

PREPARED FOR SUBMISSION TO JHEP

# Cluster Algebras and the Subalgebra Constructibility of the Seven-Particle Remainder Function

---

**John Golden<sup>1,2</sup> and Andrew J. McLeod<sup>2,3,4</sup>**

<sup>1</sup> *Leinweber Center for Theoretical Physics and Randall Laboratory of Physics, Department of Physics, University of Michigan Ann Arbor, MI 48109, USA*

<sup>2</sup> *Kavli Institute for Theoretical Physics, UC Santa Barbara, Santa Barbara, CA 93106, USA*

<sup>3</sup> *SLAC National Accelerator Laboratory, Stanford University, Stanford, CA 94309, USA*

<sup>4</sup> *Niels Bohr International Academy, Blegdamsvej 17, 2100 Copenhagen, Denmark*

*E-mail:* [amcleod@nbi.ku.dk](mailto:amcleod@nbi.ku.dk), [johngolden@gmail.com](mailto:johngolden@gmail.com)

**ABSTRACT:** We review various aspects of cluster algebras and the ways in which they appear in the study of loop-level amplitudes in planar  $\mathcal{N} = 4$  supersymmetric Yang-Mills theory. In particular, we highlight the different forms of cluster-algebraic structure that appear in this theory’s two-loop MHV amplitudes—considered as functions, symbols, and at the level of their Lie cobracket—and recount how the ‘nonclassical’ part of these amplitudes can be decomposed into specific functions evaluated on the  $A_2$  or  $A_3$  subalgebras of  $\text{Gr}(4, n)$ . We then extend this line of inquiry by searching for other subalgebras over which these amplitudes can be decomposed. We focus on the case of seven-particle kinematics, where we show that the nonclassical part of the two-loop MHV amplitude is also constructible out of functions evaluated on the  $D_5$  and  $A_5$  subalgebras of  $\text{Gr}(4, 7)$ , and that these decompositions are themselves decomposable in terms of the same  $A_4$  function. These nested decompositions take an especially canonical form, which is dictated in each case by constraints arising from the automorphism group of the parent algebra.

---

## Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
<b>2</b>	<b>A Brief Introduction to Cluster Algebras</b>	<b>5</b>
2.1	Clusters, Mutations, and Cluster $\mathcal{A}$ -coordinates	7
2.2	Cluster $\mathcal{X}$ -coordinates	9
2.3	Subalgebras and Cluster Polytopes	12
2.4	Grassmannian Cluster Algebras and Planar $\mathcal{N} = 4$ sYM Theory	15
2.5	Finite Cluster Algebras	20
2.6	Cluster Automorphisms	23
<b>3</b>	<b>Cluster Polylogarithms and MHV Amplitudes</b>	<b>27</b>
3.1	The Symbol and Cobracket	28
3.2	Cluster-Algebraic Structure at Two Loops	32
3.3	Subalgebra Structure and Cluster Polylogarithms	36
<b>4</b>	<b>Nonclassical Cluster Polylogarithms</b>	<b>39</b>
4.1	$A_2$ Functions as a Basis	39
4.2	Constructing all Cluster Polylogarithms in $\text{Gr}(4, 7)$	42
4.3	Nested Cluster Polylogarithms	44
<b>5</b>	<b>Subalgebra Constructibility and <math>R_7^{(2)}</math></b>	<b>46</b>
5.1	The $D_5$ Constructibility of $R_7^{(2)}$	47
5.2	The $A_5$ Constructibility of $R_7^{(2)}$	50
5.3	The Many Facets of $R_7^{(2)}$	52
<b>6</b>	<b>Conclusion</b>	<b>54</b>
<b>A</b>	<b>Counting Subalgebras of Finite Cluster Algebras</b>	<b>56</b>
<b>B</b>	<b>Cobracket Spaces in Finite Cluster Algebras</b>	<b>59</b>

---

# 1 Introduction

Multi-loop scattering amplitudes in the planar limit of  $\mathcal{N} = 4$  supersymmetric Yang-Mills (sYM) theory exhibit a great deal of intriguing mathematical structure. Much of this structure, at least at low loops and particle multiplicity, seems to be intimately tied to the cluster algebras associated with the Grassmannian  $\text{Gr}(4, n)$  [1–3]. This is especially true in the maximally-helicity-violating (MHV) sector, where amplitudes have been computed at high loop orders in six- and seven-particle kinematics [4–8], and algorithms exist for calculating two loop amplitudes for any number of particles [9, 10]. Remarkably, each of the branch cuts in these amplitudes ends at the vanishing locus of some cluster coordinate on  $\text{Gr}(4, n)$  [2, 11–13], and—even more strikingly—their iterated discontinuities vanishes unless sequentially taken in coordinates that appear together in a cluster of  $\text{Gr}(4, n)$  [14]. All six- and seven-particle next-to-MHV (NMHV) amplitudes that have currently been computed in this theory share these remarkable properties [6–8, 15–19], as do certain classes of Feynman integrals [14, 20–22], some of which have been computed to all loop orders [23]. While this collection of amplitudes and integrals represents the simplest this theory has to offer, it remains suggestive that cluster algebras combinatorially realize these salient aspects of their analytic structure, thereby encoding locality in a non-obvious way.

The fact that cluster algebras appear in this context is not totally surprising, given that the plabic graphs that describe the integrands of this theory to all loop orders are themselves dual to cluster algebras. In particular, the boundaries of the positive Grassmannian, where it is known that these integrands can develop physical singularities, all lie on the vanishing loci of cluster coordinates on  $\text{Gr}(4, n)$  [1]. Despite this, it’s far from obvious that the location of *all* physical singularities will be picked out by cluster coordinates in this way—and indeed, even at one loop,  $\text{N}^2\text{MHV}$  amplitudes have singularities at points involving square roots when expressed in terms of cluster coordinates [24]. This obfuscates the general connection between cluster algebras and the amplitudes of this theory, as does the eventual appearance of functions beyond polylogarithms [25–31]. Both complications point to the need for more general objects than cluster algebras to describe the analytic structure of this theory at higher loops and particle multiplicities.

There is reason, however, to be optimistic that an analogously simple characterization of this (more complicated) analytic structure might be found. This optimism stems from the observation that the infinite class of amplitudes we currently have access to—the two-loop MHV amplitudes—have properties beyond branch cuts that seem to be indelibly tied to cluster algebras. In particular, the ‘nonclassical’ part of each of these amplitudes (that is, the part that cannot be expressed in terms of classical

polylogarithms) is uniquely determined by a small set of physical and cluster-algebraic properties [13]. Seemingly unrelated, a pair of functions can be associated with the simplest cluster algebras, related to the Dynkin diagrams  $A_2$  and  $A_3$ , in terms of which these nonclassical components can be decomposed into a sum over the  $A_2$  or  $A_3$  subalgebras of  $\text{Gr}(4, n)$  [12]. Furthermore, the remaining ‘classical’ part of these amplitudes can always be written as products of classical polylogarithms involving only negative cluster coordinates as arguments [10]. It can be hoped that the pervasiveness of such cluster-algebraic structure points to the existence of a deeper and more general combinatorial structure that extends to all particle multiplicities and helicity configurations. If so, better understanding the many ways in which cluster algebras appear in these amplitudes can help us identify the features this structure must have.

