1	The first surface Symanzik	71
C.2	The second surface Symanzik	73
D	The recursion formula	74
E	Recursion examples	75
E.1	The 3-point non-planar 1-loop amplitude	75
E.2	The 2-loop vacuum at genus one	76
E.3	A comment on the 1-loop planar amplitudes	77
F	Details for the non-planar 1-loop propagator	77

1 Introduction and summary

Scattering amplitudes are perhaps the most basic and important observables in fundamental physics. The data of a scattering process — the on-shell momenta and spins of the particles — are specified at asymptotic infinity in Minkowski space. The conventional textbook formalism

for computing amplitudes “integrates in” auxiliary structures that are not present in the final amplitude, including the bulk spacetime in which particle trajectories are imagined to live, and the Hilbert space in which the continuous bulk time evolution of the wavefunction takes place. These auxiliary structures are reflected in the usual formalism for computing amplitudes, using Feynman diagrams, which manifests the rules of spacetime (locality) and quantum mechanics (unitarity). As has been increasingly appreciated over the past three decades, this comes at a heavy cost — the introduction of huge redundancies in the description of physics, from field redefinitions to gauge and diffeomorphism redundancies, leading to enormous complexities in the computations, that conceal a stunning hidden simplicity and seemingly miraculous mathematical structures revealed only in the final result [1–7].

This suggests that we should find a radically different formulation for the physics of scattering amplitudes. The amplitudes should be the answer to entirely new mathematical questions that make no reference to bulk spacetimes and Hilbert space, but derive locality and unitarity from something more fundamental. A number of concrete examples of this have already been found in special cases. The discovery of deep and simple new structures in combinatorics and geometry has led to new definitions of certain scattering amplitudes, without reference to spacetime or quantum mechanics. Notably, the amplituhedron determines the scattering amplitudes in planar $N = 4$ SYM, and associahedra and cluster polytopes determine colored scalar amplitudes at tree-level and one-loop [8–11].

Up to now, these results have been limited in how much of the perturbative expansion they describe — at all loop orders for maximally supersymmetric theories, but only in the planar limit, and only through to one loop for non-supersymmetric theories. Furthermore, the connection between combinatorial geometry and scattering amplitudes at loop level has only been made through the integrand (pre-loop integration) of the amplitudes, and not the amplitudes themselves. Both of these limitations must be transcended to understand all aspects of particle scattering in the real world.

This article is the first in a series reporting on what we believe is major new progress towards this goal. These ideas set the foundation for a number of other interrelated threads and results that will appear in various groups of papers. So we take this opportunity to give a birds-eye view of the nature of these developments and the new concepts that are driving this progress.

Our departure point is a new formulation of a simple theory, — colored scalar particles with cubic interactions, — at all loop orders and to all orders in the topological ’t Hooft expansion, in the form of what we call a *curve integral*. This approach has no hint of a sum over Feynman diagrams anywhere in sight and is instead associated with a simple counting problem defined at any order in the topological expansion. This counting problem defines a remarkable set of variables, u_C , associated with every curve, C , on a surface. The u -variables non-trivially define *binary geometries* [12] by dint of satisfying the remarkable non-linear equations [13]

$$u_C + \prod_D u_D^{n(C,D)} = 1, \quad (1.1)$$

where $n(C, D)$ is the intersection number of the curves C, D .¹ In the *positive region*, where all the u_C are non-negative, the u -equations force all the u_C to lie between 0 and 1: $0 \leq u_C \leq 1$. Of mathematical interest, this positive region is a natural and invariant compactification of *Teichmüller space*. This algebraic presentation of Teichmüller space is a counterpart to the famous synthetic compactification of Teichmüller spaces and surface-type cluster varieties given by Fock-Goncharov [14, 15]. The new compactifications defined by the u_C variables are immediately relevant for physics, and lead to the new *curve integral* formulation of all-loop amplitudes presented in this article.

The curve integral does more than reformulate the perturbative series in a new way. It also exposes basic new structures in field theory. A striking example, of this, to be explored elsewhere, is that the curve integral formula for n particles at L -loops effectively decomposes into a tree and a loop computation. In particular, the integrands for the curve integral formulas at large n can be reconstructed from low-point L -loop calculations. (Of course, the amplitudes themselves do not factorise into a tree and loop computation!) Moreover, our curve integral formulas make it possible to *define* and study an integrand for each partial amplitude, and these integrands satisfy a family of differential equations that we explore elsewhere. The solutions of these equations give novel and efficient recursion relations for all-loop amplitudes.

This article focuses on a colored scalar theory. However, the results here have extensions to other theories. New curve integral formulations have been discovered for theories of colored scalar particles with arbitrary local interactions, as well as for the amplitudes of pions and non-supersymmetric Yang-Mills theories. These formulas reveal striking inter-relations between these theories, together with surprising hidden properties of their amplitudes that are made manifest by the curve integral formalism.

Our results also have implications for the understanding of strings and UV completion. The counting problem at the heart of this paper is related not only to QFT amplitudes, but they also can be used to study amplitudes for bosonic strings, via the u -variables, u_C , mentioned above. This suggests that there is a combinatorial formulation of string amplitudes that makes no reference to worldsheet CFTs and vertex operators. This new approach to string amplitudes differs from the conventional theory in a very fundamental way. The u -variables, which are derived from a simple counting problem, have a beautiful and direct connection to the geometry of two-dimensional surfaces. But this connection is via the *hyperbolic geometry* of Teichmüller space, and *not* via the conventional picture of Riemann surfaces with a complex structure. The new string formulas are not just an exercise in passing between the complex and the hyperbolic pictures for Teichmüller space. We find that we can reproduce bosonic strings at loop level, but other choices are just as consistent, at least insofar as the field theory limit is concerned. This allows us to deform string amplitudes into a larger, but still highly constrained, space of interesting objects. This runs counter to the lore that string theory is an inviolable structure that cannot be modified without completely

¹The intersection number between two curves is defined to be the minimal number of intersections achievable as the two curves are deformed without changing their homotopy class. In the mathematical literature this is sometimes referred to as the *geometrical* intersection number. In particular, two curves are non-intersecting if they can be drawn on the fatgraph in such a way that they never cross, i.e. they are homotopic to a pair of non-crossing curves. In addition, see section 3.2 for an easier and more combinatorial way to think about intersections.

breaking it. Our larger class of string amplitudes transcends the usual strictures on spacetime dimension, as well as the famous instabilities of non-supersymmetric strings. Moreover, our new combinatorial-geometric point of view also makes it easier to recover particle amplitudes from strings in the $\alpha' \rightarrow 0$ limit. By contrast, recovering field theory from conventional string theory involves vastly (technically, infinitely!) more baggage than is needed [16].

There are several other related developments, including the discovery of a remarkable class of polytopes, *surfacehedra*, whose facet structure captures, mathematically, the intricate boundary structure of Teichmüller space, and, physically, the intricate combinatorics of amplitude singularities at all loop orders, and whose *canonical form* determines (an appropriate notion of the) loop integrand at all orders in the topological expansion.

The results of all these parallel threads of investigation will be presented in various groups of papers. We end this preview of coming attractions by explaining a quite different sort of motivation for our works that will be taken up in near-future work. The counting problem that lies at the heart of this paper has an entirely elementary definition. But the central importance of this counting problem will doubtless seem mysterious at first sight. It finds its most fundamental origin in remarkably simple but deep ideas from the “quiver representation theory” [17, 18] of (triangulated) surfaces. Arrows between the nodes of a quiver can be associated with maps between vector spaces attached to the nodes. Choosing compatible linear maps between the nodes defines a *quiver representation*. In this context, our counting problem is equivalent to counting the *sub-representations* of these quiver representations. This perspective illuminates the mathematical structure underlying all of our formulas. But these ideas also hint at a fascinating prospect. The amplitudes we study are associated with the class of surface-type quivers, which are dual to triangulated surfaces. Nothing in our formulas forces this restriction on us: we are free to consider a much wider array of quivers. *All* of these quivers can be associated with amplitude-like functions. This vast new class of functions enjoys an intricate (amplitude-like) structure of “factorisations” onto simpler functions. This amounts to a dramatic generalisation of the notion of an “amplitude”, and in a precise sense also generalises the rules of spacetime and quantum mechanics to a deeper, more elementary, but more abstract setting.

Having outlined this road map, we return to the central business of this first paper. We will study a simple theory of a single massive scalar field with values in tensor product of the fundamental and anti-fundamental representations of $SU(N)$, $\phi = \phi_J^I t_I \tilde{t}^J$. The Lagrangian, with minimal cubic coupling, is

$$\mathcal{L} = \text{Tr}(\partial\phi)^2 + m^2\text{Tr}(\phi^2) + g\text{Tr}(\phi^3), \tag{1.2}$$

in any number D of spacetime dimensions. The Feynman diagrams of this theory are fatgraphs (see section 2 for more details on the theory and its partial amplitude expansion).

This cubic scalar theory is a simpler cousin of all theories of colored particles, including Yang-Mills theories. Kinematical singularities of scattering amplitudes ultimately arise from integrating around loci in loop momentum space where propagators of Feynman diagrams go on-shell. While the “numerator structure” of a specific colored theory may suppress these singularities, the “denominator structure” of the cubic theory is universal, i.e. contains that of any other colored theory. This means that the cubic theory includes all *potential* singularities

that may be encountered in colored theories. The singularities of amplitudes are associated with some of the most fundamental aspects of their conventional interpretation in terms of spacetime processes respecting unitarity. So understanding the amplitudes for this simple theory is an important step towards attacking much more general theories.

We will show that *all* amplitudes in this theory, for any number n of external particles, and to all orders in the genus (or $1/N$) expansion [19], are naturally associated with a strikingly simple counting problem. This counting problem is what allows us to give new integral formulas for the amplitudes at all orders. These *curve integral* formulas resemble the Schwinger parameterised Feynman integrals for a single diagram.

As an example, consider the single-trace amplitude for n -point scattering at 1-loop. Let the particles have momenta p_i^μ , $i = 1, \dots, n$. The curve integral for this amplitude (pre-loop integration) is

$$\mathcal{A}_n^{1\text{-loop}} = \int d^D l \int_{\sum_i t_i \geq 0} d^n t \exp \left[- \sum_{i=1}^n \alpha_i (l + p_1 + \dots + p_i)^2 - \sum_{i,j} \alpha_{i,j} (p_i + \dots + p_{j-1})^2 \right] \quad (1.3)$$

where

$$\alpha_{i,j} = f_{i,j} + f_{i+1,j+1} - f_{i,j+1} - f_{i+1,j}, \quad (1.4)$$

$$\alpha_i = \alpha_{i,i+n},$$

$$f_{i,j} = \max(0, t_j, t_j + t_{j-1}, \dots, t_j + t_{j-1} + \dots + t_{i+2}). \quad (1.5)$$

The propagators that arise in the 1-loop Feynman diagrams are either loop propagators, with momenta $(l + p_1 + \dots + p_i)$, or tree-like propagators, with momenta $(p_i + p_{i+1} + \dots + p_{j-1})$. The exponential in (1.3) looks like a conventional Schwinger parametrisation integral, except that *all* the propagators that arise at 1-loop are included in the exponent. Instead of Schwinger parameters, we have *headlight functions*: α_i (for the loop propagators) and $\alpha_{i,j}$ (for the tree propagators). The headlight functions are piecewise linear functions of the t_i variables. The magic is that (1.3) is a *single* integral over an n -dimensional vector space. Unlike conventional Schwinger parametrisation, which is done one Feynman diagram at a time, our formulas make no reference to Feynman diagrams. Amazingly, the exponent in (1.3) breaks t -space into different cones where the exponent is linear. Each of these cones can be identified with a particular Feynman diagram, and the integral in that cone reproduces a Schwinger parameterisation for that diagram. This miracle is a consequence of the properties of the headlight functions $\alpha_i(t)$ and $\alpha_{i,j}(t)$. These special functions arise from a simple counting problem associated with the corresponding propagator.

As in conventional Schwinger parametrisation, the dependence on the loop momentum variable, l^μ , in the curve integral, (1.3), is Gaussian. We can perform the loop integration to find a second curve integral for the amplitude (post loop integration),

$$\mathcal{A}_n^{1\text{-loop}} = \int_{\sum_i t_i \geq 0} d^n t \left(\frac{\pi^L}{\mathcal{U}} \right)^{\frac{D}{2}} e^{-\frac{\mathcal{F}}{\mathcal{U}}}. \quad (1.6)$$

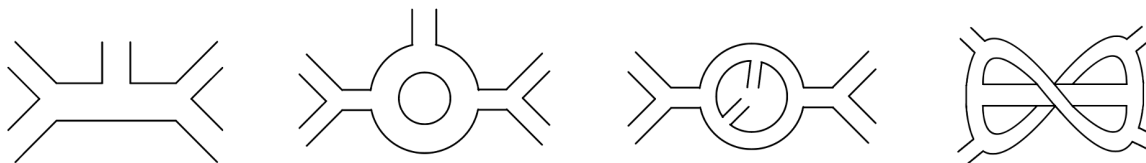


Figure 1. Fat graphs at tree-level, 1-loop single trace, 1-loop double trace, and 2-loop single trace, respectively.

In this formula, the polynomials \mathcal{U} and \mathcal{F} are given by

$$\mathcal{U} = \sum_i \alpha_i, \quad \mathcal{F} = \sum_{i,j} \alpha_i \alpha_j (p_i + \dots + p_{j-1})^2 - \left(m^2 \sum_i \alpha_i + 2 \sum_{i,j} \alpha_{i,j} p_i \cdot p_j \right) \mathcal{U}. \quad (1.7)$$

These polynomials are analogs of the familiar Symanzik polynomials, but whereas the Symanzik polynomials appear in individual Feynman integrals, this one integral above is an integral formula for the whole amplitude.

These 1-loop curve integrals generalise to all orders in perturbation theory, at any loop order and genus. In the rest of this introductory section we give a birds-eye view of the key formulas and results.

1.1 Kinematic space

To begin with, we have to define the *kinematic space* where all the action will take place. In our theory, each Feynman diagram is what is called a ‘double-line notation diagram’, ‘ribbon graph’ or ‘fatgraph’ in the literature; we will call them fatgraphs in what follows. Examples of fatgraphs are shown in figure 1. Order by order, in the ’t Hooft expansion, these Feynman diagrams get organised into partial amplitudes, labeled by their shared *color structure*. Conventionally, when we do a ’t Hooft expansion, we think of these fat graphs as ‘living on’ or ‘being drawn on’ a surface with some genus and number of boundary components. We will think of them in a different way: a *single* fat graph itself *defines* a surface. In fact, we will use a single fat graph to define all the data we need to compute an amplitude!

Take some fatgraph, Γ , at any order in the ’t Hooft expansion. Suppose that it has n external lines and E internal edges. Then this fat graph has loop order, L , with

$$E = n + 3(L - 1). \quad (1.8)$$

Let the external lines have momenta p_1, \dots, p_n , and introduce L loop variables, ℓ_1, \dots, ℓ_L . Then, by imposing momentum conservation at each vertex of Γ , we can find a consistent assignment of momenta to all edges of the fat graph in the usual way: if each edge, e , gets a momentum p_e^μ , then whenever three edges, e_1, e_2, e_3 , meet at a vertex, we have

$$p_{e_1}^\mu + p_{e_2}^\mu + p_{e_3}^\mu = 0. \quad (1.9)$$

For example, figure 2 is an assignment of momenta to the edges of a tree graph.

The amplitude itself depends on momenta only through Lorentz invariant combinations. So we want to define a collection of Lorentz invariant kinematic variables. Consider a curve,

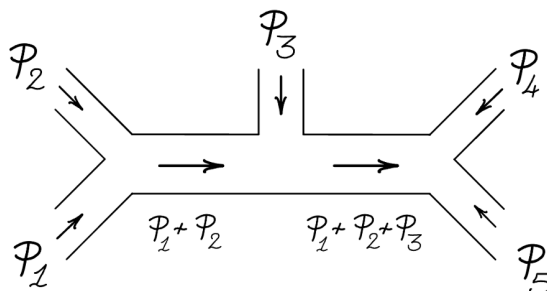


Figure 2. A tree fat graph with momenta assigned to all edges.

C , drawn on the fatgraph Γ that starts at an external line, passes through the graph and exits at another external line. For example, the curve in figure 3 starts at p_2 , and exits at p_5 . Every such curve can be assigned a unique momentum. It is given by the momentum of the first edge plus the sum of all momenta on the graph entering the curve ‘from the left’. For example, in figure 3, the curve starts with momentum p_2 , and then takes two right turns. At the first right turn, momentum p_3 enters from the left. At the second right turn, momentum p_4 enters from the left. The total momentum of the curve is then given by

$$p_C^\mu = p_2^\mu + p_3^\mu + p_4^\mu. \tag{1.10}$$

Notice that if we had gone in the opposite direction (starting at p_5), we would have got

$$-p_C^\mu = p_5^\mu + p_1^\mu. \tag{1.11}$$

But by total momentum conservation ($p_1 + \dots + p_n = 0$), it does not matter which direction we take.

For a general curve, C , on any fatgraph, this rule can be written as:

$$P_C^\mu = p_{\text{start}}^\mu + \sum_{\text{right turns}} p_{\text{from left}}^\mu. \tag{1.12}$$

This rule assigns to every curve C on our fatgraph Γ some momentum, P_C^μ . Each P_C^μ is a linear combination of external momenta, p_i , and loop variables, ℓ_a . Each curve, C , then also defines a Lorentz invariant kinematic variable

$$X_C = P_C^2 + m^2. \tag{1.13}$$

The collection of variables X_C , for *all* curves C on the fatgraph, defines a complete set of kinematic variables in our kinematic space. Modulo a small detail about how to deal with internal color loops, this completes the description of our kinematic space.

It is significant in our story that we can use the momenta of a *single* fat graph (or Feynman diagram) to define a complete set of kinematic variables X_C . As we will see in more detail in section 6, this basic idea ends up solving the long-standing problem of defining a good notion of loop integrand beyond the planar limit!

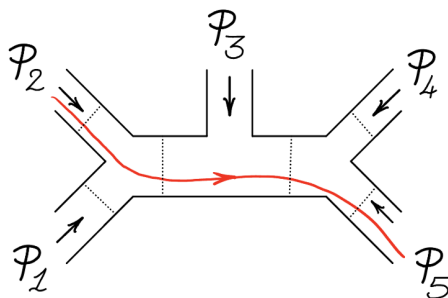


Figure 3. Every curve, C , drawn on the fat graph inherits a momentum, P_C^μ , from the momenta assigned to the fat graph itself.

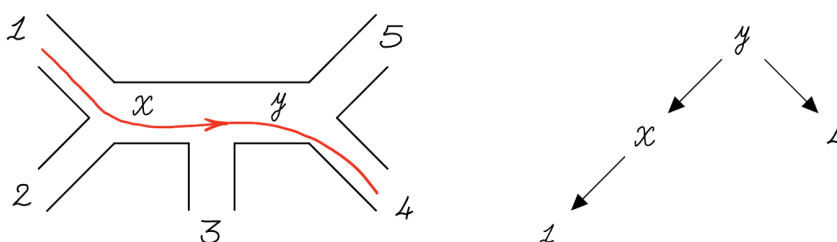


Figure 4. Describing a curve on a fatgraph (left) using a mountainscape diagram (right).

1.2 The first miracle: discovering Feynman diagrams

We now look for a question whose answer produces scattering amplitudes. We just saw how we can define all our kinematics using a single fatgraph. So with this starting point, what would make us consider *all* possible Feynman diagrams (i.e. all spacetime processes)? And why should these be added together with equal weights (as demanded by quantum mechanics)? Amazingly, the answer to both of these fundamental questions is found right under our noses, once we think about how to systematically describe all the curves on our fatgraph.

How can we describe a curve on our fat graph without drawing it? We can do this by labeling all the edges, or “roads”, on the fatgraph. Any curve passes through a series of these roads. Moreover, at each vertex, we demand that the curve must turn either left or right: we do not allow our curves to do a ‘U turn’. It follows that a curve is fully described by the order of the roads and turns it takes as it passes through the graph. For example, the curve in figure 4 enters through edge ‘1’, takes a left turn, goes down ‘ x ’, takes a left turn, goes down ‘ y ’, takes a right turn, and then exits via ‘4’.

We can represent this information graphically as a *mountainscape*, where left turns are represented by upward slopes, and right turns are represented by downward slopes. The mountainscape for the curve in figure 4 is shown in the figure.

Once again, let our fatgraph have E internal edges. To every curve C , we will associate a vector \mathbf{g}_C in *curve space*. As a basis for this vector space, take E vectors \mathbf{e}_i , associated to each internal edge. Then \mathbf{g}_C can be read off from the mountainscape for C using the following rule:

$$\mathbf{g}_C = \sum_{\text{peaks } p} \mathbf{e}_p - \sum_{\text{valleys } v} \mathbf{e}_v. \tag{1.14}$$

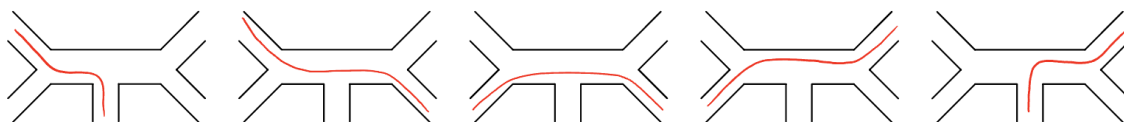


Figure 5. The five (non boundary) curves that we can draw on a 5-point tree fatgraph.

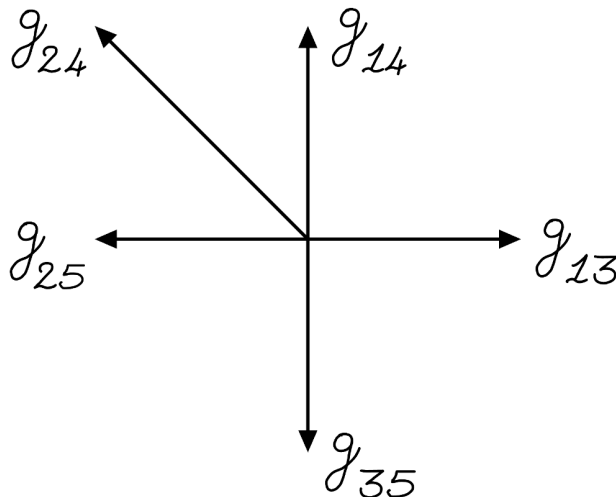


Figure 6. The collection of \mathbf{g} -vectors for the fat graph in figure 4 cuts the 2-dimensional vector space into five regions.

For example, the curve in figure 4 has a peak at ‘ y ’ and no valleys. So the g -vector for this curve is

$$\mathbf{g}_C = \mathbf{e}_y. \tag{1.15}$$

Now consider *every* curve that we can draw on the fatgraph in figure 4. There are 10 possible curves. 5 of these are ‘boundaries’, and their g -vectors end up vanishing (because their mountainscapes have no peaks or valleys). The remaining 5 curves are drawn in figure 5. If we label the external lines, each curve can be given a name C_{ij} ($i, j = 1, 2, 3, 4, 5$), where C_{ij} is the curve connecting i and j . Their g -vectors are

$$\mathbf{g}_{13} = \mathbf{e}_x, \quad \mathbf{g}_{14} = \mathbf{e}_y, \quad \mathbf{g}_{24} = -\mathbf{e}_x + \mathbf{e}_y, \quad \mathbf{g}_{25} = -\mathbf{e}_x, \quad \mathbf{g}_{35} = -\mathbf{e}_y. \tag{1.16}$$

If we draw these five g -vectors, we get figure 6. This has revealed a wonderful surprise! Our g -vectors have divided curve space into five regions or *cones*. These cones are spanned by the g -vectors for the following pairs of curves:

$$(C_{13}, C_{14}), (C_{14}, C_{24}), (C_{24}, C_{25}), (C_{25}, C_{35}), \text{ and } (C_{35}, C_{13}). \tag{1.17}$$

These pairs of curves precisely correspond to *all* the five Feynman diagrams of the 5-point tree amplitude!

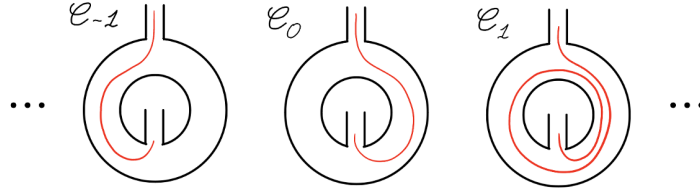


Figure 7. A double-trace 1-loop fat graph, which has infinitely many possible curves.

This is a general phenomenon. The collection of g -vectors for all the curves C on a fatgraph is called *the g -vector fan* [20–22], or *the Feynman fan*, associated to that fatgraph. Each top-dimensional cone of the fan is spanned by an E -tuple of curves, C_{a_1}, \dots, C_{a_E} , and these E -tuples of curves are precisely the propagators of Feynman diagrams. Moreover, the cones are non-overlapping, and together they densely cover the entire vector space! The g -vector fan is telling us that all the Feynman diagrams for the amplitude are combined in curve space.

Even better, each of the cones in the g -vector fan have the same size. It is natural to measure the size of a cone, bounded by some g -vectors $\mathbf{g}_1, \dots, \mathbf{g}_E$, using the determinant of these vectors: $\langle \mathbf{g}_1 \cdots \mathbf{g}_E \rangle$. Remarkably, the cones of the g -vector fan all satisfy: $\langle \mathbf{g}_1 \cdots \mathbf{g}_E \rangle = \pm 1$.

To summarise, starting with a *single* fatgraph at any order in perturbation theory, simply recording the data of the curves on the fatgraph, via their g -vectors, brings *all* the Feynman diagrams to life. Furthermore, we see why they are all naturally combined together into one object, since they collectively cover the entire curve space! This represents a very vivid and direct sense in which the most basic aspects of spacetime processes and the sum-over-histories of quantum mechanics arise as the answer to an incredibly simple combinatorial question.

1.3 An infinity of diagrams and the spectre of gravity

An important novelty appears with the first non-planar amplitudes. Consider the double-trace one-loop amplitude at 2-points. A fatgraph for this amplitude is given in figure 7. There are now infinitely many curves that we can draw on this fat graph: they differ from one another only in how many times they wind around the graph.

The g -vector fan for this infinity of curves is shown in figure 8. These g -vectors break curve space up into infinitely many cones. Each of these cones is bounded by a pair of g -vectors, g_{C_m} and $g_{C_{m+1}}$, where C_m and C_{m+1} are two curves that differ by exactly one *winding*. If we use our rule for the momenta of curves, (1.12), the momenta of these curves are

$$P_{C_m}^\mu = mk^\mu + \ell^\mu, \text{ and } P_{C_{m+1}}^\mu = (m+1)k^\mu + \ell^\mu, \quad (1.18)$$

where k^μ is the momentum of the external particles. Note that the momenta associated to each cone are related to each other by a translation in the loop variable, $\ell^\mu \mapsto \ell^\mu + k^\mu$. It follows that *every* cone in figure 8 corresponds to a copy of the *same* Feynman diagram.

What has gone wrong? The g -vector fan is telling us to include infinitely many copies of one Feynman diagram. This is a consequence of the *mapping class group* of the fat graph in figure 7. The mapping class group of this fatgraph acts by increasing the winding of

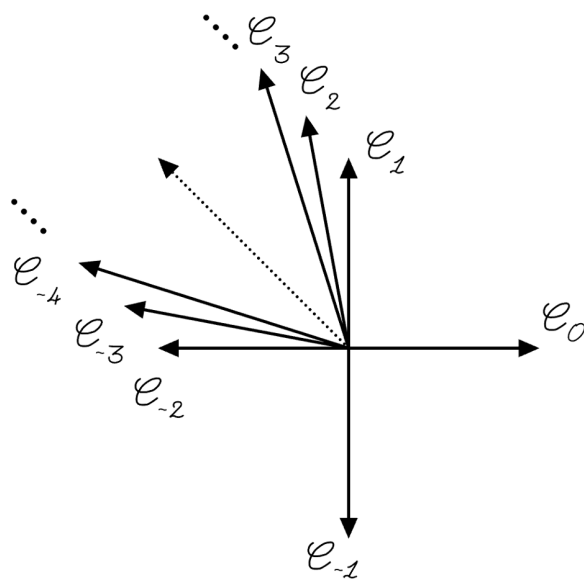


Figure 8. The g-vector fan for the 2-point double-trace 1-loop fat graph, which has infinitely many regions.

curves drawn on the fatgraph. In fact, this infinity of windings is the heart of the well-known difficulty in defining a loop integrand for non-planar amplitudes. Fortunately, as we will see, it is easy to *mod out* by the action of the mapping class group, which we will do using what we call the *Mirzakhani trick*[23]. Getting rid of these infinities using the Mirzakhani trick is the final ingredient we need in order to define amplitudes directly from the combinatorics of a single fatgraph.

