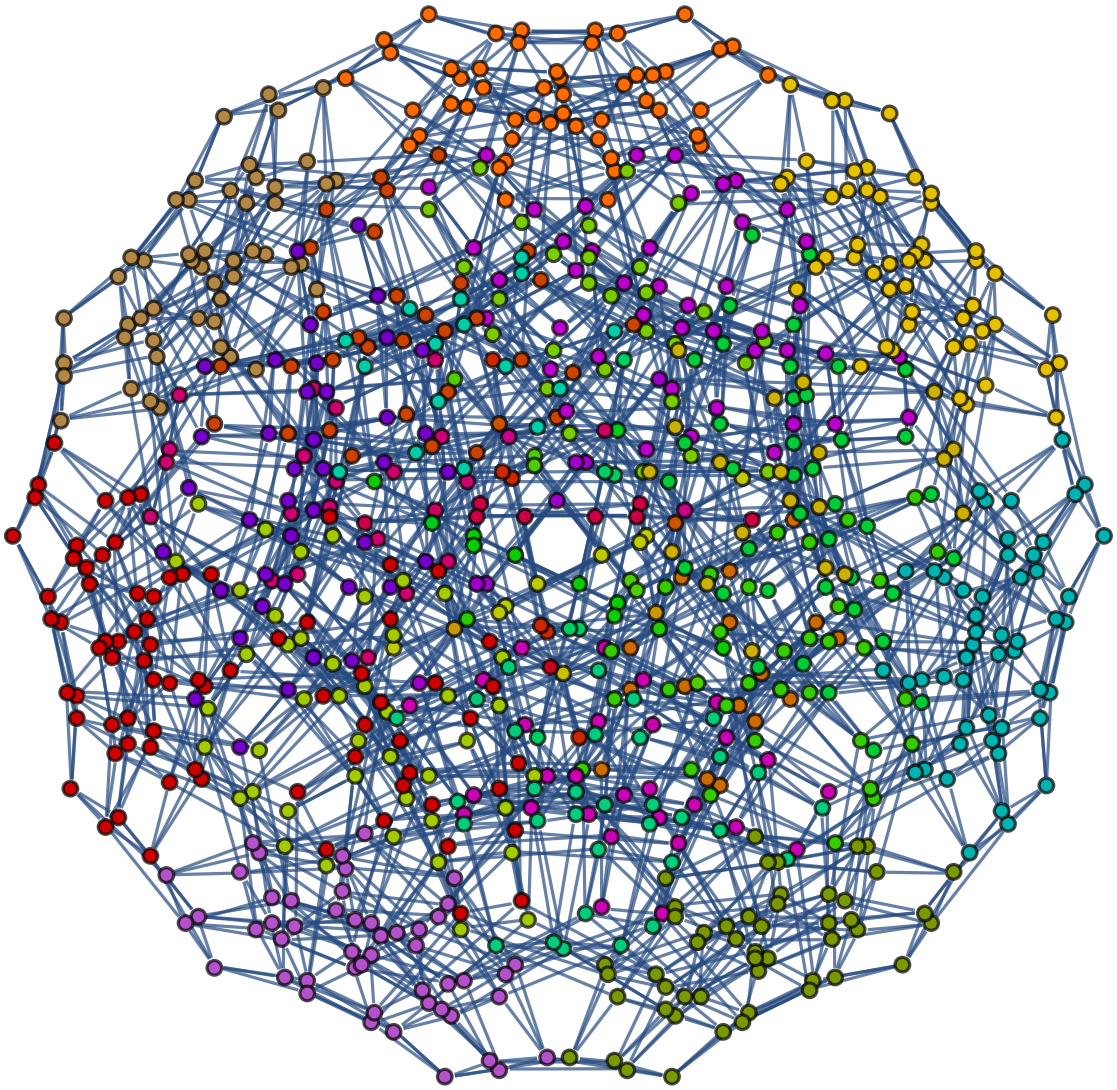
Senior Thesis

Brown University

Department of Physics

Cluster Algebra Structures for Scattering Amplitudes in $\mathcal{N} = 4$ Super Yang-Mills



Author:
Daniel Parker

Advisor:
Prof. Anastasia Volovich

April 2015

Contents

1 Introduction **4**

2 Yang-Mills Theory **7**
2.1 Motivation and the Lagrangian . . . 7
2.2 Quantization . . . 10

3 Kinematics for Scattering Amplitudes **13**
3.1 Spinor Helicity . . . 13
3.2 Twistors and Momentum Twistors . . . 14
3.3 Configurations . . . 15
3.4 Grassmannians and Plücker Coordinates . . . 16

4 Review of Amplitudes in $\mathcal{N} = 4$ SYM **17**
4.1 The Parke-Taylor Formula . . . 18
4.2 Supersymmetry . . . 18
4.3 The MHV Classification . . . 19
4.4 Loops and the BDS Ansatz . . . 19

5 Cluster Algebras **20**
5.1 What is a Cluster Algebra? . . . 21
5.2 The A_2 Cluster Algebra . . . 24
5.3 General Properties of Cluster Algebras . . . 27
5.4 More Examples of Cluster Algebras . . . 30
5.5 Cluster $\mathcal{X}$-coordinates . . . 35
5.6 Grassmannian Cluster Algebras . . . 36

6 Polylogarithms **38**
6.1 Classical Polylogarithms . . . 38
6.2 Iterated Integrals . . . 40
6.3 Shuffle Algebras and Lyndon Words . . . 41
6.4 The Symbol and Chen's Theorem . . . 44
6.5 Generalized Polylogarithms . . . 46
6.6 Examples of Symbols . . . 48

7 Results on Cluster Functions **52**
7.1 The Hopf Algebra of Polylogarithms . . . 52
7.2 Examples of the Symbol Coproduct . . . 54
7.3 Cluster Polylogarithms on A_2: The A_2 Function . . . 59
7.4 A Computationally Efficient Basis for Finding Cluster Functions at Weight 4 61
7.5 Dictionary of Cluster Polylogarithms . . . 63
7.6 Counting $\mathcal{X}$-coordinates, A_2 and A_3 Functions . . . 64
7.7 An Algorithm to Factor Products of $\mathcal{X}$-coordinates . . . 65

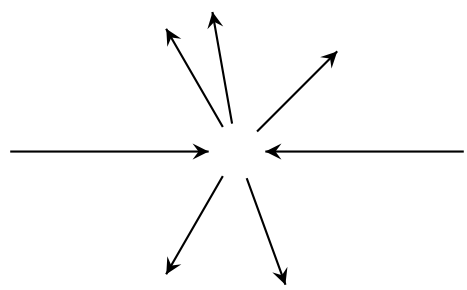
8 Cluster Bases **67**
8.1 The Basis Problem 68
8.2 The Spaces M, P, R and Q 68
8.3 Brown's Theorem 70
8.4 The Hedgehog Theorem 71

Bibliography **77**

1 Introduction

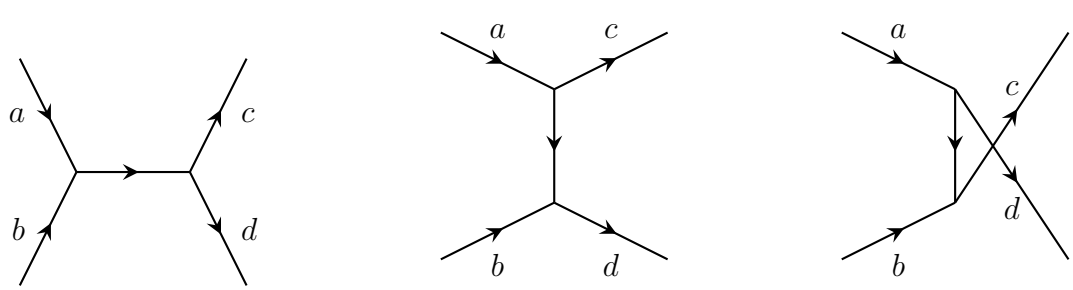
Scattering amplitudes are the heart of particle physics. They are a bridge between theory and experiment. Understanding them is necessary to interpret experimental results from particle accelerators. On the theoretical side, they provide a bountiful source of mathematical structures and connections. The study of scattering amplitudes is a well-established subject — dating back at least to the 1950s — and remains an area of active work, with many unsolved problems. Indeed, the last decade has seen the rise of "amplitudeology", a program for uncovering the hidden structure of scattering amplitudes. This thesis focuses on one aspect of scattering amplitudes: the structure of MHV remainder functions of 2-loop, n-particle, planar limit $\mathcal{N} = 4$ Super Yang-Mills amplitudes. (These terms will be defined precisely.) These have a fascinating mathematical structure, with many connections to the subjects of *cluster algebras* and *polylogarithms*.

What is a scattering amplitude? Imagine a particle collider. The collider accelerates subatomic particles nearly to light speed and smashes them into other particles. If these particles are energetic enough, they then break apart and form other particles, which fly off ("scatter") in all directions.

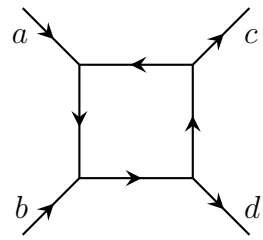


The energies and momenta of all these particles can be measured. Therefore a theorist's job is to predict and explain such measurements. A typical question is: *if* 3 *particles with known energies and momenta go in, then what is the probability of getting* 7 *particles out with definite energies and momenta?* Such probabilities are called *scattering amplitudes*.

Scattering amplitudes are calculated within the framework of quantum field theory. The most common technique for calculating them is the method of *Feynman diagrams*. Very roughly, the method works as follows. One draws a diagram with incoming particles on the left, represented by arrows, and the outgoing particles on the right, and then connects them in all possible ways. For instance, suppose particles a and b are incoming and particles c and d are outgoing. The three simplest possible diagrams are:



Other more complicated diagrams are also possible, such as this one with a *loop*.



In general, there are infinitely many such diagrams that can be drawn.

Feynman rules give the value of each Feynman diagram. The scattering amplitude is the sum of the values of all possible Feynman diagrams. In perturbation theory, the more complicated diagrams all have very small values, so the scattering amplitude can be approximated well just by looking at the simplest diagrams — ones with just a few loops. This is fortunate, because the difficulty of computing the value of a Feynman diagram increases immensely as loops are added.

Much of amplitudeology focuses on a particular quantum field theory: $\mathcal{N} = 4$ Super Yang-Mills (SYM) theory. This theory falls in a kind of "Goldilocks zone"; it is complicated enough to be very interesting, but not so complicated that it can't be understood. Although $\mathcal{N} = 4$ SYM is not a physically realistic theory, it is very similar to quantum chromodynamics, which describes the strong nuclear force. The hope is that understanding the scattering amplitudes of $\mathcal{N} = 4$ SYM will provide general principles that work in any quantum field theory.

Recent work by Golden, Goncharov, Spradlin, Volovich, *et al.* follows a single theme:

$$\boxed{\text{Particular scattering amplitudes in } \mathcal{N} = 4 \text{ SYM can be written as sums of polylogarithms in variables with a cluster algebra structure.}} \tag{1.1}$$

Polylogarithms are special functions which generalize the logarithm, and cluster algebras are a new type of algebraic structure discovered about a decade ago. Both of these mathematical structures are intriguing in their own right, with deep connections to scattering amplitudes. Exploring these connections has already been highly productive. As an example, a heroic calculation by Del Duca, Duhr and Smirnov established an formula for the most complex part of the 2-loop, 6-particle diagrams [1]. The expression is 17 pages long. Using mathematical techniques imported from the study of polylogarithms, Goncharov *et al.* reduced this expression to a single line [2]. Using additional cluster algebraic structure, Golden *et al.* gave a similar formula for 7 particles that was only slightly more complex [3]. The moral of this story is that deeply understanding the theme (1.1) will reveal simplifications and new insights into scattering amplitudes.

This thesis attempts to provide a bridge between quantum field theory, as it might be taught in a first semester graduate course, and current research in this topic. The topics discussed are as follows.

- Section 2 introduces Yang-Mills theory. This is fairly standard material which can be found in quantum field theory textbooks. Unlike the rest of the thesis, no attempt is made to be self-contained here; a semester course in field theory is probably required for this section.

- Section 3 describes kinematics for scattering amplitudes. Since amplitudes are functions of the energy and momentum of the ingoing and outgoing particles, they are usually written as equations in the four-vector momenta of those particles. However, it turns out to be much more convenient to describe amplitudes as functions on Grassmannians. (Grassmannians can be thought of as the space of k-dimensional planes inside n-dimensional space.) This section describes how one gets from four-vectors to Grassmannians.

- Section 4 is a very brief review of scattering amplitudes in $\mathcal{N} = 4$ SYM. This is an entire subject unto its own, and books have been written on the subject. The primary goal here is to introduce terminology and set out which questions are answered.

- Section 5 introduces cluster algebras. Cluster algebras are a very new subject, which is both a blessing and a curse. On the other hand, there are not yet standard pedagogical treatments of the subject, partly because the basic definitions are still in flux. This section attempts a complete explanation of cluster algebras starting "from scratch" and developing all the necessary concepts for physics. No background is assumed or, hopefully, required.

- Section 6 introduces polylogarithms. Though one can think of polylogarithms as just another special function, they have a deep theory. Research in this area — much of it by Professor Goncharov — tends to be much more mathematically technical than cluster algebras. The section gives an overview of polylogarithms at the level used in physics.

- The last two sections describe novel results by the author and current research. Much of this work was done in collaboration with Adam Scherlis. Section 7 describes "cluster functions", sums of polylogarithms with cluster algebraic structure. Section 8 gives some results towards finding bases for spaces of polylogarithms, a task complicated by the considerable number of functional identities between polylogarithms.

I would like to thank my thesis advisor Professor Volovich for being a superb mentor and for introducing me to this beautiful subject. I would also like to thank Professor Spradlin and Dr. Golden for answering all of my questions, and teaching me so much. I further thank my all friends and family for their support and help. Lastly, I thank Adam Scherlis for showing me how wonderful physics is and my father for inspiring me to study science.

2 Yang-Mills Theory

Yang-Mills theory, and Gauge theories more generally, describe how the electromagnetic, weak, and strong forces work in the Standard Model. This is a very large subject, and this section only describes a small part of it, focusing on motivation for the Lagrangian, how that Lagrangian is quantized and the associated Feynman rules..

2.1 Motivation and the Lagrangian

This discussion attempts to motivate the Yang-Mills Lagrangian from a mathematical perspective. The discussion is conceptual rather than rigorous, and follows treatments in [4, 5, 6].

It's handy to first consider a special case. Consider a wavefunction on Minkowski space $\mathbb{M}$ that is valued in the vector space $V = \mathbb{C}^n$. Two-component wavefunctions in quantum mechanics are an example of such wavefunctions, with values in $\mathbb{C}^2$. In principle, though, any vector space can be used. To construct a quantum field theory, a Lagrangian (density) is needed. This is a real-valued function of the wavefunction, $\mathcal{L}(\psi)$. The usual choice is

$$\mathcal{L}(\psi) = \|d\psi\|^2 + m\|\psi\|^2 \tag{2.1}$$

where d is the gradient and $\|\cdot\|$ is the Euclidean norm on V. If a basis $\{e_i\}_{i=1}^n$ is chosen for V, then $\psi = \psi^i e_i$ and the Lagrangian becomes

$$\mathcal{L}(\psi) = \partial^\mu \psi_i \partial_\mu \overline{\psi}^i + m\psi_i \overline{\psi}^i. \tag{2.2}$$

where $\mu = 0, \ldots, 3$ is a space-time index.

One mandates that the Lagrangian should not depend on the choice of basis; the basis of V is arbitrary, so a different choice should not yield different physics. Again in the case of spinors, there is no preferred direction for "spin up" and therefore no special basis for the space of spinors. One can choose any arbitrary direction for "spin up" and construct a basis from that choice. Regardless of the choice of "up", the physics is the same. The same is true in the general case $V = \mathbb{C}^n$.

More precisely, suppose that $U \in U(V) \cong U(n)$, the Lie group of unitary transformations of $\mathbb{C}^n$. In particular, $UU^\dagger = I$, and $\|Uv\|^2 = \|v\|$ for any $v \in V$. Then the wavefunction transforms under U as

$$\psi(x) \mapsto U\psi(x). \tag{2.3}$$

Actually, one can take V to be any vector space with a $U(n)$ representation on it. Since the gradient is a linear function on V, it transforms as

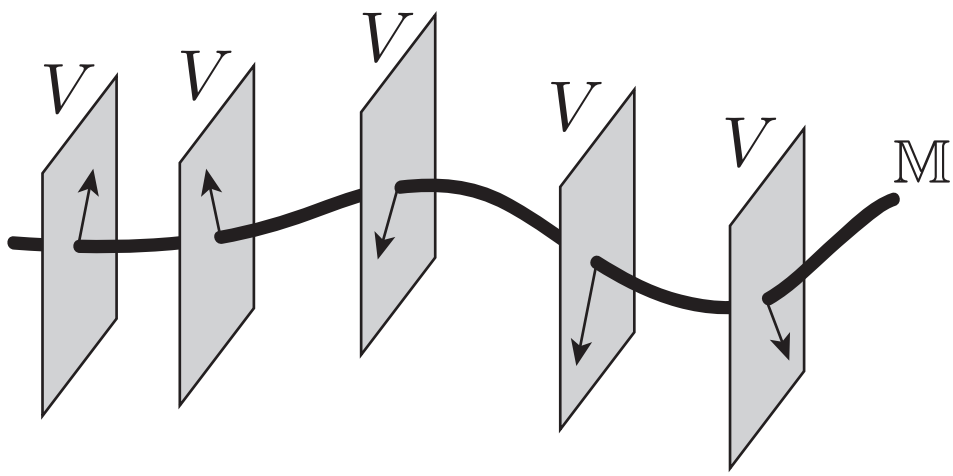
$$d \mapsto UdU^\dagger. \tag{2.4}$$

Therefore the action of U on Λ is

$$\begin{aligned} U \cdot \mathcal{L}(\psi) &= \left\|(UdU^\dagger)(U\psi)\right\|^2 + m\|(U\psi)\|^2 \\ &= \left\|Ud(U^\dagger U)\psi\right\|^2 + m\|\psi\|^2 \\ &= \|d\psi\|^2 + m\left\|\psi^2\right\| \\ &= \mathcal{L}(\psi). \end{aligned} \tag{2.5}$$

Thus $\mathcal{L}$ is $U(n)$-invariant, as stipulated.

Now consider a generalization. Consider Minkowski space with a copy of the vector space V "glued" to each point to form a vector bundle over Minkowski space. Now the wavefunction is a map that, for each point x in Minkowski space, gives a vector in the copy of V glued to x. What can the Lagrangian look like now?



As before, one mandates that the Lagrangian should not depend on the choice of basis. The difference is that now one makes a choice of basis for V *at each point in space.* A **gauge transformation** is a map

$$\psi(x) \mapsto U(x)\psi(x) \tag{2.6}$$

where $U(x)$ is a smooth, $U(n)$-valued function of x. In other words, it is a change of basis at each point in spacetime. But there's a problem: there's no longer a unique directional derivative. In fact, for any $\mathsf{u}(n)$-valued 1-form $A_\mu(x)dx^\dagger$ (where $\mathsf{u}(n)$ is the Lie Algebra of U(n)),

$$D^A = d + igA(x)dx \Longleftrightarrow D^A_\mu = \partial_\mu + igA_\mu(x) \tag{2.7}$$

is a perfectly fine directional derivative. Here , g is a real-valued coupling constant, and A is known as a **Gauge field**. "Perfectly fine" means that it obeys the product rule and is compatible with the norm on V. This D^A is known as a **covariant derivative**.[1] The superscript on D^A is often dropped for clarity.

Covariant derivatives differ from normal derivatives in one important respect: the familiar fact that "mixed partials" are equal ($\partial_\mu\partial_\nu = \partial_\nu\partial_\mu$) does not hold. The degree to which the mixed partials fail to commute is the **curvature**

$$F_{\mu\nu} := \frac{i}{g}[D_\mu, D_\nu]. \tag{2.8}$$

[1]More precisely, this is a gauge connection or gauge covariant derivative induced from a principle $U(n)$-bundle over $\mathbb{M}$ on V. For more on the elegant subject of connections on principle bundles in gauge theory, see [6] or [7].

This can be expressed in terms of the Gauge field:

$$
\begin{aligned}
&[D_\mu, D_\nu]\psi \\
&= ((\partial_\mu + igA_\mu)(\partial_\nu + igA_\nu) - (\partial_\mu + igA_\mu)(\partial_\nu + igA_\nu))\,\psi \\
&= (\partial_\mu\partial_\nu - \partial_\nu\partial_\mu)\,\psi + ig(\partial_\nu A_\mu - A_\mu\partial_\nu - \partial_\mu A_\nu + A_\nu\partial_\mu)\psi - g^2(A_\mu A_\nu - A_\nu A_\mu)\psi \\
&= ig\Big(\partial_\nu(A_\mu\psi) + A_\mu(\partial_\nu\psi) - A_\mu(\partial_\nu\psi) - \partial_\mu(A_\nu\psi) - A_\nu(\partial_\mu\psi) + A_\nu(\partial_\mu\psi)\Big) - g^2[A_\mu, A_\nu]\psi \\
&= ig(\partial_\nu A_\mu - \partial_\mu A_\nu)\psi - g^2[A_\mu, A_\nu]
\end{aligned}
$$

so

$$F = F_{\mu\nu}dx^\mu \wedge dx^\nu = (\partial_\mu A_\nu - \partial_\nu A_\mu) - ig[A_\mu, A_\nu]dx^\mu \wedge dx^\nu = dA - ig[A, A] \tag{2.9}$$

The bracket here is defined by

$$[R, S] := [R_\mu dx^\mu, S_\nu dx^\nu] = [R_\mu, S_\nu]dx^\mu \wedge dx^\nu \tag{2.10}$$

where the bracket on the right-hand side is the Lie bracket for $\mathsf{u}(n)$, i.e. is just the commutator for n by n matrices. Since both the Lie bracket and wedge product are anti-symmetric, this bracket is actually symmetric and thus $[A, A] \neq 0$. This satisfies the Jacobi Identity

$$[Q_\mu, [R_\nu, S_\rho]] + [R_\nu, [S_\rho, Q_\mu]] + [S_\rho, [Q_\mu, R_\nu]] = 0 \tag{2.11}$$

for any $\mathsf{u}(n)$-valued functions Q, R, S. It follows that,

$$[D_\mu, [D_\nu, D_\rho]] + [D_\nu, [D_\rho, D_\mu]] + [D_\rho, [D_\mu, D_\nu]] = 0. \tag{2.12}$$

Translating this into a statement about the curvature,

$$D_\mu F_{\nu\rho} + D_\nu F_{\rho\mu} + D_\rho F_{\mu\nu} = 0, \tag{2.13}$$

a statement known as the **Bianchi Identity**. This can be written in a coordinate-independent fashion as

$$D^\star F = 0. \tag{2.14}$$

in terms of the Hodge Dual ${}^\star F^{\mu\nu} := \varepsilon^{\mu\nu\rho\sigma}F_{\rho\sigma}$.

Like an ordinary derivative, the covariant derivative changes under gauge transformations according to

$$D \mapsto U(x)DU^\dagger(x). \tag{2.15}$$

Note that the product rule implies

$$(U\nabla U^\dagger)\psi = UU^\dagger\nabla\psi + U\left(\nabla U^\dagger\right)\psi = \left(\nabla + U(\nabla U^\dagger)\right)\psi \tag{2.16}$$

so

$$
\begin{aligned}
D\psi \mapsto\ & (UDU^\dagger)\psi \\
&= \Big(U(d + iA)U^\dagger\Big)\psi \\
&= (UdU^\dagger)\psi + iUAU^\dagger\psi \\
&= d\psi + U(\nabla U^\dagger)\psi + iUAU^\dagger\psi \\
&= [d + iA']\,\psi
\end{aligned}
$$

where $A' = U(dU^\dagger) + iUAU^\dagger$. In other words, A transforms like

$$A \mapsto U(dU^\dagger) + iUAU^\dagger. \tag{2.17}$$

Putting Equation (2.15) into Equation (2.8) easily shows $F \mapsto U(x)FU^\dagger(x)$, so $\|F\|^2$ is gauge-invariant. To sum up: the field and covariant derivative transform analogously to Equations (2.3) and (2.4), except now the transformations are functions of position. The curvature transforms nicely as well. However, A transforms *inhomogenously* and, in particular, includes an extra term under gauge transformations.

What Lagrangian could this more complicated theory have? As before, there is an issue. Which covariant derivative should be used? The solution is actually to sidestep this problem; if there is no unique derivative, just make it another field variable and let the Euler-Lagrange equations find the one which minimizes the action. So one can form the trial Lagrangian (density)

$$\mathcal{L}_{\text{trial}}(A, \psi) := \left\|D^A\psi\right\|^2 + m\left\|\psi\right\|^2 \tag{2.18}$$

where a superscript A has been added to show that D is actually a function of A. Using the transformation laws above, one can check that this is gauge invariant and does not depend on a specific choice of basis. (In the notation used here, this is the same calculation as Equation (2.5) with d replaced by D and U by $U(x)$.) But there's something very odd about this Lagrangian (2.18): there's no kinetic term for A.

It turns out that there is a unique gauge-invariant term that is second-order in A, namely $\left\|F^A\right\|^2 := \operatorname{Tr} F^{\mu\nu}F_{\mu\nu}$ (up to a multiplicative constant). A very sensible Lagrangian (density) for this Gauge theory is

$$\mathcal{L}_{\text{GT}}(A, \psi) := -\frac{1}{4}\left\|F^A\right\|^2 + \left\|D^A\psi\right\|^2 + m\left\|\psi\right\|^2. \tag{2.19}$$

The factor of $-\frac{1}{4}$ is a convention.

Variations of this Lagrangian are the starting point for almost all gauge theories. As an example, when the symmetry group is $U(1)$, this becomes the Lagrangian for scalar electrodynamics. $U(1)$ is an Abelian group, so its Gauge theories are called *Abelian gauge theories*. In such theories, the gauge field transforms more simply and is much easier to work with. Quantum chromodynamics uses this same framework with the symmetry group $SU(3)$, which is non-abelian.

Yang-Mills theory starts from the Lagrangian

$$\mathcal{L}_{YM}(A) := -\frac{1}{4}\left\|F^A\right\|^2 \tag{2.20}$$

where the symmetry group is taken to be $SU(n)$. This is a perfectly good classical field theory.

2.2 Quantization

Equipped with the Lagrangian for Yang-Mills theory, the next task is to quantize it. This discussion follows Srednicki and Zee[8, 9].

A first attempt at writing down a partition function for Yang-Mills theory might looks like

$$Z[A,J] \propto \int \mathcal{D}[A] \, \exp\left(i \int_{\mathbb{M}} \mathcal{L}_{YM}(A) + AJ \right) \tag{2.21}$$

where J is a vector-valued source. At this point, it is unfortunately necessary to introduce coordinates. Let $\{T^a\}_{a=1}^{n(n-1)/2}$ be a basis for $\mathsf{su}(n)$, normalized by $\mathrm{Tr}\left(T^a T^b\right) = \delta^{ab}$. In this basis, the Lie bracket of $\mathsf{su}(n)$ has the form

$$[T^a, T^b] = f^{abc} T^c \tag{2.22}$$

where f^{abc} are the (real-valued) structure constants of $\mathsf{su}(n)$. The Latin indices here run from 1 to $n(n-1)/2$, the dimension of $\mathsf{su}(n)$. Then

$$\begin{aligned} \mathcal{L}_{YM}(A) &= -\frac{1}{4} \|F\|^2 = -\frac{1}{4} \|dA + [A,A]\|^2 \\ &= -\frac{1}{4}\Big(\|dA\|^2 - 2ig \langle dA, [A,A] \rangle - g^2 \|[A,A]\|^2 \Big) \\ &= \mathcal{L}_0(A) + \mathcal{L}_1(A) + \mathcal{L}_2(A) \end{aligned}$$

where

$$\mathcal{L}_0(A) = -\frac{1}{4} \mathrm{Tr} \left(\partial_\mu A_\nu - \partial_\nu A_\mu \right)^2, \tag{2.23}$$

$$\mathcal{L}_1(A) = -\frac{1}{2} g \left(\partial_\mu A^a_\nu - \partial_\nu A^a_\mu \right) f^{abc} A^{b\mu} A^{c\nu}, \tag{2.24}$$

$$\mathcal{L}_2(A) = -\frac{1}{4} g^2 f^{abc} f^{ade} A^a_\mu A^c_\nu A^{d\mu} A^{e\nu}. \tag{2.25}$$

Note that $\mathcal{L}_0, \mathcal{L}_1, \mathcal{L}_2$ are respectively constant, linear, and quadratic in the coupling. One thinks of $\mathcal{L}_0$ and $\mathcal{L}_1$ as a "free field theory" with $\mathcal{L}_2$ as a perturbation. To apply perturbation theory, one must find the propagator, i.e. the operator inverse of $\mathcal{L}_0$. Unfortunately, due to gauge freedom, $\mathcal{L}_0$ has a non-trivial kernel and thus no inverse exists. This can be cured by "fixing" the gauge.