The motivation for this work is to review the connections already mentioned between cluster algebras, as well as to further mine the two-loop MHV amplitudes for additional cluster-algebraic structure. The review portion of this paper, comprising most of sections 2 and 3, aims to be a self-contained pedagogical introduction to cluster algebras and their appearance in loop-level amplitudes in  $\mathcal{N} = 4$  sYM. These sections in particular focus on the subalgebra structure of cluster algebras, as these subalgebras will be used in the second half of the paper to build progressively more complicated polylogarithmic functions. A central ingredient in this story is the set of automorphisms respected by a given cluster algebra, as these automorphisms constrain the space of polylogarithms that can be defined on them.

We will in particular be interested in the space of ‘cluster polylogarithms’—polylogarithms defined on (both finite and infinite) cluster algebras that exhibit interesting forms of cluster-algebraic structure at both the level of their symbol and their Lie co-bracket. Such functions naturally arise in the study of the two-loop MHV amplitudes in planar  $\mathcal{N} = 4$  sYM theory [12], but also constitute an interesting class of functions in their own right. In the case of finite cluster algebras these functions can be studied systematically, and in the second half of this paper we describe an efficient method for constructing all nonclassical cluster polylogarithms on a given finite cluster algebra. This procedure is based on what we call the ‘cluster subalgebra constructibility’ of these functions, and in particular on the conjecture that all nonclassical cluster polylogarithms (on both finite and infinite cluster algebras) can be decomposed into sum over their  $A_2$  subalgebras [12]. We illustrate this method by constructing all (nonclassical) cluster polylogarithms on  $\text{Gr}(4, 7)$  and its subalgebras (this space was also explored in [32], using a slightly different approach).

While the (nonclassical part of the) two-loop MHV amplitudes in planar  $\mathcal{N} = 4$  sYM [12] are already known to be  $A_2$ - and  $A_3$ -constructible, it is interesting to ask whether they can be constructed in terms of cluster polylogarithms on larger algebras.

In particular, using the construction described above, we can ask this question of the two-loop seven-point MHV amplitude—itsself a cluster polylogarithm associated with the Grassmannian  $\text{Gr}(4, 7)$ . We show that such decompositions not only exist over the  $D_5$ ,  $A_5$ , and  $A_4$  subalgebras of  $\text{Gr}(4, 7)$ , but that these decompositions take an especially canonical form, uniquely dictated (up to an overall scale) by the interplay of the subalgebra and automorphism structures of  $\text{Gr}(4, 7)$ . Specifically, we will find that the seven-particle remainder function can be decomposed as

$$\begin{aligned} R_7^{(2)} &= \frac{1}{20} \sum_{D_5 \subset \text{Gr}(4, 7)} \sum_{A_4 \subset D_5} f_{A_4}^{+-}(x_1 \rightarrow x_2 \rightarrow x_3 \rightarrow x_4) + \dots \\ &= -\frac{1}{20} \sum_{A_5 \subset \text{Gr}(4, 7)} \sum_{A_4 \subset A_5} f_{A_4}^{+-}(x_1 \rightarrow x_2 \rightarrow x_3 \rightarrow x_4) + \dots, \end{aligned}$$

in terms of a single function  $f_{A_4}^{+-}$  defined in section 5, where the trailing dots indicate a contribution consisting of purely classical polylogarithms.

In a forthcoming companion paper, we also use this technology to construct the full eight-point remainder function in terms of cluster polylogarithms. That is, we describe methods for searching for decompositions of the nonclassical part of  $R_n^{(2)}$  even when  $\text{Gr}(4, n)$  is infinite, and then—using such a decomposition—systematically construct the classical and beyond-the-symbol components of  $R_8^{(2)}$  using the techniques described in [10]. We also show that there exists a ‘generalized BDS-like ansatz’ that preserves all Steinmann and cluster adjacency relations in eight-particle kinematics (and for all multiplicities that are a multiple of 4), despite the nonexistence of the standard BDS-like ansatz for these numbers of particles (this normalization is also discussed in section 3.2 of the present work).

This paper is organized as follows. In section 2 we provide a self-contained introduction to cluster algebras and why they appear in  $\mathcal{N} = 4$  sYM theory. While this section will largely constitute review for those familiar with recent developments at the intersection of these topics, the discussion of automorphisms in section 2.6 will likely be new to even those familiar with the physics literature. In section 3 we discuss some of the tools relevant for working with polylogarithms, particularly their associated coaction and Lie cobracket. We then recap the ways in which the coproduct and cobracket of the two-loop MHV amplitudes in planar  $\mathcal{N} = 4$  sYM theory have been found to exhibit curious cluster-algebraic structure. In section 4 we turn to the more general space of cluster polylogarithms, and initiate a systematic exploration of the nonclassical functions in this space, showing (in the case of finite cluster algebras) that there are surprisingly few such functions. Finally, as an application of this technology, we explore the ways in which these functions can be used to express the two-loop seven-point MHV amplitude in section 5. In particular, we show that this amplitude admits

a (nested) decomposition in terms of cluster polylogarithms defined on cluster algebras of every rank smaller than that of  $\text{Gr}(4, 7)$ , and thereby identify cluster polylogarithms of physical interest on the  $D_5$ ,  $A_5$ , and  $A_4$  cluster algebras, supplementing the two that are already known to be of interest on  $A_2$  and  $A_3$ .

We also include two appendices. Appendix A tabulates all the (nested) subalgebras of  $\text{Gr}(4, 7)$ , while appendix B tabulates the number of independent functions that have a nonzero Lie cobracket on each of these finite cluster algebras.

## 2 A Brief Introduction to Cluster Algebras

Cluster algebras were first introduced by Fomin and Zelevinsky [33] as a tool for identifying which algebraic varieties come equipped with a natural notion of positivity, and what quantities determine this positivity. As a consequence, they naturally appear in the study of the positive Grassmannian  $\text{Gr}_+(k, n)$ , i.e. the space of  $k \times n$  matrices modulo the action of  $\text{GL}(k)$  where all ordered  $k \times k$  minors are positive. They correspondingly also appear in the study of planar  $\mathcal{N} = 4$  sYM theory, since the integrands in this theory are encoded to all orders by  $\text{Gr}_+(4, n)$  [1].

A simple example of the type of questions cluster algebras help address is: if one were to check just the positivity of a set of matrix minors (not their actual value), what are the minimal sets of minors that suffice to determine whether a point is in  $\text{Gr}_+(k, n)$ ? These minors are not all independent; they satisfy identities known as Plücker relations, for example

$$\langle abI \rangle \langle cdI \rangle = \langle acI \rangle \langle bdI \rangle - \langle adI \rangle \langle bcI \rangle, \quad (2.1)$$

where each Plücker coordinate  $\langle i_1, \dots, i_k \rangle$  corresponds to the minor of columns  $i_1, \dots, i_k$ , and  $I$  is a multi-index object with  $k - 2$  entries.