As an aside, note that the infinite collection of cones in figure 8 does not quite cover the entire vector space! The g-vectors asymptotically approach the direction $(-1, 1)$, but never reach it. This is the beginning of fascinating story: it turns out that the vector $(-1, 1)$ is the g-vector for the *closed* curve that loops once around the fat graph. Nothing in our story asks us to consider these closed curves, but the g-vector fan forces them on us. Physically, these new closed curves are associated with the appearance of a new *uncoloured* particle, σ . These missing parts of the fan are then seen to have a life of their own: they tell us about a theory with uncoloured self-interactions, σ^3 , that is minimally coupled to our colored particle by an interaction $\sigma \text{Tr}(\Phi)$. The appearance of σ is a scalar avatar of how the graviton is forced on us in string theory even if we begin only with open strings. From our perspective, however, this has absolutely nothing to do with the worldsheet of string theory; it emerges directly from the combinatorics defined by a fatgraph.

1.4 The amplitudes

The g-vector fan gives a beautiful unified picture of all Feynman diagrams living in an E -dimensional vector space, *curve space*. This result suggests a natural formula for the full amplitude in the form of an integral over curve space. To find this formula, we need one extra ingredient. For every curve, C , we will define a piecewise-linear *headlight function*, $\alpha_C(\mathbf{t})$. We will define the headlight function α_C so that it “lights up” curve space in the

direction \mathbf{g}_C , and vanishes in all other g-vector directions:

$$\alpha_C(\mathbf{g}_D) = \delta_{C,D} \tag{1.19}$$

This definition means that α_C vanishes everywhere, except in those cones that involve \mathbf{g}_C . Moreover, α_C is *linear* inside any given cone of the Feynman fan.

Using linear algebra, we can give an explicit expression for α_C in any cone where it is non-vanishing. Suppose that the g-vectors of such a cone are $(\mathbf{g}_C, \mathbf{g}_{D_1}, \dots, \mathbf{g}_{D_{E-1}})$. The unique linear function of \mathbf{t} which evaluates to 1 on \mathbf{g}_C and 0 on all the other g-vectors is

$$\alpha_C = \frac{\langle \mathbf{t} \mathbf{g}_{D_1} \cdots \mathbf{g}_{D_{E-1}} \rangle}{\langle \mathbf{g}_C \mathbf{g}_{D_1} \cdots \mathbf{g}_{D_{E-1}} \rangle}, \tag{1.20}$$

where $\langle \dots \rangle$ denotes the determinant of the vectors inside the brackets (i.e. the vectors give the columns of a square matrix). In what follows, imagine that we already know these functions, $\alpha_C(\mathbf{t})$.

We now define an *action*, S , given by a sum over all curves on a fatgraph:

$$S(\mathbf{t}) = \sum_C \alpha_C(\mathbf{t}) X_C, \quad \text{with } X_C = P_C^2 + m^2. \tag{1.21}$$

Recall that P_C^μ is the momentum we associate to a curve C . If we restrict $S(\mathbf{t})$ to a single cone, bounded by some g-vectors, $\mathbf{g}_{C_1}, \dots, \mathbf{g}_{C_E}$, then the only α 's that are non-zero in this cone are precisely $\alpha_{C_1}, \dots, \alpha_{C_E}$. Moreover, $S(\mathbf{t})$ is linear in this cone. It is natural to parametrise the region inside this cone by $\mathbf{t} = \rho_1 \mathbf{g}_{C_1} + \dots + \rho_E \mathbf{g}_{C_E}$, with $\rho_i \geq 0$ positive. Then we can integrate $\exp(-S)$ in this cone. The result is identical to the result of a standard Schwinger parametrisation for a single Feynman diagram:

$$\int_{\text{cone}} d^E t e^{-S} = \int_0^\infty d^E \rho | \langle g_{C_1} \cdots g_{C_E} \rangle | \prod_{i=1}^E e^{-\rho_i X_{C_i}} = \prod_{i=1}^E \frac{1}{P_{C_i}^2 + m^2}. \tag{1.22}$$

The factor $| \langle g_{X_1} \cdots g_{X_E} \rangle |$ is the Jacobian of the change of variables from (t_1, \dots, t_E) to (ρ_1, \dots, ρ_E) . As we have remarked, the cones are *unimodular* and these Jacobian factors are all equal to 1!

In order to get the full amplitude, all we have to do now is integrate $\exp(-S)$ over the whole vector space, instead of restricting it to just a single cone. However, to account for the infinity resulting from the *mapping class group*, we also need to factor out this MCG action in our integral, which we denote by writing the measure as

$$\frac{d^E t}{\text{MCG}}. \tag{1.23}$$

Before doing the loop integrations, the full amplitude is then given by a *curve integral*:

$$\mathcal{A} = \int d^D \ell_1 \cdots d^D \ell_L \int \frac{d^E t}{\text{MCG}} \exp \left(- \sum_X \alpha_X(\mathbf{t}) (P_X^2 + m^2) \right). \tag{1.24}$$

The dependence on loop momenta in this formula is Gaussian. When we integrate the loop momenta, we find the final amplitude is given by a curve integral

$$\mathcal{A} = \int \frac{d^E t}{\text{MCG}} \left(\frac{\pi^L}{\mathcal{U}(\alpha)} \right)^{\frac{D}{2}} \exp \left(-\frac{\mathcal{F}(\alpha)}{\mathcal{U}(\alpha)} \right). \tag{1.25}$$

$\mathcal{U}(\alpha)$ and $\mathcal{F}(\alpha)$ are homogeneous polynomials in the headlight functions. They are analogous to Symanzik polynomials, but are not associated with any particular Feynman diagram. We give simple formulas for \mathcal{A} and \mathcal{F} in section 7.

As discussed above, the MCG acts on curves by twists that change how much a curve winds around a closed loop.² The MCG can also be regarded as acting on the t_i variables. Each element $\gamma \in \text{MCG}$ acts on the t_i variables as a piecewise linear map, $\gamma : t_i \mapsto t'_i$. A key property of the headlight functions is that

$$\alpha_{\gamma(C)}(t) = \alpha_C(\gamma(t)), \tag{1.26}$$

where $\gamma(C)$ is the curve obtained from the action of MCG on curves, whereas $\gamma(t)$ is the image of the map defined by the MCG action on the t coordinates.³ Given this, we can check that the integrand of (1.25) is MCG invariant, in the sense that the integral takes the form

$$\mathcal{A} = \int \frac{d^E t}{\text{MCG}} f(t), \tag{1.27}$$

for some function $f(t)$ such that $f(\gamma t) = f(t)$ for all $\gamma \in \text{MCG}$. So the key to using these curve integral formulas lies in how we mod out by the MCG.

One way of modding by the MCG is to find a *fundamental domain* in \mathbf{t} -space. However, in practice (after 1 loop example) this becomes very difficult. Instead, we will use an elegant way of modding out that we call *the Mirzakhani trick*, which is analogous to the Fadeev-Popov trick familiar from field theory. As we will see, any MCG invariant function, f , can be integrated as,

$$\int \frac{d^E t}{\text{MCG}} f = \int d^E t \mathcal{K}(\alpha) f, \tag{1.28}$$

where the *Mirzakhani kernel* $\mathcal{K}(\alpha)$ is a simple rational function of the α_C 's.⁴ We will describe several formulas for these kernels. In all cases, \mathcal{K} has support on a finite region of the fan, so that only a small number of the α_C 's is ever needed to compute the amplitude. We will also show how some of our methods for producing \mathcal{K} give new systematic recursive methods for computing amplitudes.

1.5 The second miracle: the counting problem

We have given a formula, (1.25), for partial amplitudes at any order in the 't Hooft expansion of our theory. However, the meat of this formula is in the headlight functions, α_C . The problem is that headlight functions are, naively, hard to compute!

²These twist operations are known as Dehn twists.

³Within the scope of this paper, we can use (1.26) as a working *definition* for the MCG. We refer to [15] or [24] for excellent introductions to the topic.

⁴The restriction on the integration region $\sum_i t_i \geq 0$ in equation (1.6) for 1-loop amplitudes can be thought of as the smallest example of a Mirzakhani kernel. In this formula, we are modding out by a discrete Z_2 symmetry, described more in section 4.5.

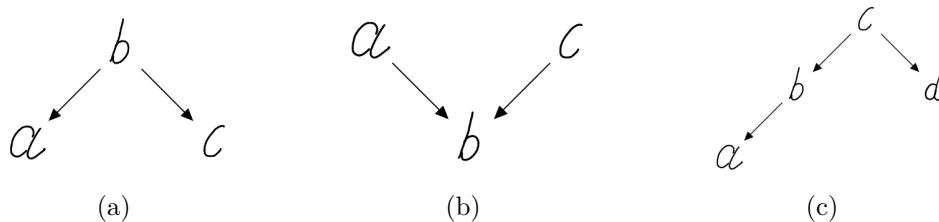


Figure 9. Three mountainscapes.

The issue can already be seen at tree level. For n -points at tree level, the number of possible curves, C , is $\sim n^2$, whereas the number of Feynman diagrams (or cones) grows exponentially as $\sim 4^n$. Each α_C restricts to a different linear function on each of the $\sim 4^n$ cones. So we would expect that it takes an exponentially-growing amount of work to compute all of the α_C , — about as much work as it would take us to just enumerate all the Feynman diagrams to begin with! So, is there an easier way to compute α_C ?

This is where a second miracle occurs. It turns out that headlight functions can be computed efficiently by matrix multiplication. In fact, the calculation is completely *local to the curve*, in the sense that we only need to know the path taken by C , and nothing else about the fatgraph it lives in. There are always many fewer curves than there are Feynman diagrams. This means that the amount of work to compute the α_C 's should grow slower than the amount of work it takes to enumerate all Feynman diagrams.

This way of computing α_C is based on a simple combinatorial problem. For a curve, C , draw its *mountainscape*. We are going to record all the ways in which we can pick a subset of letters of C , subject to a special rule: if we pick a letter y , we also have to pick any letters *downhill* of y . We will then define an *F polynomial* for the curve, $F(C)$, which records the valid subsets.

For example, for the mountainscape in figure 9(a), we get

$$F = 1 + a + c + ac + abc. \tag{1.29}$$

This is because we can choose the following subsets: no-one (“1”); just a ; just c ; a and c together; or finally we can pick b , but if we do, we must also pick a and c , which are both downhill of b . In figure 9(b), we get

$$F = 1 + b + ab + bc + abc, \tag{1.30}$$

because in this example we can choose: no-one; just b ; we can pick a , but if we do we must also pick b ; we can pick c , but we must then also pick b ; and finally we can pick both a and c , but then we must also pick b . Finally, we leave figure 9(c) as an exercise. The result is

$$F = 1 + a + d + ad + ab + abd + abcd. \tag{1.31}$$

In general, there is a fast method for computing $F(C)$ by reading the mountainscape for C from left to right. Say the leftmost letter is Y , and call the next letter y . Then write $F(C) = F_{\text{no}} + F_{\text{yes}}$, where we group the terms in $F(C)$ according to whether they include Y (F_{yes}) or not (F_{no}). Similarly write $f_{\text{no}}, f_{\text{yes}}$ for what we would get starting instead from y .

Suppose that in our mountainscape we move “up” from Y to y . Then if we do not pick Y , then we cannot pick y either, since if we choose y we must choose Y . On the other hand if we do choose Y , we can either pick or not pick y . Thus, in this case, we have

$$F_{\text{no}} = f_{\text{no}}, \quad F_{\text{yes}} = Y(f_{\text{no}} + f_{\text{yes}}). \quad (1.32)$$

Similarly if, in our mountainscape, we move down from Y to y , we find that

$$F_{\text{no}} = f_{\text{no}} + f_{\text{yes}}, \quad F_{\text{yes}} = Y f_{\text{yes}}. \quad (1.33)$$

In matrix form, we find that

$$\begin{pmatrix} F_{\text{no}} \\ F_{\text{yes}} \end{pmatrix} = M_{L,R}(Y) \begin{pmatrix} f_{\text{no}} \\ f_{\text{yes}} \end{pmatrix}, \quad (1.34)$$

where M_L and M_R are the matrices

$$M_L(Y) = \begin{pmatrix} 1 & 0 \\ Y & Y \end{pmatrix}, \quad M_R(Y) = \begin{pmatrix} 1 & 1 \\ 0 & Y \end{pmatrix}. \quad (1.35)$$

Now suppose that the curve C is given explicitly by the following series of edges and turns:

$$(y_1, \text{turn}_1, y_2, \text{turn}_2, \dots, y_{m-1}, \text{turn}_{m-1}, y_m), \quad (1.36)$$

where turn_i is either a left or right turn, immediately following y_i . Given (1.34), we find

$$\begin{pmatrix} F_{\text{no}} \\ F_{\text{yes}} \end{pmatrix} = M \begin{pmatrix} 1 \\ y_m \end{pmatrix}, \quad (1.37)$$

where

$$M(C) = M_{\text{turn}_1}(y_1) M_{\text{turn}_2}(y_2) \cdots M_{\text{turn}_{m-1}}(y_{m-1}). \quad (1.38)$$

So our counting problem is easily solved simply by multiplying a series of 2×2 matrices (equation (1.35)) associated with the left and right turns taken by the curve C .

Suppose that the initial edge of C , y_1 , and the final edge, y_m , are external lines of the fatgraph. It is natural to write $F(C)$ as a sum over four terms:

$$F(C) = F_{\text{no,no}} + F_{\text{no,yes}} + F_{\text{yes,no}} + F_{\text{yes,yes}}, \quad (1.39)$$

where we group terms in $F(C)$ according to whether they do or do not include the first and last edges: y_1 and/or y_m . Indeed, these terms are also the entries of the matrix $M(C)$,

$$M(C) = \begin{pmatrix} F_{\text{no,no}} & F_{\text{no,yes}} \\ F_{\text{yes,no}} & F_{\text{yes,yes}} \end{pmatrix}, \quad (1.40)$$

if we now set $y_m = 1$. In fact, we will also set $y = 1$ for every external line of the fatgraph, and will reserve y -variables for internal edges of the fatgraph.

Notice that $\det M_L(y) = \det M_R(y) = y$, so that

$$\det M(C) = \prod_{i=2}^{m-1} y_i. \quad (1.41)$$

In other words, we have the identity

$$F_{\text{no,no}}F_{\text{yes,yes}} = F_{\text{no,yes}}F_{\text{yes,no}} + \prod_i y_i. \tag{1.42}$$

Motivated in part by this identity, we will define u -variables for every curve,

$$u_C = \frac{F(C)_{\text{no,yes}} F(C)_{\text{yes,no}}}{F(C)_{\text{no,no}} F(C)_{\text{yes,yes}}} = \frac{M(C)_{12}M(C)_{21}}{M(C)_{11}M(C)_{22}}. \tag{1.43}$$

These u_C variables are most interesting to us in the region $y_i \geq 0$. Equation (1.42) implies that $0 \leq u_C \leq 1$ in this region. They vastly generalise the u -variables defined and studied in [25, 26].

We now define the headlight functions. We define them to capture the asymptotic behaviour of the u -variables when thought of as functions of the \mathbf{y} variables. We define

$$\alpha_C = -\text{Trop } u_C. \tag{1.44}$$

where $\text{Trop } u_C$ is the so-called *tropicalization* of u_C .

The idea of tropicalization is to look at how functions behave asymptotically in \mathbf{y} -space. To see how this works, parameterise the $y_i \geq 0$ region by writing $y_i = \exp t_i$, where the t_i are real variables. Then, as the t_i become large, $\text{Trop } u_C$ is defined such that

$$u_C(\mathbf{t}) \rightarrow \exp(\text{Trop } u_C). \tag{1.45}$$

For example, consider a simple polynomial, $P(y_1, y_2) = 1 + y_2 + y_1y_2 = 1 + e^{t_2} + e^{t_1+t_2}$. As we go to infinity in $\mathbf{t} = (t_1, t_2)$ in different directions, different monomials in P will dominate. In fact, we can write, as we go to infinity in \mathbf{t} ,

$$P \rightarrow \exp \max(0, t_2, t_1 + t_2), \tag{1.46}$$

and so $\text{Trop } (P) = \max(0, t_2, t_1 + t_2)$. If we have a product of polynomials, $F = \prod_a P_a^{c_a}$, then as we go to infinity in \mathbf{t} we have $F \rightarrow e^{\text{Trop}(F)}$, where $\text{Trop } F = \sum c_a \text{Trop } (P_a)$.

Returning to headlight functions, our definition can also be written as

$$\alpha_C = \text{Trop } (M(C)_{11}) + \text{Trop } (M(C)_{22}) - \text{Trop } (M(C)_{12}) - \text{Trop } (M(C)_{21}). \tag{1.47}$$

For example, consider again the $n = 5$ tree amplitude. Take the curve C from figure 4 (left). This curve has path $(1, L, x, R, y, R, 4)$. So it has a matrix (with $y_{23}, y_{15} \equiv 1$)

$$M(C) = M_L(1)M_R(x)M_R(y) = \begin{pmatrix} 1 & 1 + y \\ 1 & 1 + y + xy \end{pmatrix}. \tag{1.48}$$

Using this matrix, we find that its u -variable is

$$u_C = \frac{1 + y}{1 + y + xy}, \tag{1.49}$$

and so its headlight function is

$$\alpha_C = \max(0, t_y, t_x + t_y) - \max(0, t_y). \tag{1.50}$$

Amazingly, this function satisfies the key property of the headlight functions: α_C vanishes on every \mathbf{g} -vector, except for its own \mathbf{g} -vector, $\mathbf{g}_C = (1, 0)$.

1.6 Back to the amplitude!

We have now formulated how to compute all-order amplitudes in $\text{Tr}\Phi^3$ theory as a counting problem. The final integral formula for the amplitude at any order of the topological expansion associated with a surface \mathcal{S} is given as

$$\mathcal{A} = \int d^E t \mathcal{K}(\alpha) \left(\frac{\pi^L}{\mathcal{U}(\alpha)} \right)^{\frac{D}{2}} \exp\left(-\frac{\mathcal{F}(\alpha)}{\mathcal{U}(\alpha)}\right), \tag{1.51}$$

where $\mathcal{F}(\alpha), \mathcal{U}(\alpha)$ are homogeneous polynomials in the α_C 's, $\mathcal{K}(\alpha)$ is the Mirzakhani kernel that mods out by the mapping-class-group, and crucially, each α_C is determined entirely by the path of its curve, using a simple counting problem on the curve. The presence of \mathcal{K} ensures that only a finite number of α_C 's ever appear in our computations, which makes the formula easy to apply. There is no trace of the idea of ‘summing over all spacetime processes’ in this formula. Instead, small combinatorial problems attached to the curves on a fatgraph, treated completely independently of each other, magically combine to produce local and unitary physics, pulled out of the platonic thin air of combinatorial geometry.

Our goal in the rest of this paper is to describe these ideas systematically. Our focus in here will exclusively be on simply presenting the formulas for the amplitudes. This presentation will be fully self-contained, so that the interested reader will be fully equipped to find the curve integrals for the $\text{Tr}\phi^3$ theory at any order in the topological expansion. The methods can be applied at any order in the topological expansion, but there are a number of novelties that need to be digested. We illustrate these subtleties one at a time, as we progress from tree level examples through to one and two loops, after which no new phenomena occur. We begin at tree-level to illustrate the basic ideas. At one-loop single-trace, we show how to deal with *spiralling* curves. Then, as we have seen above, double-trace amplitudes at 1-loop expose the first example of the infinities associated with the mapping class group. Finally, we study the leading $1/N$ correction to single-trace at 2-loops — the genus one amplitude — to show how to deal with a non-abelian mapping class group. This non-abelian example illustrates the generality and usefulness of the Mirzakhani trick.

In all cases discussed in this paper we will use the smallest example amplitudes possible to illustrate the new conceptual points as they arise. The next paper in this series will give a similarly detailed set of formulae for amplitudes for any number of particles, n . In this sense this first pair of papers can be thought of as a “user guide” for the formalism. A systematic accounting of the conceptual framework underlying these formulae, together with an exposition of the panoply of related developments, will be given in the later papers of this series.

2 The partial amplitude expansion

Consider a single massive scalar field with values in tensor product of the fundamental and anti-fundamental representations of $\text{SU}(N)$, ϕ_J^I , and with a single cubic interaction,

$$\mathcal{L}_{\text{int}} = g \text{Tr} \left[\phi^3 \right] = g \phi_I^J \phi_J^K \phi_K^I. \tag{2.1}$$

The trace of the identity is $\text{Tr}(1) = \delta_I^I = N$. The propagator for the field ϕ can be drawn as a double line and the Feynman diagrams are *fatgraphs* with cubic vertices. The Feynman

rules follow from (2.1). To compute the n point amplitude, \mathcal{A}_n , fix n external particles with momenta k_i^μ and colour polarisations t_i^{IJ} . A fatgraph Γ with V cubic vertices contributes to the amplitude as

$$(ig)^V C_\Gamma \text{Val}(\Gamma), \tag{2.2}$$

where C_Γ is the tensorial contraction of the polarisations $(\epsilon_i)_I^J$ according to the graph, Γ . The kinematical part is given by an integral of the form

$$\text{Val}(\Gamma) = \int \prod_{i=1}^L d^D \ell_i \prod_{\text{edges } e} \frac{1}{P_e^2 + m^2}, \tag{2.3}$$

for some assignment of loop momenta to the graph. Each momentum P_e^μ is linear in the external momenta k_i^μ and in the loop momentum variables ℓ_i^μ . To find P_e^μ , the edges of Γ need to be oriented, so that momentum conservation can be imposed at each cubic vertex.

The colour factors C_Γ organise the amplitude \mathcal{A}_n into partial amplitudes. This is because C_Γ depends only on the topology of Γ regarded as a *surface*. Indeed, the Feynman rules give that we should contract the ϵ_i polarisations according to the boundary lines of the fatgraph (and any boundary loop of the fatgraph contributes $\text{Tr}(1) = N$). But this tensor depends only on the fatgraph regarded as a surface, and forgets the cubic vertices (see figure 10). Each boundary of the surface contributes a factor of the form $\text{Tr}(\epsilon_1 \cdots \epsilon_k)$ to the colour factor and each puncture of the surface, corresponding to a closed loop on the fatgraph, contributes a factor of N .

Write $S(\Gamma)$ for the surface obtained from the fatgraph Γ by ‘forgetting’ the graph. Two fatgraphs Γ_1, Γ_2 share the same colour factor, C_Σ , if they correspond to the same marked surface, $\Sigma = S(\Gamma_1) = S(\Gamma_2)$. The amplitude can therefore be expressed as

$$\mathcal{A}_n = \sum_{L=0}^{\infty} (ig)^{n-2+2L} \sum_{\substack{\Sigma \text{ s.t.} \\ h+2g=L+1}} C_\Sigma \mathcal{A}_\Sigma, \tag{2.4}$$

where we sum over marked bordered surfaces Σ having n marked points on the boundary. At loop order L , this second sum is over all surfaces Σ with h boundary components and genus g , subject to the Euler characteristic constraint: $h + 2g = L + 1$. The partial amplitudes appearing in (2.4) are

$$\mathcal{A}_\Sigma = \sum_{\substack{\Gamma \\ S(\Gamma)=\Sigma}} \text{Val}(\Gamma). \tag{2.5}$$

Examples of some ribbon graphs Γ and their corresponding surfaces are shown in figure 10.

Our aim is to evaluate \mathcal{A}_Σ . It is conventional to compute $\text{Val}(\Gamma)$ using *Schwinger parameters*. Schwinger parameters are introduced via the identity

$$\frac{1}{P^2 + m^2} = \int_0^\infty d\alpha e^{-\alpha(P^2 + m^2)}. \tag{2.6}$$

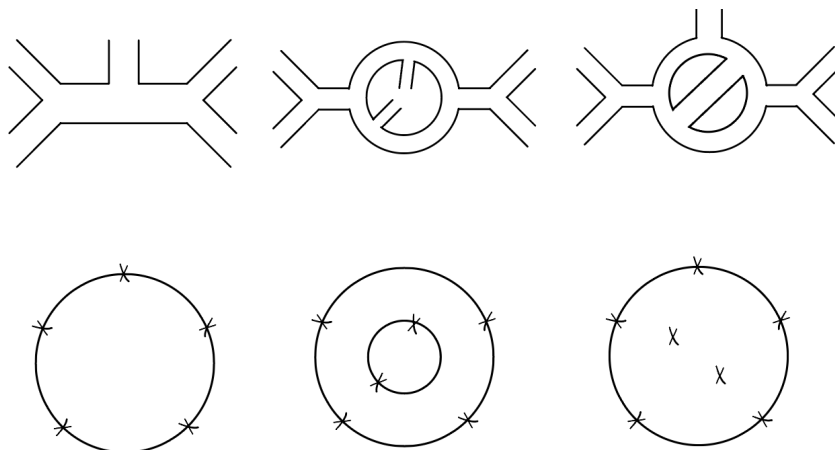


Figure 10. Feynman graphs Γ and the surfaces $S(\Gamma)$ that label their colour factors.

The integration in ℓ_i^μ loop variables then becomes a Gaussian integral, and the result can be written as

$$\text{Val}(\Gamma) = \int_{\alpha_i \geq 0} d^E \alpha \left(\frac{2\pi}{\mathcal{U}_\Gamma} \right)^{\frac{D}{2}} \exp \left(-\frac{\mathcal{F}_{\Gamma,0}}{\mathcal{U}_\Gamma} - m^2 \sum_i \alpha_i \right), \quad (2.7)$$

where \mathcal{U}_Γ and \mathcal{F}_Γ are the Symanzik polynomials of Γ . The Symanzik polynomials depend on Γ regarded as a graph (i.e. forgetting that it is a surface). The first Symanzik polynomial is given by

$$\mathcal{U}_\Gamma = \sum_T \prod_{e \notin T} \alpha_e, \quad (2.8)$$

where the sum is over all spanning trees, T , of Γ . The second Symanzik polynomial is given by a sum over all spanning 2-forests, (T_1, T_2) , which cut Γ into two tree graphs:

$$\mathcal{F}_{\Gamma,0} = \sum_{(T_1, T_2)} \left(\prod_{e \notin T_1 \cup T_2} \alpha_e \right) \left(\sum_{e \notin T_1 \cup T_2} P_e \right)^2, \quad (2.9)$$

where P_e^μ is the momentum of the edge e . It can be shown that $\mathcal{F}_{\Gamma,0}$ depends only on the external momenta, and not on the loop momentum variables.

The partial amplitudes \mathcal{A}_Σ are given by sums over integrals of this form, as in (2.5). But it is the purpose of this paper to show how \mathcal{A}_Σ can be written more compactly as a *single* Symanzik-like integral. It does not work to naively sum the integrands of $\text{Val}(\Gamma)$ for different Feynman diagrams Γ . One problem is that there is no conventional way to relate the loop momentum variables for different Feynman graphs. We will see how this is solved by basic facts from surface geometry. Moreover, a simple counting problem associated to surfaces will allow us to define tropical functions we call *headlight functions*. These simple functions allow us to evaluate the full partial amplitude without enumerating the Feynman diagrams.

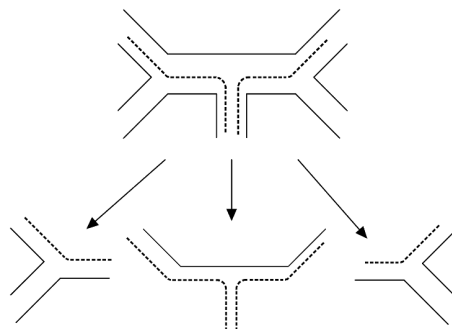


Figure 11. A triangulation of a fatgraph is a maximal set of curves that cuts the fatgraph into cubic vertices.

3 Momenta and curves

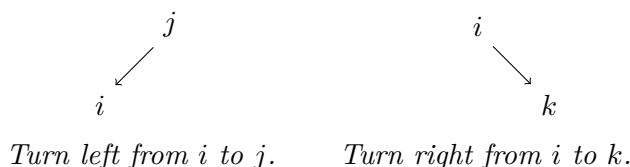
Curves on fatgraphs are the key building block for our formulation of amplitudes. In this section we show how a fatgraph can be used to assign momenta to its curves. This momentum assignment solves the problem of finding a consistent choice of momentum variables for all Feynman diagrams contributing to an amplitude. This generalizes the *dual momentum variables* that can be used for planar amplitudes.