Eliding over a long argument, one does this by introducing a functional determinant, which is computed by putting in the *Faddeev-Popov ghost fields*, two complex-scalar fields denoted c and $c^\dagger$ respectively. One should note that these fields are non-physical, and exist merely to fix the gauge. The fact that they violate spin-statistics can therefore be ignored. The new, gauge-fixed Lagrangian is

$$Z[J, \varepsilon, \varepsilon^\dagger] \propto \int \mathcal{D}[A] \, \mathcal{D}[c] \, \mathcal{D}[c^\dagger] \, \exp\left(i \int_M \mathcal{L}_{YM} + \mathcal{L}_{gf}(A) + \mathcal{L}_{gh}(c, c^\dagger) + JA + c^\dagger \varepsilon + c\varepsilon^\dagger \right) \tag{2.26}$$

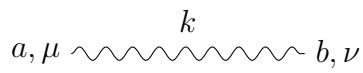
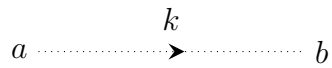
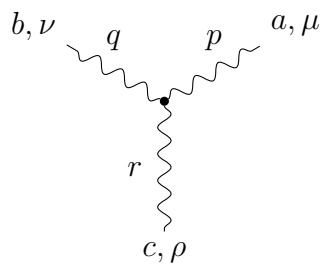
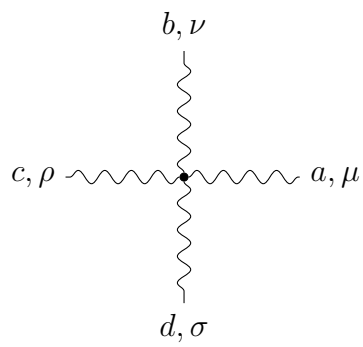
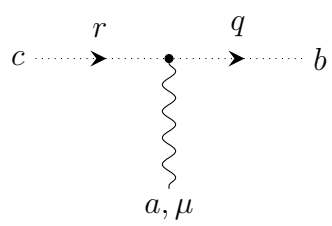
where the gauge-fixing term is

$$\mathcal{L}_{gf} = -\frac{1}{2} \xi^{-1} \partial^\mu A^a_\mu \partial^\nu A^a_\nu. \tag{2.27}$$

Here ξ is a real parameter which specifies the gauge fixing. It is non-physical and should drop out in any observable. The ghost term is

$$\mathcal{L}_{gh} = c^\dagger D c = -\partial^\mu c^{\dagger a} \partial_\mu c^a + g f^{abc} A^c_\mu \partial^\mu c^{\dagger a} c^b. \tag{2.28}$$

The particles associated to A are called **gluons** and the (non-physical) particles associated to c and $c^\dagger$ are called **ghosts**.

Component	**Graphical Representation**
Gluon propagator $\widetilde{\Delta}^{ab}_{\mu\nu}(k) = \frac{\delta^{ab}}{k^2 - i\varepsilon}\left(g_{\mu\nu} + (\xi - 1)\frac{k_\mu k_\nu}{k^2}\right)$	
Ghost propagator $\widetilde{\Delta}^{ab}(k^2) = \frac{\delta^{ab}}{k^2 - i\varepsilon}$	
Three gluon vertex $i\boldsymbol{V}^{abc}_{\mu\nu\rho}(p,q,r)$ $= gf^{abc}\left[(q-r)_\mu g_{\nu\rho} + (r-p)_\nu g_{\rho\mu} + (p-q)_\rho g_{\mu\nu}\right]$	
Four gluon vertex $i\boldsymbol{V}^{abcd}_{\mu\nu\rho\sigma} = -ig^2\Big[f^{abe}f^{cde}(g_{\mu\rho}g_{\nu\sigma} - g_{\mu\sigma}g_{\nu\rho})$ $+ f^{ace}f^{dbe}(g_{\mu\sigma}g_{\rho\nu} - g_{\mu\nu}g_{\rho\sigma})$ $+ f^{ade}f^{bce}(g_{\mu\nu}g_{\sigma\rho} - g_{\mu\rho}g_{\sigma\nu})\Big]$	
Ghost-ghost-gluon vertex $i\boldsymbol{V}^{abc}_{\mu}(q,r) = gf^{abc}q_\mu$	

The Feynman rules for the theory are given in the table. In principle, one could determine the value of any scattering amplitude just from these rules. Unfortunately, this is very impractical. Srednicki points out that just the $gg \to gg$ cross section has 12,996 terms just at tree level [8]! The methods of the next section drastically reduce the complexity of these calculations.

3 Kinematics for Scattering Amplitudes

The notion of scattering amplitude can now be made more mathematically precise. An **n-particle scattering amplitude** for gluons in Yang-Mills theory is a complex-valued function

$$A(p_1, \ldots, p_n, \varepsilon_1, \ldots, \varepsilon_n) \tag{3.1}$$

where p_i and ε_i are the four-momentum and polarization of the ith gluon respectively. The precise form of A can, in principle, be calculated by summing over all Feynman diagrams. By convention, all n particles are considered to be outgoing.

What space is A defined on? Naively, it is a function of $2n$ four-vectors, i.e. $A : \mathbb{M}^{2n} \to \mathbb{C}$. (This necessarily implies A is Lorentz-invariant.) However, energy-momentum conservation implies that

$$\sum_{i=1}^{n} p_i = 0. \tag{3.2}$$

One often makes the further restriction that the gluons are light-like, i.e.

$$p_i^2 = 0, \quad 1 \leq i \leq n. \tag{3.3}$$

These constraints carve out a complicated subset of $\mathbb{M}^{2n}$. The natural question is: *what is the best way to parameterize functions on this subset?* Alternatively, *what is the domain of A?* A sequence of maps gives better and better answers.

$$\begin{array}{c} \text{Minkowski space} \\ \downarrow \\ \text{Spinor-Helicity Parameterization} \\ \downarrow \\ \text{Momentum Twistors} \\ \downarrow \\ \text{Grassmannians} \\ \downarrow \\ \text{Cluster Algebras} \end{array} \tag{3.4}$$

The first three steps, Minkowski space to Grassmannians, are briefly reviewed in this section. The last step is taken up in Section 5.6. An extremely clear explanation of the first two steps is given by Witten in [10], whose presentation is roughly followed below. Another, more detailed explanation is given by Elvang and Huang in [11]. All four steps are used by Golden *et al.* [12].

3.1 Spinor Helicity

The complexified Lorentz group is locally isomorphic to $SL_2(\mathbb{C}) \times SL_2(\mathbb{C})$. Its representations can therefore be classified by a pair of half-integers (n, m). The representation $\left(\frac{1}{2}, 0\right)$ acts on negative chirality spinors, written as λ_a, $a = 1, 2$. These indices are raised and lowered by ε_{ab}, the Levi-Civita symbol. If λ_1 and λ_2 are both negative chirality spinors, then

$\langle \lambda_1, \lambda_2 \rangle := \varepsilon_{ab}\lambda_1^a\lambda_2^b$ is a Lorentz-invariant, anti-symmetric form. The $(0, 1/2)$ representation acts similarly on positive chirality spinors, which are denoted with dotted indices and a tilde $\widetilde{\lambda}_{\dot{a}}$. These have the invariant $[\widetilde{\lambda}_1, \widetilde{\lambda}_2] = \varepsilon_{\dot{a}\dot{b}}\widetilde{\lambda}_1^{\dot{a}}\widetilde{\lambda}_2^{\dot{b}}$.

The Lorentz group $SO(3, 1)$ is essentially the $(1/2, 1/2)$ representation. A four-vector p can then therefore be written as a *bi-spinor*, a two-by-two matrix denoted $p_{a\dot{a}}$ which is a combination of a negative and positive chirality spinor where $p^2 = \det p_{a\dot{a}}$. One finds that a light-like vector can be represented as

$$p_{a\dot{a}} = \lambda_a\widetilde{\lambda}_{\dot{a}} \tag{3.5}$$

for some positive and negative chirality spinors λ and $\widetilde{\lambda}$. Unfortunately, this is not a unique representation, since if $a \in \mathbb{C}^*$, then $\lambda \mapsto a\lambda, \widetilde{\lambda} \mapsto a^{-1}\widetilde{\lambda}$ leaves $p_{a\dot{a}}$ invariant. There is, in general, no way to determine λ and $\widetilde{\lambda}$ as functions of p. However, given a pair $(\lambda, \widetilde{\lambda})$, and the *helicity* of the gluon, one can determine the momentum p and polarization ε. (The helicity of a particle is the dot product of the spatial components of momentum and spin. For a spin-one particle such as a gluon, the helicity is ± 1, usually written as just $+$ or $-$.) This is not a unique mapping; to carry this out, two arbitrary spinors μ_a and $\widetilde{\mu}_{\dot{a}}$ must be chosen.

The upshot is that the scattering amplitude can be written as

$$A(\lambda_i, \widetilde{\lambda}_i, h_i) \tag{3.6}$$

where λ_i and $\widetilde{\lambda}_i$ are positive and negative chirality spinors, $h_i = \pm 1$ is the helicity and, of course, $1 \leq i \leq n$. Since the above discussion was specialized to light-like vectors, Equation (3.6) automatically obeys the constraint (3.3). However, the condition of total momentum conservation must still be enforced "manually". One can check that this amplitude is Lorentz invariant [10].

3.2 Twistors and Momentum Twistors

Minkowski space is often very difficult to work with, partially because the Lorentz group is non-compact. It is sometimes easier to work with a generalization of Minkowski space whose symmetries form a compact Lie group. One realization of this idea is *twistor space*, proposed by Penrose [13]. Very roughly, the idea is that complexified, compactified Minkowski space can be conformally embedded into $\mathbb{CP}^3$, three-dimensional complex projective space. Elements of twistor space are called *twistors*. A twistor Z can be written $Z = (\lambda, \mu)$ where λ and μ are spinors.

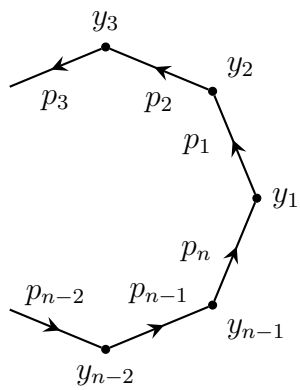
In the same way that thinking of the complex numbers as the Riemann sphere yields a much cleaner picture of Möbius transformations, twistor space simplifies Lorentz transformations. It is customary to use a metric signature $++--$, in which case the Lorentz transformations are promoted to $\mathrm{PSL}_4(\mathbb{R})$ [10].(Actually, the Lorentz transformations form a subgroup, but amplitudes are invariant under all of $\mathrm{PSL}_4(\mathbb{R})$.)

One can go between the bi-spinors of the last section and twistors easily by means of a "half-fourier transform" [10]. For an arbitrary function $f(\widetilde{\lambda})$, one makes the transformation

$$f(\widetilde{\lambda}) \mapsto \widetilde{f}(\mu) \;=\; \int \frac{d^2\widetilde{\lambda}}{(2\pi)^2} \exp(i\mu^{\dot{a}}\widetilde{\lambda}_{\dot{a}}) f(\widetilde{\lambda}). \tag{3.7}$$

Lorentz-invariant parts of $A(\lambda_i, \widetilde{\lambda}, h_i)$ are always brackets $\langle \cdot, \cdot \rangle$ or $[\cdot, \cdot]$, which are functions of either λ or $\widetilde{\lambda}$ alone, so this transformation suffices to express the amplitude as $A(\lambda_i, \mu_i, h_i) = A(Z_i, h_i)$ over twistor space. One can show that A is actually a holomorphic function of the λ_i's and μ_i's [10].

A slight variant on this approach, called *momentum twistors*, automatically enforces momentum conservation. The method is due to Hodges [14]. Consider, back in Minkowski space, the momenta p_i. Representing these as vectors laid end to end, the condition $\sum_{i=0}^{n} p_i = 0$ implies they form a closed polygon with light-like edges.



One then considers affine coordinates y_i so that $p_i = y_i - y_{i-1}$. Each of these y_i is a point on affine Minkowski space, i.e. Minkowski space with no distinguished origin. The condition on momentum conservation is simply

$$0 = \sum_{i=1}^{n} p_i = \sum_{i=1}^{n} (y_i - y_{i-1}) = y_n - y_{n-1} + \cdots + y_2 - y_1 + y_1 + y_0 = y_n - y_0 \tag{3.8}$$

or, simply, $y_0 = y_n$. The original amplitude can then be cast as $A(y_1, \ldots, y_n, \varepsilon_1, \ldots, \varepsilon_n)$, where the y's are completely unconstrained. Using the same machinery outlined above, the x's can be carried over to twistor space, where they enjoy the additional *dual conformal symmetry* [12].

3.3 Configurations

In light of this discussion, scattering amplitudes can now be considered functions

$$A(Z_1, \ldots, Z_n, h_1, \ldots, h_n) \tag{3.9}$$

where the Z_i's are (momentum) twistors in $\mathbb{CP}^3$ and $h_i = \pm 1$ are helicities. The Z's are automatically light-like and obey momentum conservation; they are unconstrained. Lorentz symmetry is still present, in the form of an overall action of $\mathrm{PSL}_4(\mathbb{C})$ acting on all of the Z_i's simultaneously. The space of n-tuples of points in $\mathbb{CP}^3$ modulo the action of $\mathrm{PGL}_4(\mathbb{C})$ is the configuration space $\mathrm{Conf}_n(\mathbb{CP}^3)$ [12]. Thus scattering amplitudes are functions

$$A : \mathrm{Conf}_n(\mathbb{CP}^3) \times \mathbb{Z}_2^n \to \mathbb{C} \tag{3.10}$$

with no constraints.

3.4 Grassmannians and Plücker Coordinates

How can one actually work with the space $\mathrm{Conf}_n(\mathbb{CP}^3)$ in practice? What sort of coordinates does it have? It turns out that $\mathrm{Conf}_n(\mathbb{CP}^3)$ is essentially the Grassmannian $\mathbb{G}(4,n)/\left(\mathbb{C}^*\right)^{n-1}$. Grassmannians are a source of many examples in algebraic geometry, and are quite well-understood by mathematicians. They are described by many books, e.g. Gekhtman *et al* [15].

The **complex Grassmannian** $\mathbb{G}(k,n)$ is the space of complex k-dimensional vector subspaces of $\mathbb{C}^n$. This can be formulated more concretely in terms of matrices. Consider a subspace $V \subset \mathbb{C}^n$, with $\dim V = k$. One can choose a basis $\{\boldsymbol{b}_1, \ldots, \boldsymbol{b}_n\}$ for $\mathbb{C}^n$ such that the span of the first k vectors is V. V can thus be specified by the matrix

$$\begin{bmatrix} \boldsymbol{b}_1 \\ \boldsymbol{b}_2 \\ \vdots \\ \boldsymbol{b}_k \end{bmatrix} = \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ b_{21} & b_{12} & \cdots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{k1} & b_{k2} & \cdots & b_{kn} \end{bmatrix} \tag{3.11}$$

Of course, any change of basis of the first k vectors spans the subspace, multiplying by a k by k invertible matrix on the left specifies the same subspace. This $k \times n$ matrix only specifies V up to a left action of $\mathrm{GL}_k(\mathbb{C})$. The Grassmannian can therefore be identified with the quotient

$$\mathbb{G}(k,n) \cong M_{k\times n}(\mathbb{C})/\mathrm{GL}_k(\mathbb{C}). \tag{3.12}$$

One important property of the Grassmannian is that $\mathbb{G}(k,n) \cong \mathbb{G}(n-k,n)$. This is because there is a bijection between k-planes and $(n-k)$-planes in n dimensions. Just as a plane in $\mathbb{R}^3$ can be specified by it's normal vector, a k-plane in $\mathbb{C}^n$ is uniquely determined by the subspace orthogonal to it.

A common choice of coordinates on the Grassmannian are *Plücker coordinates*. The Plücker Embbedding is a map of the Grassmannian into projective space. Define a map

$$\iota : \mathbb{G}(k,n) \to \mathbb{P}(\Lambda^k \mathbb{C}^n)$$

where, if $V = \mathrm{span}\,\{b_1, \ldots, b_k\} \in \mathbb{G}(k,n)$ and the b_i's are a basis for V, then

$$\iota(V) := b_1 \wedge b_2 \wedge \cdots \wedge b_k.$$

This is well-defined, because if the c_j's are another basis for V, with $c_j = A^i_j b_i$ for some change-of-basis matrix A, then

$$c_1 \wedge c_2 \wedge \cdots \wedge c_k = A^1_{j_1} b_1 \wedge A^2_{j_2} b_2 \wedge \cdots \wedge A^k_{j_k} b_k = (\det A) b_1 \wedge b_2 \wedge \cdots \wedge b_k$$

which is equal to $b_1 \wedge b_2 \wedge \cdots \wedge b_k$ in projective space. The second equality is just the normal change of basis formula for the wedge product. See, e.g. Lemma 9.11 of [16] for a proof. This map is injective, but not surjective.

Plücker Coordinates are formed as follows. Consider $V \in \mathbb{G}(k,n)$ as a $k \times n$ matrix as in Equation (3.11). Define $I := (i_1, \ldots, i_k)$ to be some multi-index of length k with $1 \le i_j \le n$.

Then $(b_{i_1} b_{i_2} \cdots b_{i_k})$ is a $k \times k$ matrix. This is known as the *minor* of V indexed by I. Since this is a square matrix, it has a well defined determinant. The Plücker coordinates for I is

$$\langle I \rangle = \langle i_1 i_2 \cdots i_k \rangle = \det \left(b_{i_1} b_{i_2} \cdots b_{i_k} \right) \tag{3.13}$$

If A is a change-of-basis matrix of V,i.e. a $k \times k$ non-degenerate matrix, then $\langle I \rangle (AV) = \det A \langle I \rangle (V)$. Plücker coordinates are therefore well-defined homogenous homogenous coordinates on the projective space $\mathbb{P}(\Lambda^k \mathbb{C}^n)$. As mentioned above, the embedding is not surjective. Indeed, the Plücker coordinates satisfy functional relations. These **Plücker relations** have the form

$$\langle ijI \rangle \langle pqI \rangle = \langle pjI \rangle \langle iqI \rangle + \langle qjI \rangle \langle piI \rangle \tag{3.14}$$

where I is a multi-index of length $k-2$ and i, j, p, q are single indeces. The homogenous coordinate ring on the Grassmannian is the ring of polynomials in the Plúcker coordinates modulo the ideal generated by all the Plücker relations.

Golden *et al* connect this to the physics[12]. Suppose Z_i are twistors. They are represented as points in $\mathbb{C}^4$ with $Z_i \sim aZ_i$ for any $a \in \mathbb{C}^*$. A configuration of n twistors is a $4 \times n$ matrix

$$\begin{bmatrix} Z_1 \\ Z_2 \\ \vdots \\ Z_n \end{bmatrix}. \tag{3.15}$$

Four-brackets

$$\langle ijk\ell \rangle := \det(Z_i Z_j Z_k Z_\ell) \tag{3.16}$$

are invariant under the symmetry $\mathrm{PGL}_4(\mathbb{C})$, and thus good coordinates on the moduli space. But these are exactly the Plücker coordinates for $\mathbb{G}(4, n)$! Because of the overall $\mathbb{C}^*$ scaling on each of the coordinates, it is necessary to further mod out by $(\mathbb{C}^*)^n$. However, multiplying by exactly the same factor for each twistor does nothing, so one really want $(\mathbb{C}^*)^{n-1}$. Precisely, in the language of algebraic geometry, there is a birational isomorphism [12]

$$\mathbb{G}(k, n) / \left(\mathbb{C}^* \right)^{n-1} \cong \mathrm{Conf}_n(\mathbb{CP}^3). \tag{3.17}$$

Thus scattering amplitudes can be considered as complex-valued maps on the Grassmannian. This may seem an unnecessarily complicated approach, but using the Grassmannian actually makes many "hidden" properties of scattering amplitudes much easier to see.

the last step in this procession of mappings — from Grassmannians to cluster algebras — is postponed to Section 5.6.

4 Review of Amplitudes in $\mathcal{N} = 4$ SYM

This section is a very brief review of amplitudes in $\mathcal{N} = 4$ SYM and should be treated as heuristic rather than authoritative. The focus is on "taxonomy": what classes of functions exist, which are known, and which are not. A comprehensive treatment of this material is given in Elvang and Huang[11]. The presentation here follows the thesis defense of Dr. John Golden.

4.1 The Parke-Taylor Formula

With the Feynman rules for Yang-Mills theory developed in Section 2 , one can in principle calculate any amplitude. In practice, this is an infinite sum over diagrams. It is typical to split up the n-particle amplitude A_n as

$$A_n = A_n^{(0)} + A_n^{(1)} + A_n^{(2)} + \cdots \tag{4.1}$$

where $A_n^{(L)}$ is contribution from diagrams with exactly L loops. The contribution of diagrams with no loops, $A_n^{(0)}$ is called **tree level**.

Even tree level results are difficult. In 1985 Parke and Taylor decided to calculate the process $gg \to gggg$ at tree level in the planar limit [17]. Even this "simple" calculation involves a sum over 220 diagrams! After adding up the contributions from all the diagrams, the final result turned out to be extremely simple! In spinor-helicity formalism, with the abbreviated notation $(\lambda_i, \widetilde{\lambda}_i, h = \pm) \leftrightarrow i^{\pm}$,

$$A_4^{(0)}(1^+, 2^+, 3^-, 4^-) = \frac{\langle 12 \rangle^4}{\langle 12 \rangle \langle 23 \rangle \langle 34 \rangle \langle 41 \rangle} \tag{4.2}$$

where $\langle ij \rangle = \langle \lambda_i \lambda_j \rangle$. The number of Feynman diagrams grown very quickly with n. For $gg \to 8g$, there are more than a million [11]. But something remarkable happens. The general form of Equation (4.2) actually works for any number of particles! Precisely,

$$A_n^{(0)}(1^+, 2^+, 3^-, 4^-, \ldots, n^-) = \frac{\langle 12 \rangle^4}{\langle 12 \rangle \langle 23 \rangle \cdots \langle n1 \rangle}. \tag{4.3}$$

(This is actually only the most complex part of the answer. The full answer is found by summing over non-cyclic permutations of $\{1, \ldots, n\}$ with a multiplicative term proportional to the trace of a product of generators of $\mathsf{su}(n)$.) This unexpectedly simple answer is taken to be evidence for a larger trend: Yang-Mills theory amplitudes are actually fairly well-behaved, but calculating them is ugly. The hope is that studying the structure of the amplitudes will give shortcuts that render the diagrams unnecessary.

$$\begin{tikzcd}
\text{Elegant Lagrangian} \arrow[d] \arrow[ddd, bend left, dotted, "?"] \\
\text{Lots of Feynman Diagrams} \arrow[d] \\
\text{Lots of complicated terms} \arrow[d] \\
\text{Simple answer}
\end{tikzcd} \tag{4.4}$$

4.2 Supersymmetry

For amplitudes above tree level, Yang-Mills theory quickly becomes very difficult. To make calculations tractable, it is standard to work in $\mathcal{N} = 4$ Super Yang-Mills theory, which offers

many additional symmetries that can be used to "tame" the amplitudes that appear. Here $\mathcal{N}$ refers to the number of supersymmetry generators in the theory. The Lagrangian for this theory is[11]

$$\mathcal{L} = Tr\left(-\frac{1}{4}\|F\|^2 - \frac{1}{2}\|D\Phi_I\|^2 + \frac{i}{2}\overline{\Psi}\not{D}\Psi + \frac{g}{2}\overline{\Psi}\Gamma^I\left[\Phi_I, \Psi\right] + \frac{g^2}{4}\left[\Phi_I, \Phi_J\right]^2\right) \tag{4.5}$$

where D is the same covariant derivative as in regular Yang-Mills for SU(N); Φ_I, $I = 1, \ldots, 6$ are real scalar fields with a global $SO(6)$ R-symmetry; Ψ are 10-dimensional Majorana-Weyl fermions; and Γ^I are gamma-matrices of the 10-dimensional Clifford algebra. All commutators are associated with the $\mathsf{su}(n)$ Lie algbera. This gives a superconformal group with 32 fermionic and 30 bosonic generators total generators [11]. In short, $\mathcal{N} = 4$ has a large number of symmetries.

Working in this framework, amplitudes are promoted to *superamplitudes*, which also include the new fermionic fields. The discussion of the kinematics remains largely similar. Further details of supersymmetry are beyond the scope of this thesis.

4.3 The MHV Classification

The Parke-Taylor formula holds in the case where two of the gluons have positive helicity, and the rest negative. This roughly corresponds to two incoming gluons and the rest outgoing. What about all the other possibilities? It turns out that in $\mathcal{N} = 4$ SYM, the amplitudes with zero or one helicity different from all the rest vanish:

$$A_n^{(L)}(1^-, 2^-, \ldots, n^-) = 0 = A_n^{(L)}(1^+, 2^-, \ldots, n^-). \tag{4.6}$$

So at least two gluons must have opposite helicity from the rest. The case when exactly two gluons have positive helicity, with the rest negative is called the **Maximally Helicity Violating (MHV)** case. The case with exactly three positive is called the **Next-to MHV (NMHV)** case. In general, the amplitude

$$A_n^{(L)}(1^+, 2^+, \ldots, (K+2)^+, (K+3)^-, \ldots, n^-) \tag{4.7}$$

is called the $\mathbf{N^K MHV}$ term. MHV terms tend to be the simplest, and grow more complicated with K. However, the difference is usually not too great. If the MHV term is known, the $\mathbf{N^K MHV}$ terms can be computed by the formula

$$A_{n;(L)}^{\mathrm{N^K MHV}}(\varepsilon) = A_{n;(0)}^{\mathrm{MHV}}\left(\mathcal{P}_{n;(0)}^{\mathrm{N^K MHV}} + \lambda\mathcal{P}_{n;(1)}^{\mathrm{N^K MHV}}(\varepsilon) + \ldots\right) \tag{4.8}$$

where $\lambda \sim g^2 N$ is the gauge coupling times the size of the gauge group SU(N) and ε is a regulator [11].

4.4 Loops and the BDS Ansatz

Using the technique of *generalized unitarity*, Bern, Dixon, Dunbar and Kosower were able to determine a general formula for $A_n^{(1)}$, the one-loop, n-particle MHV amplitude [18]. As with the Parke-Taylor formula, the final result is fairly simple.

The next step was two-loops. It was shown by Anastasiou, Bern, Dixon, and Kosower (ABDK) [19] that the 2-loop, 4-particle, planar amplitude can be expressed in terms of the one-loop amplitude by [11]

$$\mathcal{P}_{4;2}^{\mathrm{MHV}}(\varepsilon) = \frac{1}{2}\left[\mathcal{P}_{4;1}^{MHV}(\varepsilon)\right]^2 + \mathcal{P}_{4;1}^{MHV}(2\varepsilon) f^{(2)}(\varepsilon) C^{(2)} + O(\varepsilon) \tag{4.9}$$

where $f^{(2)}(\varepsilon) = -\zeta(2) - \zeta(3)\varepsilon - \zeta(4)\varepsilon^2$ and $C^{(2)} = -\zeta(2)^2/2$. The appearance of ζ, the Riemann-Zeta function is not entirely a coincidence and will be "explained" in Section 6.

This continues at three loops. Bern, Dixon, and Smirnov (BDS) found that the 3-loop, 4-particle, planar MHV amplitude can also be expressed in terms of the one-loop result [20]! Specifically,[11]

$$\mathcal{P}_{4;3}^{\mathrm{MHV}}(\varepsilon) = -\frac{1}{3}\left[\mathcal{P}_{4;1}^{\mathrm{MHV}}(\varepsilon)\right]^3 + \mathcal{P}_{4;1}^{\mathrm{MHV}}(\varepsilon)\mathcal{P}_{4;2}^{\mathrm{MHV}}(\varepsilon) + f^{(3)}(\varepsilon)\mathcal{P}_{4;1}^{\mathrm{MHV}}(3\varepsilon) + C^{(3)} + O(\varepsilon) \tag{4.10}$$

where $f^{(3)}(\varepsilon) = \frac{11}{2}\zeta(4) + O(\varepsilon)$ and $C^{(3)}$ is a constant. The coefficients that appear look very similar to the power series of e^x. This motivated the **BDS Ansatz** for the full MHV amplitude at all loops:

$$\mathcal{P}_n^{\mathrm{MHV(BDS)}}(\varepsilon) = \exp\left[\sum_{L=1}^{\infty} \lambda^L \left(f^{(L)}(\varepsilon)\mathcal{P}_{n;1}^{\mathrm{MHV}}(L\varepsilon) + C^{(L)} + O(\varepsilon)\right)\right]. \tag{4.11}$$

This is thought to be almost correct. At two loops are five particles it gives the correct amplitude. However, it begins to fail "gently" at two loops and six particles. Though the ansatz correctly predicts the infrared divergence structure of the amplitude, it does not fully produce the correct finite part. The difference between the BDS Ansatz and the actual amplitude is called the **remainder function** $\mathcal{R}_n^{(L)}$. It is in calculating these remainder function that cluster algebras and polylogarithms are effectively applied. The next two sections will develop cluster algebras and review the necessary information about polylogarithms.

5 Cluster Algebras

Cluster algebras are a new area of mathematics, introduced in 2002 by Fomin and Zelevinsky [21]. They have already been applied to many areas in mathematics. Indeed, many old and well-understood mathematical objects carry cluster algebra structure, including homogeneous coordinates of Grassmannians, polygon triangulations, Schubert varieties, and more [22, 23]. Because of this, cluster algebras are the subject of considerable current research and mathematicians are searching for them — successfully — across mathematics.

Although applications of cluster algebras can be very technical, the definition and properties of cluster algebras require only simple algebra and graph theory [22]. One can almost ignore the algebraic definitions and simply follow the algorithm set out by the examples.

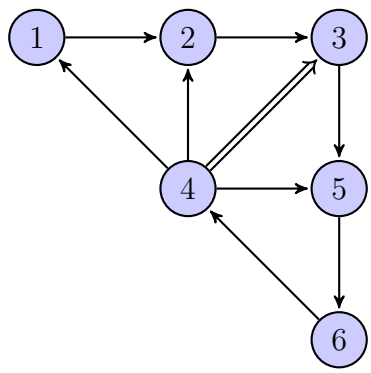
For a more in-depth introduction to cluster algebras, the textbook [24], by Gehktman, Shapiro and Vainstein is probably the most complete and detailed introductory work on cluster algebras at this time. This section follows the expository treatments in [22, 23, 25, 24].

5.1 What is a Cluster Algebra?

Cluster algebras are defined on collections of "seeds" linked by an operation known as "mutation". Seeds have two ingredients: "quivers" and "cluster variables".