To gain some intuition for this problem, let us explore the case of  $\text{Gr}_+(2, 5)$ . The five cyclically adjacent minors  $\langle 12 \rangle$ ,  $\langle 23 \rangle$ ,  $\langle 34 \rangle$ ,  $\langle 45 \rangle$ , and  $\langle 15 \rangle$  cannot be eliminated in terms of each other by Plücker relations, so each gives rise to an independent positivity constraint. However, making use of the Plücker relations

$$\begin{aligned} \langle 24 \rangle &= (\langle 12 \rangle \langle 34 \rangle + \langle 23 \rangle \langle 14 \rangle) / \langle 13 \rangle, \\ \langle 25 \rangle &= (\langle 12 \rangle \langle 45 \rangle + \langle 24 \rangle \langle 15 \rangle) / \langle 14 \rangle, \\ \langle 35 \rangle &= (\langle 25 \rangle \langle 34 \rangle + \langle 23 \rangle \langle 45 \rangle) / \langle 24 \rangle, \end{aligned} \quad (2.2)$$

we can eliminate three of the nonadjacent minors—for instance,  $\langle 24 \rangle$ ,  $\langle 25 \rangle$ , and  $\langle 35 \rangle$ —in terms of the remaining  $\binom{5}{2} - 3 = 7$  adjacent and nonadjacent ones. It can be checked that all further Plücker relations are implied by those in (2.2), telling us that seven minors must be computed to determine if a matrix is in  $\text{Gr}_+(2, 5)$ . However, as should

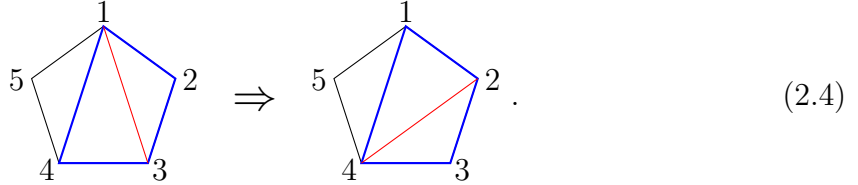
be clear, it is not sufficient to check the positivity of *any* seven (ordered) minors; only certain triples of minors can be eliminated. It would therefore be advantageous to have a method for generating all sets of minors that are sufficient to answer this question.

To motivate how cluster algebras address this problem, consider the following triangulation of the pentagon:



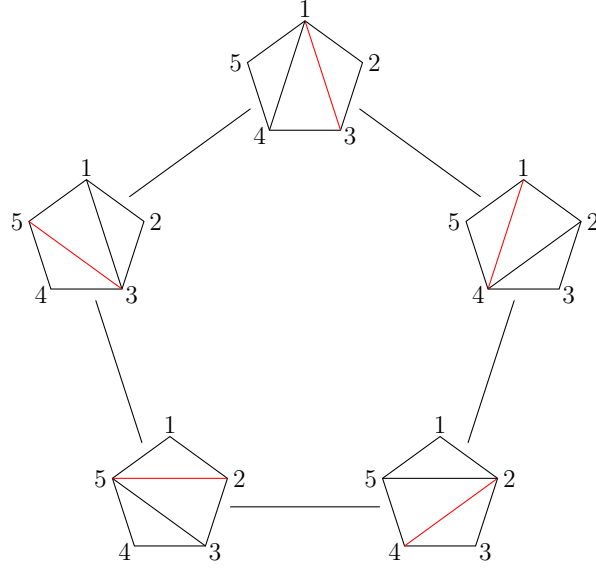
We can associate the line connecting points  $i$  and  $j$  with the Plücker coordinate  $\langle ij \rangle$ ; if we further assign these lines length  $\langle ij \rangle$ , the resulting pentagon is cyclic (in the sense that all its vertices reside on a common circle) due to Ptolemy's theorem. Conversely, all cyclic  $n$ -gons represent a point in  $\text{Gr}_+(2, n)$  [34]. Note that the length of the three diagonals that are not present in this triangulation are determined by the length of the seven lines that are present (including edges); the problem has been reduced to geometry. This makes clear why these three diagonals—the three eliminated above—are redundant for the purpose of determining whether a matrix is in  $\text{Gr}_+(2, 5)$ .

From the first relation in (2.2) we see that we could have instead chosen to check the positivity of  $\langle 24 \rangle$  rather than  $\langle 13 \rangle$ . This corresponds to choosing a different triangulation, which we get by trading the latter diagonal for the former:



We have highlighted in blue the fact that both diagonals are framed by the same quadrilateral face. More generally, we can pick any quadrilateral face and flip the diagonal it contains to generate a different triangulation. Repeatedly performing these flips generates all possible triangulations of the pentagon, as can be seen in figure 1. Each triangulation provides a set of edges/minors whose positivity ensures that a matrix is in  $\text{Gr}_+(2, 5)$ .

In an analogous way, cluster algebras answer questions about positivity for a larger class of algebraic varieties (and in particular for all  $\text{Gr}_+(k, n)$ ) by considering ‘clusters’ that can all be generated by an operation called ‘mutation’ just as all triangulations of the pentagon are generated by flipping the diagonals of quadrilateral faces. We now



**Figure 1:** All possible triangulations of the pentagon. Mutating on the red chord moves you clockwise around the figure.

turn to the definition of these objects, considering first how the example of  $\text{Gr}_+(2, 5)$  can be rephrased in this language.

## 2.1 Clusters, Mutations, and Cluster $\mathcal{A}$ -coordinates

Clusters can be defined to be quiver diagrams—namely, oriented graphs equipped with arrows connecting different nodes—in which each node is assigned a cluster coordinate.<sup>1</sup> We can form one of the clusters of  $\text{Gr}(2, 5)$  out of our original triangulation (2.3) by assigning an orientation to the pentagon and all subtriangles such as



The nodes of our cluster are then given by the lines of this triangulation (making the minors  $\langle ab \rangle$  our cluster coordinates), where an arrow is assigned from  $\langle ab \rangle$  to  $\langle cd \rangle$  if the triangle orientations in (2.5) have segment  $(ab)$  flowing into segment  $(cd)$  and these

<sup>1</sup>Here and throughout, we are implicitly restricting ourselves to skew-symmetric cluster algebras of geometric type.



segments border a common triangle. This gives us the quiver

$$\begin{array}{ccccc}
 \boxed{\langle 12 \rangle} & & & & \\
 \searrow & & & & \\
 \langle 13 \rangle & \longrightarrow & \langle 14 \rangle & \longrightarrow & \boxed{\langle 15 \rangle} \\
 \downarrow & \nearrow & \downarrow & \nearrow & \\
 \boxed{\langle 23 \rangle} & & \boxed{\langle 34 \rangle} & & \boxed{\langle 45 \rangle}
 \end{array} , \tag{2.6}$$

where the boxes around the  $\langle ii + 1 \rangle$  indicate that they are *frozen*—they can never change because they aren’t in the interior of a quadrilateral face. One can also draw arrows (with partial weight) connecting these frozen nodes (see, for instance, [1]), but we will not keep track of them in the present work.