3.1 Mountainscapes

A *curve* is a path on the fatgraph that enters from an external line, passes through the fatgraph without self-intersections, and exits on an external line. It is sometimes useful to separately consider *closed curves*, which are paths on the fatgraph that form a closed loop.

Curves are important because they define *triangulations* of fatgraphs. A triangulation is a maximal collection of pairwise non-intersecting curves. The key point is that each triangulation of Γ corresponds, by graph duality, to some fatgraph Γ' . These fatgraphs Γ' all have the same colour factor and so contribute, as Feynman diagrams, to the same amplitude.⁵ The methods in this paper can be used to automatically find all the triangulations of Γ without having to list them, using only the data of the curves on Γ .

A curve C on Γ is completely specified by reading off the order in which C passes through the edges of Γ . It is also helpful to record the *left and right turns* made by the curve. We present this information using *mountainscape diagrams*. The vertices of a mountainscape are labelled by the edges of Γ . Each left turn made by C is recorded with a left arrow (and a step up), while each right turn is written with a right arrow (and a step down):



⁵There is also a duality between triangulations of a fatgraph Γ , and triangulations of the surface $S(\Gamma)$. Defining this requires some care and is not needed for the results here.

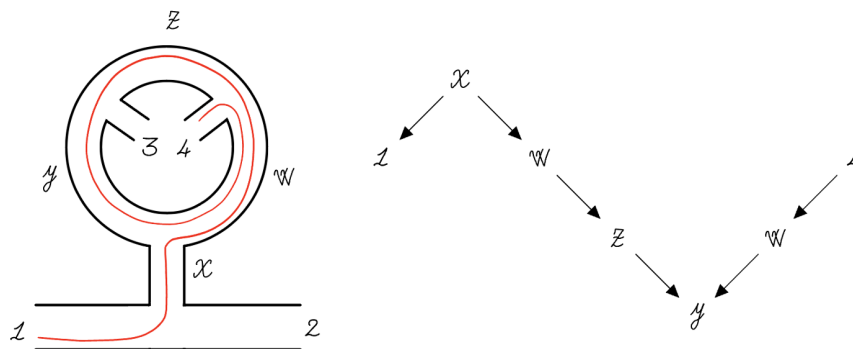


Figure 12. A curve on a fatgraph (left) and its mountainscape diagram (right).

For example, the curve in figure 12(a) passes through the edges $1, x, w, z, y, w, 4$. Its mountainscape is shown in figure 12(b). If we traverse C in the opposite direction we obtain the left-right reflection of this mountainscape. We regard these as being the *same* mountainscape. For brevity, it is convenient to write mountainscapes as a *word*, writing ‘L’ for a left turn, and ‘R’ for a right turn. The mountainscape in figure 12(b) is given by the word

$$C = 1LxRwRzRyLwL4. \tag{3.1}$$

3.2 Intersections

Mountainscape diagrams encode the intersections of curves. In fact, it is not necessary to know the whole fatgraph in order to determine if two curves intersect: their mountainscapes alone have all the data needed.

For example, consider figure 13. The two curves in figure 13(a) are

$$C = x_2RyLx_4 \quad \text{and} \quad C' = x_1LyRx_3. \tag{3.2}$$

These two mountainscapes *overlap* on the edge y , which they share in common. For C , y is a *peak*, whereas for C' , y is a *valley*. This is equivalent to the information that C and C' *intersect* at y . By contrast, the two curves in figure 13(b) are

$$C = x_1LyLx_4 \quad \text{and} \quad C' = x_2RyRx_3. \tag{3.3}$$

These curves also overlap on the edge y . But y does not appear in these curves as a peak or valley. This is equivalent to the information that C and C' do not intersect.

In general, if two curves C and C' intersect, their paths must overlap near the intersection. So suppose that C and C' share some sub-path, W , in common. Then C and C' *intersect along* W only if W is a peak for one and a valley for the other. In other words, C and C' intersect at W if they have the form

$$C = W_1RWLW_2 \quad \text{and} \quad C' = W_3LWRW_4, \tag{3.4}$$

or

$$C = W_1LWRW_2 \quad \text{and} \quad C' = W_3RWLW_4, \tag{3.5}$$

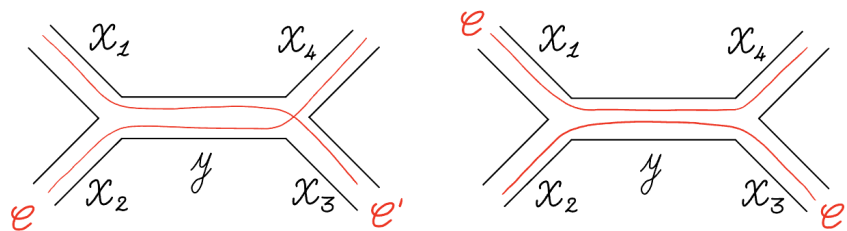


Figure 13. A pair of intersecting curves (left), and a pair of non-intersecting curves (right).

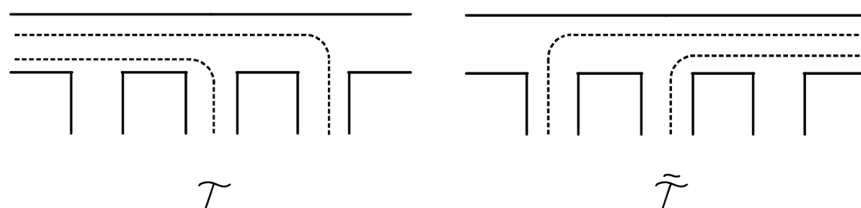


Figure 14. The two special triangulations of a fatgraph, T and \tilde{T} , are defined by curves with exactly one peak (left) and curves with exactly one valley (right).

for some sub-paths W_1, W_2, W_3, W_4 . The left/right turns are very important. If the two curves have the form, say,

$$C = W_1 R W R W_2 \quad \text{and} \quad C' = W_3 L W L W_4, \tag{3.6}$$

then they *do not intersect* at W .

Using this general rule, we can find triangulations of fatgraphs using only the data of the curves.

For every fatgraph Γ , there are two special triangulations. Suppose that Γ has edges $e_i, i = 1, \dots, E$. Let C_i be the curve that, starting from e_i , turns right in both directions away from e_i . Then

$$C_i = \dots L e L e' L e_i R e'' R e''' R \dots \tag{3.7}$$

C_i has exactly one peak, which is at e_i . The intersection rule, (3.4), shows that no pair of such curves C_i, C_j ($i \neq j$) intersect. So the C_i give E nonintersecting curves, and these form a triangulation, T . We can also consider the curves

$$\tilde{C}_i = \dots R e R e' R e_i L e'' L e''' L \dots, \tag{3.8}$$

that turn left going in both directions away from e_i . These \tilde{C}_i each have exactly one valley, at e_i , and so they are mutually nonintersecting. Together, they give another triangulation of the fatgraph, \tilde{T} . An example of these special triangulations is given in figure 14.

3.3 Momentum assignments

The edges of a fatgraph Γ are naturally decorated with momenta, induced by the *external momenta* of the graph. Let Γ have n external momenta p_1^μ, \dots, p_n^μ , directed *into* the graph (say). By imposing momentum conservation at each cubic vertex, we obtain a momentum

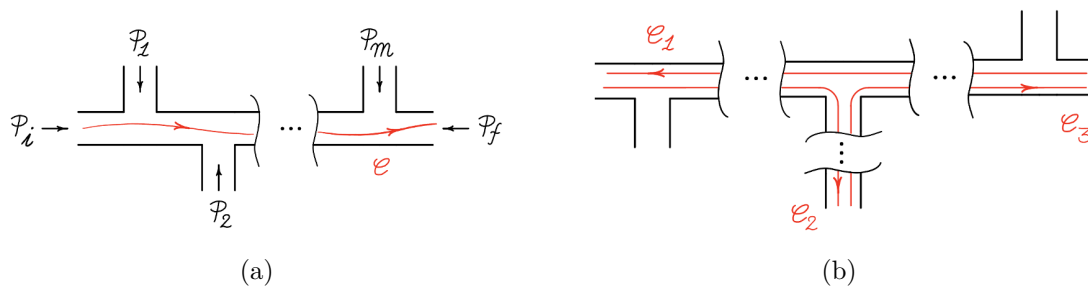


Figure 15. (a) Reversing a curve reverses its momentum assignment. (b) The momenta of three curves that cut out a cubic vertex satisfy momentum conservation.

p_e^μ for every edge. If Γ has loops (i.e. $E > n - 3$), then there is a freedom in the definition of the p_e^μ that we parametrise by some L loop momentum variables, $\ell_1^\mu, \dots, \ell_L^\mu$. This is the standard rule for assigning momenta to a fatgraph, Γ .

To go further, we now introduce a way to also assign a momentum to every *curve* on Γ . For a curve with an orientation, \vec{C} , we will assign a momentum $P_{\vec{C}}^\mu$. This momentum assignment should satisfy two basic rules. If \overleftarrow{C} is the curve with reversed orientation (figure 15(a)), then

$$P_{\overleftarrow{C}}^\mu = -P_{\vec{C}}^\mu. \quad (3.9)$$

And if three curves, $\vec{C}_1, \vec{C}_2, \vec{C}_3$, cut out a cubic vertex (figure 15(b)), then we impose momentum conservation at that vertex:

$$P_{\vec{C}_1}^\mu + P_{\vec{C}_2}^\mu + P_{\vec{C}_3}^\mu = 0. \quad (3.10)$$

The solution to satisfying both (3.9) and (3.10) is very simple, if we start with the momenta p_e^μ assigned to the edges of Γ . Suppose \vec{C} enters Γ via the external line i . Then assign this curve

$$P_{\vec{C}}^\mu = p_i^\mu + \sum_{\text{right turns}} p_{\text{left}}^\mu, \quad (3.11)$$

where p_{left}^μ is the momentum of the edge incident on C from the left, at the vertex where \vec{C} makes a right turn. The momentum assignment, (3.11), can easily be checked to satisfy (3.9) and (3.10).

For example, take the fatgraph in figure 16. The assignment of momenta to the edges of the graph is shown in the figure. The curve C_0 in figure 16 enters the graph with momentum k^μ . Then it turns left, traverses an edge, and then turns right. At the right turn, the momentum incident on the curve from the left is $-k - \ell^\mu$. So the momentum assignment of this curve is

$$P_{\vec{C}_0}^\mu = -\ell^\mu. \quad (3.12)$$

The curve C_1 in figure 16 has two right turns. At its first right turn, it gains momentum k^μ . At its second right turn, it gains momentum $-k^\mu - \ell^\mu$. So the momentum assignment of this curve is

$$P_{\vec{C}_1}^\mu = k^\mu - \ell^\mu. \quad (3.13)$$

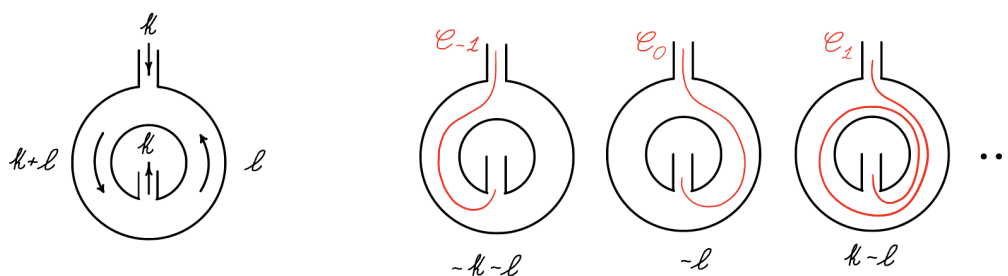


Figure 16. An assignment of momenta to the edges of a fatgraph (left) induces an assignment of momenta to curves on the fatgraph (right).

For *any* triangulation, T , the above rules assign a momentum to every curve in the triangulation. By construction, these momenta satisfy momentum conservation at each of the cubic vertices cut out by T . The upshot of this is that we can *re-use* the same loop momentum variables, ℓ_1, \dots, ℓ_L , when assigning momenta to *any* triangulation of Γ . This simple idea makes it possible to do the loop integrations for all diagrams at once, instead of one Feynman diagram at a time, which is a key step towards our formulas for amplitudes. This idea also makes it possible to compute well-defined *loop integrands*, even beyond the planar limit.

3.3.1 Aside on homology

There is a more formal way to understand the assignment of momenta to curves: these momentum assignments are an avatar of the homology of the fatgraph. Let $H_1(\Gamma, \Gamma_\infty)$ be the homology of Γ (regarded as a surface) relative to the *ends* of the external edges of the fatgraph, Γ_∞ . An oriented curve \vec{C} represents a class $[\vec{C}] \in H_1(\Gamma, \Gamma_\infty)$, and

$$[\vec{C}] + [\overleftarrow{C}] = 0 \tag{3.14}$$

in homology. Moreover, if three curves cut out a cubic vertex, their classes satisfy

$$[\vec{C}_1] + [\vec{C}_2] + [\vec{C}_3] = 0 \tag{3.15}$$

in homology. This means that a momentum assignment to curves satisfying (3.9) and (3.10) defines a linear map

$$P : H_1(\Gamma, \Gamma_\infty) \rightarrow \mathbb{R}^{1, D-1}, \tag{3.16}$$

from $H_1(\Gamma, \Gamma_\infty)$ to Minkowski space.

3.4 Spirals

The colour factor C_Γ is a product of trace factors $\text{tr}(t_1 \dots t_k)$ formed from the colour polarisations t_i^J . If Γ has a closed colour loop, this boundary contributes $\text{tr}(1) = N$ to the colour factor. For such a fatgraph, there are curves that infinitely spiral around this closed loop. These spiral curves can be treated just the same as all the other curves. In fact, the momentum assignment for spiral curves follows again from the same rule above, (3.11).

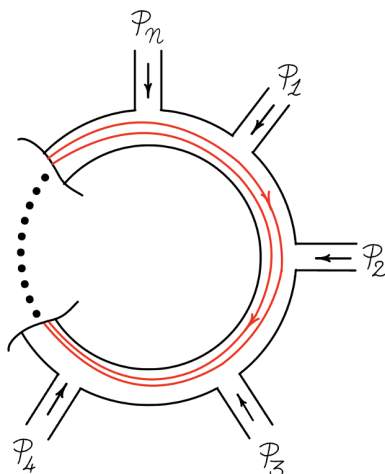


Figure 17. The momenta incident on a closed loop in a fatgraph sum to zero. This ensures that the assignment of momentum to a spiral curve is well defined.

Suppose that Γ has a closed colour loop, β . Suppose that there are some $m \geq 1$ edges incident on the loop, as in figure 17. By momentum conservation, the momenta of these edges, p_1, \dots, p_m , must sum up to zero: $\sum_{i=1}^m p_i = 0$. This ensures that (3.11) assigns a well-defined momentum to a curve that spirals around this boundary, because the contributions from the p_i^μ vanish after every complete revolution.

4 The Feynman fan

For a fatgraph Γ with E edges (e_1, \dots, e_E) , consider the E -dimensional vector space, V , generated by some vectors, $\mathbf{e}_1, \dots, \mathbf{e}_E$. To every curve C on the fatgraph, we can assign a g -vector, $\mathbf{g}_C \in V$. These simple integer vectors contain all the key information about the curves on Γ . Moreover, the g -vectors define a *fan* in V that we can use to rediscover the Feynman diagram expansion for the amplitude.

To define the g -vector of a curve, C , consider the *peaks* and *valleys* of its mountainscape. C has a *peak* at e_i if it contains

$$\dots L e_i R \dots \tag{4.1}$$

C has a *valley* at i if it contains

$$\dots R e_i L \dots \tag{4.2}$$

Let a_C^i be the number of times that C has a peak at e_i , and let b_C^i be the number of times that C has a valley at e_i . This information about the peaks and valleys is recorded by the g -vector of C ,

$$\mathbf{g}_C \equiv \sum_{i=1}^E g_C^i \mathbf{e}_i, \quad \text{where } g_C^i = a_C^i - b_C^i. \tag{4.3}$$

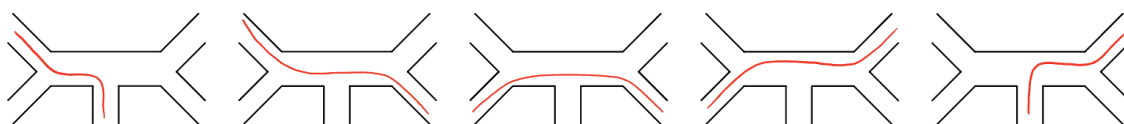


Figure 18. The five curves on the $n = 5$ tree fatgraph.

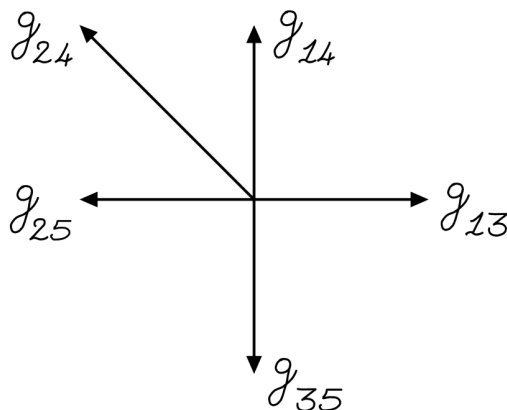


Figure 19. The Feynman fan for $n = 5$ tree level.

Each curve has a distinct g -vector. The converse is even more surprising: a curve is completely specified by its g -vector.

For example, consider the curve, C_i , in the triangulation T_Γ , which has only one peak, at e_i . The g -vector for C_i is then

$$\mathbf{g}_{C_i} = \mathbf{e}_i. \tag{4.4}$$

So the g -vectors of this triangulation T_Γ span the positive orthant of V .

4.1 Example: tree level at 5-points

Take the comb graph Γ , with edges labelled by variables x and y , as in figure 18. The five curves on Γ are

$$C_{13} = 1LxR3, \quad C_{14} = 1LxLyR4, \quad C_{24} = 2RxLyR4, \tag{4.5}$$

$$C_{25} = 2RxLyL5, \quad C_{35} = 3RyL5. \tag{4.6}$$

Counting the peaks and valleys of these mountainscapes gives

$$\mathbf{g}_{13} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \mathbf{g}_{14} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad \mathbf{g}_{24} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}, \quad \mathbf{g}_{25} = \begin{bmatrix} -1 \\ 0 \end{bmatrix}, \quad \mathbf{g}_{35} = \begin{bmatrix} 0 \\ -1 \end{bmatrix}. \tag{4.7}$$

These g -vectors are shown in figure 19. They define a *fan* in the 2-dimensional vector space. The top-dimensional cones of this fan are spanned by pairs of g -vectors, such as \mathbf{g}_{14} and \mathbf{g}_{24} , whose corresponding curves define triangulations.

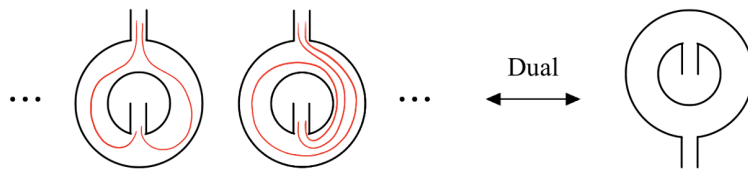


Figure 20. Triangulations (left) that are related to each other by the action of the MCG. These triangulations are all dual to the same Feynman diagram (right).

4.2 The fan

The g -vectors of all the curves on Γ together define an integer fan $\mathfrak{F} \subset V$. To define a fan, we must specify all of its *cones*. We adopt the rule that two or more g -vectors span a cone in \mathfrak{F} if and only if their curves do not intersect. The main properties of \mathfrak{F} are:

1. It is a polyhedral fan that is dense V .⁶
2. Its top dimensional cones are in 1:1 correspondence with triangulations.
3. The g -vectors of each top-dimensional cone span a parallelepiped of unit volume.

Since the top-dimensional cones of \mathfrak{F} correspond to triangulations, and hence to Feynman diagrams, we call \mathfrak{F} the *Feynman fan*, or sometimes, the *g -vector fan*.

The property that \mathfrak{F} is *polyhedral and dense* means that every rational vector $\mathbf{g} \in V$ is contained in *some* cone in the fan. This implies that every such \mathbf{g} can be *uniquely* written as a positive linear combination of g -vectors. In section 5, we solve the problem of how to do this expansion explicitly.

4.3 The mapping class group

The Feynman fan of a fat graph Γ inherits from Γ an action of a discrete, finitely generated group called the *mapping class group*, MCG. The MCG of a fatgraph, Γ , is the group of homeomorphisms of Γ , up to isotopy, that restrict to the identity on its boundaries. The action of MCG on the fatgraph can be studied by considering its action on curves. Since we only ever consider curves up to homotopy, a group element $\gamma \in \text{MCG}$ induces a map on curves

$$\gamma : C \mapsto \gamma C. \quad (4.8)$$

Since MCG acts via homeomorphisms, it does not affect curve intersections and non-intersections. If C and C' are two non-intersecting curves, then γC and $\gamma C'$ are likewise non-intersecting. Similarly, if C, C' intersect, so do γC and $\gamma C'$. This means that if some curves, C_1, \dots, C_E , form a triangulation, so do their images under MCG. Moreover, if the triangulation $\{C_1, \dots, C_E\}$ is dual to a fatgraph Γ' , then each image $\{\gamma C_1, \dots, \gamma C_E\}$ is *also* dual to the same fatgraph, Γ' .

For example, take the 2-point non-planar fatgraph Γ in figure 20. The MCG acts on Γ by *Dehn twists* that increase the number of times a curve winds around the fatgraph.

⁶A fan is *polyhedral* if the intersection of any two cones is itself, if nonempty, a cone in the fan, and the faces of each cone are cones in the fan. A fan is *dense* if any integer vector is contained in some cone of the fan. In general, irrational vectors are not always contained in our fans, but this will not play any role in this paper.

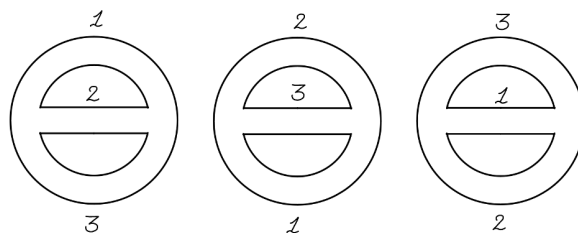


Figure 21. A fatgraph with $|\text{Aut}(\Gamma)| = 3$. Cyclic permutations of the edges leave it unchanged.

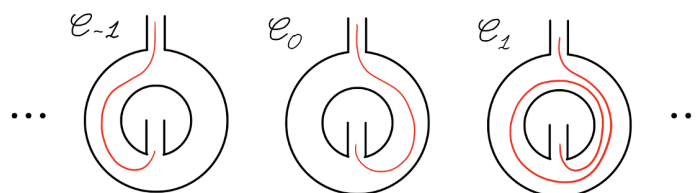


Figure 22. The infinite family of curves, C_n , for the non-planar one loop propagator.

All triangulations of Γ are related to each other by the MCG and they are all dual to the same fatgraph (right in figure 20).

In general, if Γ has loop number L , then MCG has a presentation with L generators [15]. These can be identified with Dehn twists around annuli in the fatgraph.

The MCG action on curves induces a piecewise linear action on the vector space, V ,

$$\gamma : \mathfrak{g}_C \mapsto \mathfrak{g}_{\gamma C}. \tag{4.9}$$

It follows from the above properties of the MCG action on curves that the action of MCG on V leaves the fan \mathfrak{F} invariant (if we forget the labels of the rays). Furthermore, two top-dimensional cones of the fan correspond to the same Feynman diagram if and only if they are related by the MCG action.

4.3.1 Aside on automorphisms

There is another discrete group that acts on the Feynman fan: the group of graph automorphisms, $\text{Aut}(\Gamma)$. The elements of $\text{Aut}(\Gamma)$ are permutations of the labels of the edges of Γ . A permutation is an *automorphism* if it leaves the list of fat vertices of Γ unchanged (including the vertex orientations). Each fat vertex can be described by a triple of edge labels with a cyclic orientation, (ijk) .

$\text{Aut}(\Gamma)$ has a linear action on V given by permuting the basis vectors $\mathbf{e}_1, \dots, \mathbf{e}_E$. The action of $\text{Aut}(\Gamma)$ leaves the fan invariant (again if we forget the labels of the rays).

An example of a fatgraph with nontrivial automorphisms is figure 21. In this example, cyclic permutations of the 3 edges preserve the fatgraph. Most fatgraphs that we will consider have trivial automorphism groups, and so the action of $\text{Aut}(\Gamma)$ will not play a big role in this paper.

4.4 Example: the non-planar 1-loop propagator

Take the 1-loop fatgraph Γ in figure 22, with edges labeled by variables x and y . Some of the curves on Γ , C_n , are shown in the figure. These curves are related to each other by the action of MCG, which is generated by a Dehn twist, γ . With the labelling in figure 22, the action of γ is

$$\gamma : C_n \mapsto C_{n+1}. \tag{4.10}$$

There are infinitely many such curves on the fatgraph.

The paths of the curves on Γ are

$$C_n = 1L(xLyR)^n xR2 \quad \text{for } n \geq 0, \tag{4.11}$$

$$C_n = 1Ry(RxLy)^{1+n} L2 \quad \text{for } n < 0, \tag{4.12}$$

$$\Delta = xLyR, \tag{4.13}$$

where Δ is the closed loop. Note that the curves C_n differ from one another by multiples of the closed path Δ . In this way, we can see the MCG directly in terms of the mountainscapes of the curves.

Counting peaks and valleys in the mountainscapes, the g -vectors for these curves are:

$$\mathbf{g}_n = \begin{bmatrix} -n + 1 \\ n \end{bmatrix} \quad \text{for } n \geq 0, \tag{4.14}$$

$$\mathbf{g}_n = \begin{bmatrix} n + 1 \\ -n - 2 \end{bmatrix} \quad \text{for } n < 0, \tag{4.15}$$

$$\mathbf{g}_\Delta = \begin{bmatrix} -1 \\ 1 \end{bmatrix}. \tag{4.16}$$

These g -vectors define the fan in figure 19. There are infinitely many rays of this fan. The action of MCG on curves lifts to a piecewise linear action on the fan, generated by the action of the Dehn twist γ . γ acts on the fan as

$$\mathbf{g}_{n+1} = \mathbf{g}_n + \mathbf{g}_\Delta, \quad \text{for } n \geq 0, \tag{4.17}$$

$$\mathbf{g}_0 = \mathbf{g}_{-1} + (1, 1), \tag{4.18}$$

$$\mathbf{g}_{n+1} = \mathbf{g}_n - \mathbf{g}_\Delta, \quad \text{for } n < -1. \tag{4.19}$$

This is (trivially) an isomorphism of the fan.

4.5 The delta plane

A *closed curve* is a curve Γ that forms a loop. For a closed curve Δ , consider the series of left and right turns that it makes. We can record this series of turns as a *cyclic word*, like $W_\Delta = (RRLRL)$. Whenever RL appears in W_Δ it corresponds to a *valley* in the mountainscape, which happens where the curve switches from turning right to turning left. Likewise, LR corresponds to a *peak*. If the cyclic word W_C has n occurrences of ‘ RL ’, it must also have exactly n occurrences of ‘ LR ’. For example, the cyclic word

$$(RRLRLLLLRRL), \tag{4.20}$$

switches from right-to-left 3 times, and from left-to-right 3 times.

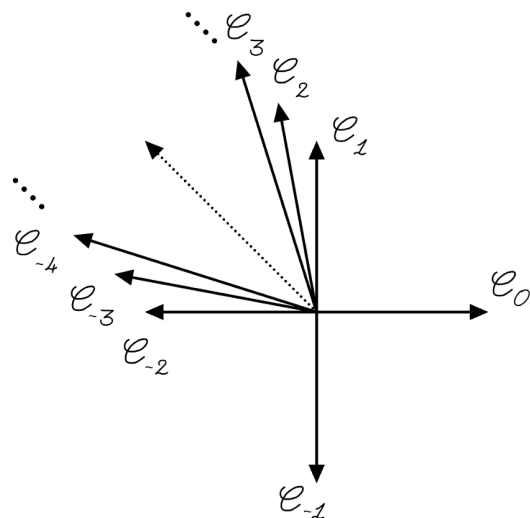


Figure 23. The Feynman fan for the non-planar 1-loop propagator.