Definition 1. *A **quiver** Q is a directed graph, possible with multiple arrow between vertices and/or loops. A quiver is represented by a square, skew-symmetric, integer-valued matrix Q_{ij} of size $n \times n$. The entry Q_{ij} is the number of arrows from vertex i to vertex j.*

As an example, the graph [25]



(5.1)

has data

$$Q = \begin{pmatrix} 0 & 1 & 0 & -1 & 0 & 0 \\ -1 & 0 & 1 & -1 & 0 & 0 \\ 0 & -1 & 0 & -2 & 1 & 0 \\ 1 & 1 & 2 & 0 & 1 & -1 \\ 0 & 0 & -1 & -1 & 0 & 1 \\ 0 & 0 & 0 & 1 & -1 & 0 \end{pmatrix}. \tag{5.2}$$

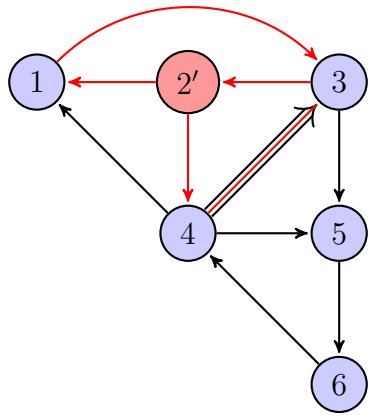
Definition 2. *If Q is a quiver, then **mutation on vertex** k is a map defined by $\mu_k(Q) = Q'$ made by applying the following rules:*

1. *reverse all arrows going into or out of vertex k;*
2. *for each path $i \to k \to j$, add in a new "shortcut" arrow $i \to j$ ("complete the triangle");*
3. *remove any two-cycles that have formed.*

Equivalently, one replaces Q_{ij} with a new matrix $Q'_{ij} = \mu_k Q_{ij}$ with entries given by

$$\begin{aligned} Q'_{ij} &= \begin{cases} -Q_{ij} & \text{if } k \in \{i,j\} \\ Q_{ij} + \dfrac{|Q_{ik}|\, Q_{kj} + Q_{ij}\, |Q_{kj}|}{2} & \text{otherwise} \end{cases} \\ &= \begin{cases} -Q_{ij} & \text{if } k \in \{i,j\} \\ Q_{ij} + B_{ik}\, |Q_{kj}| & \text{if } \operatorname{sgn} Q_{ik} = \operatorname{sgn} Q_{kj} \\ Q_{ik} & \text{otherwise.} \end{cases} \end{aligned} \tag{5.3}$$

As an example, mutating on vertex 2 of the above graph (5.1) gives



where the differences are highlighted in red. Or, expressed as matrix,

$$Q' = \begin{pmatrix} 0 & -1 & 1 & -1 & 0 & 0 \\ 1 & 0 & -1 & 1 & 0 & 0 \\ -1 & 1 & 0 & -3 & 1 & 0 \\ 1 & -1 & 3 & 0 & 1 & -1 \\ 0 & 0 & -1 & -1 & 0 & 1 \\ 0 & 0 & 0 & 1 & -1 & 0 \end{pmatrix}. \tag{5.4}$$

Now for the second ingredient. Instead of using the vertices of the quiver as mere labels, we promote them to variables.

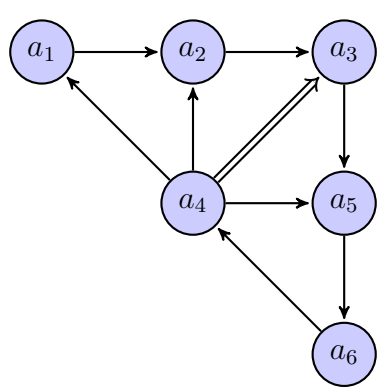
Definition 3. *Let $\{a_1, \ldots, a_n\}$ be transcendental over $\mathbb{Q}$ (independent variables with no relations). The tuple*

$$\boldsymbol{a} = (a_1, \ldots, a_n) \tag{5.5}$$

*is called a **cluster** and the a_i's are called **cluster variables**.*

*A **seed** is a pair $S = (\boldsymbol{a}, Q)$ of a cluster and a quiver where the elements of the cluster label the vertices of the quiver. Two seeds are considered equal if their labelled quivers are the same up to graph isomorphism.*

Going back to the same example, if the cluster variables are $\boldsymbol{a} = \{a_1, x_2, \ldots, a_6\}$, then the seed $S = (\boldsymbol{a}, Q)$ is represented as follows.



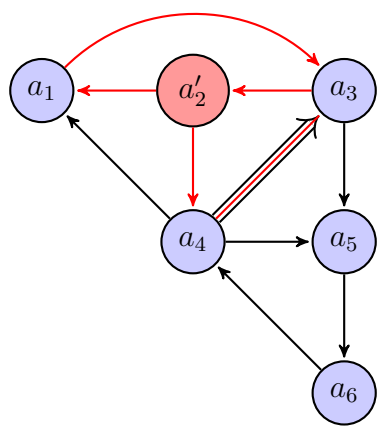
The operation of mutation changes the variables as well as the quiver.

Definition 4. *Let $S = (\boldsymbol{a}, Q)$ be a seed. Then the operation of* ***seed mutation in the direction*** *k is a map from seeds to seeds by $\mu_k(\boldsymbol{a}, Q) = (\boldsymbol{a}', Q')$ where*

$$\mu_k(a_\ell) = \begin{cases} a_\ell & \text{if } k \neq \ell \\ \dfrac{1}{a_k}\left[\displaystyle\prod_{i\to k} a_i + \prod_{k\to j} a_j\right] & \text{if } \ell = k. \end{cases} \tag{5.6}$$

The notation $i \to k$ or $k \to j$ means the product should be taken over all the arrows that are incoming to k or outgoing from k respectively. Multiple arrows should, of course, be counted multiple times. It follows from the definition that mutation on a particular vertex is an involution: $\mu_k^2 = Id$.

Taking the example of our favorite quiver, mutating on a_2 gives



where $a_2' = \frac{1}{a_2}(a_1 a_4 + a_3)$.

Now that all the ingredients have been assembled, cluster algebras can be defined. A **sequence of mutations** is simply the composition of multiple mutations, such as $\mu_1 \circ \mu_3 \circ \mu_2$.

Definition 5. *Suppose $S = (\boldsymbol{x}, Q)$ is a seed with n. Define $\mathcal{A}$, the set* ***cluster $\mathcal{A}$-coordinates*** *to be the union of all cluster variables appear under arbitrary sequences of mutations starting from S. More formally,*

$$\mathcal{A} := \{x \in \mu_{i_1}\mu_{i_2}\cdots\mu_{i_r}(\boldsymbol{x}, Q) \ : \ 1 \leq i_1, \ldots, i_r \leq n, r \in \mathbb{N}\}. \tag{5.7}$$

The ***cluster algebra*** *generated by S, $\mathcal{C}(S)$ is defined as algebraic extension of $\mathbb{Z}$ generated by $\mathcal{A}$, i.e. $\mathbb{Z}[\mathcal{A}]$. The size of each quiver, n, is called the* ***rank*** *of the cluster algebra, and the number of seeds is the* ***order*** *of the algebra.*

It is important to note that the defnition of a cluster algebra varies somewhat, depending on the source. This follows [23]. However, [22] defines them as a sub*ring* of $\mathbb{Q}(x_1, \ldots, x_n)$.

Since the definition is not yet solid, it is better to think of the cluster algebra as *the collection of seeds obtained from mutating the initial seed.* Different applications will erect various

algebraic structures on top of the seeds. But the essential features — seeds, mutations, and cluster variables — are the same in any case.

There are also various generalizations of cluster algebras built out of more general matrices Q that need not be skew-symmetric, but perhaps only skew-symmetrizable, or totally sign-skew-symmetric [24]. So far, these generalizations play no part in the physics of scattering amplitudes, and will not be discussed here. In this larger context, the cluster algebras defined here are **skew-symmetric cluster algebras of geometric type**.

One generalization that will be used, however, is the idea of *frozen variables*. Conceptually, one chooses certain vertices that may *not* be mutated. These are carried along, unchanged, in every quiver. See [26] for more details.

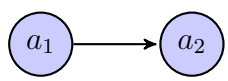
5.2 The A_2 Cluster Algebra

The definition of cluster algebras is best understood through examples. The simplest nontrivial cluster algebra is called A_2 and starts from the seed

$$S_1 = (\boldsymbol{a}_1, Q_1) = \left((a_1, a_2), \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \right) \tag{5.8}$$

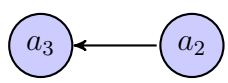
or, as a quiver,

S_1 :



There are two mutable variables, a_1 and a_2, and no frozen variables. Applying μ_1, mutation on the first vertex, gives a new seed

S_2 :



or

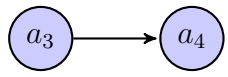
$$S_2 := \mu_1(S_1) = (\boldsymbol{a}_2, Q_2) = \left((a_3, a_2), \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \right) \tag{5.9}$$

where a_3 is given by Equation (5.6):

$$a_3 := a_1' = \frac{1}{a_1} \left[\prod_{i \to 1} a_i + \prod_{1 \to j} a_j \right] = \frac{1}{a_1} \left[a_1^0 a_2^0 + a_1^0 a_2^1 \right] = \frac{1 + a_2}{a_1}. \tag{5.10}$$

Applying μ_1 to S_2 just generates S_1 again, which gives nothing new. Applying μ_2 to S_2 gives the seed

S_3 :



or

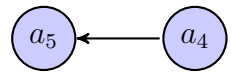
$$S_3 := \mu_2 \mu_1(S_1) = (\boldsymbol{a}_3, Q_3) = \left((a_3, a_4), \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \right) \tag{5.11}$$

where

$$a_4 := a_2' = \frac{1}{a_2}\left[\prod_{i\to 2} a_2 + \prod_{2\to j} a_2\right] = \frac{1}{a_2}\left(1+a_3\right) = \frac{1+a_1+a_2}{a_1 a_2}. \tag{5.12}$$

After this, the cluster variables start getting simpler again. Applying μ_1 to S_3 gives

S_4 :



or

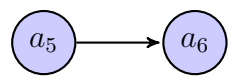
$$S_4 := \mu_1(S_3) = (\boldsymbol{a}_4, Q_4) = \left((a_5, a_4), \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\right) \tag{5.13}$$

where

$$a_5 := a_3' = \frac{1}{a_1}\left[\prod_{i\to 3} a^i + \prod_{3\to j} a^j\right] = \frac{1}{a_3}(1+a_4) = \frac{a_1 a_2 + 1 + a_1 + a_2}{a_1 a_2}\frac{a_1}{1+a_2} = \frac{1+a_1}{a_2}. \tag{5.14}$$

Applying μ_2 gives

S_5 :



or

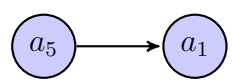
$$S_5 := \mu_2(S_4) = (\boldsymbol{a}_5, Q_5) = \left((a_5, a_6), \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}\right) \tag{5.15}$$

where

$$a_6 := a_4' = \frac{1}{a_2}\left[\prod_{i\to 4} a_i + \prod_{4\to j} a_2\right] = \frac{1}{a_4}\left(a_5+1\right) = \frac{1+a_1+a_2}{a_2}\frac{a_1 a_2}{1+a_1+a_2} = a_1. \tag{5.16}$$

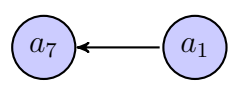
Something very strange has happened here! The sixth cluster variable is the first one again! So really S_5 has the quiver

S_5 :



Applying μ_1 one more time gives

S_6 :



or

$$S_6 := \mu_1(S_5) = (\boldsymbol{a}_6, Q_6) = \left((a_7, a_1), \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}\right) \tag{5.17}$$

where

$$a_7 := a_5' = \frac{1}{a_5}\left[\prod_{i\to 5} a_i + \prod_{5\to j} a_j\right] = \frac{1+a_1}{a_5} = \frac{1+a_1}{1}\frac{a_2}{1+a_1} = a_2. \tag{5.18}$$

So this seed is really

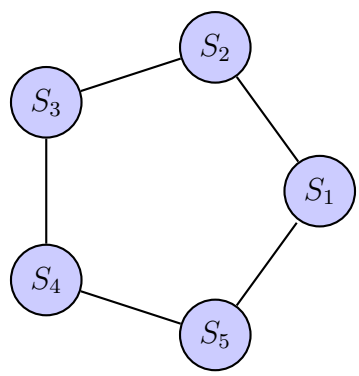
$$S_1 : \quad a_2 \longleftarrow a_1$$

which is just S_1 again!

The best way to visualize this is with an exchange graph.

Definition 6. *If $\{S_i\}$ is the set of seeds of a cluster algebra, then the* ***exchange graph*** *is the graph whose vertices are the seeds and where undirected edges are drawn between pairs of seeds linked by a single mutation. The edges are undirected because they can always be traversed back by applying the same mutation again.*

The exchange graph of A_2 is a pentagon.



There are several properties of A_2 that are worth pointing out.

- The cluster variables that appear are all rational functions in $\{a_1, a_2\}$ with positive integer coefficients.
- The denominators of all the cluster variables are monomials in $\{a_1, a_2\}$.
- The complete set of cluster variables of A_2 is

$$\mathcal{A}(A_2) = \left\{ a_1, a_2, \frac{1+a_1}{a_2}, \frac{1+a_1+a_2}{a_1 a_2}, \frac{1+a_2}{a_1} \right\}. \tag{5.19}$$

  No matter what sequence of mutations is performed on the initial seed, the variables that appear will be in this set.

- Labelling the variables as above, they satisfy a recurrence relation

$$a_k = \frac{1+a_{k-1}}{a_{k-2}} \tag{5.20}$$

  and $a_{k+5} = a_k$ for all $k \in [0, \ldots 5]$.

All the properties except the last one are actually true in more general contexts.

5.3 General Properties of Cluster Algebras

This section will discuss some general properties of the cluster algebras. The following section will return to concrete examples to elucidate them.

One of the most remarkable properties of clusters is the *Laurant Phenomenon.*

Definition 7. *A* ***Laurant Polynomial*** *over a field* $\mathbb{F}$ *is a polynomial of the form*

$$f(X) = \sum_{k\in\mathbb{Z}} c_k X^k = c_0 + c_1 X + a_{-1}\frac{1}{X} + c_2 X^2 + c_{-2}\frac{1}{X^2} + \cdots \tag{5.21}$$

where X *is a formal variable and* $c_k \in \mathbb{F}$ *are constants. In other words, a Laurant polynomial is an element of the ring* $f \in \mathbb{F}[X, X^{-1}]$.

Theorem 8 (The Laurant Phenomenon, [21, 24])**.** *In a cluster algebra of geometric type, any cluster variable may be expressed as a Laurant polynomial in the cluster variables of the initial seed. Thus if* a *is any cluster variable,*

$$a = \frac{P(a_1, \ldots, a_n)}{a_1^{d_1} \cdots a_n^{d_n}} \tag{5.22}$$

where $\{a_1, \ldots, a_n\}$ *are the cluster variables of the initial quiver,* $d_i \in \mathbb{N}$*, and* P *is a polynomial.*

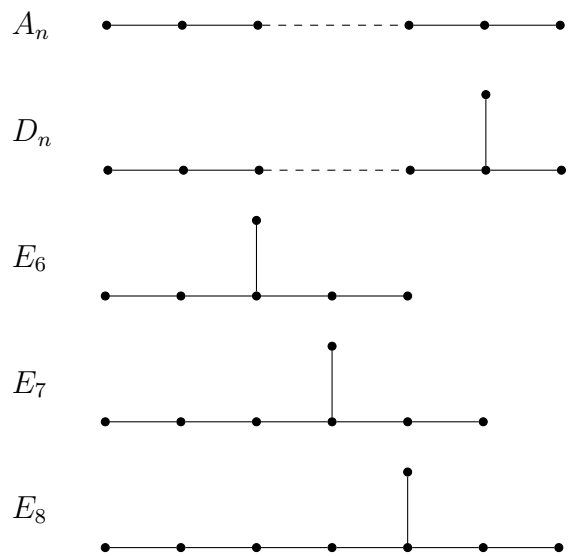
Indeed, an even stronger result is expected to be true: the so-called *Positivity Conjecture.*

Conjecture 9 (Positivity Conjecture)**.** *The Laurant polynomial expansion of any cluster variable has positive integer coefficients.*

This was recently proved in the case of skew-symmetric cluster algebras in [26]. Together, these serve to "tame" the mutation operation. No matter how much you mutate, the cluster variables that appear can be expressed in terms of the initial cluster variables.

Thinking back to the A_2 cluster algebra in Section 5.2, another surprising property was that the cluster algebra is actually finite. Why did only five distinct quivers? Why not ten, twenty, or infinitely many? The natural question to ask is: *which initial seeds generate cluster algebras with finitely many seeds?* Cluster algebras with a finite number of seeds are called **finite type**.

The answer is surprisingly beautiful. The classification of finite-type cluster algebras is in fact almost the same as the Cartan-Killing classification of semisimple Lie algberas. This is a subject quite beyond the scope of this thesis, so only a few summary remarks will be made. Semisimple Lie algebras are classified by putting them into correspondance with *Dynkin diagrams*, a type of root system represented by a graph. Some examples are given below.



(5.23)

Finite semisimple Lie algebras have Dynkin diagrams of one of the four infinite type — A_n, B_n, C_n, and D_n —or one of the 5 sporadic types — E_6, E_7, E_8, F_4, and G_2. For cluster algebras, Fomin and Zelevkinsky have the following classification[27].

Theorem 10. *For a cluster algebra $\mathcal{C}$ with cluster variables $\mathcal{A}$ whose quivers have a single connected component, the following are equivalent:*

1. *$\mathcal{C}$ is of finite type*
2. *$\mathcal{A}$ is a finite set,*
3. *For every seed $S = (\boldsymbol{a}, Q)$ of $\mathcal{C}$, the entries of the matrix satisfy $|Q_{ij}Q_{ji}| \leq 3$ for all i, j,*
4. *One of the seeds of $\mathcal{C}$ has a quiver which, viewed as an* undirected *graph, is a simply-laced Dynkin diagram, i.e. a Dynkin diagram of A, D or E-type shown in* (5.23)*.*

The finite cluster algebras are therefore as follows: A_n and D_n for any $n \in \mathbb{N}$, and E_6, E_7, E_8. The rank and order of each of these are known and can be expressed in terms of **Catalan numbers**. The nth Catalan number is [28]

$$C(n) = \frac{1}{n+1}\binom{2n}{n}. \tag{5.24}$$

Cluster Algebra	A_n	D_n	E_6	E_7	E_8
Rank	n	n	6	7	8
Order	$C(n)$	$(3n-2)C(n-1)$	833	4160	25080

Up through rank 11, the order and rank of a cluster algebra is enough to uniquely identify it. Table 1 lists finite-type cluster algebras (of geometric type) through rank six.

Algebra	Rank	Order
A_1	1	2
$A_1 \times A_1$	2	4
A_2	2	5
$A_1 \times A_1 \times A_1$	3	8
$A_1 \times A_2$	3	10
A_3	3	14
$A_1 \times A_1 \times A_1 \times A_1$	4	16
$A_1 \times A_1 \times A_2$	4	20
$A_2 \times A_2$	4	25
$A_1 \times A_3$	4	28
$A_1 \times A_1 \times A_1 \times A_1 \times A_1$	5	32
$A_1 \times A_1 \times A_1 \times A_2$	5	40
A_4	4	42
$A_1 \times A_2 \times A_2$	5	50
D_4	4	50
$A_1 \times A_1 \times A_3$	5	56
$A_1 \times A_1 \times A_1 \times A_1 \times A_1 \times A_1$	6	64
$A_2 \times A_3$	5	70
$A_1 \times A_1 \times A_1 \times A_1 \times A_2$	6	80
$A_1 \times A_4$	5	84
$A_1 \times A_1 \times A_2 \times A_2$	6	100
$A_1 \times D_4$	5	100
$A_1 \times A_1 \times A_1 \times A_3$	6	112
$A_2 \times A_2 \times A_2$	6	125
A_5	5	132
$A_1 \times A_2 \times A_3$	6	140
$A_1 \times A_1 \times A_4$	6	168
D_5	5	182
$A_3 \times A_3$	6	196
$A_1 \times A_1 \times D_4$	6	200
$A_2 \times A_4$	6	210
$A_2 \times D_4$	6	250
$A_1 \times A_5$	6	264
$A_1 \times D_5$	6	364
A_6	6	429
D_6	6	672
E_6	6	833

Table 1: All ADE-type (i.e. skew-symmetric) cluster algebras with rank one to six.

Theorem 11 (Properties of the Exchange Graph)**.** *The following are theorems in the case of finite-type cluster algebras and conjectured to be true generally [24].*

1. *The exchange graph of a cluster algebra depends only on the initial quiver Q.*

2. *Every seed is uniquely defined by it's cluster. The vertices of the exchange graph can therefore be labelled by clusters, up to a permutation of cluster variables within clusters.*

3. *In a cluster algebra of rank n, each cluster contains n cluster variables. Two clusters are adjacent in the exchange graph if, and only if, they have exactly $n-1$ variables in common.*

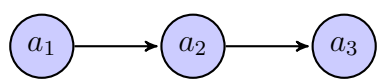
4. *This implies that: for any cluster variables x, the seeds whose clusters contain x form a connected subgraph of the exchange graph.*

This last fact can be leveraged to efficiently find subalgebras of a finite cluster algebra. The corank-1 subalgebras are exactly the connected components of the exchange graph whose vertices all contain a particular cluster variable. In general, a corank-k subalgebra is a subgraph where the same k cluster variables appear. Such subgraphs are necessarily connected. This implies there is a bijection between subsets of the set of cluster variables and subalgebras.

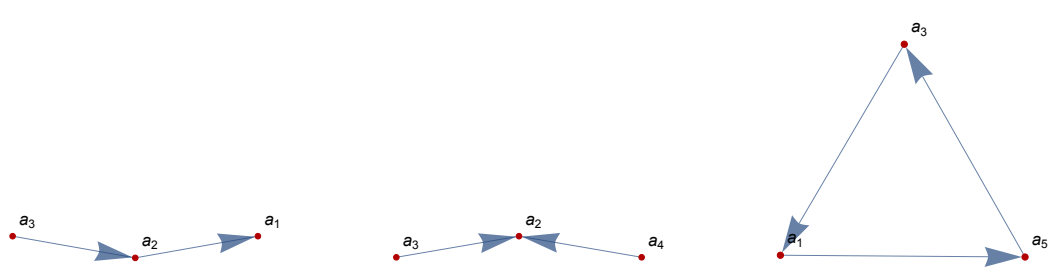
5.4 More Examples of Cluster Algebras

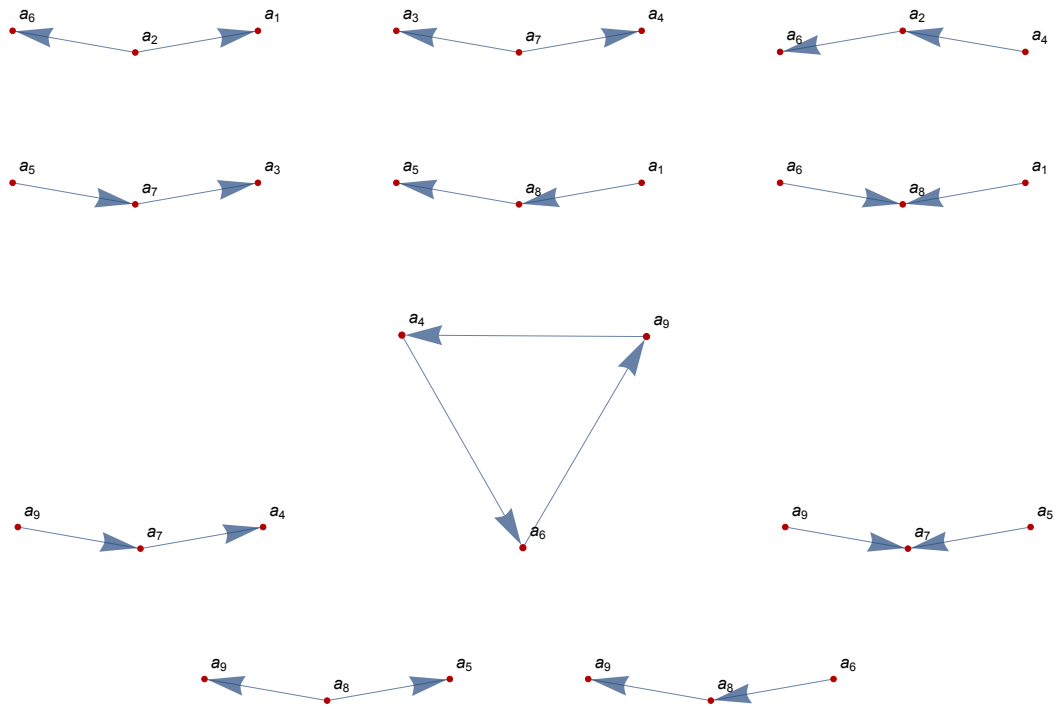
This section presents some of the common cluster algebras that turn up in connection with scattering amplitudes. In practice, the number of seeds and cluster variables quickly becomes unmangable to work with by hand. It is therefore necessary to use computer programs to generate the seeds, perform mutation, and collect the cluster variables. As part of this thesis, a `Mathematica` library to do this was developed by the author and Adam Scherlis to do these computations.

The A_3 cluster algebra starts from the initial seed



It has $C(3) = 14$ seeds.





where the cluster variables are

$$
\begin{aligned}
a_4 &= \frac{1+a_2}{a_1} \\
a_5 &= \frac{a_1+a_3}{a_2} \\
a_6 &= \frac{1+a_2}{a_3} \\
a_7 &= \frac{a_1+a_3+a_2a_3}{a_1a_2} \\
a_8 &= \frac{a_1+a_1a_2+a_3}{a_2a_3} \\
a_9 &= \frac{(1+a_2)(a_1+a_3)}{a_1a_2a_3}.
\end{aligned}
$$

The exchange graph can be seen in Figure 1. The exchange graph actually has a geometric interpretation as the *Stasheff polytope*, a three-dimensional polytope [29]. In fact, the exchange graph of a finite-type cluster algebra of rank n can be viewed as a polytope in n-dimensions. A general feature is that these polytopes have faces which are either squares or pentagons. The pentagonal faces correspond to A_2 subalgebras while the square faces correspond to $A_1 \times A_1$ subalgebras. There are 6 pentagonal faces and 3 square faces. The seeds S_3 and S_{11}, each at the intersection of 3 pentagons, are known as the "poles" of the polytope.

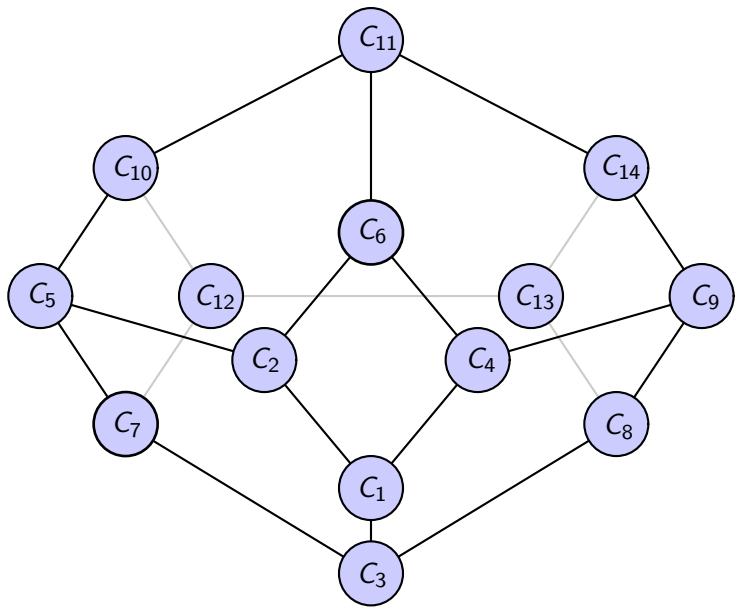
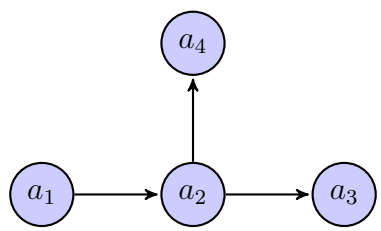


Figure 1: The exchange graph for A_3.

The D_4 cluster algebra has 50 seeds and 16 cluster variables. It's initial quiver is a directed version of the D_4 Dynkin Diagram.



The exchange graph for D_4 is displayed in Figure 2. Since this is a four-dimensional polytope, it is very difficult to visualize one two-dimensional paper. In some sense, it is "built" out of 3 A_3's stuck together. Table 2 lists its subalgebras.

The E_6 cluster algebra is even more complex, with 833 seeds and 42 cluster variables. The exchange graph for E_6 is displayed in Figure 3. The 7-fold and 14-fold symmetry visible in the graph is not a coincidence; this will be discussed in a few sections in connection with Grassmannians. The subalgebras of E_6 — of which there are 4844 — are tabulated in Table 3.

Subalgebra	**Shape**	**Number**
A_1	edge	100
$A_1 \times A_1$	square	30
A_2	pentagon	36
$A_1 \times A_1 \times A_1$	cube	4
A_3	Stasheff polytope	12

Table 2: Subalgebras of D_4.

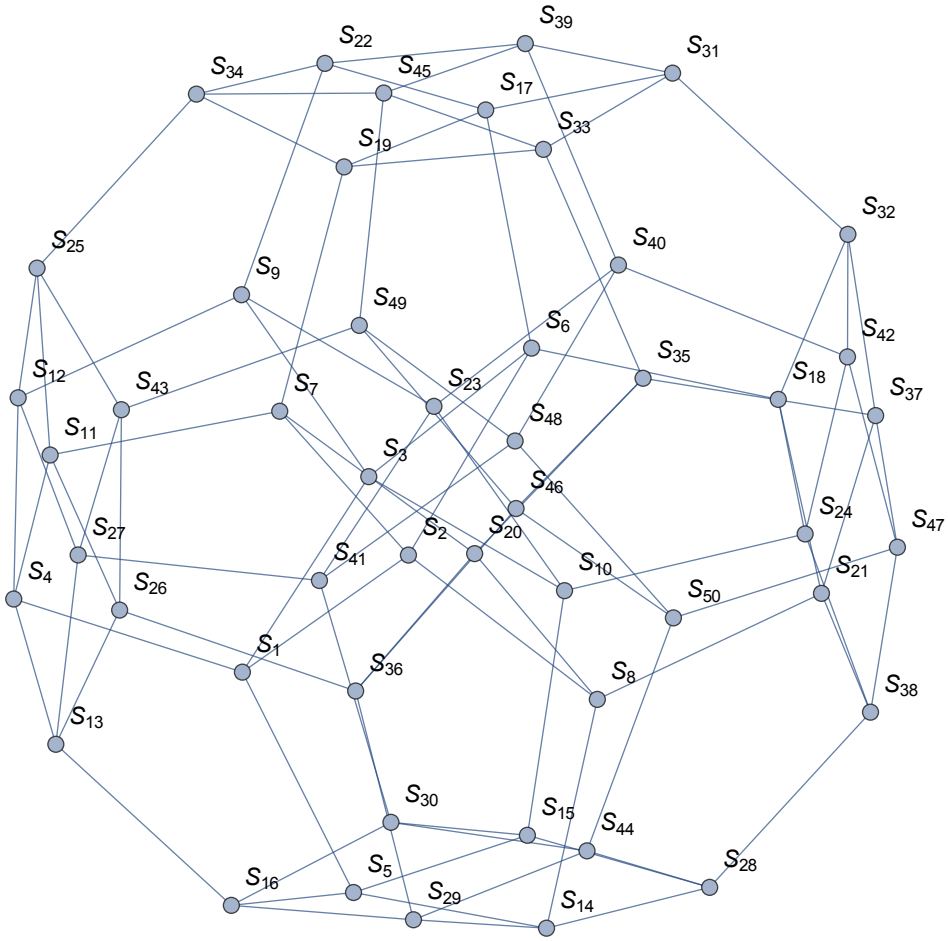


Figure 2: The exchange graph for D_4.