We have now drawn our first cluster (also sometimes called a seed). The Plücker coordinates in this cluster are referred to as cluster  $\mathcal{A}$ -coordinates, and they come in two flavors: mutable (for example,  $\langle 13 \rangle$  and  $\langle 14 \rangle$  above) and frozen ( $\langle ii + 1 \rangle$  above). In any quiver, the information encoded by the arrows can also be represented in terms of a skew-symmetric matrix

$$b_{ij} = (\# \text{ of arrows } i \rightarrow j) - (\# \text{ of arrows } j \rightarrow i) \tag{2.7}$$

called the exchange matrix (or the signed adjacency matrix).

The process of mutation, which we have described geometrically in terms of flipping diagonals, has a simple interpretation at the level of quivers. Given a quiver such as (2.6), we can choose any mutable node  $k$  to mutate on (this is equivalent to picking which diagonal to flip). Mutation gives us back a new quiver in which the  $\mathcal{A}$ -coordinate  $a_k$  has been sent to  $a'_k$ , where

$$a_k a'_k = \prod_{i|b_{ik}>0} a_i^{b_{ik}} + \prod_{i|b_{ik}<0} a_i^{-b_{ik}}, \tag{2.8}$$

(with the understanding that an empty product is set to one), while all other cluster  $\mathcal{A}$ -coordinates remain unchanged. The arrows connecting the nodes in this new quiver are also modified by carrying out the following sequence of operations:

- for each path  $i \rightarrow k \rightarrow j$ , add an arrow  $i \rightarrow j$ ,
- reverse all arrows on the edges incident with  $k$ ,
- remove any two-cycles (oppositely-oriented arrows) that may have formed.

This creates a new adjacency matrix  $b'_{ij}$  via

$$b'_{ij} = \begin{cases} -b_{ij}, & \text{if } k \in \{i, j\}, \\ b_{ij}, & \text{if } b_{ik}b_{kj} \leq 0, \\ b_{ij} + b_{ik}b_{kj}, & \text{if } b_{ik}, b_{kj} > 0, \\ b_{ij} - b_{ik}b_{kj}, & \text{if } b_{ik}, b_{kj} < 0. \end{cases} \quad (2.9)$$

Mutation is an involution, so mutating on  $a'_k$  will take you back to the original cluster (just as flipping the same diagonal twice will take you back to where you started).

In terms of these ingredients, a cluster algebra can be defined to be a set of clusters that is closed under mutation. Thus, mutating on any non-frozen node of any cluster will generate a different cluster in the same cluster algebra. In practice, one therefore constructs cluster algebras by starting from a seed such as (2.6), and iteratively mutating on all available nodes until the set of clusters closes (or it becomes clear the cluster algebra is infinite).

It is common to refer to certain cluster algebras by particularly nice representative clusters, when the mutable nodes of the corresponding quiver form an oriented Dynkin diagram. For instance, the  $\text{Gr}(2,5)$  cluster algebra is often referred to as  $A_2$ , since the mutable part of the seed (2.6) is given by  $\langle 13 \rangle \rightarrow \langle 14 \rangle$ . Thus, we will often speak interchangeably of the cluster algebras for  $\text{Gr}(2,5)$  and  $A_2$ . This is a slight abuse of notation, as the  $\text{Gr}(2,5)$  cluster algebra corresponds specifically to the cluster algebra generated by the collection of frozen and mutable nodes in eq. (2.6), whereas an  $A_2$  cluster algebra can in principle be dressed with any number of frozen nodes. We will see why this language is useful in the next section.

## 2.2 Cluster $\mathcal{X}$ -coordinates

Cluster algebras can also be formulated in terms of a different set of cluster coordinates, called Fock-Goncharov coordinates or  $\mathcal{X}$ -coordinates [35]. As we will see in future sections, cluster  $\mathcal{X}$ -coordinates play a crucial role in connecting cluster algebras to polylogarithms and scattering amplitudes. While it is always possible to phrase results involving cluster algebras directly in terms of  $\mathcal{A}$ -coordinates,  $\mathcal{X}$ -coordinates often allow for cluster-algebraic structure to be made more manifest.

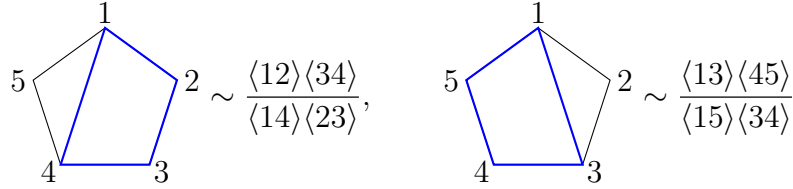
Clusters formed out of  $\mathcal{X}$ -coordinates can be directly constructed out of clusters involving  $\mathcal{A}$ -coordinates. Given a quiver equipped with  $\mathcal{A}$ -coordinates and described by the exchange matrix  $b_{ij}$ , we can compute an  $\mathcal{X}$ -coordinate to assign to each mutable node by

$$x_i = \prod_j a_j^{b_{ji}}. \quad (2.10)$$

For example, the  $\mathcal{X}$ -coordinate cluster associated with (2.6) is formed by associating cluster  $\mathcal{X}$ -coordinates with all its mutable nodes, where these  $\mathcal{X}$ -coordinates are constructed by putting all Plücker coordinates that point to that node in the numerator, and all Plücker coordinates that are pointed to by that node in the denominator. That is, we get

$$\frac{\langle 12 \rangle \langle 34 \rangle}{\langle 14 \rangle \langle 23 \rangle} \rightarrow \frac{\langle 13 \rangle \langle 45 \rangle}{\langle 15 \rangle \langle 34 \rangle}, \quad (2.11)$$

which again takes the form of the generic  $A_2$  quiver  $x_1 \rightarrow x_2$ . In the pentagon-triangulation picture, these  $\mathcal{X}$ -coordinates describe overlapping quadrilaterals, for instance



$$\begin{array}{cc} \text{Pentagon with diagonal } 14 \sim \frac{\langle 12 \rangle \langle 34 \rangle}{\langle 14 \rangle \langle 23 \rangle}, & \text{Pentagon with diagonal } 13 \sim \frac{\langle 13 \rangle \langle 45 \rangle}{\langle 15 \rangle \langle 34 \rangle}, \end{array} \quad (2.12)$$

which come in one-to-one correspondence with the diagonals in a (single) triangulation.

Mutation rules for  $\mathcal{X}$ -coordinates are different than for  $\mathcal{A}$ -coordinates, and are given by

$$x'_i = \begin{cases} x_k^{-1}, & i = k, \\ x_i(1 + x_k^{\text{sgn } b_{ik}})^{b_{ik}}, & i \neq k \end{cases}, \quad (2.13)$$

where mutation has been carried out on node  $k$ . Mutation still changes the arrows in the quiver diagram as it did in the case of  $\mathcal{A}$ -coordinates. Given just a quiver diagram, it can sometimes be unclear whether a given quiver should be mutated using the  $\mathcal{A}$ -coordinate or  $\mathcal{X}$ -coordinate rules (2.8) or (2.13). We adopt the convention in this work that if a quiver is given with no frozen nodes, it should be thought of as equipped with  $\mathcal{X}$ -coordinates.