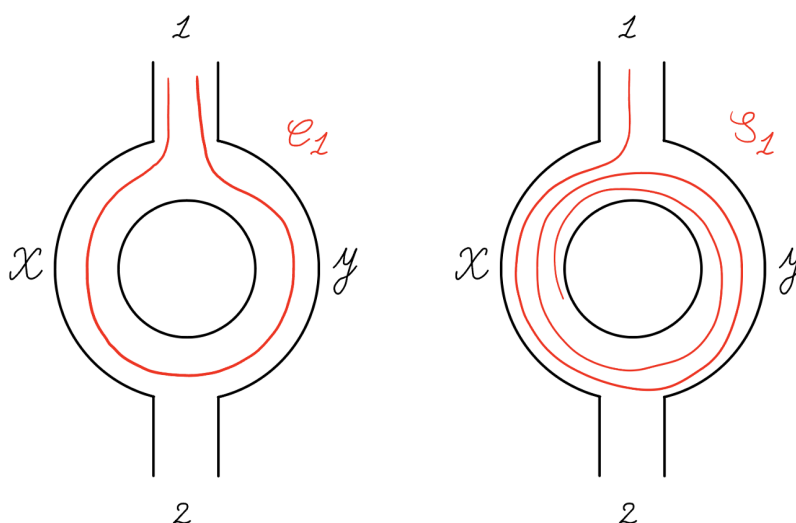


Figure 24. Curves on the bubble fatgraph.

In other words, the mountainscape for a closed curve has exactly as many peaks as valleys. It follows that the g -vector, \mathbf{g}_Δ , for any closed curve Δ is normal to the vector $\mathbf{n} = (1, 1, 1, \dots, 1)^T$. We call the plane normal to \mathbf{n} the Δ plane: $V_\Delta \subset V$.

For example, in the previous subsection, the closed curve Δ had g -vector $\mathbf{g}_\Delta = (-1, 1)$, which is normal to the vector $(1, 1)$.

Finally, note that a closed curve that makes *only* right-turns (resp. left-turns) corresponds to a path around a loop boundary of Γ . These curves have no peaks and no valleys. So these loop boundaries are assigned zero g -vector. They are also assigned zero momentum (by the reasoning in section 3.4).

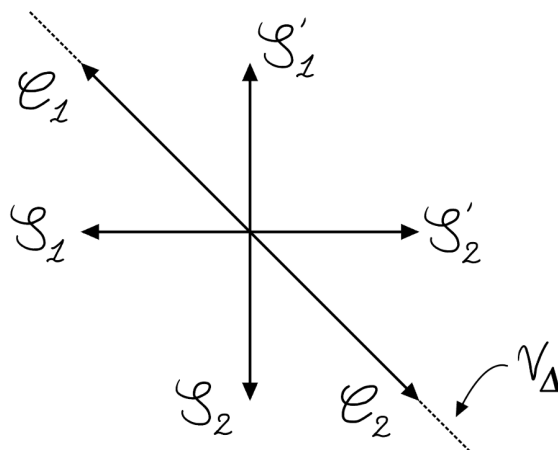


Figure 25. The Feynman Fan for the 1-loop planar propagator.

4.6 Example: the planar 1-loop propagator

Take the 1-loop bubble diagram, Γ , with edges x and y , and external edges 1 and 2, as in figure 24. Consider the curves, C_1 and S_1 that begin at edge 1 (shown in the figure), and the corresponding curves C_2 and S_2 that begin at edge 2. These have paths

$$C_1 = 1RxLyR1 \tag{4.21}$$

$$C_2 = 2RyLxR2 \tag{4.22}$$

$$S_1 = 1RxLyLxLyL \dots \tag{4.23}$$

$$S_2 = 2RyLxLyLxL \dots \tag{4.24}$$

The curves S_1, S_2 end in anticlockwise spirals around the closed loop boundary. There are also two curves, S'_1 and S'_2 , which spiral *clockwise* into the puncture:

$$S'_1 = 1LyRxRyR \dots \tag{4.25}$$

$$S'_2 = 2LxRyRxR \dots \tag{4.26}$$

Counting peaks and valleys, the g -vectors of these curves are

$$\mathbf{g}_{C_1} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}, \mathbf{g}_{S'_1} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \mathbf{g}_{S'_2} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \mathbf{g}_{C_2} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \mathbf{g}_{S_1} = \begin{bmatrix} -1 \\ 0 \end{bmatrix}, \mathbf{g}_{S_2} = \begin{bmatrix} 0 \\ -1 \end{bmatrix}. \tag{4.27}$$

These g -vectors give the fan in figure 25. Notice that the g -vectors of the curves C_1, C_2 lie on the Delta plane: $x + y = 0$.

Including the anticlockwise spirals would lead to us counting every Feynman diagram twice. This is because the triangulation with C_1, S_1 is dual to the same diagram as the triangulation by C_1, S'_1 , and so on. To prevent overcounting, it makes sense to restrict to the part of the fan that involves only C_1, S_1, S_2 , and C_2 . This part of the fan is precisely the half space, $x + y \leq 0$, cut out by the Delta plane.

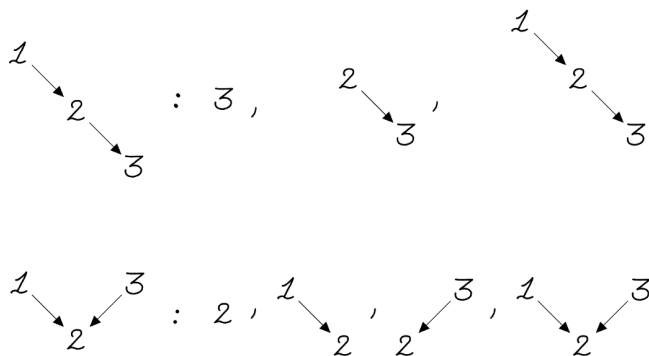


Figure 26. Two examples of mountainscapes and their sub-mountainscapes.

5 A counting problem for curves

There is a natural counting problem associated to mountainscapes, and this counting problem plays the central role in our amplitude computations.

For a mountainscape, C , the idea is to form subsets of C by *filling up* the mountainscape from the bottom. A subset is valid if it includes everything *downhill* of itself in the mountainscape.

For example, consider the curve in figure 26,

$$C = 1R2L3. \tag{5.1}$$

The valid subsets of C , shown in the figure, are $2, 1R2, 2L3$, and $1R2L3$. In other words, if 3 is in the subset, then 2 must also be included, because it is downhill of (left of) 3. Likewise, if 1 is in the subset, then 2 must also be included, because 2 is downhill of (right of) 3.

This information can be summarised using a generating function or *F-polynomial*. Introduce variables y_i , $i = 1, \dots, E$, labelled by the edges of Γ . Then the F -polynomial of a curve C is

$$F_C = 1 + \sum_{C' \subset C} \prod_{i \in C'} y_i, \tag{5.2}$$

where the sum is over all valid (non-empty) subsets of C , including C itself.

In the example, (5.1), we have four valid subsets, and the F -polynomial is

$$F_C = 1 + y_2 + y_1 y_2 + y_2 y_3 + y_1 y_2 y_3. \tag{5.3}$$

5.1 Curve matrices

Consider a curve C that starts at any edge e_i and ends at any edge e_j . It is natural to decompose its F -polynomial as a sum of four terms,

$$F_C = F_{--} + F_{-+} + F_{+-} + F_{++}, \tag{5.4}$$

where: F_{--} counts subsets that exclude the first and last edges; F_{-+} counts subsets that exclude the first edge and include the last edge; and so on.

Now consider what happens if we *extend* C along one extra edge. Let C' extend C by adding a left turn before i :

$$C' = e_k LC, \tag{5.5}$$

for some edge e_k . The F -polynomial of C' can be deduced using (5.4). Terms that involve y_i *must* contain y_k , since e_k is *downhill* of e_i in the curve. So

$$F_{C'} = (1 + y_k)F_{--} + (1 + y_k)F_{-+} + y_k F_{+-} + y_k F_{++}. \tag{5.6}$$

Similarly, if C'' is obtained from C by adding a right turn before e_i , then $C'' = e_l RC$, for some edge e_l , and we find that the new F -polynomial is

$$F_{C''} = F_{--} + F_{-+} + (1 + y_l)F_{+-} + (1 + y_l)F_{++}. \tag{5.7}$$

This equation follows because any term not containing y_i *cannot* contain y_l , since e_i is *downhill* of e_l in the curve.

Equations (5.6) and (5.7) can be used to compute the F -polynomial for any curve. It simple to do implement this is by defining a *curve matrix*, whose entries are given by the decomposition, (5.4):

$$M_C = \begin{bmatrix} F_{--} & F_{-+} \\ F_{+-} & F_{++} \end{bmatrix}. \tag{5.8}$$

The curve matrix $M_{C'}$ is obtained from the curve matrix M_C via the matrix version of (5.6):

$$M_{C'} = \begin{bmatrix} 1 & 0 \\ y_k & y_k \end{bmatrix} M_C. \tag{5.9}$$

The matrix multiplying M_C in this equation represents what happens when C is extended by adding a left turn at the start. Similarly, the matrix version of (5.7) is

$$M_{C''} = \begin{bmatrix} 1 & 1 \\ 0 & y_l \end{bmatrix} M_C, \tag{5.10}$$

which represents what happens when C is adding a right turn at the start.

It can be convenient to decompose the new matrices appearing in (5.9) and (5.10) as a product,

$$\begin{bmatrix} 1 & 0 \\ y_k & y_k \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & y_k \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 1 \\ 0 & y_l \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & y_l \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}. \tag{5.11}$$

Then, for any curve, C , we can compute its curve matrix, M_C , directly from the word specifying the curve. To do this, we just replace each turn and edge with the associated matrix:

$$L \rightarrow \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}, \quad R \rightarrow \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \quad e_i \rightarrow \begin{bmatrix} 1 & 0 \\ 0 & y_i \end{bmatrix}. \tag{5.12}$$

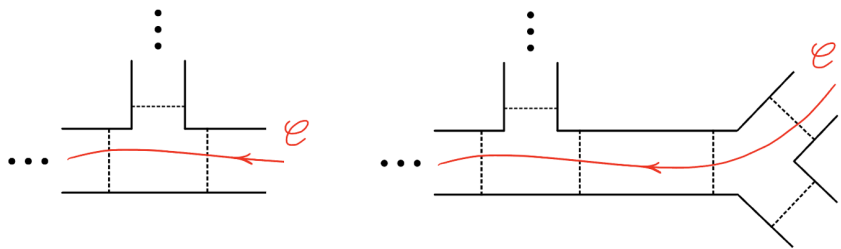


Figure 27. Getting a new fatgraph.

Every curve matrix M_C is then a product of these simple matrices.

For example, for the curve $C = 1R2L3$ considered above, its matrix is

$$M_C = \begin{bmatrix} 1 & 0 \\ 0 & y_1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & y_2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & y_3 \end{bmatrix} = \begin{bmatrix} 1 + y_2 & y_2 y_3 \\ y_1 y_2 & y_1 y_2 y_3 \end{bmatrix}. \tag{5.13}$$

The sum of the entries of this curve matrix recovers the curve’s F -polynomial, (5.3).

Curve matrices neatly factorise. If several curves all begin with the same word, W , their words can be written as $C_i = WC'_i$. Their matrices are then $M_{C_i} = M_W M_{C'_i}$, so that we only have to compute M_W once to determine all the M_{C_i} . Moreover, if we add extra legs to a fatgraph Γ , to form a larger fatgraph, Γ' , the matrices M_C for the larger fatgraph can be obtained directly from the matrices for the smaller fatgraph. In practice, this is very useful, and allows us to exploit the methods in this paper to define and study integral formulas for amplitudes at arbitrary multiplicity. [27]

5.2 Headlight functions

It follows from the definition of M_C , as a product of the matrices in (5.12), that

$$\det M_C = \prod_{e \in C} y_e. \tag{5.14}$$

Expanding the determinant, this gives

$$1 = \frac{F_{-+} F_{+-}}{F_{--} F_{++}} + \frac{\prod y_e}{F_{--} F_{++}}. \tag{5.15}$$

Motivated in part by this identity, define the u -variable of a curve C as the ratio

$$u_C = \frac{F_{-+} F_{+-}}{F_{--} F_{++}}. \tag{5.16}$$

These u -variables vastly generalise those studied in [25, 26], and it turns out that the identity (5.15) implies the surface u equations mentioned in the introduction (equation (1.1)).⁷

The *headlight function* of a curve C is the *tropicalization* of the u -variable,

$$\alpha_C = -\text{Trop } u_C. \tag{5.17}$$

⁷For our purposes in the present paper, we do not need this result, and we do not give a proof here.

For a polynomial $F(y)$, its tropicalization captures the behaviour of F at large values of y_i . Parametrise the y_i as $y_i = \exp t_i$. Then, in the large t limit,

$$F(y) \rightarrow \exp \text{Trop } F(t). \tag{5.18}$$

For example, if $F(y) = 1 + y_1 + y_1 y_2$, then $\text{Trop } F(t) = \max(0, t_1, t_1 + t_2)$. In practice, $\text{Trop } F$ is obtained from F by replacing multiplication with addition, and replacing sums with taking the maximum.

In terms of the matrix M_C , the headlight function is

$$\alpha_C = \text{Trop } M_C^{1,1} + \text{Trop } M_C^{2,2} - \text{Trop } M_C^{1,2} - \text{Trop } M_C^{2,1}. \tag{5.19}$$

Headlight functions satisfy the following remarkable property:

$$\alpha_C(\mathbf{g}_D) = \begin{cases} 1 & \text{if } C = D \\ 0 & \text{otherwise.} \end{cases} \tag{5.20}$$

This implies that headlight functions can be used to express any vector $\mathbf{g} \in V$ as a positive linear combination of the generators of a cone of the Feynman fan, by writing

$$\mathbf{g} = \sum_C \alpha_C(\mathbf{g}) \mathbf{g}_C. \tag{5.21}$$

This expansion has a geometrical interpretation. Any integer vector $\mathbf{g} \in V$ corresponds to some curve (or set of curves), L , possibly with self-intersections. Any intersections in L can be uncrossed on Γ using the *sklein relations*. Repeatedly applying skein relations, L can be decomposed on the surface into a unique set of non-self-intersecting curves, and $\alpha_C(g)$ is the number of times the curve C appears in this decomposition.

5.3 Example: tree level at 5-points

The curves for the 5-points tree level amplitude were given in section 4.1. Their curve matrices, using the replacements (5.12), are

$$C_{13} = LxR \quad \longrightarrow \quad M_{13} = \begin{bmatrix} 1 & 1 \\ 1 & 1+x \end{bmatrix}, \tag{5.22}$$

$$C_{14} = LxLyR \quad \longrightarrow \quad M_{14} = \begin{bmatrix} 1 & 1 \\ 1+x & 1+x+xy \end{bmatrix}, \tag{5.23}$$

$$C_{24} = RxLyR \quad \longrightarrow \quad M_{24} = \begin{bmatrix} 1+x & 1+x+xy \\ x & x(1+y) \end{bmatrix}, \tag{5.24}$$

$$C_{25} = RxLyL \quad \longrightarrow \quad M_{25} = \begin{bmatrix} 1+x+xy & xy \\ x+xy & xy \end{bmatrix}, \tag{5.25}$$

$$C_{35} = RyL \quad \longrightarrow \quad M_{35} = \begin{bmatrix} 1+y & y \\ y & y \end{bmatrix}. \tag{5.26}$$

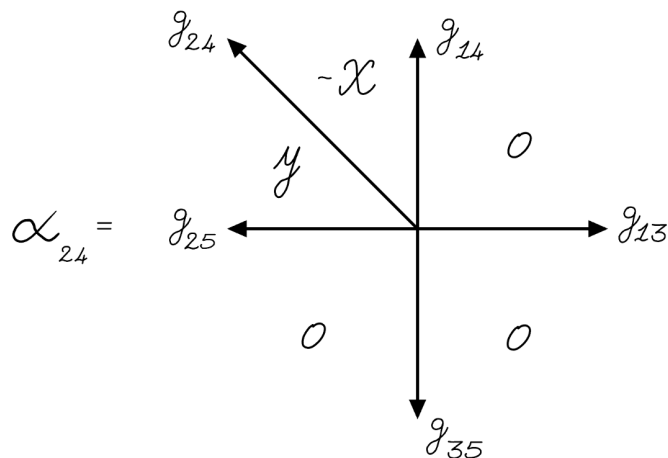


Figure 28. The Schwinger parameter α_{24} on the Feynman fan.

Given these matrices, the headlight functions are

$$\alpha_{13} = \max(0, x), \tag{5.27}$$

$$\alpha_{14} = -\max(0, x) + \max(0, x, x + y), \tag{5.28}$$

$$\alpha_{24} = -\max(0, x, x + y) + \max(0, x) + \max(0, y), \tag{5.29}$$

$$\alpha_{25} = -x - \max(0, y) + \max(0, x, x + y), \tag{5.30}$$

$$\alpha_{35} = -y + \max(0, y). \tag{5.31}$$

It can be verified that $\alpha_{ij}(\mathbf{g}_C) = 1$ if $C = C_{ij}$, and that otherwise $\alpha_{ij}(\mathbf{g}_C) = 0$. For example, the values taken by α_{24} are shown in figure 28.

5.4 Example: the non-planar 1-loop propagator

The mountainscapes for the non-planar 1-loop propagator are given in section 4.4. Using these, we can compute the headlight functions, and find:

$$\alpha_n = f_n - 2f_{n-1} + f_{n-2}, \quad n \geq 0, \tag{5.32}$$

$$\alpha_n = g_n - 2g_{n+1} + g_{n+2}, \quad n < 0. \tag{5.33}$$

where the tropical functions f_n and g_n are given by

$$f_n = \max(0, (n + 1)x, (n + 1)x + ny), \quad \text{for } n \geq 0, \tag{5.34}$$

$$g_n = \max(0, -(n + 1)x, -(n + 1)x - ny), \quad \text{for } n \leq -1, \tag{5.35}$$

with the following special cases:

$$f_{-2} = 0, \quad f_{-1} = 0, \quad g_1 = -2x - y, \quad g_0 = -x. \tag{5.36}$$

A full derivation of these functions using the matrix method is given in appendix F.

It is easy to verify that these α_n satisfy the key property:

$$\alpha_n(\mathbf{g}_m) = \begin{cases} 1 & \text{if } n = m \\ 0 & \text{otherwise.} \end{cases} \tag{5.37}$$

For example, take $n, m \geq 0$. Then we find

$$f_n(\mathbf{g}_m) = \max(0, 1 + n - m), \tag{5.38}$$

so that

$$\alpha_n(\mathbf{g}_m) = \max(0, 1 + n - m) + \max(0, -1 + n - m) - 2 \max(0, n - m). \tag{5.39}$$

This agrees with (5.37).

5.5 Spirals

Suppose C is a curve that ends in a spiral around a loop boundary of Γ . If $1, 2, \dots, m$ are the edges around that boundary, C has the form

$$C = W1L2L\dots LmL1L2L\dots, \tag{5.40}$$

for some subpath W . We can compute the transfer matrix for the infinite tail at the right end of C . The path for one loop around the boundary is

$$\Delta := 1L2L\dots LmL, \tag{5.41}$$

and the matrix for this path is

$$M_\Delta = \begin{bmatrix} 1 & 0 \\ F_* & y_* \end{bmatrix}, \tag{5.42}$$

where

$$y_* = \prod_{i=1}^m y_i, \quad \text{and} \quad F_* = y_1 + y_1y_2 + \dots + y_1y_2\dots y_m. \tag{5.43}$$

Now consider the powers, M_Δ^n ,

$$M_\Delta^n = \begin{bmatrix} 1 & 0 \\ F_*(1 + y_* + \dots + y_*^{n-1}) & y_*^n \end{bmatrix}. \tag{5.44}$$

If the path W has matrix

$$M_W = \begin{bmatrix} a & b \\ c & d \end{bmatrix}, \tag{5.45}$$

then the path $W\Delta^n$ has u -variable

$$u_{W\Delta^n} = \frac{b(c + dF_*(1 + y_* + \dots + y_*^{n-1}))}{d(a + bF_*(1 + y_* + \dots + y_*^{n-1}))} \tag{5.46}$$

and in the $n \rightarrow \infty$ limit this tends to (assuming $y_* < 1$ for convergence)

$$u_{W\Delta^\infty} = \lim_{n \rightarrow \infty} u_{W\Delta^n} = \frac{b(c(1 - y_*) + dF_*)}{d(a(1 - y_*) + bF_*)}, \tag{5.47}$$

which is the u -variable for the full spiraling curve. Instead of computing this limit, it is convenient to instead compute this u -variable by multiplying the matrix M_W with the matrix

$$\widetilde{M}_\Delta = \begin{bmatrix} 1 - y_* & 0 \\ F_* & 1 \end{bmatrix}. \tag{5.48}$$

Note that \widetilde{M}_Δ is *not* the $n \rightarrow \infty$ limit of M_Δ . We can use this matrix, (5.48), when computing the matrix for any curve that ends in a spiral: the spiraling part can be replaced by \widetilde{M}_Δ directly. Similarly, suppose a curve *begins* with a spiral around the closed path $\Delta = RmR \cdots R2R1$. Then this infinite spiral contributes a factor of \widetilde{M}_Δ^T to the beginning of the matrix product, with y_* and F_* given as before.

5.6 Example: the planar 1-loop propagator

We can put these formulas to work for the planar 1-loop propagator. The curves for this amplitude are given in section 4.6. Evaluating the curve matrices gives:

$$M_{C_1} = \begin{bmatrix} 1+x & 1+x+xy \\ x & x+xy \end{bmatrix}, \quad M_{C_2} = \begin{bmatrix} 1+y & 1+y+xy \\ y & y+xy \end{bmatrix}, \tag{5.49}$$

$$M_{S_1} = \begin{bmatrix} 1+x & 1 \\ x(1+y) & 1 \end{bmatrix}, \quad M_{S_2} = \begin{bmatrix} 1+y & 1 \\ y(1+x) & 1 \end{bmatrix}. \tag{5.50}$$

The headlight functions are

$$\alpha_{C_1} = \max(0, x) + \max(0, y) - \max(0, x, x+y), \tag{5.51}$$

$$\alpha_{C_2} = \max(0, x) + \max(0, y) - \max(0, y, x+y), \tag{5.52}$$

$$\alpha_{S_1} = -x - \max(0, y) + \max(0, x), \tag{5.53}$$

$$\alpha_{S_2} = -y - \max(0, x) + \max(0, y). \tag{5.54}$$

Once again, using the g -vectors from section 4.6, we verify that these functions satisfy

$$\alpha_C(\mathbf{g}_D) = \begin{cases} 1 & \text{if } C = D \\ 0 & \text{otherwise.} \end{cases} \tag{5.55}$$

5.7 Example: the genus one 2-loop vacuum

We now introduce a more complicated example: the 2-loop vacuum amplitude at genus one. A fatgraph for this amplitude, Γ , is given in figure 29. The colour factor of this graph has only one factor, $\text{tr}(1)$, because Γ only has one boundary. In fact, the curves on Γ must all begin and end in spirals around this one boundary. Using figure 29 we can identify the curves which have precisely *one valley* in their mountainscape: i.e. which only have one switch from turning right to turning left. These three curves are

$$C_{1/0} = (RwRzRx)^\infty R(wLxLzL)^\infty, \tag{5.56}$$

$$C_{0/1} = (RxRwRz)^\infty R(xLzLwL)^\infty, \tag{5.57}$$

$$C_{1/1} = (RzRxRw)^\infty R(zLwLxL)^\infty. \tag{5.58}$$

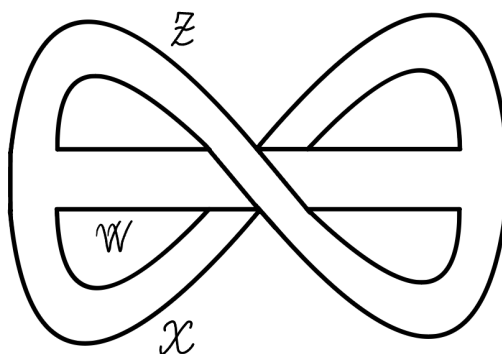


Figure 29. The 2-loop vacuum graph with genus one.

These curves are non-intersecting and form a triangulation. The surface associated to Γ is the torus with one puncture, and the labels we assign to these curves are inspired by drawing the curves on the torus, pictured as a quotient of a \mathbb{Z}^2 lattice.

Besides $C_{1/1}$, we find that the only other curve compatible with both $C_{1/0}$ and $C_{0/1}$ is

$$C_{-1/1} = (RxRwRz)^\infty RxLzR(xLzLwL)^\infty. \quad (5.59)$$

This curve has a peak at z , but no peaks at either x or w (which is what would result in an intersection with $C_{1/0}$ or $C_{0/1}$).

As we will see later, the four curves $C_{1/0}, C_{0/1}, C_{1/1}, C_{-1/1}$ are all we need to compute the 2-loop vacuum genus one amplitude. Evaluating these curves' matrices gives

$$M_{1/0} = \begin{bmatrix} w(1+x+xz)^2 + (1-xzw)^2 & 1+x+xz \\ w(1+x+xz) & 1 \end{bmatrix}, \quad (5.60)$$

$$M_{0/1} = \begin{bmatrix} x(1+z+zw)^2 + (1-xzw)^2 & 1+z+zw \\ x(1+z+zw) & 1 \end{bmatrix}, \quad (5.61)$$

$$M_{1/1} = \begin{bmatrix} z(1+w+wx)^2 + (1-xzw)^2 & 1+w+wx \\ z(1+w+wx) & 1 \end{bmatrix}, \quad (5.62)$$

$$M_{-1/1} = \begin{bmatrix} (1+x+xz)^2 + z(x+xz+xzw)^2 & 1+x+2xz+xz^2+xz^2w \\ x(1+x+2xz+xz^2+xz^2w) & x(1+z) \end{bmatrix}. \quad (5.63)$$

Because we restrict to anti-clockwise spiraling curves, all the g -vectors lie in the half-plane $x+z+w \leq 0$. Restricting to this half-plane, the headlight functions are given by

$$\alpha_{1/0} = \max(0, -w - 2 \max(0, x, x+z)), \quad (5.64)$$

$$\alpha_{0/1} = \max(0, -x - 2 \max(0, z, z+w)), \quad (5.65)$$

$$\alpha_{1/1} = \max(0, -z - 2 \max(0, w, w+x)), \quad (5.66)$$

and

$$\alpha_{-1/1} = \max(0, z) + \max(2 \max(0, x, x+z), z + 2x + 2 \max(0, z, z+w)) \quad (5.67)$$

$$- \max(0, x, x+z, x+2z, x+2z+w). \quad (5.68)$$

We again verify that these α_C are 1 on their corresponding g -vectors.

6 Integrand curve integrals

We want to compute the partial amplitudes of our theory. For some fatgraph Γ , let \mathcal{A} be the amplitude that multiplies the colour factor c_Γ .

The momentum assignment rule in section 3.3 defines one set of loop momentum variables for all propagators contributing to the amplitude, even beyond planar diagrams. This means that \mathcal{A} can be obtained as the integral of a single *loop integrand* \mathcal{I} :

$$\mathcal{A} = \int \left(\prod_{i=1}^L d^D \ell_i \right) \mathcal{I}. \tag{6.1}$$

However, beyond planar diagrams, there is a price to pay for introducing our momentum assignment. For any triangulation by curves, C_1, C_2, \dots, C_E , we associate the product of propagators

$$\frac{1}{X_{C_1} X_{C_2} \dots X_{C_E}}, \tag{6.2}$$

where X_C is given by the momentum assignment rule. If we sum over every such term, (6.2), for all triangulations of Γ , we obtain some rational function \mathcal{I}_∞ . But the loop integral of \mathcal{I}_∞ is not well defined if Γ has a nontrivial mapping class group, MCG. This is because two triangulations related by the MCG action integrate to the *same* Feynman diagram. So the loop integral of \mathcal{I}_∞ contains, in general, infinitely many copies of each Feynman integral.

Fortunately, we can compute integrands \mathcal{I} for the amplitude by ‘dividing by the volume of MCG’. As a function, \mathcal{I} is not uniquely defined. But all choices for \mathcal{I} integrate to the same amplitude.

We will compute integrands \mathcal{I} using the headlight functions, α_C . The formula takes the form of a *curve integral*,

$$\mathcal{I} = \int \frac{d^E t}{\text{MCG}} e^{-S(t)}. \tag{6.3}$$

Here, E is the number of edges of the fatgraph Γ . We call it a *curve integral* because the integral is over the E -dimensional vector space, V , whose integral points correspond to curves (or collections of curves) on Γ . As discussed in section 4.2, the mapping class group MCG has a piecewise linear action on V , and we mod out by this action in the integral. We call $S(t)$ the *curve action*. It is given by a sum

$$S(t) = \sum_C \alpha_C(t) X_C, \tag{6.4}$$

where we sum over all curves, C , on the fatgraph.⁸ For a general derivation of this curve integral formula, see appendix A. In this section, we show how to practically use (6.3) to compute some simple amplitudes.