Subalgebra	Number
$A_1 \times A_1$	1785
A_2	1071
$A_1 \times A_1 \times A_1$	357
$A_1 \times A_2$	714
A_3	476
$A_1 \times A_1 \times A_2$	119
$A_2 \times A_2$	21
$A_1 \times A_3$	112
A_4	112
D_4	35
$A_1 \times A_2 \times A_2$	7
$A_1 \times A_4$	14
A_5	7
D_5	14

Table 3: Subalgebras of E_6.

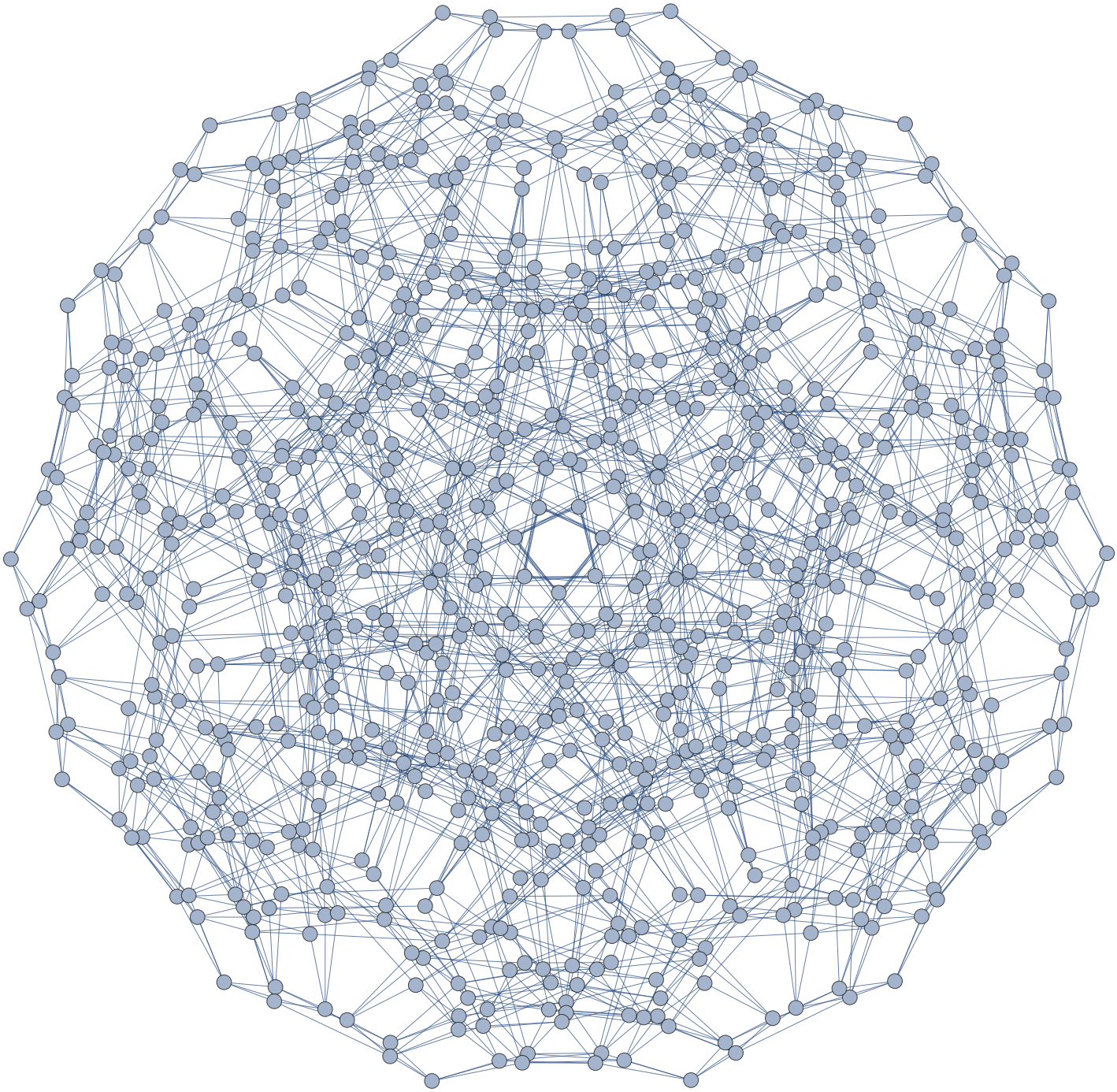


Figure 3: The exchange graph for E_6.

5.5 Cluster $\mathcal{X}$-coordinates

Cluster algebras enjoy a parallel and more geometric structure along side the cluster variables. This variously called the **cluster poisson variety** or **pre-sympletic structure** of the cluster algebra, in the terminology of Goncharov [30] or Gekhtman et al. respectively [24]. In either case, the structure is defined as follows.

Definition 12. *Suppose Q is a quiver with n mutable vertices. Label the vertices of Q with* ***cluster $\mathcal{X}$-coordinates*** $\boldsymbol{X} = \{x_1, \ldots, x_n\}$. *The cluster $\mathcal{X}$-coordinates mutate according to*

$$x_i' = \begin{cases} x_k^{-1} & i = k \\ x_i \left(1 + x_k^{\operatorname{sgn} Q_{ik}}\right)^{Q_{ik}} & i \neq k. \end{cases} \tag{5.25}$$

where Q_{ij} are the coefficients of the matrix corresponding to Q. In the language of Gekhtman et al., these are called τ-coordinates. The set of all $\mathcal{X}$-coordinates on a cluster algebra is denoted by $\mathcal{X}$. Seeds for $\mathcal{X}$-coordinates are tuples $S = (\boldsymbol{X}, Q)$.

There are two special features of the $\mathcal{X}$-coordinates. First, they reproduce the same exchange graph as the cluster variables. Second, there is a Poisson bracket $\{\cdot, \cdot\}$ defined on the $\mathcal{X}$-coordinates by

$$\{x_i, x_j\} = Q_{ij} x_i x_j. \tag{5.26}$$

Proposition 13. *This Poisson bracket is invariant under mutation of $\mathcal{X}$-coordinates. If x_i, x_j and Q mutate to x_i', x_j' and Q', then*

$$\left\{x_i', x_j'\right\} = Q_{ij}' x_i' x_j'. \tag{5.27}$$

The $\mathcal{X}$-coordinates are related to the cluster variables by a map $p : \mathcal{A} \to \mathcal{X}$ given by

$$x_i = \prod_{i \to j} a_j = \prod_j a_j^{Q_{ij}} \tag{5.28}$$

where x_i and a_j are cluster $\mathcal{X}$-coordinates and cluster variables associated with the same quiver Q. There is a natural relation between the exchange graph of the cluster algebra and the Poisson bracket described here [12]. If x_i and x_j appear together in a $A_1 \times A_1$ subalgebra, then $\{x_i, x_j\} = 0$. If they appear in adjacent quivers in an A_2 subalgebra, then $\{x_i, x_j\} = \pm 1$ depending on the ordering of x_i and x_j.

Unfortunately, the cluster $\mathcal{X}$-coordinates are not functionally independent. In many simple cases, products or quotients of two $\mathcal{X}$-coordinates is a third $\mathcal{X}$-coordinate. To choose a "functionally independent basis", one has to choose coordinates such that the corresponding rows of Q are linearly independent. The maximal number of functionally independent $\mathcal{X}$-coordinates is the corank of Q [12].

Another caveat is that $\mathcal{X}$-coordinate interact strangely with frozen vertices. If one starts with a initial quiver, finds all the $\mathcal{A}$-coordinates, and then uses Equation (5.28) to compute the $\mathcal{X}$-coordinates, then in some cases *not all the $\mathcal{X}$-coordinates will be found*. Empirically, there are two ways to cure this. First, one can add in a sufficiently large number of frozen vertices so that all the $\mathcal{X}$-coordinates can be disambiguated. Second, one can start from an initial quiver of $\mathcal{X}$-coordinates and use the $\mathcal{X}$-coordinate mutation formula instead. This confusing point will probably be eliminated in the future when the relationship between the $\mathcal{A}$- and $\mathcal{X}$-coordinates is better understood in the mathematics literature.

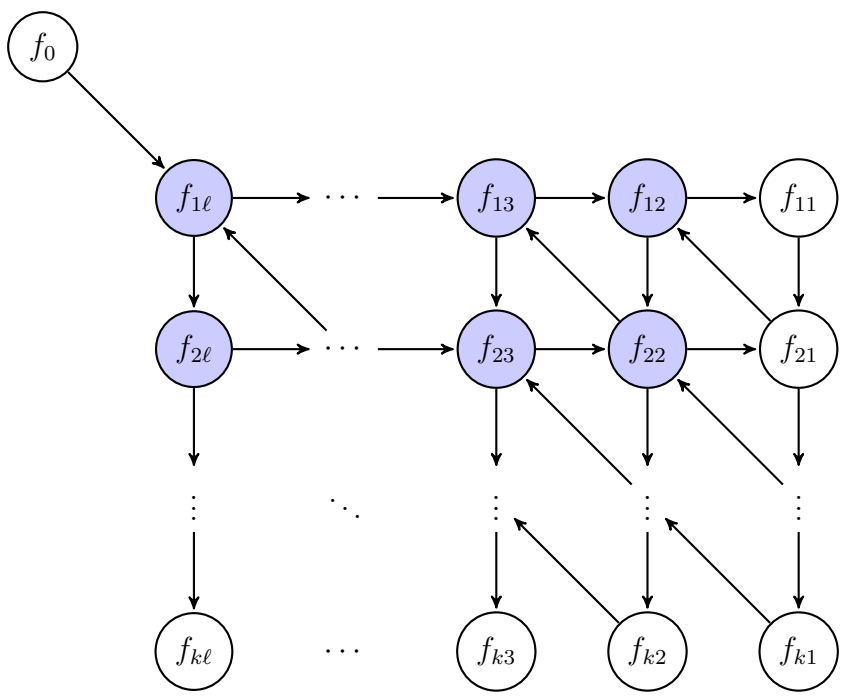
5.6 Grassmannian Cluster Algebras

Grassmannians have a natural cluster algebra structure. Recall Equation (3.13), the Plücker Relation,

$$\langle i, j, I\rangle = \frac{1}{\langle k, \ell, I\rangle}\Big[\langle i, k, I\rangle\langle j, \ell, I\rangle + \langle i, \ell, I\rangle\langle j, k, I\rangle\Big] \tag{5.29}$$

where i, j, k, ℓ are indices of minors and I is a multi-index. This bears a great superficial similarity to the mutation rule for cluster variables, Equation (5.6). Indeed, one can construct a cluster algebra corresponding to $Gr(k, n)$ starting with an initial cluster whose variables are Plücker coordinates. More precisely, for a Grassmannian $\mathbb{G}(k, n)$, the ring of homogenous coordinates $\mathbb{C}[\mathbb{G}(k, n)]$ is a cluster algebra of geometric type, a result due to Scott [31].

The initial quiver for $\mathbb{G}(k, n)$ is[12]



where blue vertices are mutable, white vertices are frozen, and the cluster variables are written in terms of Plücker coordinates as

$$f_0 := \langle 1, \ldots, k\rangle \tag{5.30}$$

and

$$f_{ij} := \begin{cases} \langle i+1, \ldots, k, k+j, \ldots, i+j+k-1\rangle & i \leq \ell - j + 1 \\ \langle 1, \ldots, i+j-\ell-1, i+1, \ldots, k, k+j, \ldots, n\rangle & i > \ell - j + 1. \end{cases} \tag{5.31}$$

The correspondence between Grassmannian cluster algebras and cluster algebras of finite type is displayed in Figure 4. Grassmannian cluster algebras with physical significance for our purposes are $\mathbb{G}(4, n)$.

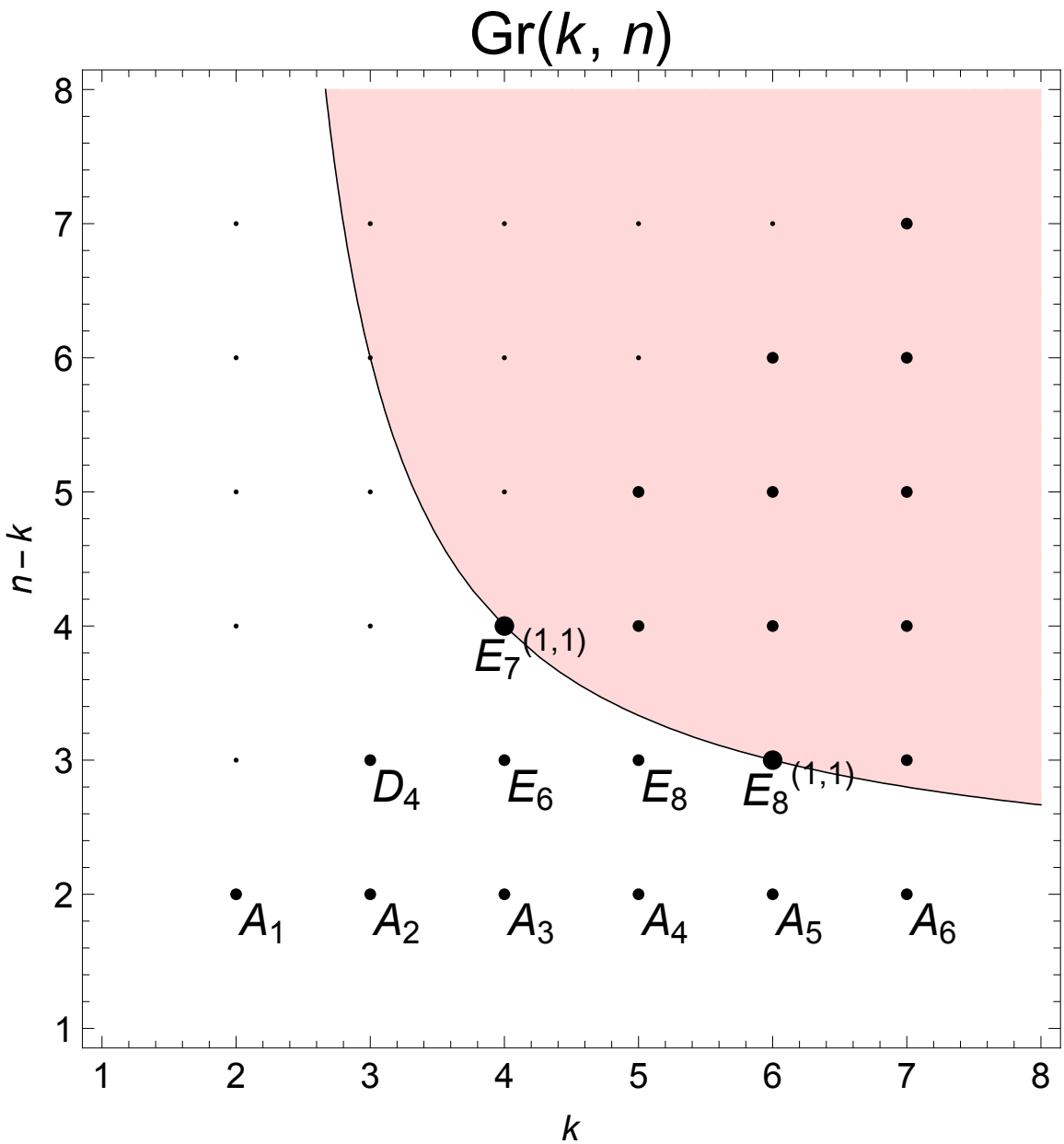


Figure 4: Grassmannian cluster algebras of finite type. Since $\mathbb{G}(k,n) = \mathbb{G}(n-k,n)$, the graph is symmetric about $y = x$. The pink region, above the curve $(x-1)(y-1) = 4$, contains Grassmannian cluster algebras of infinite type. The cluster algebras corresponding to $\mathbb{G}(4,8)$ and $\mathbb{G}(3,9)$ lie on the line and are "semi-infinite": though they have an infinite number of seeds, only a finite number of graphs appear as quivers. For instance, $E_7^{(1,1)}$ has 504 distinct quivers. $E_7^{(1,1)}$ and $E_8^{(1,1)}$ are so-called extended Dynkin diagrams that do not correspond to a finite semisimple Lie algebra nor a finite-type cluster algebra. This classification is due to Scott [31]. Credit for this figure goes to Adam Scherlis.

6 Polylogarithms

Scattering amplitudes in $\mathcal{N} = 4$ SYM are often expressed in terms of a class of special functions called polylogarithms. Several classes of polylogarithms are reviewed here, along with their properties and some tools to work with them.

6.1 Classical Polylogarithms

This section largely follows Brown [32].

The logarithm function can be expressed as a Taylor series about $z = 1$:

$$\ln(z) = \sum_{n=1}^{\infty} \frac{(-1)^n (z-1)^n}{n} \tag{6.1}$$

for $z \in \mathbb{C}$, with a radius of convergence of 1. Modifying this slightly gives a much prettier series

$$-\ln(1-z) = \sum_{n=1}^{\infty} \frac{z^n}{n} \tag{6.2}$$

which converges for $|z| < 1$. This suggests a natural generalization. Define the **kth polylogarithm** Li_n by the series

$$\mathrm{Li}_k(z) := \sum_{n=1}^{\infty} \frac{z^n}{n^k} \tag{6.3}$$

for any positive integer k. Of course, this also converges for $|z| < 1$. The polylogarithm can also by defined recursively, by

$$\begin{cases} \mathrm{Li}_1(z) := -\ln(1-z) = \int_0^z \frac{dw}{1-w} \\ \mathrm{Li}_k(z) := \displaystyle\int_0^z \mathrm{Li}_{k-1}(w) \frac{dw}{w} \end{cases} \tag{6.4}$$

where the integral is taken over any path from 0 to z in the complex plane. If a particular branch of the complex logarithm is chosen, then this definition forms an analytic continuation of the power series definition.

It is easy to check that these two definitions are equivalent on $|z| < 1$ by differentiating:

$$\frac{d}{dz} \mathrm{Li}_k(z) = \frac{d}{dz} \sum_{n=1}^{\infty} \frac{z^n}{n^k} = \sum_{n=1}^{\infty} \frac{z^{n-1}}{n^{k-1}} = \mathrm{Li}_{k-1}(z)$$

where the derivative can be pulled inside the infinite sum because the series converges absolutely on this domain. Integrating both sides of this equation yields (6.4), so the integral definition really is an analytic continuation.

Note that the differential form dw/w that appears in Equation (6.4) is "scale-invariant" in the sense that if $u = \lambda w$ for some $\lambda \in \mathbb{C}$, then

$$\frac{du}{u} = \frac{d(\lambda w)}{\lambda w} = \frac{dw}{w}.$$

Polylogarithms obey a vast number of identities, of which it is suspected that only a small fraction are known. For instance, for any $n \in \mathbb{N}$,

$$\mathrm{Li}_k(z) + \mathrm{Li}_k(-z) = 2^{1-k}\,\mathrm{Li}_k(z^2). \tag{6.5}$$

The dilogarithm, Li_2 obeys an *inversion formula*

$$\mathrm{Li}_2(z) + \mathrm{Li}_2(1/z) = -\frac{1}{2}\log(-z)^2 - \mathrm{Li}_2(1). \tag{6.6}$$

This is the starting point of a connection between polylogarithms and the Riemann-Zeta function. The Riemann-Zeta function is defined by the power series

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}, \quad \mathrm{Re}(s) > 1 \tag{6.7}$$

which is very similar to Equation (6.3). Indeed, $\mathrm{Li}_k(1) = \zeta(k)$ for all k. So

$$\mathrm{Li}_2(1) = \zeta(2) = \frac{\pi^2}{6} \tag{6.8}$$

and the inversion formula can be rewritten as

$$\mathrm{Li}_2(z) + \mathrm{Li}_2(1/z) = -\frac{1}{2}\log(-z)^2 - \frac{\pi^2}{6}. \tag{6.9}$$

The **Zeta values** $\zeta(k)$ turn up often in scattering amplitudes of many theories, including $\mathcal{N} = 4$ SYM.

Perhaps the most interesting polylogarithm identity is the **Abel Identity**:

$$\begin{aligned}
&\mathrm{Li}_2(-x) + \log x \log y \\
&+ \mathrm{Li}_2(-y) + \log y \log\left(\frac{1+y}{x}\right) \\
&+ \mathrm{Li}_2\left(-\frac{1+y}{x}\right) + \log\left(\frac{1+y}{x}\right)\log\left(\frac{1+x+y}{xy}\right) \\
&+ \mathrm{Li}_2\left(-\frac{1+x+y}{xy}\right) + \log\left(\frac{1+x+y}{xy}\right)\log\left(\frac{1+x}{y}\right) \\
&+ \mathrm{Li}_2\left(-\frac{1+x}{y}\right) + \log\left(\frac{1+x}{y}\right)\log x \\
&= -\frac{\pi^2}{2}
\end{aligned} \tag{6.10}$$

over the appropriate domain. This is actually the first instance of a connection between polylogarithms and cluster algebras. Recall from Section 5.2 that the cluster variables for the A_2 cluster algebra are

$$\mathcal{A}(A_2) = \{a_1, \ldots, a_5\} = \left\{a_1, a_2, \frac{1+a_1}{a_2}, \frac{1+a_1+a_2}{a_1 a_2}, \frac{1+a_2}{a_1}\right\}. \tag{6.11}$$

In terms of these, the Abel Identity can be recast as

$$\sum_{i=1}^{5} \Big[\mathrm{Li}_2(-a_i) + \log(a_i)\log(a_{i+1}) \Big] = -\zeta(2) \tag{6.12}$$

where indices on a_i are taken mod 5. This connection between cluster algebras and polylogarithms is explored by Goncharov in [33]. (There are several ways of writing the Abel Identity. This presentation was chosen to make cluster algebra structure clear.)

6.2 Iterated Integrals

By collapsing the recursion in Equation (6.4), the classical polylogarithm can be written as an iterated integral:

$$\begin{aligned}
\mathrm{Li}_n(z) &= \int_0^z \mathrm{Li}_{n-1}(w_n)\frac{dw_n}{w_n} \\
&= \int_0^z \int_0^{w_n} \mathrm{Li}_{n-2}(w_{n-1})\frac{dw_{n-1}}{w_{n-1}}\frac{dw_n}{w_n} \\
&= \int_0^z \int_0^{w_n} \cdots \int_0^{w_2} \frac{dw_1}{1-w_1}\frac{dw_2}{w_2}\cdots\frac{dw_n}{w_n} \\
&= \int_0^z \int_0^{w_n} \cdots \int_0^{w_2} d\log(1-\omega_1)d\log\omega_2\cdots d\log\omega_n.
\end{aligned}$$

This is an example of an iterated integral. Iterated integrals were first studied systematically by K. T. Chen [34]. Modern reviews are [35, 32, 36]. The last two have a particular emphasis on physical applications. Here is the basic definition:

Definition 14. *Suppose M is a smooth real or complex manifold and let $\gamma : [0,1] \to M$ be a piecewise smooth curve. Let $\omega_1, \ldots, \omega_n \in \Lambda^1(M)$ be smooth 1-forms. Write the pullback of ω_i by γ as $\gamma^*(\omega_i) = f_i(t)\,dt \in \Lambda^1(\mathbb{R})$. Then the* ***iterated integral*** *of $\omega_1, \ldots, \omega_n$ is defined by*

$$\int_\gamma \omega_n \circ \cdots \circ \omega_1 := \int_{0\le t_1\le\cdots\le t_n\le 1} f_n(t_n)dt_n \cdots f_1(t_1)dt_1. \tag{6.13}$$

Classical polylogarithms can therefore be thought of as iterated integrals

$$\mathrm{Li}_k(z) = \int_{\gamma_z} \omega_1 \circ \underbrace{\omega_0 \circ \cdots \circ \omega_0}_{n-1} \tag{6.14}$$

where $\gamma_z : [0,1] \to \mathbb{C}\setminus\{0,1\}$ is a curve from 0 to z, i.e. $\gamma(0) = 0$ and $\gamma(1) = z$, and

$$\gamma^*(\omega_1) = d\log(1-t), \gamma^*(\omega_0) = d\log t. \tag{6.15}$$

Usually, no notational distinction is made between the forms and their pullbacks.

Unfortunately, there is an ambiguity here. There are many paths γ_z from 0 to z, which one should be used? Chen found that there is actually a connection between the homotopy type of the path and the value of the integral. The case of a singularly-iterated integral is

familiar from multivariable calculus. If ω is a one-form, and γ_1 and γ_2 are such that γ_1 and γ_2 are homotopic then

$$\int_{\gamma_1} \omega = \int_{\gamma_2} \omega \tag{6.16}$$

if and only if ω is closed. The situation for general iterated integrals is more complicated and will be the subject of Section 6.4.

Proposition 15. *Iterated integrals enjoy the following properties.*

1. *The iterated integral $\int_\gamma \omega_1 \circ \cdots \circ \omega_n$ does not depend on the choice of parameterization of γ.*
2. *If $\gamma^{-1}(t) := \gamma(1-t)$ is the reverse path, then*

$$\int_{\gamma^{-1}} \omega_1 \circ \cdots \circ \omega_n = (-1)^n \int_\gamma \omega_n \circ \cdots \circ \omega_1. \tag{6.17}$$

3. *If $\alpha, \beta : [0,1] \to M$ are paths so that α ends where β begins , then the iterated integral over their composite path $\alpha\beta$ is given by*

$$\int_{\alpha\beta} \omega_1 \circ \cdots \circ \omega_n = \sum_{i=0}^{n} \int_\alpha \omega_1 \circ \cdots \circ \omega_i \int_\beta \omega_{i+1} \circ \cdots \circ \omega_n. \tag{6.18}$$

Furthermore, the space of iterated integrals is equipped with *shuffle product*. Intuitively, a shuffle is a permutation that interleaves two sets as if they were stacks of cards being riffle shuffled.

Definition 16. *A (n,m)-**shuffle** is a permutation $\sigma \in S_{n+m}$ (the group of permutation of size $n+m$) such that*

$$\sigma(1) < \sigma(2) < \cdots < \sigma(n) \text{ and } \sigma(n+1) < \sigma(n+2) < \cdots < \sigma(m) \tag{6.19}$$

The set of (n,m)-shuffles is denoted $S_{n,m}$.

Proposition 17. *Two iterated integrals over the same path γ may be multiplied by*

$$\left(\int_\gamma \omega_1 \circ \cdots \circ \omega_n\right)\left(\int_\gamma \omega_{n+1} \circ \cdots \circ \omega_{n+m}\right) = \sum_{\sigma \in S_{n,m}} \int_\gamma \omega_{\sigma(1)} \circ \cdots \circ \omega_{\sigma(n+m)} \tag{6.20}$$

6.3 Shuffle Algebras and Lyndon Words

Property (6.20) above is perhaps the most important. It allows iterated integrals to be thought of as part of a *shuffle algebra*. An overview of shuffle algebras is given in Chapter 12 of [37]. A much more relevant reference is by Radford [38]

Definition 18. *Let X be a set with a total ordering $\leq$. The pair $(X, \leq)$ is known as the* ***alphabet****. Let V_X be a vector space with a basis $B = \{v_\alpha \ : \ \alpha \in X\}$ over a field k. One thinks of the basis vectors as "letters" in this alphabet. The* ***free shuffle algebra*** *for X is total tensor space*

$$\mathrm{sh}(X) := \bigoplus_{n=0}^{\infty} T^n V_X \tag{6.21}$$

equipped with the shuffle product $\text{Ш} : T^n V_X \otimes T^m V_X \to T^{n+m} V_X$ given by

$$(v_1 \otimes v_2 \otimes \cdots \otimes v_n) \text{ Ш } (v_{n+1} \otimes \cdots \otimes v_m) := \sum_{\sigma \in S_{n,m}} v_{\sigma(1)} \otimes v_{\sigma(2)} \otimes \cdots \otimes v_{\sigma(n+m)}. \tag{6.22}$$

An element that can be written as a sum of k-tensor products $w = \sum_i w_{1i} \otimes \cdots \otimes w_{ki}$ is said to be ***weight*** *k.*

Since $T^0 V_X \cong k$, $\mathrm{sh}(X)$ has a unit element 1. Because of the shuffle product, many elements of $\mathrm{sh}(X)$ can be written in terms of products of elements of lower weight.

Definition 19. *A element $w \in \mathrm{sh}(X)$ is said to be* ***irreducible*** *if it* cannot *be written as the shuffle product of other elements. Precisely, for all $v, u \in sh(X), v \text{Ш} u \neq w$ whenever $u \neq 1 \neq v$. For $k \in \mathbb{N}$, define the set of elements of* ***pure weight*** *as*

$$sh^k(X) := \left\{ w \in T^k V_X \ : \ w \text{ is irreducible} \right\}. \tag{6.23}$$

This forms an $\mathbb{N}$-grading for $\mathrm{sh}(X)$.

The space $\mathrm{sh}^k(V)$ is a graded commutative Hopf algebra, equipped with a coproduct, counit, and antipode map. Some of these additional properties will be discussed in Section 7.

The next task is to find bases for $\mathrm{sh}(X)$ and $\mathrm{sh}^k(X)$. The first is easy.