Just as with  $\mathcal{A}$ -coordinate clusters, we can generate all  $\mathcal{X}$ -coordinate clusters by mutation. Mutating on alternating nodes of our  $A_2$  cluster (starting with  $x_2$ ), we

generate the following sequence of clusters:

$$\begin{aligned}
& x_1 \rightarrow x_2 \\
& x_1(1+x_2) \leftarrow \frac{1}{x_2} \\
& \frac{1}{x_1(1+x_2)} \rightarrow \frac{1+x_1+x_1x_2}{x_2} \\
& \frac{1+x_1}{x_1x_2} \leftarrow \frac{x_2}{1+x_1+x_1x_2} \\
& \frac{x_1x_2}{1+x_1} \rightarrow \frac{1}{x_1} \\
& x_2 \leftarrow x_1 \\
& \vdots
\end{aligned} \tag{2.14}$$

The series then repeats, with all arrows reversed. Note that (specifically in the case of  $A_2$ ), if we label these  $\mathcal{X}$ -coordinates by

$$\begin{aligned}
\mathcal{X}_1 &= 1/x_1, & \mathcal{X}_2 &= x_2, & \mathcal{X}_3 &= x_1(1+x_2), \\
\mathcal{X}_4 &= \frac{1+x_1+x_1x_2}{x_2}, & \mathcal{X}_5 &= \frac{1+x_1}{x_1x_2},
\end{aligned} \tag{2.15}$$

the mutation rule in (2.13) takes the simple form

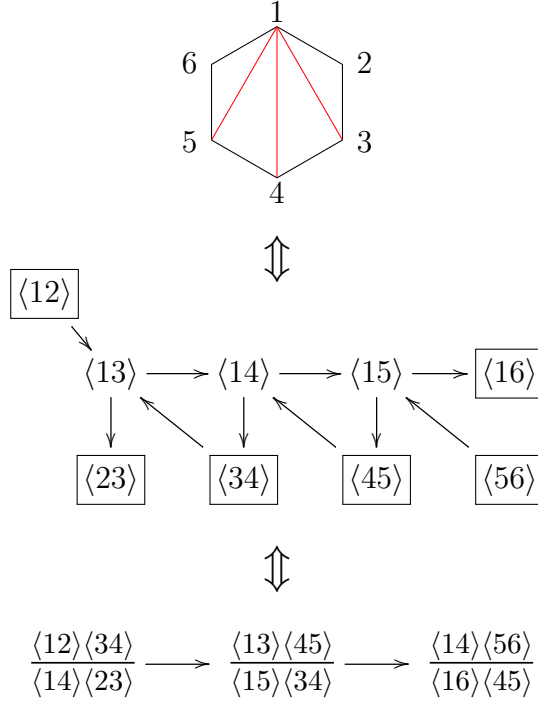
$$1 + \mathcal{X}_i = \mathcal{X}_{i-1}\mathcal{X}_{i+1}, \tag{2.16}$$

while all the clusters take the form  $1/\mathcal{X}_i \rightarrow \mathcal{X}_{i+1}$ . Eq. (2.16) is commonly referred to as the  $A_2$  exchange relation. Putting this all together, we will generically refer to an  $A_2$  cluster algebra as any set of clusters  $1/\mathcal{X}_{i-1} \rightarrow \mathcal{X}_i$  for  $i = 1 \dots 5$  where the  $\mathcal{X}_i$  satisfy eq. (2.16). We believe it is useful at this point to emphasize that one can take as input any  $x_1$  and  $x_2$ , and generate an associated  $A_2$ .

A very useful feature of cluster  $\mathcal{X}$ -coordinates is that they come equipped with a natural Poisson bracket structure, making the Grassmannian  $\text{Gr}(k, n)$  a cluster Poisson variety [36]. Namely, when two  $\mathcal{X}$ -coordinates appear together in a cluster of  $\text{Gr}(k, n)$ , there exists a Poisson bracket that evaluates to

$$\{x_i, x_j\} = b_{ij}x_ix_j. \tag{2.17}$$

This structure respects mutation, implying that the entry  $b_{ij}$  (which counts the number of arrows from  $x_i$  to  $x_j$  in a given cluster's quiver) will be the same in all clusters containing both  $x_i$  and  $x_j$ . The Poisson bracket (and associated Sklyanin bracket) will play a larger role in forthcoming work [37], so we defer further discussion of this structure (for existing discussions in the literature, see [36, 38]).



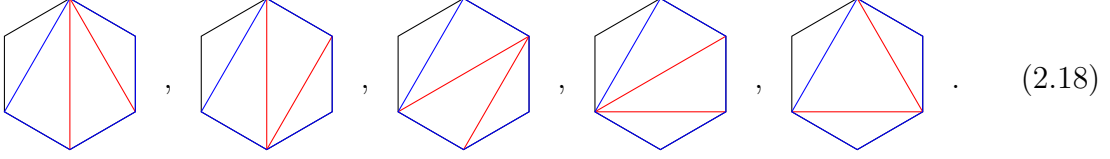
**Figure 2:** A triangulation of the hexagon along with its associated  $\mathcal{A}$ -coordinate and  $\mathcal{X}$ -coordinate seed quivers.

### 2.3 Subalgebras and Cluster Polytopes

Cluster algebras contain a rich and intricate subalgebra structure, which will play a central role in our analysis. It is simple to illustrate how these subalgebras arise by considering  $\text{Gr}(2, 6)$ , which triangulates the hexagon. In figure 2 we give the seed cluster for  $\text{Gr}(2, 6)$  in the triangulation,  $\mathcal{A}$ -coordinate, and  $\mathcal{X}$ -coordinate representations. Since the mutable nodes take the form of an  $A_3$  Dynkin diagram, we often speak of  $\text{Gr}(2, 6)$  and  $A_3$  interchangeably, just as we did with  $\text{Gr}(2, 5) \simeq A_2$ .

The  $\text{Gr}(2, 6)$  cluster algebra features 14 clusters, which can be grouped into multiple (overlapping) subalgebras. A simple example is the collection of all triangulations which involve the chord  $\langle 15 \rangle$ . This set contains 5 clusters and is itself a cluster algebra, which can be generated by treating  $\langle 15 \rangle$  as a frozen node (or in  $\mathcal{X}$ -coordinates, freezing the node  $\frac{\langle 14 \rangle \langle 56 \rangle}{\langle 16 \rangle \langle 45 \rangle}$ ). This of course is the cluster algebra corresponding to the triangulations

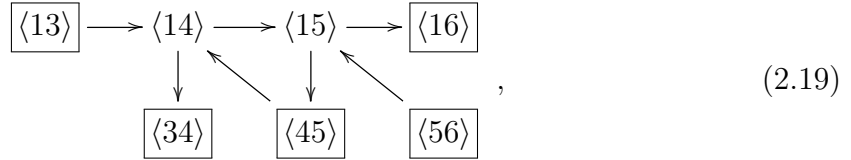
of the pentagon formed by points  $1, \dots, 5$ , outlined here in blue:



We therefore refer to the collection of clusters that involve this pentagon as an  $A_2$  subalgebra of  $\text{Gr}(2, 6)$ .