In fact, (6.3) also makes the loop integrals easy to do. This leads to a direct curve integral formula for the amplitude \mathcal{A} , which we study in section 7.

⁸We exclude *closed curves* from this sum. Including the closed curves corresponds to coupling our colored field to an uncolored scalar particle. For simplicity, we delay the discussion of uncolored amplitudes.

Later, in section 10, we also show that the integrands \mathcal{I} can be computed recursively, starting from the curve integral formula, (6.3). This result generalises the standard *forward limit* method for 1-loop amplitudes to *all* orders in the perturbation series.

6.1 Example: the tree level 5-point amplitude

Curve integrals give new and simple amplitude formulas, even at tree level. Take the same fatgraph studied in sections 4.1, 5.3 and 6.1. The kinematic variables for the curves on this graph are ($i < j - 1$)

$$X_{ij} = (k_i + \dots + k_{j-1})^2 + m^2. \tag{6.5}$$

Then the amplitude, given by (6.1), is

$$\mathcal{A}(12345) = \int dy_1 dy_2 Z, \tag{6.6}$$

where

$$-\log Z = \alpha_{13}X_{13} + \alpha_{14}X_{14} + \alpha_{24}X_{24} + \alpha_{25}X_{25} + \alpha_{35}X_{35}. \tag{6.7}$$

Using the formulas for α_{ij} from section 5.3, Z can be further simplified to

$$\log Z = X_{25}x + X_{35}y + s_{13}f_{13} + s_{14}f_{14} + s_{24}f_{24}, \tag{6.8}$$

where $s_{ij} = 2k_i \cdot k_j$ and the f_{ij} are the simple functions

$$f_{13} = \max(0, x), \quad f_{14} = \max(0, x, x + y), \quad f_{24} = \max(0, y). \tag{6.9}$$

The 5-point amplitude is then

$$\mathcal{A}(12345) = \int dx dy \exp(X_{25}x + X_{35}y + s_{13}f_{13} + s_{14}f_{14} + s_{24}f_{24}). \tag{6.10}$$

It is already interesting to note that the formula for the amplitude has been written in terms of the simple functions $f_{13}, f_{14}, f_{24}, y_1, y_2$, and the Mandelstam invariants s_{ij} . These s_{ij} are automatically summed together by the formula to form the appropriate poles of the tree level amplitude.

6.2 Example: the planar 1-loop propagator

Consider again the 1-loop planar propagator (sections 4.6 and 5.6). The amplitude is

$$\mathcal{A} = \int d^D \ell \int_{x+y \leq 0} dx dy Z, \tag{6.11}$$

where

$$-\log Z = \alpha_{C_1}X_{C_1} + \alpha_{C_2}X_{C_2} + \alpha_{S_1}X_{S_1} + \alpha_{S_2}X_{S_2}. \tag{6.12}$$

We can assign the momenta of the curves to be

$$P_{C_1} = 0, \quad P_{S_1} = \ell, \quad P_{S_2} = \ell + k, \quad P_{C_2} = 0. \tag{6.13}$$

Substituting these momenta (with $k^2 + m^2 = 0$) into the integrand gives

$$-\log Z = \ell^2(\alpha_{S_1} + \alpha_{S_2}) + 2\ell \cdot k\alpha_{S_2} + m^2(\alpha_{C_1} + \alpha_{C_2} + \alpha_{S_1}). \quad (6.14)$$

At this point, we can either integrate over $x + y \leq 0$, or do the loop integral. Doing the loop integral first is a Gaussian integral, which gives

$$\mathcal{A} = \int_{x+y \leq 0} dx dy \left(\frac{\pi}{\alpha_{S_1} + \alpha_{S_2}} \right)^{\frac{D}{2}} \exp \left(-\frac{k^2 \alpha_{S_2}^2}{\alpha_{S_1} + \alpha_{S_2}} - m^2(\alpha_{C_1} + \alpha_{C_2} + \alpha_{S_1}) \right). \quad (6.15)$$

This resembles the Symanzik formula for a single Feynman integral, but instead includes contributions from all three Feynman diagrams for this amplitude. Finally, substituting the headlight functions gives

$$\mathcal{A} = \int_{x+y \leq 0} dx dy \left(\frac{-\pi}{x+y} \right)^{\frac{D}{2}} \exp \left[m^2 \frac{(y + \max(0, x) - \max(0, y))^2}{x+y} - m^2|x| \right]. \quad (6.16)$$

It is not immediately obvious that this reproduces the Feynman integrals for this amplitude. But note that, for example, restricting the domain of the integral to the negative orthant gives

$$\int_{x,y \leq 0} dx dy \left(\frac{-\pi}{x+y} \right)^{\frac{D}{2}} \exp \left(m^2 \left(\frac{y^2}{x+y} + x \right) \right). \quad (6.17)$$

After writing

$$\frac{y^2}{x+y} + x = -\frac{xy}{x+y} + (x+y), \quad (6.18)$$

this recovers the Feynman integral for the bubble graph. By extending the integral to the full region, $x + y \leq 0$, we recover not just this bubble integral, but the full amplitude!

6.3 Example: the planar 1-loop 3-point amplitude

For a more complicated planar example, consider the 1-loop planar 3-point amplitude, with the fatgraph Γ , in figure 30. There are nine curves on this graph: three curves $C_{i,i+2}$, connecting external lines $i, i+2$; three curves $C_{i,i}$, which loop around and come back to external line i ; and three curves $C_{i,0}$ that start from the external line i and end in a spiral around the closed loop.

In the planar sector, a convenient way to assign momenta is to use *dual variables*. Let z_i^μ ($i = 1, 2, 3$) be dual variables for the external lines, and z_0 be the dual variable for the closed loop. Then curves from external lines i to j have

$$X_{i,j} = (z_j - z_i)^2 + m^2, \quad (6.19)$$

whereas a curve from i that ends in a spiral around the loop has

$$X_{i,0} = (z_i - z_0)^2 + m^2. \quad (6.20)$$

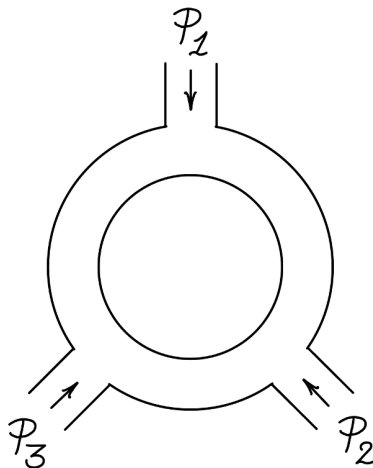


Figure 30. A fatgraph for the 3-point 1-loop planar amplitude.

If the external momenta are p_1, p_2, p_3 , then we can take $z_1 = 0, z_2 = p_1, z_3 = p_1 + p_2$. The closed loop variable, z_0 , can be used as a loop momentum variable.

The 3-point one-loop planar amplitude is then

$$\mathcal{A} = \int d^D z_0 \int_{\sum t_i \geq 0} dt Z, \tag{6.21}$$

where (taking cyclic indices mod 3)

$$-\log Z = \sum_{i=1}^3 \alpha_{i,i+2} X_{i,i+2} + \sum_{i=1}^3 \alpha_{i,i} X_{i,i} + \sum_{i=1}^3 \alpha_{i,0} X_{i,0}. \tag{6.22}$$

The headlight functions for these curves are

$$\alpha_{i,0} = t_i + g_{i+1} - g_i, \tag{6.23}$$

$$\alpha_{i,i+2} = g_i - f_i - f_{i+1}, \tag{6.24}$$

$$\alpha_{i,i} = f_{i+1} + h_i - g_i - g_{i+1}, \tag{6.25}$$

where

$$f_i = \max(0, t_i), \tag{6.26}$$

$$g_i = \max(0, t_i, t_i + t_{i+1}), \tag{6.27}$$

$$h_i = \max(0, t_i, t_i + t_{i+1}, t_i + t_{i+1} + t_{i+2}). \tag{6.28}$$

6.4 Note on factorization

The integrands defined by curve integrals factorise in the correct way. Take again the curve integral

$$\mathcal{I} = \int \frac{d^E t}{\text{MCG}} Z. \tag{6.29}$$

In appendix B, we show that the residue at $X_C = 0$ is given by

$$\text{Res}_{X_C=0} \mathcal{I} = \int \frac{d^{E-1}t}{\text{MCG}'} Z', \tag{6.30}$$

which is now the curve integral for the fatgraph Γ_C , obtained by cutting Γ along C . In this formula, MCG' is the MCG of Γ_C , and the momentum P_C^μ of the curve C is put on shell. In the fatgraph Γ_C , the curve C gives two new boundaries, which are assigned momenta $\pm P_C^\mu$.

For example, before loop integration, the non-planar 1-loop fatgraph Γ has loop integrand

$$\mathcal{I} = \int dx dy \exp\left(-\sum_{n=-\infty}^{\infty} \alpha_n X_n\right). \tag{6.31}$$

Here, the momenta of the curves are $P_n^\mu = \ell^\mu + nk^\mu$. Consider the $X_0 = 0$ pole. The parameter α_0 vanishes outside $x \geq 0$. In this region, the only non-vanishing parameters are α_1 and α_{-1} . The residue at $X_0 = 0$ is then

$$\text{Res}_{X_0=0} \mathcal{I} = \int dy \exp(-\alpha'_1 X_1 - \alpha'_{-1} X_{-1}), \tag{6.32}$$

where the restriction to $x = 0$ gives $\alpha'_1 = \max(0, y)$ and $\alpha'_{-1} = y - \max(0, y)$. This is the $n = 4$ tree level amplitude, with external momenta are $k, \ell, -k, -\ell$, and ℓ^μ . The two propagators are $X_1 = (k + \ell)^2 + m^2$ and $X_{-1} = (k - \ell)^2 + m^2$.

7 Amplitude curve integrals

Following the previous section, the curve integral formula for the full amplitude is

$$\mathcal{A} = \int \frac{d^E \mathbf{t}}{\text{MCG}} \int \left(\prod d^D \ell_a\right) \exp(-S(\mathbf{t})). \tag{7.1}$$

The loop integration variables, ℓ_a , appear quadratically in the curve action $S(\mathbf{t})$. So, if we perform the loop integral *before* performing the curve integral over the t_i , it is a Gaussian integral. The result is a curve integral

$$\mathcal{A} = \int \frac{d^E \mathbf{t}}{\text{MCG}} \left(\frac{\pi^L}{\mathcal{U}}\right)^{\frac{D}{2}} \exp\left(-\frac{\hat{\mathcal{F}}_0}{\mathcal{U}} - \mathcal{Z}\right), \tag{7.2}$$

where $\mathcal{U}, \hat{\mathcal{F}}_0$ and \mathcal{Z} are homogeneous polynomials in the α_C 's that we call *surface Symanzik polynomials*.

The curve integral (7.2) resembles the Schwinger form of a single Feynman integral, but it integrates to the full amplitude. Once again, it is important to mod out by the action of the mapping class group, to ensure that the integral does not overcount Feynman diagrams.

We now summarise how to compute the surface Symanzik polynomials, $\mathcal{U}, \mathcal{F}_0, \mathcal{Z}$. Suppose that a choice of loop momentum variables, ℓ_a^μ , has been fixed. The momentum assigned to a curve C is of the form

$$P_C^\mu = K_C^\mu + \sum h_C^a \ell_a^\mu, \tag{7.3}$$

for some integers h_C^a that follow from the momentum assignment rule (section 3.3).⁹ Using the h_C^a , define an $L \times L$ matrix

$$A^{ab} = \sum_C h_C^a h_C^b \alpha_C, \tag{7.4}$$

and a L -dimensional vector (with momentum index μ)

$$B^{a,\mu} = \sum_C h_C^a \alpha_C K_C^\mu. \tag{7.5}$$

Then, the surface Symanzik polynomials are

$$\mathcal{U} = \det A, \quad \frac{\hat{\mathcal{F}}_0}{\mathcal{U}} = -B_\mu^a (A^{-1})_{ab} B^{b,\mu}, \quad \mathcal{Z} = \sum_C \alpha_C (K_C^2 + m^2). \tag{7.6}$$

These arise in the usual way by performing the Gaussian integral, as discussed in detail in appendix C. To connect with the standard Schwinger parametrization of Feynman integrals it is also convenient to define the polynomial \mathcal{F}_0 by

$$\mathcal{F}_0 = \hat{\mathcal{F}}_0 + \mathcal{U} \left(\sum_C \alpha_C K_C^2 \right). \tag{7.7}$$

In fact, the surface Symanzik polynomials have simple expressions when expanded as a sum of monomials. For a set of curves, $\mathcal{S} = \{C_1, \dots, C_L\}$, write $\alpha_{\mathcal{S}}$ for the corresponding monomial

$$\alpha_{\mathcal{S}} = \prod_{i=1}^L \alpha_{C_i}. \tag{7.8}$$

The determinant, $\det A$, can be expanded to give

$$\mathcal{U} = \sum_{\substack{\mathcal{S} \text{ cuts } \Sigma \\ \text{to disk}}} \alpha_{\mathcal{S}}, \tag{7.9}$$

where we sum over all sets \mathcal{S} whose curves cut Γ down to a tree fatgraph. In other words, \mathcal{U} is the sum over all *maximal cuts* of the graph Γ . Moreover, using the Laplace expansion of the matrix inverse, \mathcal{F}_0 can be expanded to find

$$\mathcal{F}_0 = \sum_{\substack{\mathcal{S}' \text{ cuts } \Sigma \\ \text{to 2 disks}}} \alpha_{\mathcal{S}'} \left(\sum_{C \in \mathcal{S}'} K_C^\mu \right)^2, \tag{7.10}$$

where the sum in this formula is now over sets \mathcal{S}' of $L + 1$ curves that factorise Γ into two disjoint tree graphs. Each monomial in the sum is multiplied by the total momentum flowing through the factorisation channel. In particular, note that restricted a single cone of the Feynman fan, where only E headlight functions α_C are nonzero, \mathcal{U} and \mathcal{F}_0 become the first and second Symanzik polynomials, respectively. The formulas above are then a surface analogue of the formulas for the Symanzik polynomials in terms of spanning trees and 2-forests of a graph. Further details are collected in appendix C.

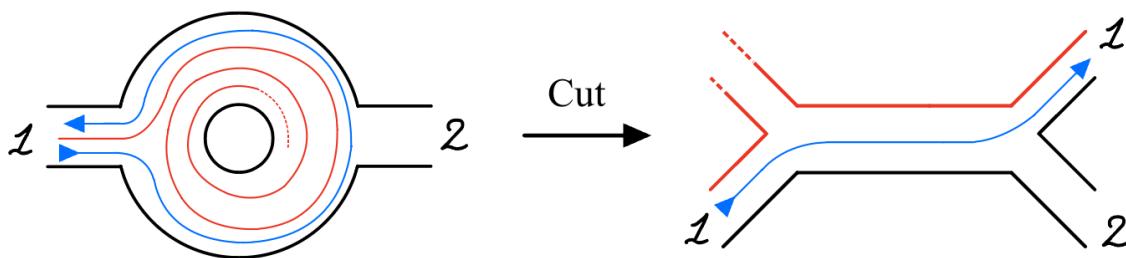


Figure 31. Cutting the annulus along the spiralling red curve produces a tree-surface.

7.1 Example: the planar 1-loop propagator

We return to the planar 1-loop propagator (sections 4.6, 5.6, 6.2). Of the four curves C_1, C_2, S_1, S_2 , only the spiralling curves, S_1 and S_2 , carry loop momentum and therefore contribute to \mathcal{U} and \mathcal{F}_0 . In particular, note that cutting along either S_1 or S_2 cuts Γ open to a 4-point tree graph (see figure 31). The first surface Symanzik polynomial is therefore

$$\mathcal{U} = \alpha_{S_1} + \alpha_{S_2}. \quad (7.11)$$

The B -vector is

$$B^\mu = \alpha_{S_2} k^\mu, \quad (7.12)$$

so that

$$\hat{\mathcal{F}}_0 = -\alpha_{S_2}^2 k^2. \quad (7.13)$$

Finally,

$$\mathcal{Z} = m^2(\alpha_{S_1} + \alpha_{S_2} + \alpha_{C_1} + \alpha_{C_2}) + \alpha_{S_2} k^2. \quad (7.14)$$

The amplitude is then given by the curve integral

$$\mathcal{A} = \int_{x+y \leq 0} dx dy \left(\frac{\pi}{\alpha_{S_1} + \alpha_{S_2}} \right)^{\frac{D}{2}} \exp \left(-\frac{\alpha_{S_2} k^2}{\alpha_{S_1} + \alpha_{S_2}} - m^2 (\alpha_{S_1} + \alpha_{C_1} + \alpha_{C_2}) \right). \quad (7.15)$$

We have used the on shell condition, $k^2 + m^2 = 0$, to simplify the expression. This again recovers the formula (6.15), which we obtained by direct integration in the previous section.

Note that we can also express the formula in terms of

$$\mathcal{F}_0 = -\alpha_{S_2}^2 k^2 + (\alpha_{S_1} + \alpha_{S_2}) \alpha_{S_2} k^2 = \alpha_{S_1} \alpha_{S_2} k^2, \quad (7.16)$$

where S_1, S_2 is precisely the one set of curves that cut the surface (resp. fatgraph) to two disks (resp. tree graphs).

⁹These h_C^c can also be understood in terms of the homology of curves on the surface. The numbers h_C^c are the coefficients of the class $[C]$ in relative homology when expanded in a basis.

7.2 Example: the non-planar 1-loop propagator

We return to the non-planar 1-loop propagator (sections 4.4 and 5.4). The momentum of the curve C_n is

$$P_n^\mu = \ell^\mu + nk^\mu. \tag{7.17}$$

Every curve C_n cuts Γ to a tree graph with 4 external legs. So the first Symanzik polynomials is

$$\mathcal{U} = \sum_{n=-\infty}^{\infty} \alpha_n, \tag{7.18}$$

where α_n is the headlight function for C_n . Moreover, we find

$$\hat{\mathcal{F}}_0 = - \sum_{n,m=-\infty}^{\infty} nm\alpha_n\alpha_mk^2, \tag{7.19}$$

and

$$\mathcal{Z} = \sum_{n=-\infty}^{\infty} \alpha_n(m^2 + n^2k^2). \tag{7.20}$$

Note also that

$$\mathcal{F}_0 = \hat{\mathcal{F}}_0 + \mathcal{U} \sum_C \alpha_C K_C^2 = \sum_{n \neq m} nm\alpha_n\alpha_mk^2. \tag{7.21}$$

This corresponds to the fact that every pair of distinct curves C_n, C_m cuts Γ into two trees.

The amplitude is then

$$\mathcal{A} = \int \frac{dxdy}{\text{MCG}} \left(\frac{\pi}{\mathcal{U}}\right)^{\frac{D}{2}} \exp\left(-\frac{\mathcal{F}_0}{\mathcal{U}} - m^2 \sum \alpha_n\right). \tag{7.22}$$

Finally, the MCG acts on the fan in this case as $\mathbf{g}_n \mapsto \mathbf{g}_{n+1}$. A fundamental domain for this action is clearly the positive orthant, spanned by $\mathbf{g}_0, \mathbf{g}_1$. In this orthant, the surface Symanzik polynomials are

$$\mathcal{U} = x + y, \tag{7.23}$$

$$\hat{\mathcal{F}}_0 = y^2k^2, \tag{7.24}$$

$$\mathcal{Z} = xm^2. \tag{7.25}$$

So we find

$$\mathcal{A} = \int_{x,y \geq 0} dxdy \left(\frac{\pi}{x+y}\right)^{D/2} \exp\left(m^2 \left(-\frac{y^2}{x+y} - x\right)\right), \tag{7.26}$$

where we have put k^μ on shell, i.e. $k^2 + m^2 = 0$. Or, equivalently,

$$\mathcal{A} = \int_{x,y \geq 0} dxdy \left(\frac{\pi}{x+y}\right)^{D/2} \exp\left(-k^2 \frac{xy}{x+y} - m^2(x+y)\right), \tag{7.27}$$

in agreement with the standard Feynman integral for the bubble.

7.3 Example: the non-planar 3-point amplitude

Even at 1-loop, it is not always easy to identify the fundamental domain of the MCG. To see the problem, consider the non-planar one-loop 3-point amplitude. Let the first trace factor have external particle p_1^μ , and the second trace factor have p_2^μ and p_3^μ . The curves, C_{ij}^n , connecting a pair of distinct start and end points, i, j , are labelled by the number of times, n , they loop around the graph. The curves C_{22} and C_{33} begin and end at the same edge, and are invariant under the MCG. Then, for a specific choice of loop momentum variable, we find the momentum assignments

$$K_{12}^n = np_1^\mu, \quad K_{13}^n = np_1^\mu - p_2^\mu, \quad K_{22} = 0, \quad K_{33} = 0. \quad (7.28)$$

We can readily give the curve integral formula for the amplitude,

$$\mathcal{A} = \int \frac{dx dy dz}{\text{MCG}} \left(\frac{\pi}{\mathcal{U}} \right)^{\frac{D}{2}} \exp \left(-\frac{\hat{\mathcal{F}}_0}{\mathcal{U}} - \mathcal{Z} \right), \quad (7.29)$$

where the surface Symanzik polynomials are

$$\mathcal{U} = \sum_{n=-\infty}^{\infty} \alpha_{13}^n + \alpha_{12}^n, \quad \hat{\mathcal{F}}_0 = -B^\mu B^\mu, \quad (7.30)$$

and

$$\mathcal{Z} = \left(\sum_{n=-\infty}^{\infty} \alpha_{12}^n (n^2 p_1^2) + \alpha_{13}^n (np_1 - p_2)^2 \right) + m^2 \left(\alpha_{22} + \alpha_{33} + \sum_{n=-\infty}^{\infty} \alpha_{12}^n + \alpha_{13}^n \right) \quad (7.31)$$

In the formula for \mathcal{F}_0 , the B -vector is

$$B^\mu = \sum_{n=-\infty}^{\infty} np_1^\mu \alpha_{12}^n + (np_1^\mu - p_2^\mu) \alpha_{13}^n. \quad (7.32)$$

However, at this point we confront the problem of quotienting by MCG. The MCG is generated by

$$\mathbf{g}_{12}^n \mapsto \mathbf{g}_{12}^{n+1}, \quad \mathbf{g}_{13}^n \mapsto \mathbf{g}_{13}^{n+1}, \quad (7.33)$$

and it leaves \mathbf{g}_{22} and \mathbf{g}_{33} invariant. Naively, we might want to quotient by the MCG by restricting the integral to the region spanned by: $\mathbf{g}_{12}^0, \mathbf{g}_{13}^0, \mathbf{g}_{22}, \mathbf{g}_{33}$. However, this region is too small. It does not include any full cones of the Feynman fan. We could also try restricting the integral to the region spanned by: $\mathbf{g}_{12}^0, \mathbf{g}_{13}^0, \mathbf{g}_{12}^1, \mathbf{g}_{13}^1, \mathbf{g}_{22}, \mathbf{g}_{33}$. But this region is too large! The amplitude has *three* Feynman diagrams, but this region contains *four* cones, so it counts one of the diagrams twice. As this example shows, it is already a delicate problem to explicitly specify a fundamental domain for the MCG action.

7.4 Example: genus-one 2-loop amplitudes

The problem of modding by MCG becomes even more acute for non-planar amplitudes. The genus one 2-loop vacuum amplitude, considered in section 5.7, is computed by a 3-dimensional curve integral. But the MCG action in this case is an action of $\text{SL}_2\mathbb{Z}$. The

action on g -vectors is of the form

$$\mathfrak{g}_{p/q} \mapsto \mathfrak{g}_{(ap+bq)/(cp+dq)}, \quad \text{for } \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \text{SL}_2\mathbb{Z}. \quad (7.34)$$

For the vacuum amplitude, a simple example of a fundamental region is the region spanned by $\mathfrak{g}_{1/0}$, $\mathfrak{g}_{0/1}$, and $\mathfrak{g}_{1/1}$. However, for the n -point genus one 2-loop amplitude, identifying a fundamental region of this $\text{SL}_2\mathbb{Z}$ -action becomes very difficult.

In the next section, we present a simple method to compute the integrals in our formulas, for any MCG action.

8 Modding out by the mapping class group

Our formulas for amplitudes and integrands take the form of integrals over \mathbb{R}^E modulo the action of the Mapping Class Group, MCG,

$$\mathcal{A} = \int \frac{d^E t}{\text{MCG}} f(t), \quad (8.1)$$

for some MCG-invariant function, $f(t)$. One way to evaluate this integral is to find a fundamental domain for the MCG action. But it is tricky to identify such a region in general. Instead, it is convenient to mod out by the MCG action by defining a kernel, \mathcal{K} , such that

$$\mathcal{A} = \int d^E t \mathcal{K}(t) f(t). \quad (8.2)$$

In this section, we find kernels, \mathcal{K} , that can be used at all orders in perturbation theory, for all Mapping Class Groups.

8.1 Warm up

Consider the problem of evaluating an integral modulo a group action on its domain. For example, suppose $f(x)$ is invariant under the group of translations, T , generated by $x \mapsto x + a$, for some constant, a . We want to evaluate an integral

$$I = \int_{\mathbb{R}/T} dx f(x). \quad (8.3)$$

One way to do this is to restrict to a fundamental domain of T :

$$I = \int_0^a dx f(x). \quad (8.4)$$

But we can alternatively find a kernel $\mathcal{K}(x)$ such that

$$I = \int_{-\infty}^{\infty} dx \mathcal{K}(x) f(x). \quad (8.5)$$

One way to find such a kernel is to take a function $g(x)$ with finite support around 0, say. Then we can write

$$1 = \frac{\sum_{n=-\infty}^{\infty} g(x - na)}{\sum_{n=-\infty}^{\infty} g(x - na)}, \tag{8.6}$$

provided that $\sum_{n=-\infty}^{\infty} g(x - na)$ is nowhere vanishing. Inserting this into (8.3),

$$I = \int_{\mathbb{R}/T} dx \frac{\sum_{n=-\infty}^{\infty} g(x - na)}{\sum_{n=-\infty}^{\infty} g(x - na)} f(x) = \int_{-\infty}^{\infty} dx \frac{g(x)}{\sum_{n=-\infty}^{\infty} g(x - na)} f(x). \tag{8.7}$$

So that we can use

$$\mathcal{K}(x) = \frac{g(x)}{\sum_{n=-\infty}^{\infty} g(x - na)} \tag{8.8}$$

as a kernel to quotient out by the translation group. For example, suppose that we take $g(x) = \Theta(x+a)\Theta(-x+a)$, where $\Theta(x)$ is the Heaviside function. Inserting this into (8.7) gives

$$I = \int_{-a}^a dx \frac{1}{2} f(x). \tag{8.9}$$

The domain of this integral contains two copies of a fundamental domain for T , but this is compensated for by the 1/2 coming from $\mathcal{K}(x)$ to give the correct answer.

8.2 A tropical Mirzakhani kernel

The headlight functions, α_C , give a very natural solution to the problem of defining an integration kernel, \mathcal{K} .

Consider the case when MCG has *one generator*. Let \mathcal{S} be the set of curves which are *not* invariant under MCG. The sum of their headlight functions,

$$\rho = \sum_{C \in \mathcal{S}} \alpha_C, \tag{8.10}$$

is itself a MCG-invariant function. Moreover, ρ does not vanish on any top-dimensional cone (because no diagram can be formed without using at least one propagator from \mathcal{S}). So we can consider inserting the function

$$1 = \frac{\rho}{\rho} \tag{8.11}$$

into our integrals.

The set \mathcal{S} is the disjoint union of cosets under the MCG action, by the Orbit-Stabilizer theorem. When MCG has a single generator, these cosets are easy to describe. MCG does not alter the endpoints of curves. So if $C_{ij} \in \mathcal{S}$ is a curve connecting external lines i and j , the orbit of C_{ij} is a coset of \mathcal{S} . By the Orbit-Stabilizer theorem, these cosets are disjoint. So ρ can be resumed as

$$\rho = \sum_{i,j} \sum_{\gamma \in \text{MCG}} \alpha_{\gamma C_{ij}}. \tag{8.12}$$

Given this, we can mod out by the MCG action by defining

$$\mathcal{K} = \sum_{i,j} \frac{\alpha_{C_{ij}}}{\rho}, \tag{8.13}$$

where we choose a distinguished representative, C_{ij} , for each coset. We call (8.13) a *tropical Mirzakhani kernel*, because it is a tropical version of the kernel introduced by Mirzakhani to compute Weil-Petersson volumes [23]. Each headlight function, $\alpha_{C_{ij}}$, is non-vanishing in a convex region $V_{C_{ij}}$ that is spanned by all the cones in the fan that contain $\mathbf{g}_{C_{ij}}$. These regions *over-count* the diagrams, but this over-counting is corrected by the kernel, \mathcal{K} .