Proposition 20. *The set*

$$\beta(X) := \{v_{\alpha_1} \otimes \cdots \otimes v_{\alpha_k} \ : \ \alpha_i \in X, k \in \mathbb{N}\} \tag{6.24}$$

is a basis for $\mathrm{sh}(X)$ as a vector space.

This is the usual basis for a tensor product space. However, since there is a shuffle product, what is really needed is an *algebraic basis*. Roughly, a minimal set of elements such that products and linear combinations generate the whole space. Finding an algebraic basis for $\mathrm{sh}^k(X)$ is a challenge, in which the combinatorics of *Lyndon words* play a role.

Definition 21. *Suppose $(X, \leq)$ is an alphabet as above. A* ***word*** *is an element*

$$x = x_1 x_2 \cdots x_k \in X^k, \tag{6.25}$$

written as a product instead of a tuple. The total ordering extends to X^k by lexicographical (alphabetical) ordering. A ***Lyndon word*** *is a word whose ordering is strictly less than all of its cyclic permutations. Let $\mathrm{Lyn}^k(X)$ denote the set of all Lyndon words of length n.*

For example, if $X = \{a, b\}$ is a two letter alphabet with ordering $a \leq b$, then the Lyndon words of length 4 are

$$\text{aaab} \quad \text{aabb and abbb.}$$

Proposition 22. *The number of Lyndon words of length n in an alphabet of length k is[39]*

$$N_k(n) = \frac{1}{n} \sum_{d|n} \mu(d) k^{n/d} \tag{6.26}$$

where μ is the ***Möbius function*** *defined on $\mathbb{N}$ by*

$$\mu(n) = \begin{cases} 1 & \text{if } n = 1 \\ (-1)^k & \text{if } n = p_1 \cdots p_k \text{ are distinct primes} \\ 0 & \text{if } n \text{ is not square free, i.e. } d^2 \mid p \text{ for some } d \in \mathbb{N}. \end{cases} \tag{6.27}$$

Computationally, Lyndon words may be efficiently generated using Duval's Algorithm [40, 41]. It is linear in the number of Lyndon words generated, which is the best that could be expected. The algorithm works as follows. Suppose $X = \{a \leq b \leq \cdots \leq z\}$ is an alphabet with n letters. The algorithm will generate all words *up to* length k. Given a Lyndon word w, define it's *successor* word $S(w)$ as follows:

1. Let v be the word formed by repeating w until it is longer than length n, then cutting off whatever is longer than n, so the length of v is n.
2. Remove any z's from the end of v.
3. Replace the last letter of v with the next letter: $a \to b, b \to c, etc.$ Since the last letter is never z, no replacement is needed for z.
4. Let this word be $S(w)$.

The algorithm is then as follows:

1. Start with the word a.
2. Generate successor words.
3. Terminate upon reaching the one-letter word z.
4. Return all words generated.

The further combinatorics of Lyndon words are interesting in their own right, and in particular are used in the Poicaré-Birkhoff-Witt theorem in Lie theory. However, nothing further is needed to understand shuffle algebras.

Theorem 23 (Radford's Theorem)**.** *Let* $\mathrm{sh}(X)$ *be a shuffle algebra as above. Then*

$$\beta^k(X) := \left\{ v_{\alpha_1} \otimes \cdots \otimes v_{\alpha_k} \ : \ \alpha_1 \alpha_2 \cdots \alpha_k \in \mathrm{Lyn}^k(X) \right\} \tag{6.28}$$

is a basis for the space of irreducible, weight k elements of $\mathrm{sh}(X)$*. This is a paraphrase of Theorem 3.1.1 of [38] in terms of Lyndon words.*

While a proof is beyond the scope of this thesis, it is not too complicated. The main difficulty is deciding exactly what "prime factorization" means in the context of words.

6.4 The Symbol and Chen's Theorem

This section will answer the following question: *under what conditions is an iterated integral* $\int_\gamma \omega_1 \circ \cdots \circ \omega_n$ *independent of the choice of path* γ*?* The answer will provide one of the most useful tools for working with special functions in scattering amplitudes, *The Symbol.* The Symbol was introduced by Goncharov, Spradlin, Vergu, and Volovich in [2]. This section partially follows the presentation in [36, 32].

Definition 24. *Suppose that* $\Omega \subset \Lambda^1 M$ *is a set (often finite) of one-forms where* M *is a smooth real or complex manifold. Suppose that* $\omega_1, \cdots, \omega_n \in \Omega$*. Define the* ***integration map*** *from 1-forms to iterated integrals by*

$$I_\gamma(\omega_1 \otimes \cdots \otimes \omega_n) = \int_\gamma \omega_1 \circ \cdots \circ \omega_n. \tag{6.29}$$

One can check that this is an homomorphism of algebras, i.e. I_γ commutes with the shuffle product Ш.

Definition 25. *Further define a linear "derivative"* $D : T^n\Omega \to T^{n+1}\Omega \oplus T^{n-2}\Omega \otimes \Lambda^2 M$ *by*

$$D\left(\omega_1 \otimes \cdots \otimes \omega_n\right) := \sum_{k=1}^{n} \omega_1 \otimes \cdots \otimes d\omega_i \otimes \cdots \otimes \omega_n + \sum_{k=1}^{n-1} \omega_1 \otimes \cdots \otimes \omega_{i-1} \otimes \omega_i \wedge \omega_{i+1} \otimes \omega_{i+2} \otimes \cdots \otimes \omega_n. \tag{6.30}$$

Any element $\xi = \sum_i \omega_{i1} \otimes \cdots \otimes \omega_{in} \in T^n\Omega$ *such that* $D\xi = 0$ *is said to be* ***integrable****. The kernel of* D *is called the set of* ***integrable words*** *on* M *and is denoted by* $B(M) := \ker D$*.*

Theorem 26 (Chen's Theorem [34])**.** *Under appropriate conditions on* Ω*,*

$$I_\gamma : B(M) \to \{\textit{homotopy invariant iterated integrals}\}. \tag{6.31}$$

is an isomorphism.

The upshot is that iterated integrals are independent of path (up to homotopy) if they are integrals of integrable words. Iterated integrals are then well-defined as functions of the endpoint of the path. But, more importantly, one can go the other way; the value of any homotopy-invariant iterated integral is determined entirely by its corresponding integrable word.

This is a tremendously useful tool for scattering amplitudes. Finding the numerous identities between polylogarithms make them difficult to work with, because one can never be certain if a given expression is unique. However, if one works instead with the corresponding integrable words, then equality is a matter of linear algebra; many functional relations are "automatically" taken into account.

In the physics literature, a map called *The Symbol* is used to go between iterated integrals and integrable words.

Definition 27. *Suppose* F *is a space of one-variable functions of the form*

$$f(z) = \int_{\gamma_z} \omega_{\alpha_1} \circ \cdots \circ \omega_{\alpha_n} \tag{6.32}$$

where γ_z is a path from 0 to $z \in \mathbb{C}$. Suppose that the differential forms ω_α are drawn from

$$\Omega = \{\omega_\alpha = d\log g_\alpha \ : \ g_\alpha \text{ is a complex-valued rational function}\}. \tag{6.33}$$

One further mandates that the one-forms are an integrable word, i.e. $\xi = \omega_{\alpha_1} \circ \cdots \circ \omega_{\alpha_n} \in B(\mathbb{C})$, so that the functions $f(z)$ are well-defined.

The Symbol *is a map from iterated integrals into a shuffle algebra of rational functions*

$$\mathcal{S}(f(z)) := \mathcal{S}\left(\int_{\gamma_z} \omega_{\alpha_1} \circ \cdots \circ \omega_{\alpha_n}\right) = \log g_{\alpha_1} \otimes \cdots \otimes \log g_{\alpha_n}. \tag{6.34}$$

It is standard notation to drop the log's on the right hand side. An alternative definition is given in [2] that extends The Symbol to functions of many variables.

Definition 28. *If $F_k(x_1, \ldots, x_n)$ is a transcendental function of weight k in $x_1, \ldots x_x$, its differential may be written as*

$$dF_k = \sum_i F_{i,k-1} d\log R_i \tag{6.35}$$

where $F_{i,k-1}$ are transcendental functions of weight $k-1$ and the R_i's are rational functions in $\{x_1, \ldots, x_n\}$. The Symbol is then defined recursively by

$$\mathcal{S}\left(F_k\right) = \sum_i \mathcal{S}\left(F_{i,k-1}\right) \otimes R_i, \tag{6.36}$$

where the log *'s were dropped on the right hand side as is customary.*

In both of these definitions, a precise set of functions is not specified. In any specific application, some class of polylogarithms is usually used. This lack of precision rarely results in any confusion.

At this point, an example is badly needed. From Equation (6.14),

$$\mathrm{Li}_k(z) = \int_{\gamma_z} -d\log(1-w) \circ d\log w \circ \cdots \circ d\log w. \tag{6.37}$$

From Equation (6.34), it is immediate that

$$\mathcal{S}\left(\mathrm{Li}_k(z)\right) = -\log(1-z) \otimes \underbrace{\log z \otimes \cdots \otimes \log z}_{k-1} = -(1 - z \otimes \underbrace{z \otimes \cdots \otimes z}_{k-1}). \tag{6.38}$$

This reveals a shortcoming of the notation of dropping the log's — constant factors at the front must be kept track of separately, and do not distribute:

$$-\left(1 - z \otimes \underbrace{z \otimes \cdots \otimes z}_{k-1}\right) \neq (-1+z) \otimes \underbrace{z \otimes \cdots \otimes z}_{k-1}. \tag{6.39}$$

This can also be done from the other definition, (6.36). Equation (6.4) implies

$$\mathrm{Li}_k(z) = \begin{cases} \displaystyle\int_0^z \frac{dw}{1-w} = \int_0^z -d\log(1-w) & k = 1 \\ \displaystyle\int_0^z \mathrm{Li}_{k-1}(w)\, d\log w & k > 1 \end{cases} \tag{6.40}$$

so

$$d\,\mathrm{Li}_k = \begin{cases} -d\log(1-z) & k=1 \\ \mathrm{Li}_{k-1}(z)\,d\log z & k>1. \end{cases} \tag{6.41}$$

Therefore

$$\mathcal{S}(\mathrm{Li}_k(z)) = \begin{cases} -(1-z) & k=1 \\ \mathcal{S}\left(\mathrm{Li}_{k-1}(z)\right)\otimes z & k>1 \end{cases} = -(1-z)\otimes \underbrace{z\otimes\cdots\otimes z}_{k-1}, \tag{6.42}$$

which of course matches the result using the other definition. Many further examples of symbols will follow in latter sections. Reviews on the properties of symbols and their mathematical properties have been made by Duhr *et al.* in [42, 43].

Proposition 29. *The symbol satisfies the following properties.*

1. *Symbols are additive with respect to multiplication:*

$$a_1\otimes\cdots a_n\otimes XY\otimes b_1\otimes\cdots\otimes b_n = a_1\otimes\cdots a_n\otimes X\otimes b_1\otimes\cdots\otimes b_n + a_1\otimes\cdots a_n\otimes Y\otimes b_1\otimes\cdots\otimes b_n. \tag{6.43}$$

   This is immediate from the multilinearity of tensor products if the implicit logs are put back in.

2. *This also implies*

$$a_1\otimes\cdots a_n\otimes X^n\otimes b_1\otimes\cdots\otimes b_n = n\left(a_1\otimes\cdots a_n\otimes X\otimes b_1\otimes\cdots\otimes b_n\right) \tag{6.44}$$

3. *The Symbol commutes with shuffle products:*

$$\mathcal{S}\left(F \text{ ш } G\right) = \mathcal{S}(F) \text{ ш } \mathcal{S}(G). \tag{6.45}$$

   This follows from the fact that the integration map is an algebra homomorphism.

Section 6.6 will give more examples of symbols. It is first necessary to define some broader classes of polylogarithms.

6.5 Generalized Polylogarithms

The classical polylogarithms of Section 6.1 are the simplest special functions that arise in this subject. This section introduces the full "cast" of polylogarithms. Reviews for physicists are given in [42, 43, 44] and are studied more fully by Goncharov in [45].

Definition 30. *Define a* ***generalized polylogarithm*** $I:\mathbb{C}^{k+2}\to\mathbb{C}$ *as an iterated integral by*

$$I(a_0;a_1,\ldots,a_n;a_{n+1}) = \int_{a_0}^{a_{k+1}} \frac{dt}{t-a_1}\circ\frac{dt}{t-a_2}\circ\cdots\circ\frac{dt}{t-a_k}. \tag{6.46}$$

Proposition 31. *The generalized polylogarithm is invariant under affine transformation* $a_i\mapsto\alpha a_i+\beta$ *for all* $0\le i\le k+1$ *with* $0\neq\alpha,\beta\in\mathbb{C}$.

In particular, this means that the first index isn't really necessary:

$$I(a_0; a_1 \ldots, a_k; a_{k+1}) = I(0; a_1 - a_0, \ldots, a_k - a_0; a_{k+1} - a_0). \tag{6.47}$$

This motivates the following definition.

Definition 32. *The* ***Goncharov polylogarithm*** $G : \mathbb{C}^{k+1} \to \mathbb{C}$ *is defined by*

$$G(a_1, \ldots, a_k; a_{k+1}) = \begin{cases} \displaystyle\int_0^{a_{k+1}} \frac{dt}{t-a} & \text{if } k = 1 \\ \displaystyle\int_0^{a_{k+1}} G(a_2, \ldots, a_k; t) \frac{dt}{t-a_1} & \text{if } k > 1 \end{cases} \tag{6.48}$$

$$= \int_0^{a_{k+1}} \frac{dt}{t-a_1} \circ \frac{dt}{t-a_2} \circ \cdots \circ \frac{dt}{t-a_k}.i \tag{6.49}$$

As a notational shorthand, it is common to use $G(\boldsymbol{a}; x) = G(a_1, \ldots, a_k; x)$*. The* ***weight*** *of* $G(\boldsymbol{a}; x)$ *is defined as* $k = |\boldsymbol{a}|$*.*

Generalized polylogarithms and Goncharov polylogarithms contains equivalent information, but with one fewer argument. The Goncharov polylogarithms have a very rich structure, with many beautiful properties and relations. However, it is sometime more symmetrical and convenient to use the I notation instead of the G notation. As a caveat, there is some notational variation between papers for these functions. The order of the arguments of G is sometimes reversed from what is shown here.

Proposition 33. *Here are a few of the identities the Goncharov polylogarithm satisfies.*

1. *Closed form expressions in special cases.*

$$\begin{aligned} G(\mathbf{0}_k; x) &= \frac{1}{k!} [\log(x)]^k \\ G(\boldsymbol{a}_k; x) &= \frac{1}{k!} \log^k \left(1 - \frac{x}{a}\right) \\ G(\mathbf{0}_{k-1}, a; x) &= -\operatorname{Li}_k \left(\frac{x}{a}\right) \end{aligned}$$

   where $\mathbf{0}_k = (\underbrace{0, \ldots, 0}_{k})$ *and* $\boldsymbol{a}_k = (\underbrace{a, \ldots, a}_{k})$*.*

2. *Scale invariance. If* $k \in \mathbb{C}^*$ *and* $\boldsymbol{a} = (a_1, \ldots, a_k)$ *then*

$$G(k\boldsymbol{a}; kx) = G(\boldsymbol{a}; x) \tag{6.50}$$

   whenever $a_k \neq 0$*.*

3. *Hölder Convolution. Whenever $a_1 \neq 1$ and $a_k \neq 0$, for all $p \in \mathbb{C}^*$,*

$$G(a_1, \ldots, a_k; 1) = \sum_{i=0}^{k} (-1)^i G\left(1 - a_i, \ldots, 1 - a_1; 1 - \frac{1}{p}\right) G\left(a_{i+1}, \ldots, a_k; \frac{1}{p}\right). \quad (6.51)$$

In particular, in the limiting case $p \to \infty$,

$$G(a_1, \ldots, a_k; 1) = (-1)^n G(1 - a_k, \ldots, 1 - a_1; 1). \quad (6.52)$$

4. *Shuffle product:*

$$G(\boldsymbol{a}; x) G(\boldsymbol{b}; x) = G\left(\boldsymbol{a} \text{ Ш } \boldsymbol{b}; x\right). \quad (6.53)$$

Equation (6.53) implies that for a fixed last argument x, not everything of the form $G(-; x)$ is functionally independent, but that there are polynomial relations between them. This will be revisited in Section 8.

Another common class of polylogarithms are the *multiple polylogarithms.*

Definition 34. *As a power series, a* ***multiple polylogarithm*** *is a function*

$$\mathrm{Li}_{n_1, \ldots, n_\ell} : \mathbb{C}^k \to \mathbb{C}$$

defined by

$$\mathrm{Li}_{n_1, \ldots, n_\ell}(z_1, \ldots, z_\ell) := \sum_{1 \leq k_1 \leq \cdots < k_\ell} \frac{z_1^{k_1}}{k_1^{n_1}} \cdots \frac{z_\ell^{k_\ell}}{k_\ell^{n_\ell}} \quad (6.54)$$

for $n_i \in \mathbb{Z}$.

It turns out that these are actually provide the same class of functions as the Goncharov polylogarithms, due to the following identity:[44]

$$\mathrm{Li}_{n_1, \ldots, n_\ell}(z_1, \ldots, z_\ell) = (-1)^\ell G\left(\mathbf{0}_{n_\ell - 1}, \frac{1}{z_\ell}, \ldots, \frac{1}{z_2 \cdots z_\ell}, \mathbf{0}_{n_1 - 1}, \frac{1}{z_1 z_2 \cdots z_\ell}; 1\right). \quad (6.55)$$

6.6 Examples of Symbols

Thus equipped with many more functions defined by iterated integrals, the next thing to do is compute their symbols. The following theorem of Goncharov is useful.

Theorem 35 (Goncharov [45])**.** *The differential of the generalized polylogarithms is given by*

$$\begin{aligned} &dI(a_0; a_1, \ldots, a_k; a_{n+1}) \\ &= \sum_{i=1}^{k} I(a_0; a_1, \ldots, \widehat{a_i}, \ldots, a_k; a_{n+1}) \left[d \log\left(a_{i+1} - a_i\right) - d \log\left(a_{i-1} - a_i\right)\right]. \end{aligned} \quad (6.56)$$

The proof is just a computation. Applying Equation (6.36) shows the symbol of the generalized polylogarithms is (logs added for clarity)

$$\mathcal{S}\Big(I(a_0;a_1,\dots,a_k;a_{k+1}\Big) = \begin{cases} \log(a_1-a_2)-\log(a_1-a_0) & \text{if } k=1 \\ \displaystyle\sum_{i=1}^{k} \mathcal{S}\left(I(a_0;a_1,\dots,\widehat{a_i},\dots a_k;a_{k+1})\right) \otimes \left[\log(a_i-a_{i-1})-\log\left(a_i+a_{i+1}\right)\right] & \text{if } k>1. \end{cases} \tag{6.57}$$

As an example, the symbol of $\mathrm{Li}_{n,m}$ can be computed from fundamental definitions. By definition,

$$\mathrm{Li}_{n,m}(x,y) := \sum_{i=1}^{\infty}\sum_{j=1}^{\infty}\frac{x^i y^j}{i^n j^m}; \quad |x|<1, |y|<1. \tag{6.58}$$

One method to compute the symbol of an arbitrary function $f(x)$ is to note

$$f(x) = \int_0^x \frac{df}{dt}dt = \int_0^x t\frac{df}{dt}\frac{dt}{t} = \int_0^x t\frac{df}{dt}d\log t$$

so its symbol is recursively

$$\mathcal{S}(f(x)) = \mathcal{S}\left(x\frac{df}{dx}\right)\otimes x, \tag{6.59}$$

so long the right-hand side is well-defined. This doesn't always work, and in fact fails spectacularly for $f(x)=x$, but works for polylogarithms. For functions of multiple variables this generalizes to

$$\mathcal{S}(f(x_1,\dots,x_n)) = \sum_i \mathcal{S}\left(x_i\frac{df}{dx_i}\right)\otimes x_i. \tag{6.60}$$

To apply this to $\mathrm{Li}_{2,2}(x,y)$, we need to compute its differential:

$$d\,\mathrm{Li}_{2,2}(x,y) = \sum_{i=1}^{\infty}\sum_{j=1}^{\infty}\frac{x^{i-1}y^j}{i^{n+1}j^m}dx + \sum_{i=1}^{\infty}\sum_{j=1}^{\infty}\frac{x^i y^{j-1}}{i^n j^{m+1}}dy = \mathrm{Li}_{n-1,m}(x,y)\frac{dx}{x} + \mathrm{Li}_{n,m-1}(x,y)\frac{dy}{y}.$$

Therefore this symbol is

$$\mathcal{S}(\mathrm{Li}_{n,m}(x,y)) = \mathcal{S}(\mathrm{Li}_{n-1,m}(x,y))\otimes x + \mathcal{S}(\mathrm{Li}_{n,m-1}(x,y))\otimes y. \tag{6.61}$$

To apply this recursively, some edge cases need to be understood. If $m=0$, then using the partial sum of the geometric series $\sum_{k=r}^{n} a^n = \frac{a^{n+1}-a^r}{a-1}$,

$$\mathrm{Li}_{n,0}(x,y) = \sum_{i=1}^{\infty}\sum_{j=1}^{i-1}\frac{x^i y^j}{i^n} = \sum_{i=1}^{\infty}\frac{x^i}{i^n}\frac{y^{i+1}-y}{y-1} = \sum_{i=1}^{\infty}\left(\frac{(xy)^i}{i^n(y-1)} - \frac{y}{y-1}\frac{x^i}{i^n}\right) = \frac{y\,\mathrm{Li}_n(x)-\mathrm{Li}_n(xy)}{1-y}$$

and similarly for $n=0$,

$$\mathrm{Li}_{0,m}(x,y) = \sum_{i=1}^{\infty}x^i\sum_{j=1}^{i-1}\frac{y^j}{j^m} = \sum_{j=1}^{\infty}\sum_{i=j+1}^{\infty}\frac{x^i y^i}{j^m} = \sum_{j=1}^{\infty}\frac{y^j}{j^m}\frac{x^{j+1}}{1-x} = \frac{x}{1-x}\sum_{j=1}^{\infty}\frac{(xy)^j}{j^m} = \frac{x}{1-x}\mathrm{Li}_m(xy).$$

Applying (6.61) with $n = m = 1$,

$$\begin{aligned} dL_{1,1}(x,y) &= \mathrm{Li}_{0,1}\,\frac{dx}{x} + \mathrm{Li}_{1,0}\,\frac{dy}{y} \\ &= \frac{x}{1-x}\,\mathrm{Li}_1(xy)\frac{dx}{x} + \frac{y\,\mathrm{Li}_1(x) - \mathrm{Li}_1(xy)}{1-y}\frac{dy}{y} \\ &= -\,\mathrm{Li}_1(xy)d\ln(1-x) - \mathrm{Li}_1(x)d\ln(1-y) - \mathrm{Li}_1(xy)\frac{dy}{y(1-y)}. \end{aligned}$$

Using the fact that $d\log\frac{y}{1-y} = \frac{dy}{y(1-y)}$, the symbol becomes

$$\mathcal{S}(\mathrm{Li}_{1,1}(x,y)) = -\mathcal{S}(\mathrm{Li}_1(xy)) \otimes (1-x) - \mathcal{S}(\mathrm{Li}_1(x)) \otimes (1-y) - \mathcal{S}(\mathrm{Li}_1(xy)) \otimes \frac{y}{1-y}.$$

For convenience, define $\mathcal{S}_{n,m} := \mathcal{S}\left(\mathrm{Li}_{n,m}(x,y)\right)$. So

$$\mathcal{S}_{1,1} = (1-xy) \otimes (1-x) + (1-x) \otimes (1-y) + (1-xy) \otimes \frac{y}{1-y}.$$

Similar calculations give

$$\mathcal{S}_{n,1} = S_{n-1,1} \otimes x - \mathcal{S}(\mathrm{Li}_n(x)) \otimes (1-y) - \mathcal{S}(\mathrm{Li}_n(xy)) \otimes \frac{y}{1-y}$$

and

$$\mathcal{S}_{1,m} = -\mathcal{S}(\mathrm{Li}_m(xy)) \otimes (1-x) + \mathcal{S}_{1,m-1} \otimes y.$$

For further convenience, define $\widehat{\mathcal{S}}_k(z) := \mathcal{S}\left(\mathrm{Li}_k(z)\right)$. This is enough to write a concise recursive definition for $\mathcal{S}_{n,m}$. Combining the above formulae,

$$\mathcal{S}_{n,m} = \begin{cases} (1-xy) \otimes (1-x) + (1-x) \otimes (1-y) + (1-xy) \otimes \frac{y}{1-y} & n = m = 1 \\ \mathcal{S}_{n-1,1} \otimes x + \widehat{\mathcal{S}}_n(x) \otimes (1-y) + \widehat{\mathcal{S}}_n(xy) \otimes \frac{y}{1-y} & m = 1, n > 1 \\ \mathcal{S}_{1,m-1} \otimes y + \widehat{\mathcal{S}}_m(xy) \otimes (1-x) & n = 1, m > 1 \\ \mathcal{S}_{n-1,m} \otimes x + \mathcal{S}_{n,m-1} \otimes y & n > 1, m > 1. \end{cases} \tag{6.62}$$

It's actually possible to put this into a closed form. For the $n = 1$ or $m = 1$ cases, it turns into just a sum from 1 to $n + m$. For larger n and m's, it turns into a sum over paths from the chosen (n, m) to "edge" points on the lattice (i.e. where $n = 1$ or $m = 1$), such that the paths are monotonically decreasing in both n and m. It's probably easier just to use the recursive version in practice.

As another example, Abel's identity can be proven with symbols. It is easier to use a slightly different form that was used above:

$$\mathrm{Li}_2\,\frac{x}{1-y} + \mathrm{Li}_2\,\frac{y}{1-x} - \mathrm{Li}_2\,\frac{xy}{(1-x)(1-y)} = \mathrm{Li}_2\,x + \mathrm{Li}_2\,y + \ln(1-x)\ln(1-y). \tag{6.63}$$

If the symbols of the right and left hand side are the same, then the identity holds, possibly up to additive constants.

The first task is to find the symbol for all the summands. By Equation (6.38),

$$\begin{aligned}
\mathcal{S}\left(\operatorname{Li}_2 \frac{x}{1-y}\right) &= -\left(1-\frac{x}{1-y}\right)\otimes\frac{x}{1-y} = -\left(\frac{1-x-y}{1-y}\right)\otimes\frac{x}{1-y}\\
\mathcal{S}\left(\operatorname{Li}_2 \frac{y}{1-x}\right) &= -\left(1-\frac{y}{1-x}\right)\otimes\frac{y}{1-x} = -\left(\frac{1-y-x}{1-x}\right)\otimes\frac{y}{1-x}\\
\mathcal{S}\left(-\operatorname{Li}_2 \frac{xy}{(1-x)(1-y)}\right) &= \left(1-\frac{xy}{(1-x)(1-y)}\right)\otimes\frac{xy}{(1-x)(1-y)}\\
&= \left(\frac{1-x-y+xy-xy}{(1-x)(1-y)}\right)\otimes\frac{xy}{(1-x)(1-y)}.
\end{aligned}$$

The first three terms all have a suspiciously similar factor of $1-x-y$, which suggests some cancellations will occur. Using the commutativity and power properties from Proposition 29, the third term can be expanded as

$$\begin{aligned}
&\left(\frac{1-x-y}{(1-x)(1-y)}\right)\otimes\frac{xy}{(1-x)(1-y)}\\
&= \left(\frac{1-x-y}{(1-x)(1-y)}\right)\otimes\frac{y}{1-x} + \left(\frac{1-x-y}{(1-x)(1-y)}\right)\otimes\frac{x}{1-y}\\
&= \frac{1}{1-y}\otimes\frac{y}{1-x} + \left(\frac{1-x-y}{1-x}\right)\otimes\frac{y}{1-x} + \frac{1-x-y}{1-y}\otimes\frac{x}{1-y} + \frac{1}{1-x}\otimes\frac{x}{1-y}\\
&= \frac{1-x-y}{1-x}\otimes\frac{y}{1-x} + \frac{1-x-y}{1-y}\otimes\frac{x}{1-y} - (1-y)\otimes\frac{y}{1-x} - (1-x)\otimes\frac{x}{1-y}\\
&= \frac{1-x-y}{1-x}\otimes\frac{y}{1-x} + \frac{1-x-y}{1-y}\otimes\frac{x}{1-y}\\
&\quad - (1-y)\otimes y + (1-y)(1-x) - (1-x)\otimes x + (1-x)(1-y).
\end{aligned}$$

The first two terms exactly cancel out $\mathcal{S}\operatorname{Li}_2\frac{x}{1-y} + \mathcal{S}\operatorname{Li}_2\frac{y}{1-x}$, so the left-hand side of (6.63) reduces to

$$-(1-y)\otimes y + (1-y)(1-x) - (1-x)\otimes x + (1-x)(1-y).$$

Meanwhile, the right hand side is given by

$$\begin{aligned}
&\mathcal{S}\operatorname{Li}_2(x) + \mathcal{S}\operatorname{Li}_2(y) + \mathcal{S}\left(\ln(1-x)\ln(1-y)\right)\\
&= -(1-x)\otimes x - (1-y)\otimes y + \mathcal{S}(\ln(1-x))Ш\mathcal{S}(\ln(1-y))\\
&= -(1-x)\otimes x - (1-y)\otimes y + (1-x)Ш(1-y)\\
&= -(1-x)\otimes x - (1-y)\otimes y + (1-x)\otimes(1-y) + (1-x)\otimes(1-y),
\end{aligned}$$

which is exactly what the left-hand side reduced to. This proves Abel's identity, up to a constant.

7 Results on Cluster Functions

Now that all the background has been assembled, the theme (1.1) can be stated more precisely

Remainder functions $\mathcal{R}_n^{(L)}$ of MHV scattering amplitudes in the planar limit of $\mathcal{N}=4$ SYM tend to be linear combinations of generalized polylogarithms of weight $2L$ whose symbols are composed of $\mathcal{X}$-coordinates of the the cluster algebra $\mathbb{G}(4,n)$.	(7.1)

This has been shown explicitly to be true for the n-particle remainder function at 2 loops, for $n=6,7$ at 3 loops, and is generally expected to be true in a much wider range of cases. Moreover, there is a cluster algebra structure that informs which linear combinations of polylogarithms can appear in the amplitudes. Such combinations are called *cluster polylogarithms*. The full amplitude, at least in specific cases, can be written as a sum of these cluster polylogarithms.