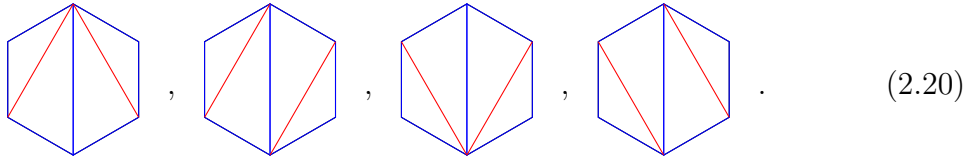
It should be clear, upon referring back to the mutation rules (2.8) and (2.13), that this  $A_2$  subalgebra is truly identical to what we have been calling  $\text{Gr}(2, 5)$ . That is, neither the  $\mathcal{A}$ - or  $\mathcal{X}$ -coordinate mutation rule (or the rule for constructing the  $\mathcal{X}$ -coordinate cluster out of the  $\mathcal{A}$ -coordinate one) depends on nodes further than a single arrow away from the node on which one is mutating. Correspondingly, this subalgebra doesn't know about the existence of nodes involving point/column 6. (In the  $\mathcal{X}$ -coordinate case, the coordinate associated with the newly frozen node will change when one mutates on the node it is connected to, but the presence of this frozen node does not effect the coordinates appearing in the  $A_2$  subalgebra itself.) We consider two subalgebras to be identical when the clusters they appear in only differ by nodes that have no effect on the mutable nodes of the subalgebra.

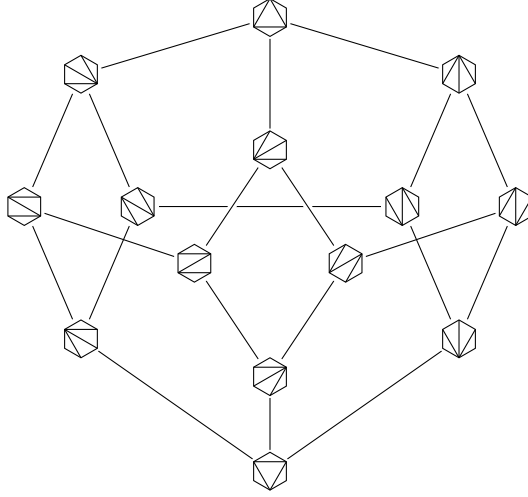
What if we instead disallow mutation on the chord  $\langle 13 \rangle$  (and the corresponding  $\mathcal{X}$ -coordinate node  $\frac{\langle 12 \rangle \langle 34 \rangle}{\langle 14 \rangle \langle 23 \rangle}$ )? Dropping the nodes that play no role in any of the mutations that remain gives rise to the effective quiver



where we have put a box around  $\langle 13 \rangle$  to make clear we are now treating it as frozen. We have also dropped the arrow from  $\langle 34 \rangle$  to  $\langle 13 \rangle$  since we are ignoring arrows between frozen nodes. The comparison to (2.6) should be clear; this just represents a re-labeled version of  $\text{Gr}(2, 5)$ .

Similarly, if we disallow mutation on the chord  $\langle 14 \rangle$  (and  $\frac{\langle 13 \rangle \langle 45 \rangle}{\langle 15 \rangle \langle 34 \rangle}$ ), we generate an  $A_1 \times A_1$  subalgebra, since the chord  $\langle 14 \rangle$  divides the hexagon in to two non-overlapping squares, each of which are triangulated by  $A_1$  (or really  $\text{Gr}(2, 4)$ ):





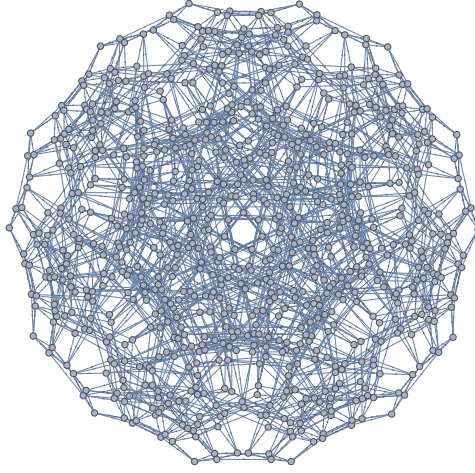
**Figure 3:** The cluster polytope of  $A_3 \simeq \text{Gr}(2, 6)$ , in which each cluster is represented by a triangulation of the hexagon.

In appendix A we have tabulated the number of such  $A_2$  and  $A_1 \times A_1$  subalgebras in  $A_3$ , as well as the subalgebras of other cluster algebras that appear in  $\text{Gr}(4, 7)$ . There it will be found that there are in fact six  $A_2$  subalgebras and three  $A_1 \times A_1$  subalgebras of  $A_3$ .

This subalgebra structure can be nicely visualized by constructing an object known as the cluster polytope of a given cluster algebra. The vertices of this polytope each represent a cluster, while its edges represent the mutations that map these clusters into each other. For instance, figure 1 corresponds to the  $\text{Gr}(2, 5) \simeq A_2$  cluster polytope, which also coincidentally takes the form of a pentagon. Note that every node has valency two since each cluster has two mutable vertices.

Similarly, the cluster polytope of  $\text{Gr}(2, 6) \simeq A_3$  is given in figure 3. It has 14 vertices, corresponding to the 14 clusters of  $\text{Gr}(2, 6)$ , each with valency three. These vertices assemble into three square faces and six pentagonal faces—corresponding exactly to the three  $A_1 \times A_1$  subalgebras and the six  $A_2$  subalgebras of  $A_3$ . This makes it easy to read off the subalgebra structure of  $A_3$  directly.

In  $A_3$ , it turns out that each of the faces corresponds to a distinct subalgebra. Cluster polytopes of larger cluster algebras become quite complicated, and it is often the case that distinct faces (or higher-dimensional polytopes) correspond to identical subalgebras. As an example, the cluster polytope of  $\text{Gr}(4, 7)$ , which will be the focus of much of the rest of this paper, is shown in figure 4. It has 833 vertices, each of valence 6, and 1071 pentagons corresponding to  $A_2$  subalgebras; however, only 504 of these  $A_2$  subalgebras are distinct.



**Figure 4:** The cluster polytope of  $\text{Gr}(4, 7) \simeq E_6$ .

While cluster polytopes give us a nice pictorial way to think about the relations between different clusters in a cluster algebra, their interiors can also be identified with the ‘positive region’ where all (cyclically ordered)  $\mathcal{A}$ -coordinates (and consequently  $\mathcal{X}$ -coordinates) are positive [1, 19]. Thus, in the case of Grassmannian cluster algebras, the positive Grassmannian  $\text{Gr}_+(k, n)$  (restricted to real values) consists of all points in the interior of the cluster polytope. Each cluster, moreover, can be thought of as providing a coordinate system in which the full positive region is spanned by allowing each  $\mathcal{X}$ -coordinate to range from 0 to  $\infty$ . For a more in-depth discussion of this way of thinking about cluster polytopes, see [19].