8.3 Example: the non-planar 1-loop propagator

As a sanity check, let us repeat the calculation of the non-planar 1-loop propagator from section 7.2, but now using the tropical Mirzakhani kernel. The MCG has one generator, and no curves are MCG-invariant. So take the set \mathcal{S} to be the set of all curves, C_n , and write

$$\rho = \sum_{n=-\infty}^{\infty} \alpha_n. \tag{8.14}$$

Choose C_0 , say, as the coset representative (all other curves are in the orbit of C_0). Then the tropical Mirzakhani kernel, (8.13), is

$$\mathcal{K} = \frac{\alpha_0}{\rho}. \tag{8.15}$$

Using this kernel, we find a pre-loop-integration integrand,

$$\mathcal{I} = \int dx dy \mathcal{K}(x, y) \exp \left(- \sum_{i=-\infty}^{\infty} \alpha_i X_i \right). \tag{8.16}$$

The headlight functions for this example were given in (5.33). In particular, $\alpha_0 = \max(0, x)$, which is vanishing outside of the region $x \geq 0$. In this region, the only other non-vanishing headlight functions are

$$\alpha_{-1} = \max(0, y) \quad \text{and} \quad \alpha_1 = -y + \max(0, y). \tag{8.17}$$

The formula is therefore

$$\mathcal{I} = \int_{x \geq 0} dx dy \frac{x}{x + |y|} \exp(-\alpha_{-1} X_{-1} - \alpha_0 X_0 - \alpha_1 X_1). \tag{8.18}$$

We can now perform the loop integral. Recall that $X_n = (\ell + nk)^2 + m^2$. Using this, the exponent in (8.18) is

$$S = -\rho \ell^2 - 2\ell \cdot k(\alpha_1 - \alpha_{-1}) - m^2 \alpha_0. \tag{8.19}$$

The Gaussian integral gives

$$\mathcal{A} = \int_{x \geq 0} dx dy \frac{x}{x + |y|} \left(\frac{\pi}{x + |y|} \right)^{\frac{D}{2}} \exp \left(k^2 \frac{|y|^2}{x + |y|} - m^2 x \right). \tag{8.20}$$

This doesn't immediately look like the Feynman integral for the 1-loop bubble, because of the Mirzakhani kernel pre-factor. However, writing

$$\frac{x}{x+y} = \frac{1}{2} + \frac{1}{2} \frac{x-y}{x+y}, \tag{8.21}$$

we find that the formula becomes

$$\mathcal{A} = \int_{x,y \geq 0} dx dy \left(\frac{\pi}{x+y} \right)^{\frac{D}{2}} \exp \left(k^2 \frac{y^2}{x+y} - m^2 x \right), \tag{8.22}$$

where we have used that the integrand over $x, y \geq 0$ is even under $x \leftrightarrow y$, whereas $x - y$ is odd. This is still not exactly the same as the conventional integral, because of the form of the exponent. However, note that the exponent can be rewritten as

$$k^2 \frac{y^2}{x+y} - m^2 x = -k^2 \frac{xy}{x+y} - (x+y)m^2, \tag{8.23}$$

reproducing the standard form of the Schwinger parametrized Feynman integral.

8.4 General tropical Mirzakhani kernels

Tropical Mirzakhani kernels can be defined to *any* mapping class group, with more than one generator. Fix some fatgraph Γ , with mapping class group MCG.

A conceptually simple way to define a kernel is to consider the set of L -tuples of curves that cut Γ to a tree graph. These define the *first Symanzik polynomial*,

$$\mathcal{U} = \sum_{\substack{S \\ \text{cuts to tree}}} \alpha_S, \tag{8.24}$$

which can also be computed as a determinant of a matrix (section 7). This function does not vanish on top-dimensional cones of the Feynman fan, since every diagram contains a subset of propagators that cut Γ to a tree. We can therefore insert

$$1 = \frac{\mathcal{U}}{\mathcal{U}} \tag{8.25}$$

into our integrals. Under the MCG action, the set of L -tuples appearing in \mathcal{U} is partitioned into cosets. Each coset represents an MCG-inequivalent way of cutting Γ down to a tree. By choosing a representative L -tuple for each such loop cut, we arrive at a kernel

$$\mathcal{K} = \sum_{\text{distinct loop cuts}} \frac{\alpha_S}{\mathcal{U}}. \tag{8.26}$$

Our integrals can then be computed as a sum over maximal cuts:

$$\mathcal{A} = \int \frac{d^E y}{\text{MCG}} \mathcal{I} = \sum_{\text{distinct loop cuts}} \int d^E y \frac{\alpha_S}{\mathcal{U}} \mathcal{I}. \tag{8.27}$$

The disadvantage of this formula is that it can be difficult to systematically identify a set of MCG-inequivalent maximal cuts.

8.5 The general iterative method

A more systematic way to quotient out by MCG is to break the MCG-action one generator at a time. This iterative method has the advantage of being completely algorithmic.

To apply the method, pick a trace-factor of Γ , β , which has some external particles, $1, \dots, m$. Let \mathcal{S}_β be the set of curves that have at least one endpoint in β , excluding any curves that are MCG-invariant, and write

$$\rho_\beta = \sum_{C \in \mathcal{S}_\beta} \alpha_C. \tag{8.28}$$

ρ_β is MCG-invariant. This is because the MCG action does not alter the endpoints of a curve. The set \mathcal{S}_β therefore has a coset decomposition. For each MCG orbit in \mathcal{S}_β , pick a representative curve, so that

$$\rho_\beta = \sum_{i=1}^k \sum_{\gamma \in \text{MCG}(\Sigma)} \alpha_{\gamma C_i}, \tag{8.29}$$

for some $k = |\mathcal{S}_\beta / \text{MCG}(\Sigma)|$ coset representatives C_1, \dots, C_k . We give more details about how to pick a set of coset representatives below.

Every top-dimensional cone is generated by at least *one* curve from the set \mathcal{S}_β , because otherwise that cone would not correspond to a complete triangulation of Γ . This means that ρ_β is non-vanishing everywhere, except on some lower-dimensional cones. Away from this vanishing locus, we can write

$$1 = \frac{\rho_\beta}{\rho_\beta}. \tag{8.30}$$

Given this, we define a tropical Mirzakhani kernel

$$\mathcal{K}_\beta = \sum_{i=1}^k \frac{\alpha_{C_i}}{\rho_\beta}. \tag{8.31}$$

This has the effect of breaking the MCG symmetry of the integrand, and reducing us to evaluating simpler integrals. In particular, we have

$$\mathcal{A} = \int \frac{d^E t}{\text{MCG}} \mathcal{I} = \sum_{i=1}^k \int \frac{d^E t}{\text{Stab}(C_i)} \frac{\alpha_{C_i}}{\rho_\beta} \mathcal{I}, \tag{8.32}$$

where $\text{Stab}(C_i) \leq \text{MCG}$ is the *stabilizer subgroup* for C_i . The factor

$$\frac{\alpha_{C_i}}{\rho_\beta} \tag{8.33}$$

is *itself* invariant under $\text{Stab}(C_i)$. So the integrals,

$$\int \frac{d^E t}{\text{Stab}(C_i)} \frac{\alpha_{C_i}}{\rho} \mathcal{I}, \tag{8.34}$$

can themselves be evaluated by finding a Mirzakhani kernel for the new group, $\text{Stab}(C_i)$. This iterative method ultimately yields an integral with no group action,

$$\mathcal{A} = \int \frac{d^E y}{\text{MCG}} \mathcal{I} = \int d^m y \mathcal{K} \mathcal{I}, \tag{8.35}$$

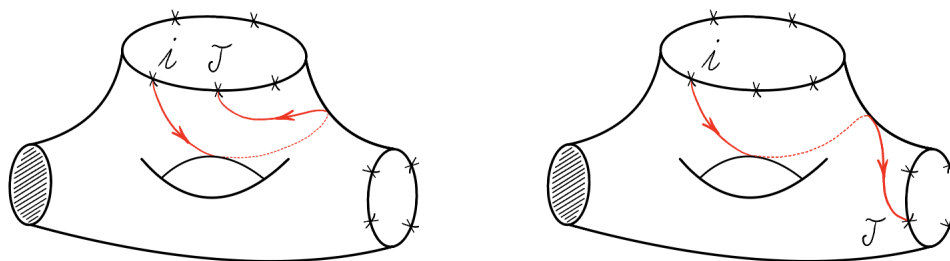


Figure 32. The two types of curves that are not invariant under the MCG, drawn here on the surface $S(\Gamma)$ associated to a fatgraph: curves connecting distinct trace factors (right), and topologically nontrivial curves that begin and end on the same trace factor (left).

where \mathcal{K} is a sum of products of kernels of the form (8.33).

To complete the description of the iterative method, we describe how to choose coset representatives from the set \mathcal{S}_β . The curves in this set break into two subsets, as in figure 32:

1. Curves C whose endpoints lie in two distinct trace factors. These curves cut Γ to a fatgraph Γ_C which has one fewer trace factors.
2. Curves C with both endpoints in the same trace factor. These curves cut Γ to a fatgraph Γ_C with one lower genus.

Both of these subsets have decompositions into cosets specified by the endpoints of the curves. So, for every pair of particles, i, j (with i in trace factor β), pick *any* curve C_{ij}^0 connecting them. These can be taken as coset representatives. The caveat is that, if i, j are both in trace factor β , we must choose a curve C_{ij}^0 which is not MCG-invariant. An MCG-invariant curve generates a trivial coset. The first step to break the MCG is then to insert the kernel

$$\sum_{i \in \beta} \sum_j \frac{\alpha_{ij}^0}{\sum_{\mathcal{S}_\beta} \alpha_C}. \tag{8.36}$$

For amplitudes involving a large number of external particles, this iterative method naively requires a lot of work (growing like n^L with the number of particles, n). However, this apparent complexity goes away completely if we choose an appropriate fatgraph, Γ , for our calculation. We use this to obtain simple formulas for amplitudes at all- n in a separate paper, [27]. But for now we will focus on low-point amplitudes, to illustrate the method in its simplest form.

8.6 Example: the genus one 2-loop vacuum amplitude

As an example, we briefly describe what happens for the genus one 2-loop vacuum amplitude (sections 5.7 and 7.4). The MCG is now $SL_2\mathbb{Z}$. In this case, there is only *one* coset to consider, since every curve is related to every other by

$$\mathfrak{g}_{p/q} \mapsto \mathfrak{g}_{(ap+bq)/(cp+dq)}, \quad \text{for } \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in SL_2\mathbb{Z}. \tag{8.37}$$

For the first step of the iteration, we can take any curve, say $C_{1/0}$, as a coset representative. The kernel for the first step is

$$\mathcal{K}_{1/0} = \frac{\alpha_{1/0}}{\sum_C \alpha_C}. \tag{8.38}$$

The subgroup that leaves $C_{1/0}$ invariant is

$$\text{Stab } C_{1/0} = \left\{ \begin{bmatrix} 1 & n \\ 0 & 1 \end{bmatrix} : n \in \mathbb{Z} \right\} < \text{SL}_2\mathbb{Z}. \tag{8.39}$$

The curves compatible with $C_{1/0}$ form a single coset for the action of this subgroup. So, for the second step, we can choose just one of them, $C_{0/1}$, say, as a coset representative. The kernel for the second step is

$$\mathcal{K}_{0/1} = \frac{\alpha_{0/1}}{\sum_{C'} \alpha_{C'}}, \tag{8.40}$$

where we sum only over curves, C' , that are non-intersecting with $C_{1/0}$. The final kernel is simply

$$\mathcal{K} = \frac{\alpha_{1/0}}{\alpha_{1/0} + \alpha_{0/1} + \alpha_{1/1} + \alpha_{-1/1}} \frac{\alpha_{0/1}}{\alpha_{0/1} + \alpha_{1/1} + \alpha_{-1/1}}, \tag{8.41}$$

where the simplification arises because $C_{1/1}$ and $C_{-1/1}$ are the only curves compatible with both $C_{1/0}$ and $C_{0/1}$.

9 Exemplitudes

We now show how to use the tropical Mirzakhani kernels to evaluate curve integrals. We give detailed low-dimensional examples of amplitudes up to 3 loops.

9.1 The non-planar 1-loop 3-point amplitude

The formula for the 1-loop non-planar 3-point amplitude was given in section 7.3. However, we did not show how to quotient by the MCG. Using the tropical Mirzakhani kernel, we now find the formula

$$\mathcal{A} = \int d^3t \mathcal{K} \left(\frac{\pi}{\mathcal{U}} \right)^{\frac{D}{2}} \exp \left(-\frac{\hat{\mathcal{F}}_0}{\mathcal{U}} - \mathcal{Z} \right), \tag{9.1}$$

where the Mirzakhani kernel is

$$\mathcal{K} = \frac{\alpha_{12}^0 + \alpha_{13}^0}{\rho}, \tag{9.2}$$

with ρ the sum over all α_C (except for those curves which are invariant under the MCG, namely C_{22} , C_{33}). The surface Symanzik polynomials are, as before,

$$\mathcal{U} = \sum_{n=-\infty}^{\infty} \alpha_{13}^n + \alpha_{12}^n, \quad \hat{\mathcal{F}}_0 = -B_\mu B^\mu, \quad \text{with } B^\mu = \sum_{n=-\infty}^{\infty} np_1^\mu \alpha_{12}^n + (np_1^\mu - p_2^\mu) \alpha_{13}^n. \tag{9.3}$$

Moreover, we have

$$\mathcal{Z} = \left(\sum_{n=-\infty}^{\infty} \alpha_{12}^n (n^2 p_1^2) + \alpha_{13}^n (n p_1 - p_2)^2 \right) + m^2 \left(\alpha_{22} + \alpha_{33} + \sum_{n=-\infty}^{\infty} \alpha_{12}^n + \alpha_{13}^n \right). \quad (9.4)$$

Alternatively we can write the curve integral in terms of

$$\mathcal{F}_0 = \sum_{n \neq m} n m p_1^2 \alpha_{12}^n \alpha_{12}^m + \sum_{n \neq m} (n p_1 - p_2) \cdot (m p_1 - p_2) \alpha_{13}^n \alpha_{13}^m + \sum_{n, m} n p_1 \cdot (m p_1 - p_2) \alpha_{12}^n \alpha_{13}^m, \quad (9.5)$$

which is precisely a sum over the sets of curves that cut the surface (resp. fatgraph) into two disks (resp. trees). Then

$$\mathcal{A} = \int d^3 t \mathcal{K} \left(\frac{\pi}{\mathcal{U}} \right)^{\frac{D}{2}} \exp \left(-\frac{\mathcal{F}_0}{\mathcal{U}} - m^2 \sum_C \alpha_C \right). \quad (9.6)$$

In either case, the key task is to understand how we quotient by the MCG using the kernel \mathcal{K} .

Let us first see why (9.2) is a Mirzakhani kernel. The MCG has one generator. It leaves C_{22} and C_{33} invariant, but acts non-trivially on the set $\{C_{12}^n, C_{13}^n\}$ of all curves that connect the first trace factor to the second trace factor. ρ is the sum of α_C for all these curves,

$$\rho = \sum_{n=-\infty}^{\infty} (\alpha_{12}^n + \alpha_{13}^n). \quad (9.7)$$

This set has two MCG cosets, labelled by the start and end points of the curves. We can take C_{12}^0 and C_{13}^0 as the two coset representatives. C_{12}^0 , for instance, represents the coset of all curves that begin at 1 and end at 2. (Recall section 8.)

Naively, it looks as if (9.6) involves infinitely many α_C , which it would be laborious to compute. However, the Mirzakhani kernel ensures that only a few α_C are needed. To see how this works, consider, say, the first term in the kernel,

$$\mathcal{K}_{12} = \frac{\alpha_{12}^0}{\rho}. \quad (9.8)$$

In the region where $\alpha_{12}^0 \neq 0$, all other α_C are vanishing, except for:

$$\alpha_{12}^{-1}, \alpha_{12}^1, \alpha_{13}^0, \alpha_{13}^1, \alpha_{22}. \quad (9.9)$$

So in this region, \mathcal{U} and B^μ simplify to

$$\mathcal{U} = \alpha_{12}^0 + \alpha_{12}^1 + \alpha_{12}^{-1} + \alpha_{13}^0 + \alpha_{13}^1, \quad (9.10)$$

$$B^\mu = -k_1^\mu \alpha_{12}^{-1} - k_2^\mu \alpha_{13}^0 + (k_1^\mu - k_2^\mu) \alpha_{13}^1. \quad (9.11)$$

When we compute these α 's, using the matrix method, we find that they become simple functions in the region $x > 0$, where α_{12}^0 is non-zero. In this region, we have $\alpha_{12}^0 = x$. Moreover, the remaining 5 headlight functions become

$$\alpha_{13}^1 = -\max(0, y) + \max(0, y, y + z), \quad \alpha_{13}^0 = \max(0, y), \quad (9.12)$$

$$\alpha_{12}^1 = -y - \max(0, z) + \max(0, y, y + z), \quad \alpha_{12}^{-1} = -z + \max(0, z), \quad (9.13)$$

$$\alpha_{22} = -\max(0, y, y + z) + \max(0, y) + \max(0, z). \quad (9.14)$$

These are precisely the headlight functions for the 5-point tree amplitude! We could have anticipated this, because cutting Γ along C_{12}^0 yields a 5-point tree graph. Using these tree-like headlight functions, we can compute the contribution of \mathcal{K}_{12} to the curve integral, (9.6). The contribution from the second term in the Mirzakhani kernel is similar.

In this example, we find that we only need to know the headlight functions α_C for *tree level* amplitudes, in order to compute the full 1-loop amplitude! In fact, we can prove that this happens *in general*. Suppose a monomial, α_S (for some set of L curves S), appears in the numerator of the kernel \mathcal{K} . In the region where $\alpha_S \neq 0$, all remaining α_C 's simplify to become headlight functions for the tree-fatgraph obtained by cutting Γ along all the curves in S . This general phenomenon is computationally very useful, and we study it in greater detail elsewhere.

9.2 The genus one 2-loop vacuum amplitude

We have already mentioned the 2-loop genus one vacuum computation in sections 5.7 and 7.4. We now have all the tools to compute it properly. The result is the following simple integral

$$\mathcal{A} = \int_R dx dz dw \mathcal{K} \left(\frac{\pi^2}{\mathcal{U}} \right)^{\frac{D}{2}} \exp(-\mathcal{Z}), \tag{9.15}$$

where the kernel is (as given in section 8.6)

$$\mathcal{K} = \frac{\alpha_{1/0}}{\alpha_{1/0} + \alpha_{0/1} + \alpha_{1/1} + \alpha_{-1/1}} \frac{\alpha_{0/1}}{\alpha_{0/1} + \alpha_{1/1} + \alpha_{-1/1}}, \tag{9.16}$$

and now with surface Symanzik polynomials

$$\mathcal{U} = \det A, \tag{9.17}$$

$$\mathcal{Z} = m^2(\alpha_{1/0} + \alpha_{0/1} + \alpha_{1/1} + \alpha_{-1/1}). \tag{9.18}$$

Note that the region, R , where $\alpha_{1/0}\alpha_{0/1} \neq 0$ is, in the coordinates of section 5.7, given by $w, x \leq 0$ and $x + 2z \leq 0$. The curve integral is restricted by the Mirzakhani kernel to this region.

To see how this curve integral comes about, we need to understand how to assign momenta to the curves. The easiest way to assign momenta is to use the homology of curves on the torus, section 3.3.1. Assign the A-cycle momentum ℓ_1 and the B-cycle momentum ℓ_2 . The curve $C_{p/q}$ wraps the A-cycle q times and the B-cycle p times, and so it has momentum $p\ell_1 + q\ell_2$ giving

$$X_{p/q} = (p\ell_1 + q\ell_2)^2 + m^2. \tag{9.19}$$

With this momentum assignment, the matrix A , which records the dependence on the chosen basis of loops, is

$$A^{ab} = \begin{bmatrix} \alpha_{1,0} + \alpha_{1,1} + \alpha_{-1,1} & \alpha_{1,1} - \alpha_{-1,1} \\ \alpha_{1,1} - \alpha_{-1,1} & \alpha_{0,1} + \alpha_{1,1} + \alpha_{-1,1} \end{bmatrix}. \tag{9.20}$$

Moreover, the momentum assigned to the curves has no non-loop part, so that

$$\mathcal{Z} = m^2 \sum_C \alpha_C, \tag{9.21}$$

which restricts to (9.18) in the region R .

We now evaluate the amplitude. Once again, we will be aided by a striking simplification of the headlight parameters. The headlight parameters were given in section 5.7. But in the region R , $\alpha_{1/1}$ and $\alpha_{-1/1}$ simplify to become tree-like headlight functions:

$$\alpha_{1/1} = \max(0, -z) \quad \text{and} \quad \alpha_{-1/1} = \max(0, z). \tag{9.22}$$

This corresponds to the fact that cutting Γ along $C_{1/0}$ and $C_{0/1}$ gives a 4-point tree graph. Moreover, in this region,

$$\alpha_{1/0} = -w, \quad \text{and} \quad \alpha_{0/1} = -x - 2 \max(0, z), \tag{9.23}$$

and so

$$\alpha_{1/0} + \alpha_{0/1} + \alpha_{1/1} + \alpha_{-1/1} = -w - x - z \tag{9.24}$$

$$\alpha_{0/1} + \alpha_{1/1} + \alpha_{-1/1} = -x - z. \tag{9.25}$$

Moreover, \mathcal{U} and \mathcal{Z} become

$$\mathcal{U} = \det A = wx + zw - (x + 2 \max(0, z))|z|, \quad \text{and} \quad \mathcal{Z} = -m^2(w + x + z). \tag{9.26}$$

So the vacuum amplitude is

$$\mathcal{A} = \int_R dw dx dz \frac{w(x + 2 \max(0, z))}{(w + x + z)(x + z)} \left(\frac{\pi^2}{\mathcal{U}} \right)^{\frac{D}{2}} \exp \left(m^2(w + x + z) \right). \tag{9.27}$$

It is not obvious that this is the correct answer. In the conventional calculation, the amplitude receives just a single contribution: the vacuum sunset Feynman diagram. Our formula resembles, but is not the same, as the Schwinger parameterisation for this diagram. To see that they are equivalent, note that R can be divided into two regions: one where $z \geq 0$ and another where $z \leq 0$. By a simple change of variables, the integral over either one of these regions can be written as

$$I = \int_{a,b,c \geq 0} dadbdc \frac{ab}{(a + b + c)(b + c)} \left(\frac{\pi^2}{ab + bc + ca} \right)^{\frac{D}{2}} \exp \left(-m^2(a + b + c) \right). \tag{9.28}$$

However, summing over the 6 possible permutations of a, b, c , note that

$$\frac{ab}{b + c} + (\text{permutations of } a, b, c) = a + b + c. \tag{9.29}$$

It follows that I is 1/6 times the sunset integral with the usual parameterisation. Integrating over all of R , we then recover that

$$\mathcal{A} = \frac{1}{3} \int_{x,y,z \geq 0} dx dy dz \left(\frac{\pi^2}{xy + y|z| + |z|x} \right)^{\frac{D}{2}} \exp \left(-m^2(x + y + |z|) \right). \tag{9.30}$$

This is 1/3 times the vacuum sunset integral. The factor of 1/3 corresponds to the fact that the fatgraph has $|\text{Aut}(\Gamma)| = 3$.

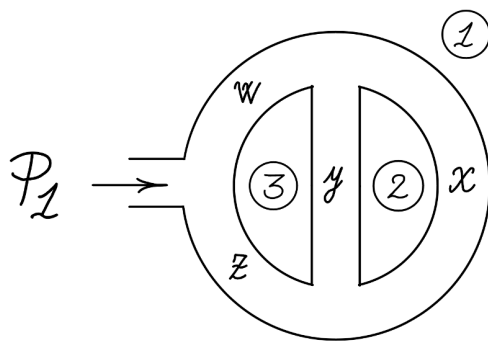


Figure 33. A planar 2-loop tadpole graph.

9.3 The planar 2-loop tadpole

We can compute the planar 2-loop tadpole amplitude using the fatgraph Γ in figure 33. The curves on this fatgraph can be labelled by their endings. We have two loop boundaries, labelled 2, 3 in the figure. The curves are then $C_{23}, C_{22}, C_{33}, C_{12}^n, C_{13}^n$, where n indexes how many times the curves C_{12}^n, C_{13}^n loop around before beginning their spiral. As usual, we will only need a small number of these curves to compute the amplitude.

Because Γ is planar, we can introduce dual variables $z_1^\mu, z_2^\mu, z_3^\mu$ to parametrise the momenta of the curves. The propagator factors are then

$$X_{12}^n = (z_2 - z_1)^2 + m^2, \quad X_{13}^n = (z_3 - z_1)^2 + m^2, \quad X_{23} = (z_3 - z_2)^2 + m^2. \quad (9.31)$$

It is convenient to take $z_3 - z_1$ and $z_2 - z_1$ as our loop momentum variables.

The curve integral for the amplitude is then

$$\mathcal{A} = \int d^4t \mathcal{K} \left(\frac{\pi^2}{\mathcal{U}} \right)^{\frac{D}{2}} \exp(-\mathcal{Z}), \quad (9.32)$$

where

$$\mathcal{U} = \det A, \quad \text{and} \quad \mathcal{Z} = m^2 \left(\alpha_{23} + \alpha_{22} + \alpha_{33} + \sum_n (\alpha_{12}^n + \alpha_{13}^n) \right). \quad (9.33)$$

Moreover, using the momenta assignments from the dual variables, (9.31), A is the 2×2 matrix

$$A = \begin{bmatrix} \alpha_{23} + \sum_{n=-1}^1 \alpha_{12}^n & \alpha_{23} \\ \alpha_{23} & \alpha_{23} + \sum_{n=-1}^1 \alpha_{13}^n \end{bmatrix}. \quad (9.34)$$

\mathcal{U} is the determinant of A , and each monomial in this determinant corresponds to a pair of curves that cut Γ to a 5-point tree graph. Using the fact that $\alpha_C \alpha_D = 0$ if C, D intersect, we find

$$\mathcal{U} = \sum_{n=-\infty}^{\infty} \left(\alpha_{23} \alpha_{12}^n + \alpha_{23} \alpha_{13}^n + \alpha_{12}^n \alpha_{13}^n + \alpha_{12}^n \alpha_{13}^{n+1} \right). \quad (9.35)$$

Here, we have chosen a convention for the index n such that C_{12}^n, C_{13}^{n+1} are compatible, but C_{12}^n, C_{13}^{n-1} intersect. The MCG has one generator, which acts on the index n . So it is clear

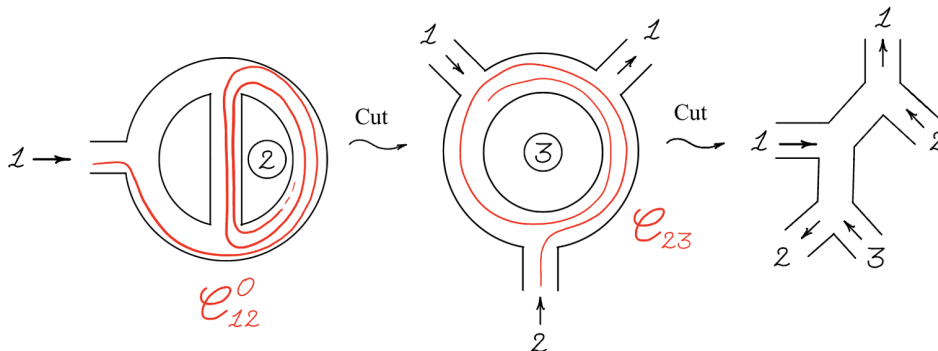


Figure 34. A maximal cut of the planar 2-loop tadpole graph. The curve C_{12}^0 cuts Γ to a 3-point 1-loop graph, and the curve C_{23} cuts this further to a 5-point tree graph.

that the monomials in \mathcal{U} can be decomposed into four cosets (corresponding to the four terms in the sum). We therefore get a Mirzakhani kernel (of the type discussed in section 8.4)

$$\mathcal{K} = \frac{\mathcal{U}_0}{\mathcal{U}}, \quad (9.36)$$

with

$$\mathcal{U}_0 = \alpha_{23}\alpha_{12}^0 + \alpha_{23}\alpha_{13}^0 + \alpha_{12}\alpha_{13}^0 + \alpha_{12}\alpha_{13}^1. \quad (9.37)$$

In the region where $\mathcal{U}_0 \neq 0$, only 12 α_C 's are non-vanishing. In fact, each monomial in \mathcal{U}_0 defines a maximal cut of Γ , which cuts Γ to a 5-point tree graph. See figure 34. \mathcal{A} is the sum of four terms,

$$\mathcal{A} = \mathcal{A}_{C_{23}, C_{12}^0} + \mathcal{A}_{C_{23}, C_{13}^0} + \mathcal{A}_{C_{12}^0, C_{13}^0} + \mathcal{A}_{C_{12}^0, C_{13}^1}, \quad (9.38)$$

each corresponding to a different maximal cut of the fatgraph.