This section will review some of the main definitions of cluster polylogarithms, show the results of a few novel computations, and present some new facts about their structure on finite cluster algebras.

7.1 The Hopf Algebra of Polylogarithms

The space of polylogarithms has the structure of a graded, commutative Hopf algebra. A Hopf algebra, very roughly, is an algebra equipped with a product, coproduct, and antipode map. A coproduct can be thought of as "dismantling" an element of the algebra into simpler pieces. For a full treatment of Hopf algebras see [37]. A treatment suitable for physicists is given in [44]. For Hopf algebras and polylogarithms, a technical treatment is given in [45] and some of the physically relevent results are given in Appendix A of [46].

Definition 36. *Define the* ***space of generalized polylogarithms of weight*** k *to be*

$$\mathcal{A}_k = \left\{ I(a_0; a_1, \ldots, a_k; a_{k+1}) : \mathbb{C}^{k+2} \to \mathbb{C} \right\}. \tag{7.2}$$

Further define

$$\mathcal{A}_\bullet := \bigoplus_{k=0}^{\infty} \mathcal{A}_k. \tag{7.3}$$

$\mathcal{A}$ inherits the shuffle product from (6.22). In this notation, $ш : \mathcal{A}_k \otimes \mathcal{A}_\ell \to \mathcal{A}_{k+\ell}$ by

$$I(a; a_1, \ldots, a_k; b) I(a; a_{k+1}, \ldots, a_{k+\ell}; b) = \sum_{\sigma \in S_{k,\ell}} I(a; a_{\sigma(1)}, \ldots, a_{\sigma(k+\ell)}; b) \tag{7.4}$$

$$= I\Big(a; (a_1, \ldots, a_k) ш (a_{k+1}, \ldots, a_{k+\ell}); b\Big). \tag{7.5}$$

$\mathcal{A}_\bullet$ is also equipped with a coproduct $\Delta : \mathcal{A}_n \to \bigoplus_{k=1}^{n} \mathcal{A}_k \otimes \mathcal{A}_{n-k}$ by

$$\Delta\left(I(a_0; a_1, \ldots, a_n; a_{n+1})\right) = \sum_{0<i_1<\cdots<i_k=n} I(a_0; a_{i_1}, \ldots, a_{i_k}; a_{n+1}) \otimes \left[\prod_{q=0}^{k} I(a_{i_q}; a_{i_q+1}, \ldots, a_{i_{q+1}-1}; a_{i_{q+1}}\right]. \tag{7.6}$$

These are compatible in the sense that:

$$\Delta(I_1 \sqcup\!\sqcup I_2) = \Delta(I_1)\Delta(I_2). \tag{7.7}$$

Furthermore, the coproduct is coassociative: $(\mathrm{id} \otimes \Delta)\Delta = (\Delta \otimes \mathrm{id})\Delta$, which implies that applying Δ multiple times gives a unique result. Therefore $\Delta \circ \cdots \Delta = \Delta^n : \mathcal{A}_n \to \bigotimes_{i=1}^{n} \mathcal{A}_1$ is well-defined. This coproduct provides a *third* definition of The Symbol on $\mathcal{A}_n$ as

$$\mathcal{S} = \Delta^n \pmod{\pi\mathbb{Z}}. \tag{7.8}$$

It is often useful to work with polylogarithms of *pure weight* that cannot be written as products of terms of lower weight.

Definition 37. *Analogously to* (6.23)*, define the* ***space of polylogarithms of pure weight*** *to be*

$$\mathcal{L}_k = \{I \in \mathcal{A}_k \ : \ I \neq RS \text{ for any } R, S \in \mathcal{A}_\bullet\}. \tag{7.9}$$

Again, set $\mathcal{L}_\bullet := \bigoplus_k \mathcal{L}_k$.

On the level of symbols, several related spaces are used frequently. First, define $A_\bullet^{\mathcal{S}}$ and $\mathcal{L}_\bullet^{\mathcal{S}}$ as the images of $\mathcal{A}_\bullet$ and $\mathcal{L}_\bullet$ under the Symbol map, respectively. Further define $B_k^{\mathcal{S}}$ as the quotient of $\mathcal{L}_k^{\mathcal{S}}$ by the ideal generated by functional equations for the classical polylogarithms Li_k. A very useful fact is that for[12]

$$\mathcal{L}_k^{\mathcal{S}} = B_k^{\mathcal{S}}, k < 4. \tag{7.10}$$

Meaning that all polylogarithms of weight three or less can be written as linear combinations of classical polylogarithms. This does *not* hold at weight four or above.

The spaces $A_\bullet^{\mathcal{S}}$ and $\mathcal{L}_\bullet^{\mathcal{S}}$ are also commutative graded Hopf algebras. However, a *completely different* coproduct is used on them. (This is called δ instead of Δ to distinguish the two.) See [3] for details.

Definition 38. *Define the* ***Symbol coproduct*** $\delta : \mathcal{L}_\bullet^{\mathcal{S}} \to \Lambda^2 \mathcal{L}_\bullet^{\mathcal{S}}$ *by*

$$\delta(a_1 \otimes \cdots \otimes a_k) = \sum_{n=1}^{k-1} (a_1 \otimes \cdots \otimes a_n) \bigwedge (a_{n+1} \otimes \cdots \otimes a_k). \tag{7.11}$$

One should note that the large wedge is in a different space and is not the wedge associated with the tensor product $\otimes$. The coproduct satisfies $\delta^2 = 0$ and gives $\mathcal{L}_K^{\mathcal{S}}$ the structure of a Lie coalgebra.

One often wants to restrict an arbitrary symbol to one of pure weight. This can be done by the following projection map.

Definition 39. *Define the projection* $\rho : \mathcal{A}_k^{\mathcal{S}} \to \mathcal{L}_k^{\mathcal{S}}$ *by* $\rho(a_1) = a_1$ *and*

$$\rho\left(a_1 \otimes \cdots \otimes a_k\right) = \frac{k-1}{k}\left[\rho(a_1 \otimes \cdots \otimes a_{k-1}) \otimes a_k - \rho\left(a_2 \otimes \cdots \otimes a_k\right) \otimes a_1\right]. \tag{7.12}$$

One can check that $\rho^2 = \rho$, so it is indeed a projection.

7.2 Examples of the Symbol Coproduct

In this section, the action of the Symbol coproduct δ on $\mathrm{Li}_{2,2}$ will be computed as an example. The calculation is long, though not hard. Notation from Golden *et al.* [3] is used freely in this section; consult it for definitions of $\Lambda^2 B_2$ and $B_3 \otimes \mathbb{C}^*$, as well as a more in-depth discussion of δ and how it can be applied to polylogarithms of low weight. As a shortcut notation, let $\{x\}_k := \mathcal{S}(\mathrm{Li}_k(x))$. By Equation (6.62), the symbol of $\mathrm{Li}_{2,2}(x,y)$ is

$$\begin{aligned}
\mathcal{S}(\mathrm{Li}_{2,2}(x,y)) = {} & (1-y) \otimes (1-x) \otimes x \otimes y \\
& + (1-y) \otimes (1-x) \otimes y \otimes x \\
& + (1-y) \otimes y \otimes (1-x) \otimes x \\
& - ((1-xy) \otimes (1-x) \otimes x \otimes y) \\
& - ((1-xy) \otimes (1-x) \otimes y \otimes x) \\
& - ((1-xy) \otimes x \otimes (1-x) \otimes x) \\
& + (1-xy) \otimes x \otimes x \otimes x \\
& + (1-xy) \otimes x \otimes x \otimes y \\
& + (1-xy) \otimes x \otimes (1-y) \otimes y \\
& + (1-xy) \otimes x \otimes y \otimes x \\
& + (1-xy) \otimes (1-y) \otimes x \otimes y \\
& + (1-xy) \otimes (1-y) \otimes y \otimes x \\
& - ((1-xy) \otimes y \otimes (1-x) \otimes x) \\
& + (1-xy) \otimes y \otimes x \otimes x \\
& + (1-xy) \otimes y \otimes (1-y) \otimes y.y.
\end{aligned} \tag{7.13}$$

At weight 4, the Symbol coproduct is a map

$$\delta : \mathcal{L}_4 \to \Lambda^2 \mathcal{L}_2 \oplus (\mathcal{L}_3 \wedge \mathcal{L}_1) \cong \Lambda^2 B_2 \oplus (B_3 \otimes C^*). \tag{7.14}$$

The wedge product in the second term can be dropped because elements of $\mathcal{B}_3$ and B_1 are in different spaces and cannot be anti-symmetrized. The action of δ is given by

$$\delta\left(a \otimes b \otimes c \otimes d\right)\big|_{\Lambda^2 B_2} = \rho\left(a \otimes b\right) \bigwedge \rho(c \otimes d) \tag{7.15}$$

$$\delta\left(a \otimes b \otimes c \otimes d\right)\big|_{B_3 \otimes \mathbb{C}^*} = \rho\left(a \otimes b \otimes c\right) \bigotimes d - \rho(b \otimes c \otimes d) \bigotimes a \tag{7.16}$$

The outsized $\bigwedge$ and $\bigotimes$ symbols denote formal tensor products between different spaces; these are not the same as the $\otimes$ tensor products from the symbol.

Applying Equation (7.15) to (7.13) and re-writing in terms of the notation $\{z\}_k$ from [3], gives

$$\begin{aligned}
\delta(\mathcal{S}(\mathrm{Li}_{2,2}(x,y))|_{\Lambda^2 B_2} &= -(1-xy)\wedge y \bigwedge \{x\}_2 && (7.17)\\
&+ (1-xy)\wedge y \bigwedge \{x\}_2 && (7.18)\\
&- (1-xy)\wedge x \bigwedge \{x\}_2 && (7.19)\\
&+ (1-xy)\wedge x \bigwedge x\wedge x && (7.20)\\
&+ (1-xy)\wedge x \bigwedge x\wedge y && (7.21)\\
&+ (1-xy)\wedge x \bigwedge \{y\}_2 && (7.22)\\
&+ (1-xy)\wedge x \bigwedge y\wedge x && (7.23)\\
&- (1-xy)\wedge x \bigwedge \{y\}_2 && (7.24)\\
&+ (1-xy)\wedge x \bigwedge \{y\}_2 && (7.25)\\
&- (1-xy)\wedge y \bigwedge \{x\}_2 && (7.26)\\
&+ (1-xy)\wedge y \bigwedge x\wedge x && (7.27)\\
&+ (1-xy)\wedge y \bigwedge \{y\}_2 && (7.28)\\
&+ \{y\}_2 \bigwedge \{x\}_2 && (7.29)\\
&- \{y\}_2 \bigwedge \{x\}_2 && (7.30)\\
&+ \{y\}_2 \bigwedge \{x\}_2 . && (7.31)
\end{aligned}$$

There are many cancellations and combinations here. The combinations

$$(7.17)+(7.18) \;=\; 0, (7.24)+(7.25)=0, \text{ and } (7.30)+(7.31)=0$$

all vanish. Since $a\wedge a = 0$, Equations (7.20), (7.21), (7.23), and (7.27) all vanish individually. Lastly, we can combine two pairs

$$\begin{aligned}
(7.19)+(7.26) &= -(1-xy)\wedge x\wedge\{x\}_2 - (1-xy)\wedge y\wedge\{x\}_2 = -(1-xy)\wedge xy\wedge\{x\}_2\\
&= -\{xy\}_2\wedge\{x\}_2\\
(7.22)+(7.28) &= (1-xy)\wedge x\wedge\{y\}_2 + (1-xy)\wedge y\wedge\{y\}_2 = (1-xy)\wedge xy\wedge\{y\}_2\\
&= \{xy\}_2\wedge\{y\}_2 .
\end{aligned}$$

Altogether, then, this gives

$$\delta(\mathcal{S}(\mathrm{Li}_{n,m}(x,y))|_{\Lambda^2 B_2} = -\{xy\}_2\wedge\{x\}_2 + \{xy\}_2\wedge\{y\}_2 + \{y\}_2\wedge\{x\}_2 .$$

Now for the $B_3\otimes\mathbb{C}^*$ component. Applying Equation (7.16) to the symbol of $\mathrm{Li}_{2,2}(x,y)$ gives $15\times 2\times 2 = 60$ terms, These can be grouped by their last factor: everything is of the form $X\bigotimes z$ for $z\in\{x,y,1-y,1-xy\}$. Collecting terms like this gives:

$$\delta(\mathcal{S}(\mathrm{Li}_{2,2}(x,y)))|_{B_3\otimes\mathbb{C}^*} = A\bigotimes(1-y) + B\bigotimes(1-xy) + C\bigotimes y + D\bigotimes x \tag{7.32}$$

$$\begin{aligned}
&= \Big[(1-x)\wedge(y\otimes x) - x\wedge(y\otimes(1-x)) \\
&\qquad + y\wedge((1-x)\otimes x) - y\wedge(x\otimes(1-x))\Big]\bigotimes(1-y) \\
&+ \Big[-((1-x)\wedge(x\otimes x)) - (1-x)\wedge(y\otimes x) + x\wedge((1-x)\otimes x) + x\wedge((1-y)\otimes y) \\
&\qquad + x\wedge(y\otimes(1-x)) - x\wedge(y\otimes(1-y)) + (1-y)\wedge(x\otimes y) + (1-y)\wedge(y\otimes y) \\
&\qquad - y\wedge((1-x)\otimes x) + y\wedge(x\otimes(1-x)) - y\wedge(x\otimes(1-y)) - y\wedge((1-y)\otimes y)\Big]\bigotimes(1-xy) \\
&+ \Big[-((1-x)\wedge(x\otimes(1-y))) + (1-x)\wedge(x\otimes(1-xy)) - x\wedge(x\otimes(1-xy)) \\
&\qquad - x\wedge((1-y)\otimes(1-xy)) + (1-y)\wedge((1-x)\otimes x) - (1-y)\wedge(x\otimes(1-xy)) \\
&\qquad - y\wedge((1-y)\otimes(1-xy)) - (1-xy)\wedge((1-x)\otimes x) + (1-xy)\wedge(x\otimes x) \\
&\qquad + (1-xy)\wedge(x\otimes(1-y)) + (1-xy)\wedge((1-y)\otimes x) + (1-xy)\wedge(y\otimes(1-y))\Big]\bigotimes y \\
&+ \Big[-((1-x)\wedge(y\otimes(1-y))) + (1-x)\wedge(y\otimes(1-xy)) + x\wedge((1-x)\otimes(1-xy)) \\
&\qquad - x\wedge(x\otimes(1-xy)) - x\wedge(y\otimes(1-xy)) + (1-y)\wedge((1-x)\otimes y) \\
&\qquad + (1-y)\wedge(y\otimes(1-x)) - (1-y)\wedge(y\otimes(1-xy)) - y\wedge((1-x)\otimes(1-y)) \\
&\qquad + y\wedge((1-x)\otimes(1-xy)) - y\wedge(x\otimes(1-xy)) - (1-xy)\wedge((1-x)\otimes y) \\
&\qquad - (1-xy)\wedge(x\otimes(1-x)) + (1-xy)\wedge(x\otimes x) + (1-xy)\wedge(x\otimes y) \\
&\qquad + (1-xy)\wedge((1-y)\otimes y) - (1-xy)\wedge(y\otimes(1-x)) + (1-xy)\wedge(y\otimes x)\Big]\bigotimes x.
\end{aligned} \tag{7.33}$$

The symbol-wedge is supposed to be associative with the symbol, so the placement of parentheses doesn't actually matter here.

Equation (7.32) simplifies considerably. Using the fact that $a\wedge b = -b\wedge a$, all four terms for $A\bigotimes(1-y)$ vanish. For $B\bigotimes(1-xy)$, there are 12 terms. Terms 1 and 3 are both equal to $-\{x\}_3$ and terms 8 and 12 both equal $\{y\}_3$. Furthermore, the following pairs cancel: 2 and 9; 4 and 7; 5 and 10; 6 and 11. This leaves

$$B\bigotimes(1-xy) = (2\{y\}_3 - 2\{x\}_3)\bigotimes(1-xy). \tag{7.34}$$

Next consider the 12 terms $C\bigotimes y$. Using the $a\wedge b = -b\wedge a$ shows two of them cancel immediately. Using $\{z\}_2 = 1 - z\wedge z$ and $\{z\}_3 = (1-z)\wedge z\otimes z$, we can rewrite the remaining terms as:

$$\begin{aligned}
C = &\{xy\}_2\otimes(1-y) - \{x\}_2\otimes(1-y) \\
&+ \{x\}_2\otimes(1-xy) + \{y\}_2\otimes(1-xy) \\
&+ (1-xy)\wedge x\otimes x + (1-y)\wedge(1-x)\otimes x \\
&-(1-xy)\wedge(1-x)\otimes x + (1-xy)\wedge(1-y)\otimes x
\end{aligned}$$

Now apply the five term Abel identity

$$0 = \{x\}_2 + \{y\}_2 - \{xy\}_2 + \left\{\frac{1-x}{1-xy}\right\}_2 + \left\{\frac{1-y}{1+xy}\right\}_2. \tag{7.35}$$

twice: once multiplied by $\otimes(1-y)$ and again with $\otimes(1-xy)$ at the end. Hence

$$\begin{aligned} C = &\{y\}_2 \otimes (1-y) + \left\{\frac{1-x}{1-xy}\right\}_2 \otimes (1-y) + \left\{\frac{1-y}{1-xy}\right\}_2 \otimes (1-y) \\ &+ \{xy\}_2 \otimes (1-xy) - \left\{\frac{1-x}{1-xy}\right\}_2 \otimes (1-xy) - \left\{\frac{1-y}{1-xy}\right\}_2 \otimes (1-xy) \\ &+ (1-xy) \wedge x \otimes x + (1-y) \wedge (1-x) \otimes x - (1-xy) \wedge (1-x) \otimes x + (1-xy) \wedge (1-y) \otimes x. \end{aligned}$$

Then note that

$$\begin{aligned} &\{y\}_2 \otimes (1-y) = -\{(1-y)\}_3 \\ &\{xy\}_2 \otimes (1-xy) = -\{1-xy\}_3 \\ &\left\{\frac{1-y}{1-xy}\right\}_2 \otimes (1-y) - \left\{\frac{1-y}{1-xy}\right\}_2 \otimes (1-xy) = \left\{\frac{y-1}{xy-1}\right\}_3. \end{aligned}$$

Also using $\{z\}_2 = -\{1-z\}_2$,

$$\begin{aligned} &\left\{\frac{1-x}{1-xy}\right\}_2 \otimes (1-y) - \left\{\frac{1-x}{1-xy}\right\}_2 \otimes (1-xy) \\ &= -\left\{\frac{x(1-y)}{1-xy}\right\}_2 \otimes (1-y) + \left\{\frac{x(1-y)}{1-xy}\right\} \otimes (1-xy) \\ &= -\left\{\frac{x(1-y)}{1-xy}\right\}_2 \otimes \frac{1-y}{1-xy} \\ &= -\left\{\frac{x(y-1)}{xy-1}\right\}_3 + \left\{\frac{x(y-1)}{xy-1}\right\}_2 \otimes x. \end{aligned}$$

Recombining everything we have so far, C has been reduced to

$$\begin{aligned} C = &-\{1-y\}_3 - \{1-xy\}_3 + \left\{\frac{y-1}{xy-1}\right\}_3 - \left\{\frac{x(y-1)}{xy-1}\right\}_3 + \left\{\frac{x(y-1)}{xy-1}\right\}_2 \otimes x \\ &+ (1-xy) \wedge x \otimes x + (1-y) \wedge (1-x) \otimes x - (1-xy) \wedge (1-x) \otimes x + (1-xy) \wedge (1-y) \otimes x \end{aligned} \tag{7.36}$$

However,

$$\begin{aligned} \left\{\frac{x(y-1)}{xy-1}\right\}_2 \otimes x &= \frac{x-1}{xy-1} \wedge \frac{y-1}{xy-1} \otimes x + \frac{x-1}{xy-1} \wedge x \otimes x \\ &= (x-1) \wedge (y-1) \otimes x - (x-1) \wedge (xy-1) \otimes x \\ &\quad - (xy-1) \wedge (y-1) \otimes x + 0 + \{x\}_3 - (xy-1) \wedge x \otimes x. \end{aligned}$$

The terms aside from $\{x\}_3$ exactly cancel terms in (7.36). Combining everything, C becomes simply

$$C = -\{1-y\}_3 - \{1-xy\}_3 + \left\{\frac{y-1}{xy-1}\right\}_3 - \left\{\frac{x(y-1)}{xy-1}\right\}_3 + \{x\}_3 .$$

It remains only to simplify the term D in (7.32).Cancelling the things that are clearly zero and combining terms gives

$$\begin{aligned} D = &-\{x\}_2 \otimes (1-xy) + (1-y) \wedge (1-x) \otimes y \\ &+ \{y\}_2 \otimes (1-x) - \{y\}_2 \otimes (1-xy) \\ &- (1-xy) \wedge (1-x) \otimes y + (1-xy) \wedge x \otimes xy \\ &+ (1-xy) \wedge (1-y) \otimes y - \{xy\}_2 \otimes (1-x) \\ &+ (1-xy) \wedge y \otimes x. \end{aligned}$$

Using the 2-term identity twice turns this into:

$$\begin{aligned} &\left\{\frac{1-x}{1-xy}\right\}_2 \otimes (1-xy) + \left\{\frac{1-y}{1-xy}\right\}_2 \otimes (1-xy) - \{xy\}_2 \otimes (1-xy) - \{x\} \otimes (1-x) \\ &- \left\{\frac{1-x}{1-xy}\right\}_2 \otimes (1-x) - \left\{\frac{1-y}{1-xy}\right\}_2 \otimes (1-x) + (1-y) \wedge (1-x) \otimes y - (1-xy) \wedge (1-x) \otimes y \\ &+ (1-xy) \wedge x \otimes x + (1-xy) \wedge x \otimes y - (1-xy) \wedge (1-y) \otimes y + (1-xy) \wedge y \otimes x \\ &= \\ &- \left\{\frac{1-x}{1-xy}\right\}_3 + \{x-1\}_3 + \left\{\frac{(1-x)y}{1-xy}\right\}_3 + \left\{\frac{1-y}{1-xy}\right\}_2 \otimes y \\ &- (1-xy) \wedge xy \otimes (1-xy) + (1-y) \wedge (1-x) \otimes y - (1-xy) \wedge (1-x) \otimes y \\ &+ (1-xy) \wedge x \otimes x + (1-xy) \wedge x \otimes y \\ &- (1-xy) \wedge (1-y) \otimes y + (1-xy) \wedge y \otimes x. \end{aligned}$$

Expanding gives

$$\begin{aligned} \left\{\frac{1-y}{1-xy}\right\}_2 \otimes y = &-(1-y) \wedge y \otimes y - (1-y) \wedge (1-x) \otimes y \\ &+ (1-y) \wedge (1-xy) \otimes y + (1-xy) \wedge (1-x) \otimes y + (1-xy) \wedge y \otimes y. \end{aligned}$$

These exactly cancel out several of the terms above, leaving

$$\begin{aligned} D = &-\left\{\frac{1-x}{1-xy}\right\}_3 + \{x-1\}_3 + \left\{\frac{(1-x)y}{1-xy}\right\}_3 + \{y\}_3 - (1-xy) \wedge xy \otimes (1-xy) \\ &+ (1-xy) \wedge y \otimes y + (1-xy) \wedge x \otimes y + (1-xy) \wedge y \otimes x + (1-xy) \wedge x \otimes x \\ = &-\left\{\frac{1-x}{1-xy}\right\}_3 + \{x-1\}_3 + \left\{\frac{(1-x)y}{1-xy}\right\}_3 + \{y\}_3 + (1-xy) \wedge xy \otimes \frac{xy}{1-xy} \\ = &-\left\{\frac{1-x}{1-xy}\right\}_3 + \{x-1\}_3 + \left\{\frac{(1-x)y}{1-xy}\right\}_3 - \{y\}_3 + (1-xy) \wedge \frac{xy}{1-xy} \otimes \frac{xy}{1-xy} \\ = &-\left\{\frac{1-x}{1-xy}\right\}_3 + \{x-1\}_3 + \left\{\frac{(1-x)y}{1-xy}\right\}_3 - \{y\}_3 + \left\{\frac{xy}{1-xy}\right\}_3 . \end{aligned}$$

Putting all of the parts together, the coproduct is

$$
\begin{aligned}
&\delta(\mathcal{S}(\mathrm{Li}_{2,2}(x,y))|_{B_3\otimes\mathbb{C}^*} \\
&= \Big(2\{y\}_3 - 2\{x\}_3\Big)\bigotimes(1-xy) \\
&\quad + \left(-\left\{\frac{1-x}{1-xy}\right\}_3 + \{x-1\}_3 + \left\{\frac{(1-x)y}{1-xy}\right\}_3 - \{y\}_3 + \left\{\frac{xy}{1-xy}\right\}_3\right)\bigotimes x \\
&\quad + \left(-\{1-y\}_3 - \{1-xy\}_3 + \left\{\frac{y-1}{xy-1}\right\}_3 - \left\{\frac{x(y-1)}{xy-1}\right\}_3 + \{x\}_3\right)\bigotimes y.
\end{aligned}
$$

It is advisable to automate such calculations in `Mathematica` instead of doing them by hand.

7.3 Cluster Polylogarithms on A_2: The A_2 Function

Cluster polylogarithms are defined in Golden *et al.* [3], whose notation will be employed in this section.

The A_2 function is the unique function in $\mathcal{L}_4$ whose Symbol an element of $\mathrm{sh}^4(\mathcal{X}_{A_2})$ that satisfies the integrability condition

$$0 = \delta^2 f = \delta(\delta_{2,2} f) + \delta(\delta_{3,1} f). \tag{7.37}$$

The goal of this section is to find an explicit formula for f. That is to say, to explicitly write the coproduct of the symbol of f, from which f can be deduced.

The action of δ on $\Lambda^2 B_2$ and $B_3 \otimes \mathbb{C}^*$ is given by

$$\delta\left(\{x_i\}_2 \bigwedge \{x_j\}_2\right) = \{x_i\}_2 \bigotimes (1+x_j) \bigwedge x_j - \{x_j\}_2 \bigotimes (1+x_i)) \bigwedge x_i \tag{7.38}$$

$$\delta\left(\{x_i\}_3 \bigwedge x_j\right) = \{x_i\}_2 \bigotimes x_i \bigwedge x_j. \tag{7.39}$$

$$\tag{7.40}$$

Additionally, due to Abel's identity and the properties of A_2,

$$0 = \sum_{i=1}^{5} \{x_i\}_2 \tag{7.41}$$

$$(1+x_i) = x_{i-1}x_{i+1} \tag{7.42}$$

where in the last Equation, indices are understood mod 5. Since $x_i, \{x_i\}_2$, and $\{x_i\}_3$ where $x_i, i = 1, \ldots, 5$ the $\mathcal{X}$-coordinates of A_2, are bases for $\mathbb{C}^*$, B_2 and B_3 respectively, $\delta_{2,2} f \in \Lambda^2 B_2$ can be written as

$$\delta_{2,2} f = \sum_{i<j}^{5} a_{ij} \left(\{x_i\}_2 \wedge \{x_j\}_2\right) \tag{7.43}$$

and

$$\delta_{3,1} f = \sum_{i\neq j}^{5} d_{ij} \left(\{x_i\}_3 \wedge x_j\right) \tag{7.44}$$

for some coefficients a_{ij} and d_{ij}.

Applying the Abel Identity to eliminate $\{x_5\}_2$, and then Equation (7.38) to $\delta_{2,2}f$:

$$\begin{aligned}
\delta(\delta_{2,2}f) &= \sum_{i<j}^{4} c_{ij}\left(\{x_i\}_2 \otimes 1 + x_j \wedge x_j - \{x_j\}_2 \otimes 1 + x_i \wedge x_i\right) \\
&= \sum_{i<j}^{4} c_{ij}\left(\{x_i\}_2 \otimes (x_{j+1} \wedge x_j + x_{j-1} \wedge x_j) - \{x_j\}_2 \otimes (x_{i+1} \wedge x_i + x_{i-1} \wedge x_i)\right) \\
&= \sum_{i<j}^{4} c_{ij}\left(-w(i,j,j+1) + w(i,j-1,j) + w(j,i,i+1) - w(j,i-1,i)\right)
\end{aligned}$$

where $w(i,j,k) := \{x_i\}_2 \otimes x_j \wedge x_k$. The c_{ij}'s are linear combinations of the a_{ij}'s that come about from applying the Abel Identity. Note that $w(i,j,k) = -w(i,k,j)$. By convention, the preferred notation is $j < k$. The w's are linearly independent and in fact form a basis for a much larger space than $B_2 \bigotimes \Lambda^2\mathbb{C}^*$. With this notation, the Abel identity becomes

$$w(5,i,j) = \sum_{k=1}^{4} -w(k,i,j).$$

Written out in full, the $2,2$ component is

$$\begin{aligned}
\delta\delta_{2,2}f &= c_{12}e_{12} + c_{13}e_{13} + c_{14}e_{14} + c_{23}e_{23} + c_{24}e_{24} + c_{34}e_{34} \\
&= c_{12}\left(-w(1,2,3) + w(1,1,2) + w(2,1,2) + w(2,1,5)\right) \\
&\quad + c_{13}\left(-w(1,3,4) + w(1,2,3) + w(3,1,2) + w(3,1,5)\right) \\
&\quad + c_{14}\left(-w(1,4,5) + w(1,3,4) + w(4,1,2) + w(4,1,5)\right) \\
&\quad + c_{23}\left(-w(2,3,4) + w(2,2,3) + w(3,2,3) - w(3,1,2)\right) \\
&\quad + c_{24}\left(-w(2,4,5) + w(2,3,4) + w(4,2,3) - w(4,1,2)\right) \\
&\quad + c_{34}\left(-w(3,4,5) + w(3,3,4) + w(4,3,4) - w(4,2,3)\right)
\end{aligned}$$

Now consider $\delta_{3,1}f$. To eliminate the $\{x_5\}_3$ components, it makes sense to let $E_{ij} := w(i,i,j)$. Then

$$E_{5j} = w(5,5,j) = \sum_{i=1}^{4} -w(i,5,j) = \sum_{i=1}^{4} w(i,j,5).$$

Hence

$$\delta(\delta_{3,1})f = \sum_{i\neq j}^{5} d_{ij}E_{ij}.$$

The condition (7.37) that $\delta^2 = 0$ now can be written in this explicit w-basis as

$$0 = \delta\left(\delta_{2,2}f\right) + \delta\left(\delta_{3,1}f\right) \tag{7.45}$$

$$= \sum_{i<j}^{4} c_{ij}\left[-w(i,j,j+1) + w(i,j-1,j) + w(j,i,i+1) - w(j,i-1,i)\right] + \sum_{i\neq j}^{5} d_{ij}w(i,i,j). \tag{7.46}$$

By Equation (7.37) and the linear-independence of the w's, the coefficient for each $w(i,j,k)$ must be zero separately. This is now just a linear algebra problem. Moreover, some groups of terms are also independent. Only terms of the form $w(i,i,j)$ or $w(i,j,5)$ can cancel between the $2,2$ and the $3,1$ components. Everything else must sum to zero separately.