## 2.4 Grassmannian Cluster Algebras and Planar $\mathcal{N} = 4$ sYM Theory

So far we have leaned heavily on the correspondence between the triangulations of an  $n$ -gon and the cluster algebra for  $\text{Gr}(2, n)$ . Based on the examples of  $\text{Gr}(2, 5)$  and  $\text{Gr}(2, 6)$ , it is not hard to write down a generic seed cluster for  $\text{Gr}(2, n)$  corresponding to the triangulation consisting of all chords  $\langle 13 \rangle, \dots, \langle 1n-1 \rangle$ :

$$\begin{array}{c}
 \begin{array}{ccc}
 & 1 & \\
 n & \diagdown & 2 \\
 & \text{---} & \\
 n-1 & \diagup & 3
 \end{array}
 & \Leftrightarrow &
 \begin{array}{ccccccc}
 \boxed{\langle 12 \rangle} & & & & & & \\
 \downarrow & & & & & & \\
 \langle 13 \rangle & \longrightarrow & \dots & \longrightarrow & \langle 1 \ n-1 \rangle & \longrightarrow & \boxed{\langle 1n \rangle} \\
 \downarrow & \swarrow & \downarrow & \swarrow & \downarrow & \swarrow & \\
 \boxed{\langle 23 \rangle} & & \dots & & \boxed{\langle n-2 \ n-1 \rangle} & \boxed{\langle n-1 \ n \rangle}
 \end{array}
 \end{array} \quad (2.21)$$

Here one sees that  $\text{Gr}(2, n) \simeq A_{n-3}$ .



For  $\text{Gr}(k > 2, n)$ , there is no longer a simple connection with triangulations or Dynkin diagrams. However, there exists a generalization of eq. (2.21) valid for all  $\text{Gr}(k, n)$  [39]:

$$\begin{array}{ccccccc}
 \boxed{f_{11}} & \leftarrow f_{12} & \leftarrow f_{13} & \leftarrow \cdots & \leftarrow f_{1l} & \leftarrow \boxed{\langle 1, \dots, k \rangle} \\
 & \nearrow & \downarrow & \nearrow & \downarrow & \nearrow & \downarrow \\
 \boxed{f_{21}} & \leftarrow f_{22} & \leftarrow f_{23} & \leftarrow \cdots & \leftarrow f_{2l} & & \\
 & \nearrow & \downarrow & \nearrow & \downarrow & \nearrow & \downarrow \\
 \vdots & \leftarrow \vdots & \leftarrow \vdots & \leftarrow \vdots & \leftarrow \vdots & & \\
 & \nearrow & \downarrow & \nearrow & \downarrow & \nearrow & \downarrow \\
 \boxed{f_{k1}} & \boxed{f_{k2}} & \boxed{f_{k3}} & \cdots & \boxed{f_{kl}} & & 
 \end{array} , \tag{2.22}$$

where  $l = n - k$  and

$$f_{ij} = \begin{cases} \langle i+1, \dots, k, k+j, \dots, i+j+k-1 \rangle, & i \leq l-j+1, \\ \langle 1, \dots, i+j-l-1, i+1, \dots, k, k+j, \dots, n \rangle, & i > l-j+1. \end{cases} \tag{2.23}$$

(Note that evaluating the above expression for  $k = 2$  will not directly give (2.21); the two are equivalent after cyclically rotating the indices in (2.21) and flipping the direction of the arrows.) The cluster algebra on  $\text{Gr}(k, n)$  is therefore of rank  $(n - k - 1)(k - 1)$ , i.e. the number of mutable nodes in (2.22).

The cluster algebra on  $\text{Gr}(4, n)$  naturally appears in planar  $\mathcal{N} = 4$  sYM theory, where it parametrizes the space of  $n$ -particle kinematics. To make this connection, one first decomposes the external momenta  $p_i$  (which are endowed with a natural ordering in the planar limit) into a set of spinors  $\lambda_i, \tilde{\lambda}_i$  or into a set of dual coordinates  $x_i$  as

$$p_i^\mu \sigma_\mu^{\alpha\dot{\alpha}} = \lambda_i^\alpha \tilde{\lambda}_i^{\dot{\alpha}} = x_i^{\alpha\dot{\alpha}} - x_{i+1}^{\alpha\dot{\alpha}}, \tag{2.24}$$

where  $x_{n+1}^{\alpha\dot{\alpha}} \equiv x_1^{\alpha\dot{\alpha}}$  (for more background on these spinors and dual coordinates, see for example [40, 41]). The dual coordinates describe the cusps of a light-like Wilson loop dual to the amplitude; as a result, the original amplitude respects an additional superconformal symmetry that is associated with these dual coordinates (up to an anomaly associated with the cusps of the Wilson loop, which is accounted for by the BDS ansatz) [42–50]. In terms of the quantities in (2.24), we can define momentum twistors

$$Z_i^R = (\lambda_i^\alpha, x_i^{\beta\dot{\alpha}} \lambda_{i\beta}), \tag{2.25}$$

where  $R = (\alpha, \dot{\alpha})$  is an  $SU(2, 2)$  index. Momentum twistors are invariant under the little group, which acts as an overall rescaling  $Z_i^R \rightarrow t_i Z_i^R$ , and as such represent points in  $\mathbb{CP}^3$ .

If we assemble these momentum twistors into a  $4 \times n$  matrix in which the  $i^{\text{th}}$  column corresponds to the four  $SU(2, 2)$  components of  $Z_i^R$ , invariance under the dual conformal group becomes invariance under  $SL(4)$ . The overall rescaling symmetry of one of the momentum twistors can be combined with this  $SL(4)$  invariance to identify this matrix as a point in the (not necessarily positive) Grassmannian  $\text{Gr}(4, n)$ , modulo the rescaling invariance of the remaining  $n - 1$  columns. Thus, the kinematic data of an  $n$ -point scattering process is encoded in a momentum twistor matrix

$$Z \in \text{Gr}(4, n)/\text{GL}(1)^{n-1}. \quad (2.26)$$

For more details regarding this correspondence, see [1, 2].

To relate the dual-conformal invariants encoded in  $Z_n$  to more familiar kinematic quantities, we can translate the (cyclically ordered) Mandelstam invariants into squared differences of dual coordinates,

$$s_{i, \dots, j-1} \equiv (p_i + \dots p_{j-1})^2 = \det(x_i^{\alpha\dot{\alpha}} - x_j^{\alpha\dot{\alpha}}) \equiv x_{ij}^2. \quad (2.27)$$

Dual conformal invariants can be constructed out of these objects by putting together combinations that are invariant under the dual conformal inversion generator, which acts on these coordinates as

$$I(x_i^{\alpha\dot{\alpha}}) = \frac{x_i^{\alpha\dot{\alpha}}}{x_i^2}, \quad I(x_{ij}^2) = \frac{x_{ij}^2}{x_i^2 x_j^2}. \quad (2.28)$$