For instance, $\mathcal{A}_{C_{23}, C_{12}^0}$ is given by the curve integral over the region $\alpha_{23}\alpha_{12}^0 \neq 0$. In this region, only 5 other α_C 's are non-vanishing. The curves correspond to the five curves on the 5-point tree graph obtained by cutting along C_{23}, C_{12}^0 . The 5 curves compatible with C_{23}, C_{12}^0 are

$$C_{12}^1, C_{12}^{-1}, C_{13}^0, C_{13}^1, C_{22}. \quad (9.39)$$

In this region, the headlight functions simplify, so that, similar to the previous examples, the curve integral only sees the headlight functions of the 5-point tree-level problem. Explicitly, in coordinates, we can take (in this region) $\alpha_{23} = w$, $\alpha_{12}^0 = x$, and the remaining headlight functions take the form of tree-level headlight functions (see figure 34):

$$\alpha_{13}^1 = f_2 - f_1 - z, \quad \alpha_{13}^0 = f_1 - y, \quad (9.40)$$

$$\alpha_{22} = f_2 - f_3, \quad \alpha_{12}^1 = f_3, \quad (9.41)$$

$$\alpha_{12}^{-1} = f_1 + f_3 - f_2. \quad (9.42)$$

where

$$f_1 = \max(0, y), \quad f_2 = \max(0, z, y + z), \quad f_3 = \max(0, z). \quad (9.43)$$

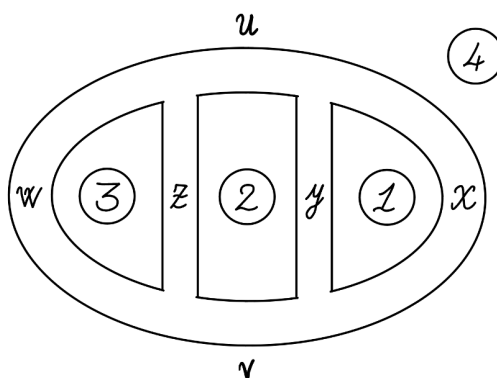


Figure 35. Three loop.

So, in this region, the A matrix restricts to

$$A' = \begin{bmatrix} w + f_1 - f_2 + 2f_3 & w \\ w & w + f_2 - y - z \end{bmatrix}, \tag{9.44}$$

and \mathcal{Z} restricts to

$$\mathcal{Z}' = m^2(w + x - y - z + f_1 + f_2 + f_3). \tag{9.45}$$

The contribution of this term to the amplitude is then

$$\mathcal{A}_{C_{23}, C_{12}^0} = \int_{w,x \geq 0} dw dx dy dz \frac{wx}{\det A'} \left(\frac{\pi^2}{\det A'} \right)^{\frac{D}{2}} \exp(-\mathcal{Z}'). \tag{9.46}$$

The other 3 cuts are similarly computed.

9.4 The planar 3-loop vacuum amplitude

We now consider a 3-loop example. The 3-loop vacuum amplitude can be computed using the 3-loop fatgraph, Γ , in figure 35. The curves on Γ all begin and end in a spiral. There are four loop boundaries, labelled $a = 1, 2, 3, 4$ in the figure, that the curves can spiral around. Let C_{ab}^δ be the curves that begin spiralling around a , and end spiralling around b . There are infinitely many such curves, all related by the action of the MCG. In fact, the MCG action in this case is quite complicated: it is an action of the braid group B_3 . However, using a tropical Mirzakhani kernel, we can still compute the amplitude.

The momentum assignment to the curves is easy to describe, because Γ is a planar graph. Introduce dual momentum variables, z_a^μ , associated to the four boundaries, $a = 1, 2, 3, 4$. Then the propagator for C_{ab}^δ is just

$$X_{ab} = (z_b^\mu - z_a^\mu)^2 + m^2. \tag{9.47}$$

We can choose any three z_a to be our loop momentum variables.

Our formula for the amplitude is then

$$\mathcal{A} = \int d^6t \mathcal{K} \left(\frac{\pi^3}{\mathcal{U}} \right)^{\frac{D}{2}} \exp(-\mathcal{Z}), \tag{9.48}$$

where the surface Symanzik polynomials are

$$\mathcal{U} = \det' \tilde{A}, \quad \mathcal{Z} = m^2 \sum \alpha_{ab}^\delta. \tag{9.49}$$

Here, we take a slightly different approach to presenting \mathcal{U} , adapted to the planar case, by using a reduced determinant, \det' , which excludes a row and column. The 4×4 matrix \tilde{A} is (for $a \neq b$)

$$\tilde{A}_{ab} = \sum_{\delta} \alpha_{ab}^\delta, \quad \tilde{A}_{aa} = - \sum_{c \neq a} \tilde{A}_{ac}. \tag{9.50}$$

By the matrix-tree theorem, the reduced determinant, $\det' \tilde{A}$, turns into a sum over all maximal cuts of the fatgraph Γ . In this case, a maximal cut is given by any three non-intersecting curves, $\{C_{ab}^\delta, C_{cd}^{\delta'}, C_{ef}^{\delta''}\}$, such that the pairs, ab, cd, ef , span a tree on the set $\{1, 2, 3, 4\}$. So $\det' \tilde{A}$ indeed recovers the definition of \mathcal{U} as the sum over maximal cuts of the fatgraph. Explicitly, it takes the form

$$\mathcal{U} = \sum_{\delta, \delta', \delta''} \sum_{\text{trees}} \alpha_{ab}^\delta \alpha_{cd}^{\delta'} \alpha_{ef}^{\delta''} \tag{9.51}$$

We can now use this formula for \mathcal{U} to define a Mirzakhani kernel, \mathcal{K} . This set of triples appearing in \mathcal{U} can be decomposed as a sum of cosets under the MCG. The MCG-action leaves the starts and ends of each curve unchanged. So we find that there are 16 MCG-inequivalent maximal cuts of Γ , corresponding to the 4^2 distinct labelled trees in the set $\{1, 2, 3, 4\}$. For each such labelled tree, we choose a coset representative.

$$\alpha_{ab}^0 \alpha_{cd}^0 \alpha_{ef}^0, \tag{9.52}$$

where the pairs ab, cd, ef define the tree, and $C_{ab}^0, C_{cd}^0, C_{ef}^0$ is some choice of 3 non-intersecting curves. Let \mathcal{U}_0 be the sum of monomials for these 16 coset representatives. It has the form

$$\mathcal{U}_0 = \sum_{12 \text{ perms}} \alpha_{12} \alpha_{23} \alpha_{34} + \sum_{4 \text{ perms}} \alpha_{14} \alpha_{24} \alpha_{34}. \tag{9.53}$$

Then

$$\mathcal{K} = \frac{\mathcal{U}_0}{\mathcal{U}} \tag{9.54}$$

is our Mirzakhani kernel.

An exercise in the intersection rules for mountainscapes shows that the following 6 curves are sufficient to build each of the 16 maximal cuts:

$$C_{14}^0 = (xRyR)^\infty x(LvLwLuLx)^\infty, \tag{9.55}$$

$$C_{24}^0 = (uRyRvRzR)^\infty u(LxLvLwLu)^\infty, \tag{9.56}$$

$$C_{34}^0 = (wRzR)^\infty w(LuLxLvLw)^\infty, \tag{9.57}$$

$$C_{12}^0 = (yRxR)^\infty y(LuLzLvLy)^\infty, \tag{9.58}$$

$$C_{23}^0 = (zRuRyRvR)^\infty z(LwLz)^\infty, \tag{9.59}$$

$$C_{13}^0 = (RyRx)^\infty LvR(zLwL)^\infty. \tag{9.60}$$

This is because all of these curves are pairwise compatible. Using these curves, we can define a restricted matrix (for $a \neq b$)

$$\tilde{A}_{ab}^0 = \alpha_{ab}^0, \quad \tilde{A}_{aa}^0 = - \sum_{c \neq a} \tilde{A}_{ac}^0 \tag{9.61}$$

so that, by the matrix-tree theorem, $\mathcal{U}^0 = \det' \tilde{A}^0$. Our Mirzakhani kernel is then

$$\mathcal{K} = \frac{\det' \tilde{A}^0}{\det' \tilde{A}}. \tag{9.62}$$

For each of the 16 monomials in \mathcal{U}^0 we get a contribution to \mathcal{A} . For instance, take the monomial

$$\alpha_{12}^0 \alpha_{23}^0 \alpha_{34}^0, \tag{9.63}$$

corresponding to the tree $1 - 2 - 3 - 4$. The associated contribution to \mathcal{A} only involves α_C for curves C compatible with this maximal cut. This maximal cut gives a tree fatgraph, with colour ordering (123432).¹⁰ So this contribution to the amplitude involves only the 9 headlight functions for this 6-point tree fatgraph.

Finally, note that by permutation symmetry (with respect to the dual variables z_a), we only really need to evaluate two of the maximal cuts in our formula, say:

$$\alpha_{12}^0 \alpha_{23}^0 \alpha_{34}^0 \quad \text{and} \quad \alpha_{14}^0 \alpha_{24}^0 \alpha_{34}^0. \tag{9.64}$$

Then

$$\mathcal{A} = 12 \mathcal{A}_{12,23,34} + 4 \mathcal{A}_{14,24,34}, \tag{9.65}$$

where each of $\mathcal{A}_{12,23,34}$ and $\mathcal{A}_{14,24,34}$ can be computed knowing only the headlight functions for a 6-point tree graph.

10 A first look at recursion

The tropical Mirzakhani kernels dramatically simplify the task of evaluating our amplitudes. Using these kernels, our formulas for amplitudes at L loops end up expressed in terms of the headlight functions, α_C , that we have already computed for lower loop level amplitudes. In this section, we show an alternative way to apply the Mirzakhani kernels to compute amplitudes, by using them to define a powerful recursion relation for the integrands, \mathcal{I} .

Fix a fatgraph Γ . Its associated (pre-loop-integration) integrand is given by the curve integral

$$\mathcal{I} = \int \frac{d^n t}{\text{MCG}} Z, \quad Z = \exp \left(- \sum_C \alpha_C X_C \right). \tag{10.1}$$

To evaluate the curve integral, we introduce a tropical Mirzakhani kernel, as above. Take, for example, some trace factor β . The non-separating curves with endpoints on Γ form a set \mathcal{S}_β , and which can be partitioned into MCG orbits with some coset representatives

¹⁰Cutting a curve that ends in a spiral around a loop boundary creates a new external line on that boundary.

C_1, \dots, C_k . Each of these curves, C_i , cuts Γ to a fat graph Γ_{C_i} with a smaller number of loops. The Mirzakhani kernel \mathcal{K}_β then gives

$$\mathcal{I} = \sum_{i=1}^k \int \frac{d^n t}{\text{MCG}} \frac{\alpha_{C_i}}{\rho} Z. \tag{10.2}$$

Introducing an auxiliary parameter, ξ , the $1/\rho$ can be incorporated into the exponential using

$$\frac{1}{\rho} = \int_0^\infty d\xi e^{-\rho\xi}. \tag{10.3}$$

Equation (10.2) then implies the following recursion formula:

$$\mathcal{I} = \int_0^\infty d\xi \sum_{i=1}^k \frac{-1}{(X_{C_i} + \xi)^2} \mathcal{I}_{\Gamma_{C_i}}(X'_C), \tag{10.4}$$

where the new dual variables X'_C appearing in the integrand $\mathcal{I}_{\Gamma_{C_i}}(X'_C)$ are given by

$$X'_C = \begin{cases} X_C + \xi & \text{if } C \in \mathcal{S}_\beta \\ X_C & \text{else.} \end{cases} \tag{10.5}$$

This formula, (10.4), is a completely recursive way to obtain the rational functions \mathcal{I} to all orders in the perturbation series. A detailed derivation of (10.4) is given in appendix D.

For example, consider again the 1-loop non-planar propagator computed in section 7.2. The curves on Γ are $\mathcal{S} = \{C_n\}$ as before, and their associated dual variables are

$$X_n = (\ell + nk)^2. \tag{10.6}$$

The MCG has just one generator, and so we will only need to apply the global forward limit once. Taking C_0 as our coset representative, (10.4) gives

$$\mathcal{I}_\Gamma = \int_0^\infty d\xi \frac{-1}{(X_0 + \xi)^2} \mathcal{I}_{\Gamma_{C_0}}(X_1 + \xi, X_{-1} + \xi), \tag{10.7}$$

where Γ_{C_0} is the 4-point tree graph obtained by cutting Γ along C_0 . The curves C_1 and C_{-1} become the two possible propagators of Γ_{C_0} : on Γ , C_1 and C_{-1} are the only two curves that do not intersect C_0 . So we have,

$$\mathcal{I}_\Gamma = - \int_0^\infty d\xi \left(\frac{1}{(X_0 + \xi)^2} \frac{1}{X_1 + \xi} + \frac{1}{(X_0 + \xi)^2} \frac{1}{X_{-1} + \xi} \right). \tag{10.8}$$

Evaluating the ξ integral gives the following formula for the integrand:

$$\mathcal{I}_\Gamma = \frac{1}{X_0(X_1 - X_0)} + \frac{1}{X_0(X_{-1} - X_0)}. \tag{10.9}$$

Here we see the appearance of *linearised propagators*, of the form $1/(X_C - X_{C_i})$. Such linearised propagators have arisen in previous studies of forward limit [28–33]. In the full sum,

these linearised propagators sum to give back the ordinary loop integrand after identifications made using shifts of the loop momenta. In our current example, the loop momentum shift $\ell \mapsto \ell + k$ shifts the dual variables by $X_n \mapsto X_{n+1}$. Applying this shift to the second term in (10.9) gives

$$\mathcal{I}'_{\Gamma} = \frac{1}{X_0(X_1 - X_0)} + \frac{1}{X_1(X_0 - X_1)} = \frac{1}{X_0 X_1}. \tag{10.10}$$

For higher loop integrands, we can use multiple iterations of (10.4) to write \mathcal{I} as a sum over some tree amplitudes, with various shifts in the kinematic variables.

Note that the recursion, (10.4), continues to hold even when the X_C variables are not all distinct. For example, if all X_C are set equal to a constant, $X_C = X$, then $\mathcal{I}_{\Gamma} = C_{\Gamma}/X^E$, where C_{Γ} is the number of Feynman diagrams contributing to the amplitude. In this case, (10.4) can be used to recursively compute the number of diagrams. Moreover, the recursion (10.4) also holds when there are higher poles in the integrand, arising from diagrams like bubbles. We give a more complete analysis of these recursions elsewhere.

11 Outlook

The new representation of all-loop amplitudes we have studied in this paper has implications far beyond our understanding of scalar amplitudes, and has consequences for the understanding of particle and string scattering generally. We highlight a number of directions that are especially primed for immediate development.

The magic of the *curve integral* formulas is that integrals over an $O(n)$ dimensional space, of an action built from $O(n^2)$ piecewise linear functions, automatically reproduces the full amplitudes, which are conventionally sums over $O(4^n)$ Feynman diagrams. The novelty of this formalism over conventional field theory must therefore become most manifest in the limit $n \rightarrow \infty$ of a large number of particles. In examples, we have found evidence that the external kinematical data can be chosen so that the large- n limits of the curve integrals are well defined, leading to formulas for amplitudes in the large- n limit in terms of *tropical path integrals*. Studying this limit might lead to a new understanding of the emergence of strings from colored particles at strong coupling. At strong coupling, the scattering for a small number of particles is exponentially small, and the amplitude is instead dominated by the emission of a huge number of particles, approximating field configurations that should more continuously connect to a string worldsheet picture.

Even at finite n the curve integral formalism offers radically new methods to compute amplitudes. For instance, it allows to evaluate amplitudes numerically by direct integration, thus avoiding the generation of Feynman diagrams altogether. The geometric properties of the fan suggest a new search for an optimal numerical integration strategy, uplifting recent breakthroughs in the numerical evaluation of Feynman integrals in parametric form to entire amplitudes [34, 35].

A second frontier ripe for immediate investigation is an understanding of gravity and gravity-like amplitudes. Just as the $\text{tr}\phi^3$ theory is a model for general colored amplitudes, a scalar model for gravity is given by an uncolored scalar σ with cubic self-interaction σ^3 . In special cases, it is now standard to think of uncolored and colored theories as related by

double-copy or ‘gravity = gauge²’ formulas [36]. The stringy origin of these formulas, the KLT relations, is deeply connected to thinking about the string worldsheet in a fundamentally *complex* fashion as a Riemann surface with a complex structure. But there are many reasons why our formulation of uncolored amplitudes will involve a very different sort of treatment. As we alluded to in the introduction, the existence of σ is forced on us in the most elementary way by the structure of the Feynman fan, which has lower-dimensional ‘holes’ that are beautifully completed by adding in new vectors corresponding to σ particles. This does not remotely have the flavor of ‘gravity = gauge²’. Moreover, as alluded to in the introduction, the u -variables central to our story are deeply connected to the string worldsheet (and Teichmüller space), but via *hyperbolic geometry* and *not* through the conventional picture of Riemann surfaces with complex structure. All of this dovetails nicely with the many observations, in examples of gravity amplitudes, that there is vastly more structure to gravity amplitudes than is suggested by the ‘gravity=gauge²’ slogan. The striking way in which σ is forced on us in our story is a new departure point for uncovering more of this hidden structure.

Finally, our results here strongly suggest that there is way to describe fundamental particle physics in the real world from a more elementary starting point, with spacetime and quantum mechanics appearing as emergent principles. We believe that we have taken a major new step in this direction with the results we have begun to introduce in this paper. A number of major challenges remain before we can reach this goal. The first is to understand how fermions arise from this new perspective, which has so far only been applied to bosonic scattering. For Standard Model physics, describing chiral fermions will be especially interesting and important. Another challenge is that the key structures in our formulas stem from a fatgraph, which is most immediately connected to the adjoint representation of $U(N)$ gauge theories. But the quantum numbers of the Standard Model are more interesting. For instance, in the $SO(10)$ grand unified theory, the matter lives in ten fundamentals (higgses) together with three $\mathbf{16}$ ’s for the fermions. How might the amplitudes for matter in these representations emerge from elementary combinatorial foundations?

Acknowledgments

We especially thank Song He and Thomas Lam for countless stimulating conversations on the topics of this paper over many years. We also thank Sebastian Mizera and Hofie Hannesdottir for many discussions, and Song He, Carolina Figueiredo, Daniel Longenecker, Qu Cao and Jin Dong for ongoing interactions related to the material of this paper over the past year. NAH is supported by the DOE under grant DE-SC0009988; further crucial contributions to his work were made possible by the Carl B. Feinberg cross-disciplinary program in innovation at the IAS. NAH also expresses sincere thanks to HF, PGP, GS and HT for restraining themselves from strangling him during the completion of this work. PGP is supported by ANR grant CHARMS (ANR-19-CE40-0017) and by the Institut Universitaire de France (IUF). PGP worked on this project while participating in *Representation Theory: Combinatorial Aspects and Applications* at the Centre for Advanced Study, Oslo. HF is supported by Merton College, Oxford. During this project HF received additional support from ERC grant GALOP (ID: 724638). During this project GS was supported by Brown University, Providence, the Perimeter Institute, Waterloo, and the Institute for Advanced Study, Princeton. GS was

also funded by the European Union’s Horizon 2020 research and innovation programs *Novel structures in scattering amplitudes* (No. 725110) of Johannes Henn. GS thanks the groups of C. Anastasiou and N. Beisert at ETH Zurich for hospitality during the worst phase of the COVID-19 pandemic. HT was supported by NSERC Discovery Grant RGPIN-2022-03960 and the Canada Research Chairs program, grant number CRC-2021-00120.

A Deriving the curve integral formula

To see why (6.3) is correct, let us write the amplitude explicitly. Write

$$X_C = P_C^2 + m^2 \tag{A.1}$$

for the propagator factor associated to curve C (with momentum P_C^μ). Fix some fatgraph Γ with some color factor C_Γ . The associated partial amplitude can be expressed with just one overall loop integration as

$$\mathcal{A} = \int \prod_{i=1}^L d^D \ell_i \left(\sum_{\Gamma'} \prod_C \frac{1}{X_C} \right), \tag{A.2}$$

where sum over exactly one of every fatgraph Γ' that has color factor $C_{\Gamma'} = C_\Gamma$. The integrand in this formula can be written as an integral over *curve space*, V . To do this, recall that every top dimensional cone of the Feynman fan corresponds to some triangulation of Γ . Any vector $\mathbf{g} \in V_\Gamma$ can be expanded as a sum of the generators of the cone that it is in using

$$\mathbf{g} = \sum_C \alpha_C(\mathbf{g}) \mathbf{g}_C, \tag{A.3}$$

where α_C are the headlight functions and \mathbf{g}_C are the g -vectors of the curves, C . Consider the function on V given by

$$Z = \exp \left(- \sum_C \alpha_C(\mathbf{t}) X_C \right), \tag{A.4}$$

where the sum in the exponent is over all open curves C . Let T be a triangulation corresponding to some top-dimensional cone, with curves C_1, \dots, C_E . Restricting Z to this cone gives

$$Z|_{\text{cone}} = \exp \left(- \sum_{i=1}^E \alpha_{C_i}(\mathbf{t}) X_{C_i} \right), \tag{A.5}$$

which follows from (A.3). Moreover, the generators of this top dimensional cone span a parallelepiped of unit volume, so there exist corresponding coordinates y'_1, \dots, y'_E such that $d^E y = d^E y'$ and so that any vector in this cone can be written as

$$\mathbf{g} = \sum_{i=1}^E y'_i \mathbf{g}_{C_i}. \tag{A.6}$$

The integral of Z over this cone is then

$$\int_{\text{cone}} d^E y Z = \int_{\geq 0} d^E y' \exp \left(\sum_{i=1}^E -y'_i X_{C_i} \right) = \prod_{i=1}^E \frac{1}{X_{C_i}}. \tag{A.7}$$

It follows from this that the partial amplitude (A.2) can be written as a curve integral over curve space:

$$\mathcal{A} = \int \frac{d^E \mathbf{t}}{\text{MCG}} \int \prod_{i=1}^L d^D \ell_i Z. \tag{A.8}$$

In this formula, we integrate over curve space modulo the action of the mapping class group. This ensures that we count each fatgraph Γ only once. We explain how to compute these curve integrals, with non-trivial MCG actions, in section 8.

B Factorization in detail

In the text, the factorization of the curve integral formula for integrands \mathcal{I} is stated in (6.30). This formula gives the residue of the pole $1/X_C$. To derive the formula, there are two possible cases to consider: either C is MCG-invariant, or not.

B.1 MCG invariant curve

Suppose C is MCG-invariant. The X_C pole arises from the part of the integral over the region of curve space where $\alpha_C > 0$. Since $\text{Stab}(C) = \text{MCG}(\Gamma)$, the MCG action has a well-defined restriction to this region and we have a well-defined curve integral

$$\mathcal{I}' = \int_{\alpha_C > 0} \frac{d^E t}{\text{MCG}} Z. \tag{B.1}$$

To compute \mathcal{I}' , take a triangulation containing C , with curves C, D_1, \dots, D_{E-1} . Take coordinates adapted to this cone:

$$\mathbf{g} = t_C \mathbf{g}_C + \sum_{i=1}^{n-1} t'_i \mathbf{g}_{D_i}. \tag{B.2}$$

By the unit volume property, the integration measure is

$$d^E t = dt_C d^{E-1} t'. \tag{B.3}$$

In these coordinates, the restriction of Z to this region is

$$Z|_{t_C > 0} = e^{-t_C X_C} \exp\left(-\sum_{D|C} \alpha_D X_D\right), \tag{B.4}$$

where the sum is over D that do not intersect C . For these curves, $\alpha_D(\mathbf{g} + \mathbf{g}_C) = \alpha_D(\mathbf{g})$, so that the only t_C -dependence is in the $\exp(-t_C X_C)$ factor. Write $\alpha'_D = \alpha_D|_{t_C=0}$, for the headlight functions restricted to $t_C = 0$. α'_D is the headlight function of D considered as a curve on the cut fatgraph Γ_C .

The t_C integral gives

$$\mathcal{I}' = \frac{1}{X_C} \int \frac{d^{E-1} t'}{\text{MCG}} Z_C, \tag{B.5}$$

where

$$Z_C = \exp\left(-\sum_{D|C} \alpha'_D X_D\right). \tag{B.6}$$

The full curve integral \mathcal{I} is $\mathcal{I} = \mathcal{I}' + \dots$, where the \dots has no X_C pole. So

$$\text{Res}_{X_C=0} \mathcal{I} = \int \frac{d^{E-1}t'}{\text{MCG}} Z_C, \tag{B.7}$$

where, on the r.h.s., P_C^μ is put on shell ($X_C \rightarrow 0$).

B.2 MCG non-invariant curve

If $\text{Stab}(C) < \text{MCG}$, we can use a Mirzakhani kernel to evaluate the $1/X_C$ pole. We choose C as one of the coset representatives, so that the Mirzakhani kernel is

$$\mathcal{K} = \frac{\alpha_C}{\rho} + \dots \tag{B.8}$$

Then

$$\int \frac{d^E t}{\text{MCG}} Z = \int \frac{d^E t}{\text{Stab}C} \frac{\alpha_C}{\rho} Z + \dots, \tag{B.9}$$

where the \dots are all terms without a $1/X_C$ pole. To guarantee that X_C only appears in the first term, we can choose the other coset representatives C_1, \dots, C_{L-1} so that all of these are curves that intersect C . We can put the $1/\rho$ in the numerator, by introducing an auxiliary integration variable ξ :

$$\int \frac{d^E t}{\text{MCG}} Z = \int_0^\infty d\xi \int \frac{d^E t}{\text{Stab}(C)} \alpha_C e^{-\xi \rho} Z + \dots \tag{B.10}$$

Changing variables as before, and integrating over t_C gives

$$\int \frac{d^E t}{\text{MCG}} Z = \int_0^\infty d\xi \frac{-1}{(X_C + \xi)^2} \int \frac{d^{E-1}t'}{\text{Stab}(C)} Z' + \dots, \tag{B.11}$$

where Z' is obtained from Z by shifting $X_D \mapsto X_D + \xi$ for all D in the Mirzakhani set. Finally, integrating over ξ , and using

$$\prod_{i=1}^m \frac{1}{X_i + \xi} = \sum_{i=1}^m \frac{1}{X_i + \xi} \prod_{j \neq i} \frac{1}{X_j - X_i}, \tag{B.12}$$

we find

$$\int \frac{d^E t}{\text{MCG}} Z \rightarrow \frac{1}{X_C} \int \frac{d^{E-1}t'}{\text{Stab}(C)} Z_C + \dots, \tag{B.13}$$

where $-\log Z_C$ is the curve action given by summing over all curves, D , compatible with C :

$$-\log Z_C = \sum_D \alpha_D X_D. \tag{B.14}$$

Note that this calculation does not apply if the integrand has higher poles in X_C , such as if X_C is a bubble propagator for a planar diagram.

C The surface Symanzik polynomials

Fixing an assignment of momenta to the curves gives explicit formulas for the all the propagator factors

$$X_C = \left(K_C^\mu + \sum_{a=1}^L h_C^a \ell_a^\mu \right)^2, \quad (\text{C.1})$$

in terms of one set of loop momentum variables ℓ_a^μ . In terms of these loop variables, the curve action,

$$S = \sum_C \alpha_C X_C, \quad (\text{C.2})$$

becomes

$$S = \ell_a^\mu A^{ab} \ell_b^\mu + 2B_\mu^a \ell_a^\mu + \mathcal{Z}, \quad (\text{C.3})$$

where A, B, \mathcal{Z} are all linear functions in the generalised Schwinger parameters:

$$A^{ab} = \sum_C h_C^a h_C^b \alpha_C \quad (\text{C.4})$$

$$B_\mu^a = \sum_C h_C^a \alpha_C K_{C\mu} \quad (\text{C.5})$$

$$\mathcal{Z} = \sum_C \alpha_C (K_C^2 + m^2) \quad (\text{C.6})$$

Performing the Gaussian integral over the ℓ_a variables, in D dimensions, gives

$$\mathcal{A} = \int \frac{d^E \mathbf{t}}{\text{MCG}} \left(\frac{\pi^L}{\det A} \right)^{\frac{D}{2}} \exp \left(B^T A^{-1} B - \mathcal{Z} \right). \quad (\text{C.7})$$

We define the surface Symanzik polynomials:

$$\mathcal{U} = \det A, \quad \text{and} \quad \frac{\hat{\mathcal{F}}_0}{\mathcal{U}} = -B^T A^{-1} B. \quad (\text{C.8})$$

These are the formulas used in the main text. Moreover, for the comparison with Schwinger parametrized Feynman integrals, it is useful to also define

$$\mathcal{F}_0 = \hat{\mathcal{F}}_0 + \mathcal{U} \sum_C \alpha_C K_C^2. \quad (\text{C.9})$$

In this appendix, we consider the explicit expansions of \mathcal{U} and \mathcal{F}_0 in monomials.