Because the coefficients of $w(1,23), w(1,3,4), w(3,1,2), w(4,1,2), w(2,3,4)$ and $w(4,2,3)$ must sum to zero separately,

$$c_{12} = c_{13} = c_{14} = c_{23} = c_{24} = c_{34}.$$

By matching coefficient on the terms left without any 5's present,

$$c_{12} = d_{21} = -d_{12} = d_{32} = -d_{23} = d_{43} = -d_{34}.$$

Furthermore,

$$w(2,1,5)+w(3,1,5)+w(4,1,5) = E_{51}-E_{15} \text{ and } -w(1,4,5)-w(2,4,5)-w(3,4,5) = -(E_{54}-E_{45}),$$

which further sets $a_{12} = d_{54} = -d_{45} = d_{15} = -d_{51}$. So to summarize,

$$c_{12} = b_{(i+1)i} = -b_{i(i+1)} \quad i = 1, \ldots, 5.$$

This means all the coefficients c_{ij} are equal to c_{12} and

$$d_{ij} \doteq c_{12} \begin{pmatrix} 0 & -1 & 0 & 0 & 1 \\ 1 & 0 & -1 & 0 & 0 \\ 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 & -1 \\ -1 & 0 & 0 & 1 & 0 \end{pmatrix}.$$

Because the A_2 function should be non-trivial, $c_{12} \neq 0$. Therefore, up to a multiplicative constant c_{ij}, there is a unique function f whose coproduct has these coefficients. These results match Equation 3.3 of [3] up to a minus sign on the $2,2$ component with $c_{12} = 5$.

7.4 A Computationally Efficient Basis for Finding Cluster Functions at Weight 4

The material in this section is joint work with Adam Scherlis.

Suppose C is a cluster algebra with $\mathcal{A}$-coordinates $\mathcal{A}$ and $\mathcal{X}$-coordinates $\mathcal{X}$. Let β_1, β_2 and β_3 be (multiplicative) bases for $B_1 = \mathcal{X}, B_2$ and B_3 respectively.

Any $\mathcal{X}$-coordinate can, of course, be written as a product of elements of β_1. It follows from the rule for $\mathcal{X}$-coordinate mutation that in any finite-type cluster algebra, $1+x_i$ can be written as a product of $\mathcal{X}$-coordinates for any $x_i \in \mathcal{X}$. Therefore define two "factorization functions"

$$P : \mathcal{X} \to \mathrm{sh}^1(\beta_1) \text{ by } P(x_i) = y_1^{p_i^1} \cdots y_k^{p_i^k} \text{ where } y_i \in \beta_1 \text{ and } p_i^j \in \mathbb{Z} \tag{7.47}$$

and

$$R : \mathcal{X} \to \mathrm{sh}^1(\beta_1) \text{ by } R(1+x_i) = y_1^{r_i^1} \cdots y_k^{r_i^k} \text{ where } y_i \in \beta_1 \text{ and } r_i^j \in \mathbb{Z}. \tag{7.48}$$

In terms of these, there is an isomorphism $\mathrm{id} \otimes R \wedge P : B_2 \otimes \Lambda^2 \mathbb{C}^* \to \mathrm{sh}^2(\beta_1) \otimes \Lambda^2 \mathrm{sh}^1(\beta_1)$. The latter space has a basis

$$\{w(x_i, x_j, x_k) := \{x_i\}_2 \otimes x_j \wedge x_k \ : \ \{x_i\}_2 \in \beta_2, x_j, x_k \in \beta_1, j < k\}.$$

To find the cluster functions satisfying $\delta^2 f_4 = 0$, the goal is to express δ^2 as a linear map $T : \Lambda^2 B_2 \oplus B_3 \otimes \mathbb{C}^* \to B_2 \otimes \Lambda^2 \mathcal{X}$ explicitly in coordinates. The cluster functions should correspond to elements of $\ker T$.

Let's look at $\Lambda^2 B_2$ first. By definition of the β's above,

$$\Lambda^2 B_2 = \mathrm{span}\left\{\{x_i\}_2 \wedge \{x_j\}_2 \ : \ i < j; \ x_i, x_j \in \beta_2\right\}.$$

By 2.9 of [3], with the corrected sign,

$$\delta(\{x\}_2 \wedge \{y\}_2) = \{y\}_2 \otimes (1+x) \wedge x - \{x\}_2 \otimes (1+y) \wedge y.$$

Therefore

$$\delta(\{x_i\}_2 \wedge \{x_j\}_2) \ = \ \{x_j\}_2 \otimes 1 + x_i \wedge x_i - \{x_i\}_2 \otimes 1 + x_j \wedge x_j$$

So

$$\begin{aligned}
v_{ij}^{2,2} &:= (\mathrm{id} \otimes R \wedge P)\delta(\{x_i\}_2 \wedge \{x_j\}_2) \\
&= \{x_j\}_2 \otimes R(1+x_i) \wedge P(x_i) - \{x_i\}_2 \otimes R(1+x_j) \wedge P(x_j) \\
&= \{x_j\}_2 \otimes \sum_{x_k \in \beta_1} r_i^k x_k \wedge \sum_{x_\ell \in \beta_1} p_i^\ell x_\ell - \{x_i\}_2 \otimes \sum_{x_m \in \beta_1} r_j^m x_m \wedge \sum_{x_n \in \beta_1} p_j^n x_n \\
&= \left(\sum_{x_k, x_\ell \in \beta_1} r_i^k p_i^\ell w(x_j, x_k, x_\ell) - \sum_{x_m, x_n \beta_1} r_j^m p_j^n w(x_i, x_m, x_n) \right) \in B_2 \otimes \Lambda^2 \mathcal{X}.
\end{aligned}$$

The situation is largely similar for $B_3 \otimes \mathbb{C}^*$. The space can be written as

$$B_3 \otimes \mathbb{C}^* = \mathrm{span}\left\{\{x_i\}_3 \wedge x_j \ : \ x_i \in \beta_3, x_j \in \beta_1\right\}.$$

The action of δ on this is

$$\delta(\{x\}_3 \wedge y) = \{x\}_2 \otimes x \wedge y,$$

so

$$\begin{aligned}
v_{ij}^{3,1} := (id \otimes R \wedge P)\delta(\{x_i\}_3 \wedge x_j) \ &= \ \{x_i\}_2 \otimes R(x_i) \wedge P(x_j) \\
&= \sum_{x_k \in \beta_2} c_i^k \{x_k\}_2 \otimes \sum_{x_\ell \in \beta_1} r_i^\ell a_\ell \wedge \sum_{x_m \in \beta_1} p_j^m a_m \\
&= \sum_{\substack{x_k \in \beta_2 \\ x_\ell, x_m \in \beta_1}} c_i^k r_i^\ell p_j^m w(x_k, x_\ell, x_m).
\end{aligned}$$

We can now explicitly write T by

$$T(\{x_i\}_2 \wedge \{x_j\}_2) = v_{i,j}^{2,2} \text{ and } T(\{x_i\}_3 \wedge x_j) = v_{i,j}^{3,1}$$

The kernel of T is then the space spanned by vectors of the form

$$\sum_{i,j} a^{ij} v_{i,j}^{2,2} + \sum_{k,\ell} b^{k\ell} v_{k,\ell}^{3,1} = 0.$$

Writing this out in terms of the basis of w's, finding the kernel of T is reduced to a linear algebra problem which can be solved by explicit row-reduction on a computer for any finite cluster algebra.

7.5 Dictionary of Cluster Polylogarithms

The last section was implemented in `Mathematica`. The code was run on several cluster algebras of Grassmannian type: $A_2, A_3, A_4, A_5, A_6, D_4$ and E_6. Since the outputs are very large in general, the Table below to summarizes some of the pertinent information. It is possible to extend this to non-Grassmannian cluster algebras. However, it is much easier to determine the maps P and R in the Grassmannian case. In the Grassmannian case, this can be done with the algorithm in Section 7.7. It is also possible to do so by tropicalizing the Laruant polynomials for the cluster variables. Explicit testing in a few cases suggests that there is no more information from looking at the direct product of two cluster algebras than from studying each algebra separately.

Algebra	A3	A4	A5	A6	D4	E6
$\lvert\mathcal{A}-\text{coordinates}\rvert$	9	14	20	27	16	42
$\lvert\mathcal{A}-\text{coordinates including frozen}\rvert$	15	21	28	36	22	49
$\lvert\mathcal{X}-\text{coordinates}\rvert$	15	35	70	126	52	385
$\lvert\text{Vertices}\rvert$	14	42	132	429	50	833
$\lvert\text{Edges}\rvert$	21	84	330	1287	100	2499
$\lvert\text{SA}_2\rvert$	9	56	300	1485	66	2856
$\lvert\text{SA}_3\rvert$	0	14	120	825	16	1547
$\lvert\text{SA}_4\rvert$	0	0	20	225	0	399
$\lvert\text{SA}_5\rvert$	0	0	0	27	0	42
$\dim B_1$	9	14	20	27	16	42
$\dim B_2$	10	20	35	56	27	132
$\dim B_3$	15	35	70	126	51	363
$\dim B_2 \otimes \Lambda^2\mathbb{C}^*$	360	1820	6650	19656	3240	113652
$\dim B_{2,2} \oplus B_{3,1}$	180	680	1995	4942	1167	23892
$\dim \ker T$	21	56	126	252	86	833

Table 4: Statistics for various cluster algebras. SA_k is the number of subalgebras of rank k. The results for E_6 are novel and had *not* previously been computed by Golden *et al.*

7.6 Counting $\mathcal{X}$-coordinates, A_2 and A_3 Functions

Table 4 indicates a few patterns in the sizes of the spaces.

Proposition 40. *On A_n,*

1. $\dim B_1 = \binom{n+2}{2} - 1$
2. $\dim B_2 = \binom{n+2}{3}$
3. $\dim B_3 = \binom{n+3}{4}$

Proposition 40 was proven by Adam Scherlis, assuming some reasonable properties of the Abel identity. Additionally, the following conjecture suggests itself.

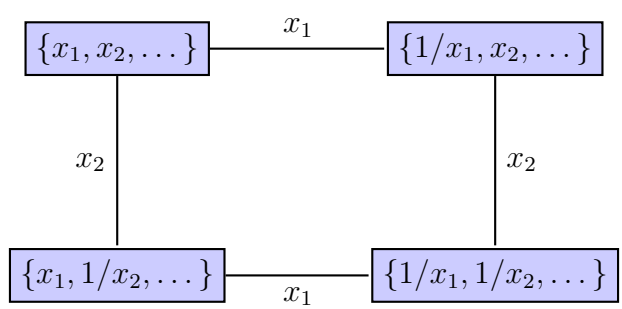
Conjecture 41. *For any cluster algebra,*

$$\dim \mathcal{L}_4^{\mathcal{S}} = \dim \{A_2 \textit{ functions}\} + \dim \{\mathrm{Li}_4 \textit{ functions}\} \tag{7.49}$$

In other words, A_2 is the only non-classical cluster polylogarithm at weight four.

This conjecture was verified by explicit computation for A_n through $n = 8$, as well as on D_4 and E_6.

There is also a relation between the geometry of the exchange graph and the number of cluster functions that appear. Consider a square inside the exchange graph. where the edges are labelled by the $\mathcal{X}$-coordinate that changes across that edge.



The $\mathcal{X}$-coordinates across the square are the same. This suggests an equivalence relation on edges. If e_1 and e_2 are edges in the exchange graph, then define

$$e_1 \sim e_2 \text{ if and only if } e_1 \text{ and } e_2 \text{ are on opposite sides of a square in the exchange graph.} \tag{7.50}$$

Colloquially, e_1 and e_2 are "square separated". The above diagram shows that $e_1 \sim e_2$ implies that the $\mathcal{X}$-coordinates associated with the edges are the same.

Conjecture 42. *There is a bijection between the number of $\mathcal{X}$-coordinates and the equivalence classes of edges.*

Given this conjecture, it was proven that

$$\mathcal{X}(A_n) = \binom{n+3}{4} \tag{7.51}$$

on any A_n-type cluster algebra. This conjecture has been verified on D_4 and E_6 by explicit computation.

This has further consequences in terms of A_2 and A_3 functions. If the exchange graph contains a pentaprism, i.e. a $A_1 \times A_2$ subalgebra, then by the above reasoning, the two pentagons have the same $\mathcal{X}$-coordinates in the same order. The A_2 function defined on these pentagons are therefore identical.

Conjecture 43. *The dimension of the space of A_2 functions on a cluster algebra C is no more than the number of "square separated" A_2 subalgebras. In addition, A_2 functions are linearly dependent if, and only if, all the pentagons are together inside some D_4 subalgebra.*

This has been tested through $n = 6$ on A_n, as well as on D_4 and E_6 by the author. On A_n, assuming that there are no Li_2 identities between "far away" parts of the polytope, Adam Scherlis has proven that

$$\dim\{A_2 \text{ functions}\} = \binom{n+3}{5}. \tag{7.52}$$

Similarly, $A_1 \times A_3$ subalgebras give rise to a single A_3 function and

$$\dim\{A_3 \text{ functions}\} = \binom{n+3}{6}. \tag{7.53}$$

7.7 An Algorithm to Factor Products of $\mathcal{X}$-coordinates

Suppose X is an alphabet, and $\mathrm{sh}\,(X)$ is the shuffle algebra of symbols over X. An expression $g(x_1, \ldots, x_k)$ can be a term in a symbol if, and only if,

$$g(x_1, \ldots, x_k) = \pm x_1^{d_1} \cdots x_k^{d_k} \tag{7.54}$$

for some coefficients $d_i \in \mathbb{Z}$. However, when X is, for example, the set of cluster $\mathcal{X}$-coordinates of some cluster, there are many non-trivial relations between the coordinates, so it is very difficult to tell when Equation (7.54) is satisfied. For instance, it is known by explicit calculation that $1 + x_i$ factors as a product of $\mathcal{X}$-coordinates on $A_2, \ldots, A_7$, D_4 and E_6. However, finding which coordinates it factors into is quite difficult. This section presents a novel numerical algorithm to check condition (7.54) and determine the coefficients n_i. A mathematical description is given first to show why it is correct, and then a practical algorithm.

Suppose $C := \mathbb{G}(k, n)$ is some Grassmannian cluster algebra. Let $X := \beta_1(C)$ be a "multiplicative basis" for $\mathcal{X}_C$, as described in Section 7.3 and put $N := |\beta_1(C)|$. Pick $p^1, \ldots, p^N \in \mathbb{G}^+(k, n)$, the positive Grassmannian, i.e. the subset of the Grassmannian where all the Plücker coordinates are positive real numbers. This can be done efficiently using an algorithm due to Postinikov [47].

Define *evaluation maps* $\varepsilon^j : \mathbb{C}(X) \to \mathbb{R}_{\geq 0}$ by $x_i \mapsto x_i(p^j)$. All $\mathcal{X}$-coordinates are rational functions in the Plücker coordinates with positive coefficients (mostly monomials, in fact). Therefore $\varepsilon^j(x_i)$ is in fact a positive, real number for all ε^j and all x_i.

Define a matrix A by

$$A := \sum_{i,j=1}^{N} \log \varepsilon^j(x_i)\, e_j \otimes e^i \tag{7.55}$$

where the $\{e_j\}$'s are the standard basis for $\mathbb{R}^N$ and the $\{e^i\}$'s are the dual basis. By the positivity of the $\mathcal{X}$-coordinates, the log's are well-defined, and thus A is a real-valued, $N \times N$ matrix. Let $g : \mathbb{C}[X] \to \mathbb{C}$ with $g(x_1, \ldots, x_N) = x_1^{d_1} \cdots x_N^{d_N}$. In other words, let it satisfy Equation (7.55), but with a positive sign. Then define $\boldsymbol{g} \in \mathbb{R}^N$ by

$$\boldsymbol{g} := \sum_{j=1}^{N} \log \varepsilon^j(g)\, e_j \tag{7.56}$$

and

$$\boldsymbol{d} := \sum_{i=1}^{N} d_i e_i. \tag{7.57}$$

Then

$$\begin{aligned}
\boldsymbol{g} &= \sum_{j=1}^{N} \log \varepsilon^j(g) e_j \\
&= \sum_{j=1}^{N} \log \left(\varepsilon^j \left(x_1^{d_1} \cdots x_N^{d_N} \right) \right) e_j \\
&= \sum_{j=1}^{N} \log \left(\varepsilon^j(x_1)^{d_1} \cdots \varepsilon^j(x_N)^{d_N} \right) e_j \\
&= \sum_{j=1}^{N} \sum_{i=1}^{N} d_i \log \varepsilon^j(x_i)\, e_j \\
&= \sum_{i,j=1}^{N} d_i A_i^j\, e_j \\
&= A\boldsymbol{d}.
\end{aligned}$$

Hence $\boldsymbol{d} = A^{-1}\boldsymbol{g}$. Therefore the coefficients d_i can be found simply from the numerical values of $\boldsymbol{g}$ and A.

In practice, one could implement this as follows. Suppose $g_1, \ldots, g_M \in \mathbb{C}[X]$ are given functions with notation as above.

1. Choose $\varepsilon^j : \mathbb{C}[X] \to \mathbb{C}$, $1 \leq j \leq k$ by Postnikov's algorithm.

2. Compute A^{-1} as above. For almost any choice of the ε^j's, A will be invertible. If it is not, find more ε^j's and try again.

3. For each g_ℓ, $1 \leq \ell \leq M$:

   (i) Compute $\varepsilon^j(g)$, $1 \leq j \leq N$. If these are all negative, put $g \mapsto -g$ and recompute. If they are all positive, continue. If they have mixed signs, then g cannot factor.

   (ii) Compute $\boldsymbol{g}$.

   (iii) Compute $\widetilde{\boldsymbol{d}} := A^{-1}\boldsymbol{g}$. If Equation (7.54) holds, then $\widetilde{\boldsymbol{d}} \in \mathbb{Z}^N$. In practice, however, there will be some numerical error, so the real condition should be

$$\sum_{i=1}^{N} \left| \widetilde{d}_i - \text{round}\left(\widetilde{d}_i\right) \right| < \varepsilon \tag{7.58}$$

   for some small $\varepsilon \geq 0$. If this condition is satisfied, then $g = \prod_i x_i^{d_i}$ where

$$d_i := \text{round}(\widetilde{d}_i).$$

This is computationally efficient: the ε^j need only be found once for each cluster algebra and, for each g_ℓ, the most difficult step is merely matrix multiplication.

8 Cluster Bases

The theme (7.1) described in the beginning of Section 7 — that remainder functions are combination of polylogarithms whose symbols are cluster $\mathcal{X}$-coordinates — greatly restricts the space of possible functions. Indeed, even without the cluster algebra structure, this idea is enough to uniquely determine $R_6^{(3)}$ [48]. Recent work by Drummond *et al.*, using this full strength of this theme, gives powerful constraints on 7-particle remainder functions at 2 and 3 loops [46].

The general strategy of these works is to consider the space of (symbols of) polylogarithms that could possible be in the symbol, then apply linear constraints based on the properties the amplitude must obey. Unfortunately, when this strategy is successful, it does not produce a simple result. The formula for $R_6^{(2)}$ from [46] is a 17-page linear combination of generalized polylogarithms. By using a judicious choice of variables and using the functional identities of polylogarithms to simplify the expression, Goncharov, Spradlin, Vergu and Volovich produced an expression for $R_6^{(2)}$ that fits on a single line [2]. It was later realized that this judicious choice of variables is intimately related to the cluster algebra structure of $\mathbb{G}(4,6)$, and this made it possible to write a similarly compact formula for $R_7^{(2)}$ [12, 3].

The moral of this story, if there is one, is that a through understanding of the mathematical properties of the space of polylogarithms with $\mathcal{X}$-coordinate symbols leads directly to an understanding of scattering amplitudes. Since this is a vector space, a natural question is: what is a basis for it? Amazingly, this is unknown. The numerous polylogarithm identities make it easy to find spanning sets, but very hard to find true bases. This section describes general strategies for finding bases for such spaces, and gives an explicit basis in the case where the cluster algebra in question is A_n.

8.1 The Basis Problem

Before presenting the solution, it is useful to make a clear and precise statement of the problem. Let C be a cluster algebra with a set of $\mathcal{X}$-coordinates $\mathcal{X}_C$. Let β_C be a multiplicative basis for $\mathcal{X}_C$. Then, using the notation from Section 6.3, the space of symbols in the $\mathcal{X}$-coordinate alphabet is the Shuffle algebra $\mathrm{sh}\left(\mathcal{X}_C\right) = \mathrm{sh}(\beta_C)$.

Definition 44. *Suppose X is a set of $\mathbb{C}$-valued functions. Then define*

$$G^n[X] = \{G(x_1; x_2, \ldots, x_n; x_{n+1}) \ : \ x_i \in X\}. \tag{8.1}$$

The Symbol is then a map $\mathcal{S} : G[X] \to \mathrm{sh}(\widetilde{X})$, where $\widetilde{X} = X \cup \{x_i - x_j \ : \ x_i, x_j \in X\}$, because the Symbol of a Goncharov polylogarithm, Equation (6.57), contains differences in the arguments as well as the arguments themselves. Moreover, not all of the symbols in $\mathrm{sh}(\widetilde{X})$ actually come from some function in $G[X]$, but only the *integrable* ones (See Equation 6.30.) There is actually an exact sequence

$$G^k[X] \xrightarrow{\mathcal{S}} \mathrm{sh}^k\left(\widetilde{X}\right) \xrightarrow{D} \mathrm{sh}^{k+1}\left(\widetilde{X}\right) \ \oplus \ \Lambda^2\widetilde{X} \otimes \mathrm{sh}^{k-2}\left(\widetilde{X}\right).$$

The integrable symbols are elements of $B(\widetilde{X}) = \ker D$.

With this notation, the **basis problem** is

$$\boxed{\begin{array}{l}\text{Given a set of } \mathcal{X}\text{-coordinates } \mathcal{X}_C, \\ \quad 1.\ \text{What is the set } X \text{ such that } \mathcal{S}\left(G^k[X]\right) = B(\mathcal{X}_C)? \\ \quad 2.\ \text{What is a basis for } G^k[X]?\end{array}} \tag{8.2}$$

8.2 The Spaces M, P, R and Q

A vague procedure for finding X for a given $\widetilde{X}$ is described by Duhr *et al.* in [42]. This section describes a more precise version for the case where $\widetilde{X} = \mathcal{X}_C$ for some cluster algebra C. Let β_C be the multiplicative basis for $\mathcal{X}_C$ as above.

Definition 45. *Define the* ***multiplicative span*** *of $\mathcal{X}_C$ as*

$$M_C := \{z = \pm x_1^{n_1} x_2^{n_2} \cdots x_k^{n_k} \ : \ x_i \in \beta \text{ and } n_i \in \mathbb{Z}\}. \tag{8.3}$$

This set has the virtue that any symbol in the alphabet M_C can be rewritten as a symbol in the alphabet β_C. Moreover, the algorithm described in Section 7.7 is exactly what is needed to check if a general rational function in the $\mathcal{X}$-coordinates of C is an element of M_C.

As a first attempt towards finding X, consider the case of classical polylogarithms. Recall that

$$\mathrm{Li}_k(z) = G(\underbrace{0, \ldots, 0}_{k-1}, 1; z) \tag{8.4}$$

and

$$\mathcal{S}(\mathrm{Li}_k(z)) = -(1-z) \otimes z \otimes \cdots \otimes z. \tag{8.5}$$

So, for this special case, a set of possible arguments of the Goncharov polylogarithm is $X = \{0, 1, z\}$ for some $z \in M_C$ such that $1 - z \in M_C$ as well. The last condition is actually a useful restriction on the element of M_C of interest. Define

$$R_C := \{z \in M_c \ : \ 1 - z \in M_c\}. \tag{8.6}$$

There is an action of S_3, the symmetric group on three letters, on R_C, generated by

$$z \mapsto \frac{1}{z} \text{ and } z \mapsto 1 - z. \tag{8.7}$$

R_C is closed under this action because $1-(1/z) = (z-1)/z \in M_C$ and $1-(1-z) = z \in M_c$. In general, R_C is infinite, and thus not very useful. However, for any A, D, or E-type cluster algebra, $1 + x \in M_C$ where x is an $\mathcal{X}$-coordinate. Therefore define the subset

$$\overline{R}_C := \{-x \ : \ x \in \mathcal{X}_C\} \subset R_C. \tag{8.8}$$

This is a much smaller set — often finite — and thus amenable to explicit computation. So if $z \in P_C$, then $X = \{0, 1, z\}$ is a set of possible arguments for Goncharov polylogarithms whose symbols are in $\mathcal{X}_C$.

Returning to the general case, one needs elements $\{z_i\}$ of M_C such that $\mathrm{Li}_k(z_i)$ make sense, as well generalized polylogarithms with arguments drawn from the z's. Therefore, one should have that $z_i - z_j$ can be written as a product of $\mathcal{X}$-coordinates, i.e. $z_i - z_j \in M_C$. The space which fulfills this property is

$$P_C := \left\{(z_i, z_j) \in R_C^2 \ : \ z_i - z_j \in M_c\right\}. \tag{8.9}$$

For $(z_1, z_2) \in P_C$, the symbol of anything in

$$G\left[\{0, 1, z_1, z_2\}\right] \tag{8.10}$$

can be written completely in terms of $\mathcal{X}$-coordinates. Again, it is useful to make a restriction to a finite case:

$$\overline{P}_C := \left\{(-x_i, -x_j) \in \overline{R}_C^2 \ : \ x_i - x_j \in M_c;\ x_i, x_j \in \mathcal{X}_C\right\}. \tag{8.11}$$

One can interpret this as a graph whose vertices are the $\mathcal{X}$-coordinates of a space, where edges are drawn whenever $x_i - x_j$ factors as a product of $\mathcal{X}$'s.

Of course, this alone is not good enough. Not all generalized polylogarithms can be written with 4 unique arguments; arbitrarily many are needed. How can larger candidates for X be assembled? Actually, the graph P_C contains all the information necessary. The following terminology from graph theory is useful.

Definition 46. *Suppose G is a graph. A subgraph H of G is called a* ***clique*** *if H is isomorphic to a complete graph, i.e. there is an edge connecting every vertex of H.*

Proposition 47. *Suppose $q = (q_1, \ldots, q_n)$ are the vertices of a clique of P_c. Put $X = \{0, 1\} \cup q$. Then $\mathcal{S}(G^k[X]) \subseteq \ker D^k(\mathcal{X}_C)$.*

Proof. Consider $G = G(z_0; z_1, \ldots, z_k; z_{k+1})$ where $z_i \in X$. The symbol of G is in $\mathrm{sh}^k(\mathcal{X}_C)$ whenever $z_i - z_j \in M_C$ for all $0 \leq i, j \leq k+1$. There are four cases:

1. If $z_i = 0$, then $z_i - z_j = z_j \in M_C$.
2. If $z_i = 1$, and $z_j = q_j$, then $z_i - z_j = 1 - q_j \in M_c$, since $q_j \in R_C$.
3. If $z_i = q_i$ and $z_j = q_j$, then $z_i - z_j = q_i - q_j \in M_C$ because q is a clique, and there is an edge between q_i and q_j in P_C.

□

It turns out that knowing the cliques of P_C is all that is necessary to find X. Later sections will show when the inclusion in the proposition becomes an equality. Define Q_C as the set of cliques q of P_C with maximal size, i.e. so that no other clique contains them. Likewise, defined $\overline{Q}_C$ for $\overline{P}_C$.

8.3 Brown's Theorem

Extensive work on bases for spaces of polylogarithms has been done by the mathematician Francis Brown. Reference [49] is a complete reference and [32, 36] are less rigorous and less complete summaries written for physicists.

Two results will be very useful to us. They are paraphrased below in notation used here. As usual, $\mathbb{P}^1$ denotes the complex projective line, otherwise known as the Riemann sphere.

Theorem 48 (Brown's Basis Theorem for One-Variable Polylogarithms)**.** *For each finite set*

$$\Sigma := (0, 1, \sigma_2, \sigma_3, \ldots, \sigma_n) \subset \mathbb{P}^1,$$

there are associated one-forms $\omega_i := d\log(z - \sigma_i)$ for $0 \leq i \leq n$. Then the space of integrable words in ω_i is

$$B\left(\mathbb{P}^1 \setminus \Sigma\right) = \mathrm{sh}(\{\omega_i\}_{i=0}^n) \tag{8.12}$$

where sh *is the free shuffle algebra over the letters $\{\omega_0, \ldots, \omega_n\}$.*

By Chen's Theorem, there is an isomorphism from $B\left(\mathbb{P}^1 \setminus \Sigma\right)$ to (homotopy invariant) polylogarithms given by

$$\omega_{i_1} \otimes \omega_{i_2} \otimes \cdots \otimes \omega_{i_k} \mapsto \int_{\gamma(z)} \omega_{i_1} \circ \omega_{i_2} \circ \cdots \circ \omega_{i_k} = G(\sigma_{i_1}, \sigma_{i_2}, \ldots, \sigma_{i_k}; z) \tag{8.13}$$

where the equality is from the definition of the Goncharov polylogarithms, Equation (6.49). Radford's Theorem says that a free shuffle algebra has a basis of Lyndon words, so

$$L(\mathbb{P}^1 \setminus \Sigma) = \{G(\sigma_{i_1}, \sigma_{i_2}, \ldots, \sigma_{i_k}; z) \ : \ \sigma_{i_1}\sigma_{i_2}\ldots\sigma_{i_k} \text{ is a Lyndon word}\} \tag{8.14}$$

is a basis for Goncharov polylogarithms with arguments drawn from Σ, considered as functions of a variable z.