C.1 The first surface Symanzik

Since X^{ij} is linear in the parameters α_C , the determinant $\det X$ is homogeneous of degree L . For a set of curves $S = \{C_1, \dots, C_L\}$, let us find the coefficient in $\det A$ of the monomial

$$\alpha_S = \prod \alpha_{C_i}. \quad (\text{C.10})$$

By the definition of the determinant, this coefficient is

$$\det A = \dots + \alpha_S (\det h|_S)^2 + \dots, \tag{C.11}$$

where

$$\det h|_S = \epsilon_{i_1 \dots i_L} h_{C_1}^{i_1} \dots h_{C_L}^{i_L}. \tag{C.12}$$

Note that the ordering of the curves C_1, \dots, C_L does not matter, because this determinant only enters the formula for $\det A$ as a square.

We now make two observations: (1) $\det h|_S$ is only non-zero if the curves in S cut Γ to a tree graph; (2) for any conventional choice of loop variables (defined below), the determinants $\det h|_S$ are all either 0 or ± 1 . So the result is that \mathcal{U} is given by

$$\mathcal{U} = \sum_{\substack{S \text{ cuts } \Gamma \\ \text{to tree}}} \alpha_S. \tag{C.13}$$

For the first statement, consider $L = 1$. Then all curves have momenta of the form

$$P_C = h_C^1 \ell_1 + K_C^\mu. \tag{C.14}$$

If $h_C^1 = 0$, cutting Σ along C breaks it into two parts: one part with $L = 1$, and a second part with $L = 0$ (i.e. a disk). Whereas, if $h_C^1 \neq 0$, cutting Γ along C cuts the loop open, giving a new surface with $L = 0$ (i.e. a disk). So at 1-loop the first Symanzik polynomial is

$$\mathcal{U} = \sum_{\substack{C \text{ cuts } \Gamma \\ \text{to tree}}} \alpha_C (h_C^1)^2. \tag{C.15}$$

For $L > 1$, the determinant $\det h|_S$ is nonzero if and only if the linear transformation (in $H_1(\Gamma, \partial\Gamma)$ from $[L_1], \dots, [L_L]$ to $[C_1], \dots, [C_L]$) is invertible. By induction from the $L = 1$ case, this means that the curves in S cut Γ to a disk. So

$$\mathcal{U} = \sum_{\substack{S \text{ cuts } \Gamma \\ \text{to tree}}} \alpha_S (\det h|_S)^2. \tag{C.16}$$

Secondly, it turns out that $(\det h|_S)^2$ is either 0 or 1. We sketch how to prove this by fixing any genus g fatgraph with h trace-factor components. The loop order of such a fatgraph is

$$L = 2g + h - 1. \tag{C.17}$$

A natural basis of loop-carrying curves can be given by picking some $2g$ curves A_i, B_i wrapping the A, B -cycles of the graph, and $h - 1$ curves C_i connecting the h trace factors. These give a set, S , of L curves that cut Γ to a tree, so $(\det h|_S)^2 = 1$. Moreover, we can choose our momentum assignment such that

$$P_{A_i} = \ell_{2i-1}, \quad P_{B_i} = \ell_{2i}, \quad P_{C_i} = \ell_{2g+i}. \tag{C.18}$$

Now consider the momenta of Dehn twists of these curves. For instance, taking one of the C_i , a Dehn twist γ around one of its trace-factors gives a new curve

$$P_{\gamma C_i} = P_{C_i} \pm k_{\text{tf}}, \tag{C.19}$$

where k_{tf} is the total momentum of the trace factor. Moreover, any product of Dehn twists acting on a pair of A,B-cycles acts on their momenta as $\text{SL}_2\mathbb{Z}$:

$$\begin{bmatrix} \ell_{2i-1} \\ \ell_{2i} \end{bmatrix} \mapsto X \begin{bmatrix} \ell_{2i-1} \\ \ell_{2i} \end{bmatrix}, \tag{C.20}$$

for some $X \in \text{SL}_2\mathbb{Z}$. In this way, we find that the momenta of any set, S' , that cuts Γ to a tree, is obtained from the momenta of S via translations by non-loop momenta, and $\text{SL}_2\mathbb{Z}$ transformations. Both of which leave the determinant unchanged:

$$(\det h|_{S'})^2 = (\det h|_S)^2 = 1. \tag{C.21}$$

We conclude with some remarks about the expansion, (C.13), of \mathcal{U} into a sum of monomials α_S . We have described the sets of curves S as those which cut the fatgraph Γ to a tree fatgraph. It is equivalent to view the curves as classes on the corresponding surface. Then the sum is over sets of curves S that cut the surface to a disk. Moreover, when restricted to a top dimensional cone of the Feynman fan, corresponding to some particular Feynman graph Γ' , only the α_C for C an edge of Γ' are nonzero. It can be checked that the sum over sets S then becomes a sum over the complements of spanning trees of Γ' . In other words, the restriction of \mathcal{U} to any single cone is the first Symanzik polynomial of the corresponding graph.

C.2 The second surface Symanzik

The second surface Symanzik polynomial is

$$\mathcal{F}_0 = -\mathcal{U} B^T A^{-1} B + \mathcal{U} \sum_C \alpha_C K_C^2. \tag{C.22}$$

In this section we sketch how \mathcal{F}_0 can be expanded in monomials, in line with the calculation above for \mathcal{U} .

The second term of (C.22) can be expanded using (C.13) as

$$\mathcal{U} \sum_C \alpha_C K_C^2 = \sum_{\substack{S' \text{ cuts} \\ \text{to two disks}}} \alpha_{S'} \sum_{C \in S'} K_C^2 + \sum_{\substack{S \text{ cuts} \\ \text{to disk}}} \alpha_C \sum_{C \in S} \alpha_C K_C^2. \tag{C.23}$$

The first sum is over sets of $L + 1$ curves S' that cut Γ to two trees, or equivalently cut the surface to two disks. The second sum is over sets of L curves S that cut Γ to a tree, or equivalently cut the surface to a disk.

To understand $\hat{\mathcal{F}}_0$, the first term of (C.22), we can use the Laplace formula:

$$\left(A^{-1}\right)^{ij} = \frac{(-1)^{i+j}}{\det A} |A|^{ij}, \tag{C.24}$$

where $|A|^{ij}$ is the i, j minor. Since $\mathcal{U} = \det A$,

$$\hat{\mathcal{F}}_0 = -2 \sum_{C \neq D} \alpha_C \alpha_D K_C \cdot K_D \sum_{i,j} (-1)^{i+j} h_C^i h_D^j |A|_{ij} - \sum_C \alpha_C^2 K_C^2 h_C^i h_C^j |A|_{ij}. \tag{C.25}$$

The minors of A are

$$|A|_{ij} = \sum_S \sum_{C \in S} \frac{\alpha_S}{\alpha_C} |h_S|_C^i |h_S|_C^j, \tag{C.26}$$

where $|h_S|_C^i$ is the (i, C) minor of the matrix $h|_S = [h_{C_1}^i | \dots | h_{C_L}^i]$. By the definition of the determinant,

$$\sum_{i=1}^L (-1)^i h_D^i |h_S|_C^i = \det h_{S_{C \rightarrow D}}, \tag{C.27}$$

where $S_{C \rightarrow D}$ is the set obtained from S by replacing C with D . We recall also from the previous section that $\det h_S = \pm 1$. It then follows that the second term of (C.25) cancels the second term of (C.23).

We are left with

$$\mathcal{F}_0 = \sum_{\substack{S' \text{ cuts} \\ \text{to two disks}}} \alpha_{S'} \left(\sum_{C \in S'} K_C^2 - 2 \sum_{C, D \in S'} (\det h_{S' \setminus C}) (\det h_{S' \setminus D}) K_C \cdot K_D \right). \tag{C.28}$$

Again, these determinants are $\det h_{S' \setminus C} = \pm 1$. The exact sign depends on the choice of how to orient C . Note that both the vectors h_C^i and the momenta K_C^μ change sign under orientation reversal of C . For a fixed ordered subset S' , these orientations can be chosen so that the coefficients are all positive, say, so that

$$\mathcal{F}_0 = \sum_{\substack{S' \text{ cuts } \Gamma \\ \text{to 2 trees}}} \alpha_{S'} \left(\sum_{C \in S'} K_C^\mu \right)^2, \tag{C.29}$$

where the signs have been absorbed into a choice of orientation of the momenta K_C .

Once again, restricting to a top dimensional cone of the Feynman fan, \mathcal{F}_0 restricts to a sum of 2-forests spanning the corresponding graph, and the monomial is the product over the complement.

D The recursion formula

For a fatgraph Γ , the curve integral for integrands is

$$\mathcal{I} = \int \frac{d^E t}{\text{MCG}} Z, \tag{D.1}$$

with

$$-\log Z = \sum_C \alpha_C X_C. \tag{D.2}$$

For some trace factor β of Γ , we have the set of curves \mathcal{S} that have one or two endpoints in β . Under the MCG, this set has some, say k , coset representatives, C_i ($i = 1, \dots, k$). Then

$$\mathcal{I} = \int \frac{d^E t}{\text{MCG}} Z = \sum_{i=1}^k \int \frac{d^E t}{\text{Stab}(C_i)} \frac{\alpha_{C_i}}{\rho} Z, \tag{D.3}$$

where

$$\rho = \sum_{C \in \mathcal{S}} \alpha_C. \tag{D.4}$$

Introducing an auxiliary parameter, ξ , we re-write this as

$$\mathcal{I} = \sum_{i=1}^k \int_0^\infty d\xi \int \frac{d^E t}{\text{MCG}} \alpha_{C_i} Z(\xi). \tag{D.5}$$

where the new integrand is

$$-\log Z(\xi) = \sum_{C \in \mathcal{S}} \alpha_C(X_C + \xi) + \sum_{D \notin \mathcal{S}} \alpha_D X_D. \tag{D.6}$$

Integrating over the α_{C_i} direction in each term, the curve integral gives

$$\mathcal{I} = \sum_{i=1}^k \int_0^\infty d\xi \frac{-1}{(X_{C_i} + \xi)^2} \int \frac{d^{n-1} t'}{\text{Stab}(C_i)} Z'(\xi), \tag{D.7}$$

where

$$-\log Z'(\xi) = \sum_{C \in \mathcal{S}, C \neq C_i} \alpha'_C(X_C + \xi) + \sum_{D \notin \mathcal{S}} \alpha'_D X_D, \tag{D.8}$$

and α'_C are the headlight functions obtained after integrating out the \mathbf{g}_{C_i} direction. These are the headlight functions for the fatgraph Γ_{C_i} obtained by cutting along C_i .

Note that we can evaluate the ξ integral using identities such as

$$\prod_{i=1}^m \frac{1}{X_i + t} = \sum_{i=1}^m \frac{1}{X_i + t} \prod_{j \neq i} \frac{1}{X_j - X_i}. \tag{D.9}$$

When all the X_C propagator factors are distinct (i.e. there are no higher poles), we can perform the integral to find

$$\mathcal{I} = \sum_{i=1}^k \frac{1}{X_{C_i}} \int \frac{d^{n-1} t'}{\text{Stab}(C_i)} Z'(-X_{C_i}), \tag{D.10}$$

E Recursion examples

E.1 The 3-point non-planar 1-loop amplitude

Take Γ to be the 3-point non-planar 1-loop diagram considered in section 9.1. The curves are $C_{12}^n, C_{13}^n, C_{22}, C_{33}$. For the Mirzakhani method, we have two cosets, with representatives C_{12}^0, C_{13}^0 . Cutting Γ along C_{12}^0 gives a 5-point tree fatgraph $\Gamma_{C_{12}^0}$. The curves compatible with C_{12}^0 are

$$C_{12}^1, C_{13}^0, C_{12}^{-1}, C_{13}^{-1}, C_{22}. \tag{E.1}$$

The global forward limit then computes I_Γ as

$$I_\Gamma = \frac{1}{X_{12}^0} I_{\Gamma_{C_{12}^0}}(X_{12}^1 - X_{12}^0, X_{13}^0 - X_{12}^0, X_{12}^{-1} - X_{12}^0, X_{13}^{-1} - X_{12}^0, X_{22}) + (2 \leftrightarrow 3). \tag{E.2}$$

But the 5-point tree amplitude is

$$I(X_1, X_2, X_3, X_4, X_5) = \sum_{i=1}^5 \frac{1}{X_i X_{i+1}}. \tag{E.3}$$

So the integrand is

$$I_\Gamma = \frac{1}{X_{12}^0(X_{12}^1 - X_{12}^0)(X_{13}^0 - X_{12}^0)} + \frac{1}{X_{12}^0(X_{13}^0 - X_{12}^0)(X_{12}^{-1} - X_{12}^0)} + \frac{1}{X_{12}^0(X_{12}^{-1} - X_{12}^0)(X_{13}^{-1} - X_{12}^0)} \\ + \frac{1}{X_{12}^0(X_{13}^{-1} - X_{12}^0)X_{22}} + \frac{1}{X_{12}^0 X_{22}(X_{12}^1 - X_{12}^0)} + (2 \leftrightarrow 3). \tag{E.4}$$

The momenta are explicitly

$$P_{12}^n = \ell + nk_1, \quad P_{13}^n = \ell + k_2 + nk_1, \quad P_{22} = k_1, \quad P_{33} = k_1 + k_2. \tag{E.5}$$

E.2 The 2-loop vacuum at genus one

The 2-loop genus 1 vacuum amplitude has already been computed in section 9.2. Take again Γ to be the 2-loop genus one vacuum diagram. The curves of Γ are $C_{p/q}$, with momentum

$$P_{p/q} = p\ell + q\ell'. \tag{E.6}$$

Every curve $C_{p/q}$ is in the same MCG-orbit. Pick, say, $C_{0/1}$ as the coset representative. The curves compatible with $C_{0/1}$ are $C_{1/n}$ for $n \in \mathbb{Z}$. Cutting Γ along $C_{0/1}$ gives a 1-loop non-planar diagram $\Gamma_{C_{0/1}}$, and the curves $C_{1/n}$ can be identified with the curves we called ‘ C_n ’ in the previous example. Applying the global forward limit once gives

$$I_\Gamma = \frac{1}{X_{0/1}} I_{\Gamma_{C_{0/1}}}(X_{1/n} - X_{0/1}). \tag{E.7}$$

However, we have already computed the 1-loop non-planar integrand, and found, up to loop-momentum shifts, that it is given by

$$I_{\Gamma_{C_{0/1}}}(X_n) = \frac{1}{X_0 X_1}. \tag{E.8}$$

Using this result in (E.7) gives

$$I_\Gamma = \frac{1}{X_{0/1}(X_{1/0} - X_{0/1})(X_{1/1} - X_{0/1})}. \tag{E.9}$$

Loop re-definitions of ℓ and ℓ' can be used to cyclically permute the labels $0/1, 1/0, 1/1$. Summing over the possible three cyclic permutations (and dividing by 3) gives

$$I_\Gamma = \frac{1}{3} \frac{1}{X_{0/1} X_{1/0} X_{1/1}}. \tag{E.10}$$

The $1/3$ factor is expected because $|\text{Aut}(\Gamma)| = 3$. We therefore recover $1/3$ of the Feynman integral of the sunrise vacuum diagram.

E.3 A comment on the 1-loop planar amplitudes

Our formula for the 1-loop planar amplitudes can be computed directly, without topological recursion. The global Schwinger formula gives a well defined loop integrand for these amplitudes, without linearized propagators. However, we can arrive at a forward-limit-like formula for the 1-loop integrand by inserting the ‘trivial’ Mirzakhani kernel

$$1 = \sum_{i=1}^n \frac{\alpha_{0i}}{\sum_{i=j}^n \alpha_{0j}} \tag{E.11}$$

into the curve integral. Here, α_{0i} is the headlight function of C_{0i} , the curve from i to the internal loop boundary, 0. Equation (E.11) then allows us to write the 1-loop planar n -point amplitude as a sum of n disk amplitudes, with linearized propagators. Evaluating the integral, using the recursion (10.4), the integrand is

$$I_{1\text{loop}}(1, \dots, n) = \sum_{i=1}^n \frac{-1}{X_{0i}} A(12 \dots i 0 i \dots n) \Big|_{X_{0j} \mapsto X_{0j} - X_{0i}}, \tag{E.12}$$

where $A(12 \dots n)$ are the tree-level partial amplitudes, but now with linearized propagators.

F Details for the non-planar 1-loop propagator

The matrix for the curves C_n with $n \geq 0$ is

$$M_n = LD_x(LD_yRD_x)^n R. \tag{F.1}$$

Taking the transpose, we see that $M_n^T = M_n$. In particular,

$$M_0 = \begin{bmatrix} 1 & 1 \\ 1 & 1+x \end{bmatrix}. \tag{F.2}$$

Given M_0 , we can compute M_n using

$$M_{n+1} = M_n B_{+1}, \quad \text{where } B_{+1} = R^{-1}LD_yRD_xR = \begin{bmatrix} 0 & -xy \\ 1 & 1+x+xy \end{bmatrix}. \tag{F.3}$$

It follows that we can write

$$M_n = \begin{bmatrix} F_{n-2} & F_{n-1} \\ F_{n-1} & F_n \end{bmatrix}, \tag{F.4}$$

where

$$F_{n+2} = (1+x+xy)F_{n+1} - xyF_n, \tag{F.5}$$

with initial conditions $F_{-2} = 1, F_{-1} = 1$. The first few examples are

$$F_0 = 1+x, \tag{F.6}$$

$$F_1 = 1+2x+x^2+x^2y, \tag{F.7}$$

$$F_2 = 1+3x+3x^2+x^3+2x^2y+2x^3y+x^3y^2. \tag{F.8}$$

Similarly, the matrix for the curves C_n with $n < 0$ is given by

$$M_n = RD_y(RD_xLD_y)^{-n-1}L, \quad n < 0. \quad (\text{F.9})$$

These matrices are again symmetric, and

$$M_{-1} = \begin{bmatrix} 1+y & y \\ y & y \end{bmatrix}. \quad (\text{F.10})$$

We can evaluate M_n using

$$M_{n-1} = M_n B_{-1}, \quad \text{where } B_{-1} = L^{-1}RD_xLD_yL = \begin{bmatrix} 1+x+xy & xy \\ -1 & 0 \end{bmatrix}. \quad (\text{F.11})$$

This implies that M_n ($n < 0$) has the form,

$$M_n = \begin{bmatrix} G_n & xyG_{n+1} \\ xyG_{n+1} & (xy)^2G_{n+2} \end{bmatrix}, \quad (\text{F.12})$$

where the polynomials G_n are determined by the recursion

$$G_n = (1+x+xy)G_{n+1} - xyG_{n+2}, \quad (\text{F.13})$$

with initial condition $G_1 = 1/(x^2y)$ and $G_0 = 1/x$. The first few polynomials are

$$G_{-1} = 1+y, \quad (\text{F.14})$$

$$G_{-2} = 1+x+2xy+xy^2, \quad (\text{F.15})$$

$$G_{-3} = (1+x+xy)^2 + x^2y(1+y)^2. \quad (\text{F.16})$$

We now need to compute the tropicalizations of the polynomials F_n ($n \geq -2$) and G_n ($n \leq 1$). Write

$$f_n = \text{Trop } F_n, \quad \text{and} \quad g_n = \text{Trop } G_n. \quad (\text{F.17})$$

Then, for $n \geq 0$, we find

$$f_n = \max(0, (n+1)x, (n+1)x+ny), \quad (\text{F.18})$$

which follows by induction using that

$$f_{n+2} = \max(\max(0, x, x+y) + f_{n+1}, \max(0, x+y) + f_n). \quad (\text{F.19})$$

Similarly, for $n \leq -1$,

$$g_n = \max(0, -(n+1)x, -(n+1)x-ny). \quad (\text{F.20})$$

We also have that

$$f_{-2} = 0, \quad f_{-1} = 0, \quad g_1 = -2x-y, \quad g_0 = -x. \quad (\text{F.21})$$

The headlight functions are

$$\alpha_n = -f_n + 2f_{n-1} - f_{n-2}, \quad n \geq 0, \quad (\text{F.22})$$

$$\alpha_n = -g_n + 2g_{n+1} - g_{n+2}, \quad n < 0. \quad (\text{F.23})$$

Data Availability Statement. This article has no associated data or the data will not be deposited.

Code Availability Statement. This article has no associated code or the code will not be deposited.

Open Access. This article is distributed under the terms of the Creative Commons Attribution License ([CC-BY4.0](https://creativecommons.org/licenses/by/4.0/)), which permits any use, distribution and reproduction in any medium, provided the original author(s) and source are credited.

References

- [1] S.J. Parke and T.R. Taylor, *An amplitude for n gluon scattering*, *Phys. Rev. Lett.* **56** (1986) 2459 [[INSPIRE](#)].
- [2] Z. Bern, L.J. Dixon, D.C. Dunbar and D.A. Kosower, *One loop n point gauge theory amplitudes, unitarity and collinear limits*, *Nucl. Phys. B* **425** (1994) 217 [[hep-ph/9403226](#)] [[INSPIRE](#)].
- [3] Z. Bern, L.J. Dixon, D.C. Dunbar and D.A. Kosower, *Fusing gauge theory tree amplitudes into loop amplitudes*, *Nucl. Phys. B* **435** (1995) 59 [[hep-ph/9409265](#)] [[INSPIRE](#)].
- [4] E. Witten, *Perturbative gauge theory as a string theory in twistor space*, *Commun. Math. Phys.* **252** (2004) 189 [[hep-th/0312171](#)] [[INSPIRE](#)].
- [5] F. Cachazo, P. Svrcek and E. Witten, *MHV vertices and tree amplitudes in gauge theory*, *JHEP* **09** (2004) 006 [[hep-th/0403047](#)] [[INSPIRE](#)].
- [6] R. Britto, F. Cachazo and B. Feng, *New recursion relations for tree amplitudes of gluons*, *Nucl. Phys. B* **715** (2005) 499 [[hep-th/0412308](#)] [[INSPIRE](#)].
- [7] 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 [[hep-th/0501052](#)] [[INSPIRE](#)].
- [8] N. Arkani-Hamed and J. Trnka, *The amplituhedron*, *JHEP* **10** (2014) 030 [[arXiv:1312.2007](#)] [[INSPIRE](#)].
- [9] N. Arkani-Hamed, Y. Bai, S. He and G. Yan, *Scattering forms and the positive geometry of kinematics, color and the worldsheet*, *JHEP* **05** (2018) 096 [[arXiv:1711.09102](#)] [[INSPIRE](#)].
- [10] G. Salvatori, *1-loop amplitudes from the halohedron*, *JHEP* **12** (2019) 074 [[arXiv:1806.01842](#)] [[INSPIRE](#)].
- [11] N. Arkani-Hamed, S. He, G. Salvatori and H. Thomas, *Causal diamonds, cluster polytopes and scattering amplitudes*, *JHEP* **11** (2022) 049 [[arXiv:1912.12948](#)] [[INSPIRE](#)].
- [12] N. Arkani-Hamed, S. He and T. Lam, *Stringy canonical forms*, *JHEP* **02** (2021) 069 [[arXiv:1912.08707](#)] [[INSPIRE](#)].
- [13] N. Arkani-Hamed, H. Frost, G. Salvatori, P.-G. Plamondon and H. Thomas, *The curvy world of curves*, in preparation.
- [14] V.V. Fock and A.B. Goncharov, *Moduli spaces of local systems and higher Teichmüller theory*, [math/0311149](#) [[INSPIRE](#)].
- [15] R.C. Penner, *Decorated Teichmüller theory*, volume 1, European Mathematical Society (2012).
- [16] N. Arkani-Hamed, H. Frost, G. Salvatori, P.-G. Plamondon and H. Thomas, *The combinatorical string*, in preparation.

- [17] H. Derksen and J. Weyman, *An introduction to quiver representations*, American Mathematical Society, U.S.A., November 2017.
- [18] N. Haupt, *Euler characteristics of quiver Grassmannians and Ringel-Hall algebras of string algebras*, *Algebr. Represent. Theory* **15** (2010) 755.
- [19] G. 't Hooft, *A planar diagram theory for strong interactions*, *Nucl. Phys. B* **72** (1974) 461 [[INSPIRE](#)].
- [20] V.V. Fock and A.B. Goncharov, *Dual Teichmüller and lamination spaces*, [math/0510312](#) [[INSPIRE](#)].
- [21] S. Fomin and A. Zelevinsky, *Cluster algebras IV: coefficients*, *Compos. Math.* **143** (2007) 112 [[math/0602259](#)] [[INSPIRE](#)].
- [22] S. Fomin and D. Thurston, *Cluster algebras and triangulated surfaces. Part II: Lambda lengths*, *Mem. Amer. Math. Soc.* **255** (2018) [[arXiv:1210.5569](#)].
- [23] M. Mirzakhani, *Simple geodesics and Weil-Petersson volumes of moduli spaces of bordered Riemann surfaces*, *Invent. Math.* **167** (2006) 179 [[INSPIRE](#)].
- [24] B. Farb and D. Margalit, *A primer on mapping class groups*, Princeton University Press, Princeton, NJ, U.S.A. (2012).
- [25] F.C.S. Brown, *Multiple zeta values and periods of moduli spaces $\overline{\mathfrak{M}}_{0,n}$* , *Ann. Sci. École Norm. Sup.* **42** (2009) 371.
- [26] N. Arkani-Hamed, S. He and T. Lam, *Cluster configuration spaces of finite type*, *SIGMA* **17** (2021) 092 [[arXiv:2005.11419](#)] [[INSPIRE](#)].
- [27] N. Arkani-Hamed, H. Frost, G. Salvatori, P.-G. Plamondon and H. Thomas, *All n paper*, in preparation.
- [28] S. He and E.Y. Yuan, *One-loop scattering equations and amplitudes from forward limit*, *Phys. Rev. D* **92** (2015) 105004 [[arXiv:1508.06027](#)] [[INSPIRE](#)].
- [29] J. Agerskov, N.E.J. Bjerrum-Bohr, H. Gomez and C. Lopez-Arcos, *One-loop Yang-Mills integrands from scattering equations*, *Phys. Rev. D* **102** (2020) 045023 [[arXiv:1910.03602](#)] [[INSPIRE](#)].
- [30] F. Porkert and O. Schlotterer, *One-loop amplitudes in Einstein-Yang-Mills from forward limits*, *JHEP* **02** (2023) 122 [[arXiv:2201.12072](#)] [[INSPIRE](#)].
- [31] Y. Geyer and R. Monteiro, *Gluons and gravitons at one loop from ambitwistor strings*, *JHEP* **03** (2018) 068 [[arXiv:1711.09923](#)] [[INSPIRE](#)].
- [32] Y. Geyer and R. Monteiro, *Two-loop scattering amplitudes from ambitwistor strings: from genus two to the nodal Riemann sphere*, *JHEP* **11** (2018) 008 [[arXiv:1805.05344](#)] [[INSPIRE](#)].
- [33] Y. Geyer, R. Monteiro and R. Stark-Muchão, *Two-loop scattering amplitudes: double-forward limit and colour-kinematics duality*, *JHEP* **12** (2019) 049 [[arXiv:1908.05221](#)] [[INSPIRE](#)].
- [34] M. Borinsky, *Tropical Monte Carlo quadrature for Feynman integrals*, *Ann. Inst. H. Poincaré D Comb. Phys. Interact.* **10** (2023) 635 [[arXiv:2008.12310](#)] [[INSPIRE](#)].
- [35] M. Borinsky, H.J. Munch and F. Tellander, *Tropical Feynman integration in the Minkowski regime*, *Comput. Phys. Commun.* **292** (2023) 108874 [[arXiv:2302.08955](#)] [[INSPIRE](#)].
- [36] Z. Bern, J.J.M. Carrasco and H. Johansson, *Perturbative quantum gravity as a double copy of gauge theory*, *Phys. Rev. Lett.* **105** (2010) 061602 [[arXiv:1004.0476](#)] [[INSPIRE](#)].