But polylogarithms are not really functions of one variable, but of many variables. Considering them as many variable functions, the situation as more complicated. Define $\mathfrak{M}_{0,n+3}$ to be the space of configurations of $n+3$ points in $\mathbb{P}^1$, modulo the action of $\mathrm{PSL}_2(\mathbb{C})$, the group of Möbius transformations. Let $\Sigma \in \mathfrak{M}_{0,n+3}$. Then it can be represented by $\Sigma = (0, 1, \sigma_2, \ldots, \sigma_{n+1})$ with $\sigma_i \neq \sigma_j$ and $\sigma_i \neq 0, 1$. Cross-ratios are $\mathrm{PSL}_2(\mathbb{C})$ invariant functions on $\mathfrak{M}_{0,n}$ given by $[ij|k\ell] : \mathfrak{M}_{0,n} \to \mathbb{P}^1$ with

$$[ij|k\ell] := \frac{(\sigma_i - \sigma_k)(\sigma_j - \sigma_\ell)}{(\sigma_i - \sigma_\ell)(\sigma_j - \sigma_k)}, \quad 1 \leq i, j, k, l \leq n. \tag{8.15}$$

These are very similar to cross-ratios of Plücker coordinates. Indeed, $\mathfrak{M}_{0,n+3}$ is very similar to the space of $\mathcal{X}$-coordinates of the cluster algebra A_n, but the exact relation is unclear at present. Again, define $\omega_i := d\log(z - \sigma_i)$.

Theorem 49 (Brown's Basis Theorem for Multi-variable Polylogarithms)**.** *Using notation as above, let $B(\mathfrak{M}_{0,n+3})$ be the space of integrable words in $\{\omega_0, \ldots, \omega_{n+1}\}$. Then there is a canonical isomorphism*

$$B(\mathfrak{M}_{0,n+3}) \cong \bigotimes_{i=1}^{n} B\left(\mathbb{P}^1 \setminus \Sigma_i\right). \tag{8.16}$$

where $\Sigma_i = \{\sigma_2, \sigma_2, \ldots, \sigma_{n-i+1}\}$. Therefore, employing Chen's Theorem and Radford's Theorem, a basis for Goncharov polylogarithms with arguments drawn from $\mathfrak{M}_{0,n+3}$ is of the form

$$\bigcup_{k=1}^{n} \{G(\sigma_{i_1}, \sigma_{i_2}, \ldots, \sigma_{i_r}; \sigma_{n-i+2}) \ : \ \sigma_{i_1}\sigma_{i_2}\ldots\sigma_{i_k} \text{ is a Lyndon word over the set } \Sigma_i\}. \tag{8.17}$$

At first glance, it might look like this answers the "basis problem", (8.2). However, the are several crucial differences:

- The arguments of the polylogarithms should be cluster $\mathcal{X}$-coordinates. The precise relation between the σ_i's and $\mathcal{X}$-coordinates is unclear at present.
- The polylogarithms should have Symbols which can be expressed entirely in terms of $\mathcal{X}$-coordinates. In general, the Symbol of an element of an element of (8.17) in terms of $\sigma_i - \sigma_j$.

One needs a choice of Σ so that $\sigma_i - \sigma_j$ always factors in terms of $\mathcal{X}$-coordinates. In other words, one needs a clique. The last section gave an algorithmic method for finding them, but does not guarantee the existence of cliques of the right size, nor does it provide an analytic method for finding them.

8.4 The Hedgehog Theorem

It turns out that the cliques of A_n are intimately related to the A_n polytope. The proof is joint work with Adam Scherlis.

Definition 50. *Suppose A is an A_n cluster algebra, $n \geq 2$, and B is a subalgebra isomorphic to the A_{n-1} cluster algebra. Define the A-B* ***hedgehog***

$$\begin{aligned} q(A,B) &= \{\textit{the set of } \mathcal{X}\textit{-coordinate whose edges start in } B \textit{ and end in } B \setminus A\} \\ &= \{x \in \mathcal{X}_A \ : \ x \in \mathcal{X}_B \textit{ and } 1/x \notin \mathcal{X}_B\}\,. \end{aligned} \tag{8.18}$$

Similarly, the inverse Hedgehog is defined as $q^{-1}(A,B) = \left\{\frac{1}{x} \ : \ x \in q(A,B)\right\}$. In contrast to the usual convention, the $\mathcal{X}$-coordinates of a hedgehog are oriented; x and $1/x$ are not considered the same in this situation.

A-B Hedgehogs are well-defined because each vertex in B has n edges coming out of it. Exactly $n-1$ of these connect with vertices in B; the last edge connects to a vertex in $B \setminus A$. The hedgehog is exactly the union of all such edges that leave B. An example is shown in Figure 5.

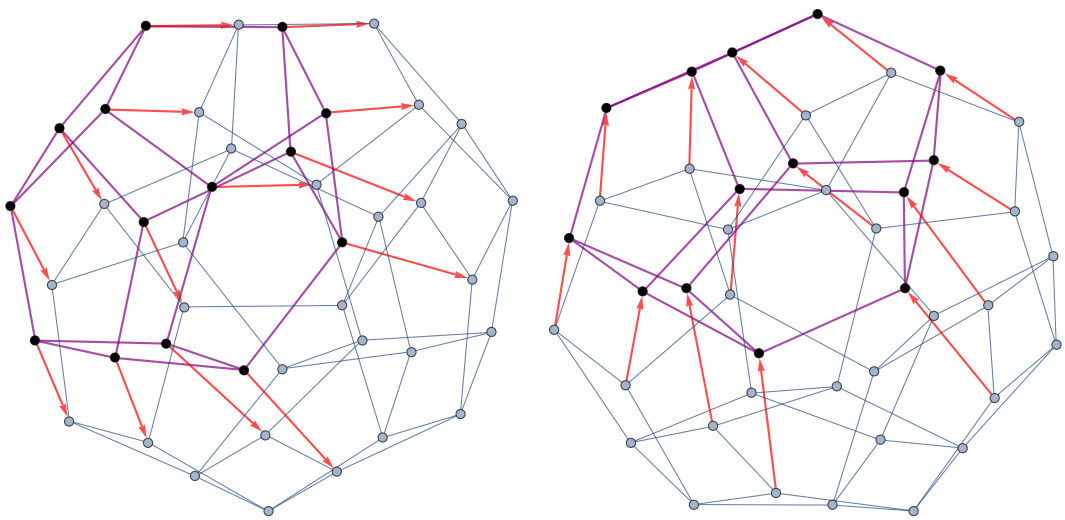


Figure 5: The hedgehog $q(A_4, A_3)$ (left) and $q^{-1}(A_4, A_3)$ (right). The A_3 subalgebra is highlighted in purple and the hedgehogs are red.

Theorem 51 (The Hedgehog Theorem)**.** *Suppose $q = q(A,B)$ is a hedgehog for A_n. Define*

$$\beta^k(q) := \bigcup_{i=0}^{n-1} \{G(-a_1, \ldots, -a_n; -q_{i+1}) \ : \ a_1 \cdots a_n \in \mathrm{Lyn}_k\{0, 1, q_1, \ldots, q_i\}\}\,. \tag{8.19}$$

Then

A. *The hedgehog q is a clique and $|q| = n$.*

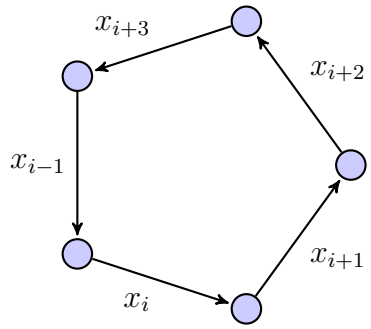
B. *Whenever $b \in \beta^k(q)$, $\mathcal{S}(b)$ can be written in terms of $\mathcal{X}$-coordinates, i.e. $\mathcal{S}(\beta^k(q)) \subset \mathrm{sh}^k(\mathcal{X}_A)$. Moreover,*

$$\mathcal{S}(\beta^k(q)) = \ker D^k(\mathcal{X}_A) \tag{8.20}$$

and thus $\beta^k(\mathcal{X}_A)$ is a basis for Goncharov polylogarithms of pure weight k whose Symbols can be written entirely in terms of $\mathcal{X}$-coordinates.

Proof. Proving A is actually the heart of the proof. The rest follows from the previous sections and a counting argument.

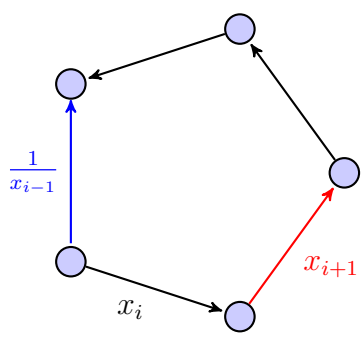
First consider the case $n = 2$. The general case can be reduced to the $n = 2$ case, so it is worth doing in detail. Let the $\mathcal{X}$-coordinates of $A \cong A_2$ be $\{x_{i-1}, x_i, x_{i+1}, x_{i+2}, x_{i+3}\}$ and let $B \cong A_1$ have $\mathcal{X}$-coordinate $\{x_i\}$.



For A_2, the $\mathcal{X}$-coordinates are actually recursive with the same pattern as the $\mathcal{A}$-coordinates:

$$x_{i+1} = \frac{1 + x_i}{x_{i-1}}. \tag{8.21}$$

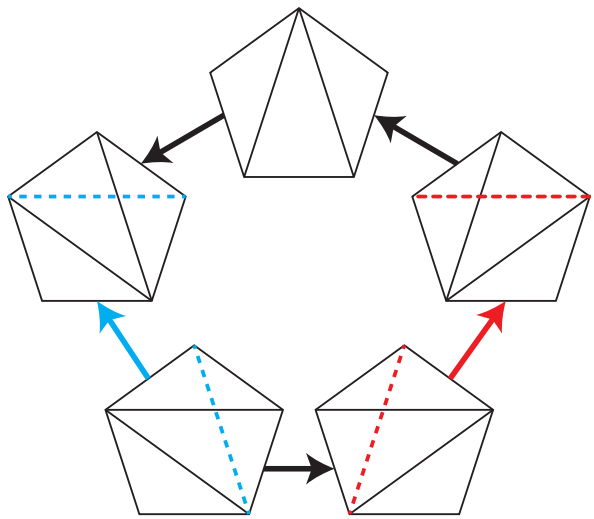
So here $q(A, B) = q(\{x_{i-1}, x_i, x_{i+1}, x_{i+2}, x_{i+3}\}, \{x_i\}) = \left\{x_{i+1}, \frac{1}{x_{i-1}}\right\}$. Pictorially, the hedgehog is the red and blue edges.



This is a clique because

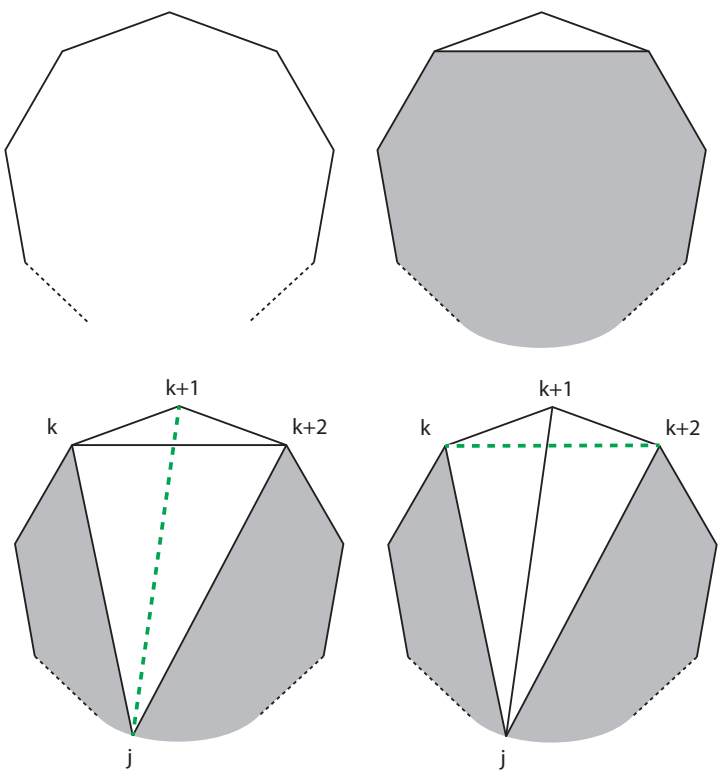
$$x_{i+1} - \frac{1}{x_{i-1}} = \frac{1 + x_i}{x_{i-1}} - \frac{1}{x_{i-1}} = \frac{x_i}{x_{i+1}}. \tag{8.22}$$

Recasting this in terms of polygon triangulation is very useful. In general, the A_n cluster algebra can be modelled by triangulations of a regular $n+3$-gon. Briefly, each triangulation corresponds to a vertex of the exchange graph, with edges drawn between triangulation that differ by exactly one chord. See [15] for more details.



This is the same picture as above in terms of triangulations. The red dashed lines are the chords that change as the red edge is traversed, and similarly for the blue. What we have shown then is that the difference between the $\mathcal{X}$-coordinates for the red and blue edges can be written in terms of products of $\mathcal{X}$-coordinates.

Now for the general case. Consider a regular $n+3$-gon. Choose three adjacent vertices $k, k+1, k+2$ and draw the chord between $k \leftrightarrow k+2$. Any sub-triangulation of the entire polygon that includes that chord is an A_{n-1} subalgebra. Said differently, if you chop off a triangle on the outer edge, then you get a $(n-1)+3$-gon, the triangulations of which correspond to A_{n-1}.

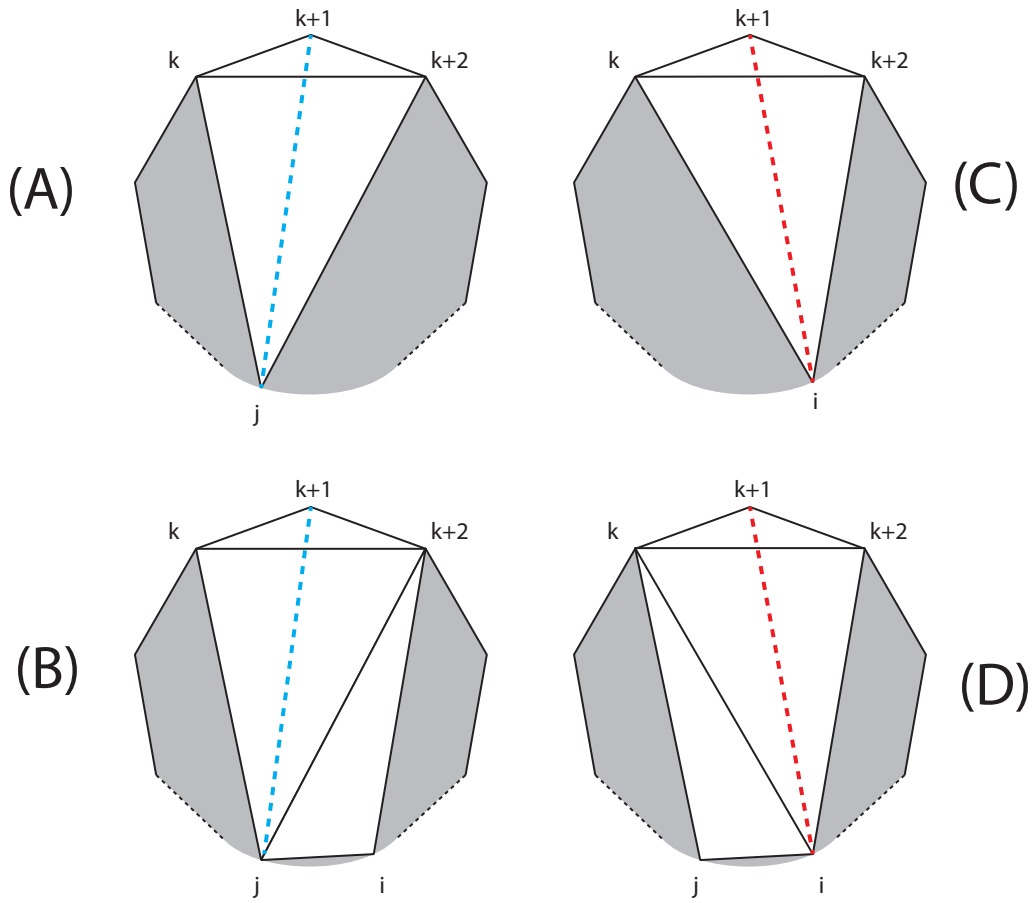


For any sub-triangulation of the $(n-1)+3$-gon, there must be some triangle

$$j \leftrightarrow k \leftrightarrow (k+2) \leftrightarrow j$$

and thus a quadrilateral with vertices $j, k, k+1, k+2$. Switching the interior chord $k \leftrightarrow k+2$ of this quadrilateral to $k+1 \leftrightarrow j$ corresponds to one of the edges in the hedgehog. This is indicated with a green dashed line in the diagram above. Note that this is not dependent on the sub-triangulation of the grey region. Therefore, any sub-triangulation of the grey region gives the same $\mathcal{X}$-coordinate in the hedgehog. Since there are exactly $(n+3)-3 = n$ choices of j, this means there are precisely n distinct $\mathcal{X}$-coordinates of $q(A, B)$.

Take q_i, q_j to be any two arbitrary elements of the hedgehog. Then q_i corresponds to flipping $k \leftrightarrow k+2$ to $i \leftrightarrow k+1$ and q_j corresponds to flipping $k \leftrightarrow k+2$ to $j \leftrightarrow k+1$. These are indicated in the diagram below in (C) and (A) respectively.



Any other sub-triangulation of the gray region preserves the q_i, so in particular one can choose the sub-triangulation with the pentagon $\{k, k+1, k+2, i, j\}$, shown in (B) of the diagram. Similarly, one can go from (C) to (D) and q_j will still be accessible by the red chord flip. But now notice that this is exactly the situation from the A_2 case! There is an embedded pentagon with exactly the same triangulations that appeared above. Therefore $q_i - q_j$ factors as a product of $\mathcal{X}$-coordinates, i.e $q_i - q_j \in M_{A_n}$. Since i and j were arbitrary, q is a hedgehog. This proves part 1.

For part 2, consider $\beta^k(q)$. Each function $g \in \beta^k(q)$ has arguments drawn from $X = \{0, 1, q_1, \ldots, q_n\}$. Therefore $\mathcal{S}(g)$ has Symbols in the alphabet $\widetilde{X} = X \cup \{1 + q_i, q_i - q_j\}$. But q_i is an $\mathcal{X}$-coordinate of a simply-laced Dynkin diagram, so $1 + q_i \in M_{A_n}$. And q is a clique, so $q_i - q_j \in M_{A_n}$. Therefore $\mathcal{S}(g) \in \mathrm{sh}^k(\mathcal{X}(A_n))$. Hence $\mathcal{S}(\beta^k(q)) \subset \mathrm{sh}^k(\mathcal{X}(A_n))$. In fact, $\mathcal{S}(\beta^k(q)) \subset \ker D^k(\mathcal{X}(A_n))$, since the Symbol of a function is always integrable.

Lastly, note that $\beta^k(q)$ is of the same form as Equation (8.17) at weight k. Specifically, they both have size

$$\sum_{i=0}^{n-1} N_k(i+2) \tag{8.23}$$

where $N_k(i+2)$ is the number of Lyndon words of length k in an alphabet with $i+2$ letters, defined in Equation (6.26). Therefore, they are bases for spaces of the same size. But Equation (8.16) is a basis for all Goncharov polylogarithms of weight k, so $\beta^k(q)$ must be as well. Taking Symbols, this implies $\mathcal{S}\left(\beta^k(q)\right) = \ker D^k(\mathcal{X}(A_n))$. □

There are a few useful corollaries.

Corollary 52. *If q and $\widetilde{q}$ are two A_n -A_{n-1} hedgehogs, then* $\operatorname{span} \beta^k(q) = \operatorname{span} \beta^k(\widetilde{q})$. *Thus every hedgehog generates the entire space of polylogarithms over an A_n.*

Corollary 53. *Suppose x_i and x_j are both $\mathcal{X}$-coordinates of A_n. Then $x_i - x_j \in M_{A_n}$ if, and only if, they are both in a A_n - A_{n-1} hedgehog.*

It is also useful to note that finding hedgehogs is computationally simple. Once the exchange graph is known, it is straightforward to find the A_{n-1} subalgebras, and thence the hedgehogs.

This is expected to generalize to other cluster algebras besides A_n. Slight modifications of this approach have been shown to work on D_4 at lower weights by explicit computation. The hope is that the space $\mathfrak{M}_{0,n+3}$ could instead be replaced by the space of configurations of points on $\mathbb{P}^3$ instead of $\mathbb{P}^1$, which would permit a similar basis to be written down for all $\mathbb{G}(4, n)$ cluster algebras.

References

[1] Vittorio Del Duca, Claude Duhr, and Vladimir A Smirnov. The one-loop one-mass hexagon integral in D= 6 dimensions. *Journal of High Energy Physics*, 2011(7):1–11, 2011.

[2] Alexander B. Goncharov, Marcus Spradlin, C. Vergu, and Anastasia Volovich. Classical Polylogarithms for Amplitudes and Wilson Loops. *Phys.Rev.Lett.105:151605,2010*, October 2010. doi:10.1103/PhysRevLett.105.151605. URL `http://arxiv.org/pdf/1006.5703v2`.

[3] John Golden, Miguel F Paulos, Marcus Spradlin, and Anastasia Volovich. Cluster Polylogarithms for Scattering Amplitudes. *arXiv preprint arXiv:1401.6446*, 2014.

[4] Michael Francis Atiyah. Geometry of Yang-Mills Fields. 1979.

[5] David Bleecker. *Gauge Theory and Variational Principles*. Addison-Wesley Publishing Company, 1981.

[6] Gregory L. Naber. *Topology, Geometry, and Gauge Fields: Interactions*, volume 141 of *Applied Mathematical Science*. Springer.

[7] Shoshichi Kobayashi and Katsumi Nomizu. *Foundations of Differential Geometry*, volume 1. New York, 1963.

[8] Mark Srednicki. *Quantum Field Theory*. Cambridge University Press, 2007.

[9] Anthony Zee. *Quantum Field Theory in a Nutshell*. Princeton university press, 2010.

[10] Edward Witten. Perturbative gauge theory as a string theory in twistor space. *Commun. Math. Phys.*, 252:189–258, 2004. doi:10.1007/s00220-004-1187-3.

[11] Henriette Elvang and Yu-tin Huang. Scattering amplitudes. *arXiv preprint arXiv:1308.1697*, 2013.

[12] John Golden, Alexander B. Goncharov, Marcus Spradlin, Cristian Vergu, and Anastasia Volovich. Motivic Amplitudes and Cluster Coordinates. August 2013. URL `http://arxiv.org/pdf/1305.1617v2`.

[13] Roger Penrose. Twistor algebra. *Journal of Mathematical physics*, 8(2):345–366, 1967.

[14] Andrew Hodges. Eliminating spurious poles from gauge-theoretic amplitudes. *Journal of High Energy Physics*, 2013(5):1–23, 2013.

[15] M. Gekhtman, M. Shapiro, and A. Vainshtein. *Cluster Algebras and Poisson Geometry*. Mathematical surveys and monographs. American Mathematical Soc., 2010. ISBN 9780821875483. URL `http://books.google.com/books?id=g0HfXaND8zYC`.

[16] John Lee. *Introduction to Smooth Manifolds*, volume 218. Springer, 2012.

[17] Stephen J Parke and TR Taylor. Amplitude for n-gluon scattering. *Physical Review Letters*, 56(23):2459, 1986.

[18] Zvi Bern, Lance Dixon, David C Dunbar, and David A Kosower. One-loop n-point gauge theory amplitudes, unitarity and collinear limits. *Nuclear Physics B*, 425(1):217–260, 1994.

[19] C Anastasiou, L Dixon, Z Bern, and DA Kosower. Planar amplitudes in maximally supersymmetric Yang-Mills theory. *Physical review letters*, 91(25):251602, 2003.

[20] Zvi Bern, Lance J Dixon, and Vladimir A Smirnov. Iteration of planar amplitudes in maximally supersymmetric Yang-Mills theory at three loops and beyond. *Physical Review D*, 72(8): 085001, 2005.

[21] Sergey Fomin and Andrei Zelevinsky. Cluster algebras. I. Foundations. *J. Amer. Math. Soc.*, 15(2):497–529 (electronic), 2002. ISSN 0894-0347.

[22] Andrei Zelevinsky. What is a cluster algebra. *Notices of the AMS*, 54(11):1494–1495, 2007.

[23] L. K. Williams. Cluster algebras: an introduction. *ArXiv e-prints*, December 2012.

[24] Michael Gekhtman, Michael Shapiro, and Alek Vainshtein. *Cluster algebras and Poisson geometry*. Number 167. American Mathematical Soc., 2010.

[25] P. DiFrancesco, R. Kedem, and R. Soto Garrido. Physics & Combinations of the Octahedron Equation: From Cluster Algebras to Arctic Curves. In *Michigan State University Mathematics Department Seminar*, April 2014.

[26] Kyungyong Lee and Ralf Schiffler. Positivity for Cluster Algebras. *Annals of Mathematics*, 2013.

[27] Sergey Fomin and Andrei Zelevinsky. Cluster algebras. II. Finite type classification. *Invent. Math.*, 154(1):63–121, 2003. ISSN 0020-9910.

[28] Sergey Fomin. Cluster algebras: basic notions and key examples (Slides). In *Spring School "Tropical Geometry and Cluster Algebras"*, April 2012. URL `http://www.msri.org/workshops/595/schedules/15561/documents/1438/assets/16934`.

[29] Sergey Fomin and Nathan Reading. Root systems and generalized associahedra. *arXiv preprint math/0505518*, 2005.

[30] V. V. Fock and A. B. Goncharov. Cluster ensembles, quantization and the dilogarithm. August 2009. URL `http://arxiv.org/pdf/math/0311245v7`.

[31] Joshua S. Scott. Grassmannians and Cluster Algebras. November 2003. URL `http://arxiv.org/pdf/math/0311148v1`.

[32] Francis Brown. Iterated integrals in quantum field theory. page 188. Cambridge University Press, 2013. URL `http://www.ihes.fr/~brown/ColombiaNotes7.pdf`.

[33] A. B. Goncharov. A simple construction of Grassmannian polylogarithms. March 2013. URL `http://arxiv.org/pdf/0908.2238v4`.

[34] Kuo-tsai Chen. Algebras of iterated path integrals and fundamental groups. *Transactions of the American Mathematical Society*, 156:359–379, 1971.

[35] Richard Hain. Iterated integrals and algebraic cycles: examples and prospects. *Contemporary trends in algebraic geometry and algebraic topology (Tianjin, 2000)*, 5:55–118, 2002. URL `http://arxiv.org/abs/math/0109204`.

[36] Christian Bogner and Francis Brown. Symbolic integration and multiple polylogarithms. Loops and Legs in Quantum Field Theory - 11th DESY Workshop on Elementary Particle Physics, April 2012.

[37] M.E. Sweedler. *Hopf Algebras*. Mathematics lecture note series. Benjamin, 1969. ISBN 9780805392555.

[38] David E Radford. A natural ring basis for the shuffle algebra and an application to group schemes. *Journal of Algebra*, 58(2):432–454, 1979.

[39] Amy Glen. Combinatorics of Lyndon words. Mini-Conference "Combinatorics, representations, and structure of Lie type", February 2012. URL `https://amyglen.files.wordpress.com/2012/03/melbourne_talk_feb2012.pdf`.

[40] Jean-Pierre Duval. Génération d'une section des classes de conjugaison et arbre des mots de Lyndon de longueur bornée. *Theoretical Computer Science*, 60(3):255–283, 1988. ISSN 0304-3975. doi:10.1016/0304-3975(88)90113-2. URL `http://www.sciencedirect.com/science/article/pii/0304397588901132`.

[41] J. Berstel and M. Pocchiola. Average Cost of Duval's Algorithm for Generating Lyndon Words. *Theor. Comput. Sci.*, 132(1-2):415–425, September 1994. ISSN 0304-3975. doi:10.1016/0304-3975(94)00013-1. URL `http://dx.doi.org/10.1016/0304-3975(94)00013-1`.

[42] Claude Duhr, Herbert Gangl, and John R. Rhodes. From polygons and symbols to polylogarithmic functions. October 2011. URL `http://arxiv.org/abs/1110.0458v1;http://arxiv.org/pdf/1110.0458v1`.

[43] Claude Duhr. Hopf algebras, coproducts and symbols: an application to Higgs boson amplitudes. March 2012. URL `http://arxiv.org/abs/1203.0454v1;http://arxiv.org/pdf/1203.0454v1`.

[44] Christian Vergu. Polylogarithms and physical applications. Notes for the Summer School "Polylogarithms as a Bridge between Number Theory and Particle Physics", July 2013. URL `http://www.maths.dur.ac.uk/lms/2013/PNTPP13/talks/0190vergu.pdf`.

[45] Alexander B Goncharov. Multiple polylogarithms and mixed Tate motives. *arXiv preprint math/0103059*, 2001.

[46] Lance J. Dixon, James M. Drummond, Matt von Hippel, and Jeffrey Pennington. Hexagon functions and the three-loop remainder function. December 2013. URL `http://arxiv.org/abs/1308.2276v2;http://arxiv.org/pdf/1308.2276v2`.

[47] A. Postnikov. Positive Grassmannian. Technical report. URL `http://www.thales.math.uqam.ca/~ahmorales/18.318lecs/lectures.pdf`.

[48] James M. Drummond, Georgios Papathanasiou, and Marcus Spradlin. A Symbol of Uniqueness: The Cluster Bootstrap for the 3-Loop MHV Heptagon. February 2015. URL `http://arxiv.org/abs/1412.3763v2;http://arxiv.org/pdf/1412.3763v2`.

[49] Francis Brown. Multiple zeta values and periods of moduli spaces $\mathfrak{M}_{0,n}$. *arXiv preprint math/0606419*, 2006.