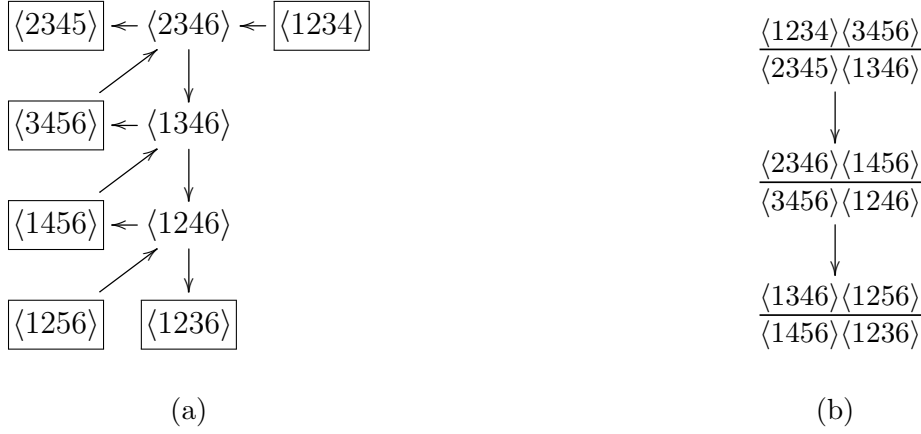
Thus, (regulated) amplitudes in this theory depend only on ratios of squared differences in which the same dual indices appear in both the numerator and denominator. The quantities  $x_{ij}^2$  can be translated into momentum twistors using the relation

$$x_{ij}^2 = \frac{\det(Z_{i-1} Z_i Z_{j-1} Z_j)}{(\epsilon_{\alpha\beta} \lambda_{i-1}^\alpha \lambda_i^\beta)(\epsilon_{\gamma\delta} \lambda_{j-1}^\gamma \lambda_j^\delta)}, \quad (2.29)$$

where  $\epsilon_{\alpha\beta}$  is the Levi-Civita tensor. In dual-conformally invariant quantities, the spinor products  $\epsilon_{\alpha\beta} \lambda_{i-1}^\alpha \lambda_i^\beta$  all cancel, leaving only determinants of four-tuples of momentum twistors. These are just minors of the momentum twistor matrix (2.26), which we recognize as the cluster  $\mathcal{A}$ -coordinates

$$\langle ijkl \rangle = \det(Z_i Z_j Z_k Z_l). \quad (2.30)$$

Note that the two-particle Mandelstams  $s_{i, i+1}$  correspond to the frozen nodes of (2.22), while higher-particle Mandelstams and more general (polynomials of) Plücker coordinates can appear as mutable nodes.



**Figure 5:** The  $\mathcal{A}$ -coordinate seed quiver (a) and  $\mathcal{X}$ -coordinate seed quiver (b) for  $\text{Gr}(4,6)$ .

By construction, the  $\mathcal{X}$ -coordinates on  $\text{Gr}(4,n)$  derived from the seed (2.22) respect dual-conformal invariance. Both mutation rules (2.8) and (2.13) preserve this property (and commute with the translation (2.10)), ensuring that all  $\mathcal{X}$ -coordinates are dual conformal invariants. Such invariants cannot be formed in four- or five-particle kinematics, due to an insufficient number of non-lightlike separated points (since we are in massless kinematics,  $x_{i+1}^2 = 0$  for all  $i$ ). This fact shows up in the seed (2.22) as  $\text{Gr}(4, n < 6)$  having no mutable nodes (and therefore no  $\mathcal{X}$ -coordinates). For  $n > 5$ , there are  $3(n - 5)$  mutable nodes in  $\text{Gr}(4, n)$ , matching the number of algebraically independent dual conformal invariants that can be formed out of  $n$  massless particles.

The  $\mathcal{A}$ -coordinate and  $\mathcal{X}$ -coordinate seed clusters of the first nontrivial example,  $\text{Gr}(4,6)$ , are shown in figure 5. As discussed above, the three  $\mathcal{X}$ -coordinates in this cluster furnish us with a chart that covers the space of (dual-conformally invariant) six-particle kinematics. Moreover, we can generate new charts by mutation—every  $\mathcal{X}$ -coordinate cluster of  $\text{Gr}(4, n)$  provides a valid chart for  $n$ -particle kinematics. As explored in great depth in [1], these charts are especially well suited to describing the boundaries of the positive Grassmannian  $\text{Gr}_+(4, n)$ , where the integrands of  $n$ -particle amplitudes can develop physical singularities. In particular, every such boundary occurs at the vanishing locus of an  $\mathcal{A}$ -coordinate of  $\text{Gr}(4, n)$ , which implies it also occurs at the vanishing locus of some set of  $\mathcal{X}$ -coordinates.

This fact is especially propitious for loop-level amplitudes (and integrals) that only have branch points on the boundaries of the positive Grassmannian. In such cases, the symbol alphabet encoding the polylogarithmic part of these amplitudes is naturally given in terms of cluster coordinates. We defer discussion of the coaction and symbol

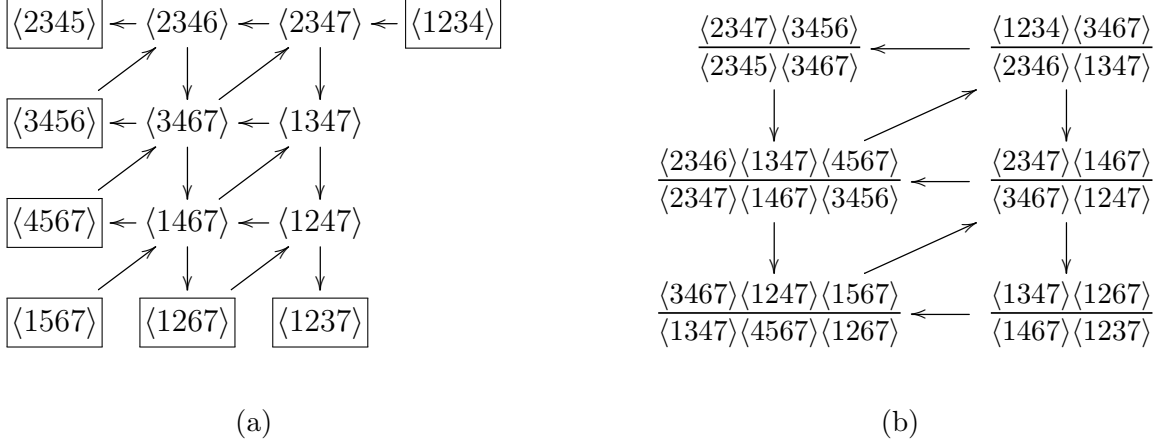
alphabets to section 3.1, but here note that it is multiplicative independence, rather than algebraic independence, that is relevant in the context of symbol alphabets. Thus, while it is not possible to realize all boundaries of the positive Grassmannian as the vanishing loci of either type of cluster coordinate in a single chart [1], all boundaries are exposed as the vanishing of some symbol letter if cluster  $\mathcal{A}$ -coordinates or  $\mathcal{X}$ -coordinates on  $\text{Gr}(4, n)$  are adopted as a symbol alphabet.

While amplitudes in planar  $\mathcal{N} = 4$  are not generically expected to have this property (and indeed, certain Feynman integrals have been computed that do not [21, 22]), an infinite class of amplitudes do—namely, all two-loop MHV amplitudes [9], and all six- and seven-particle amplitudes computed to date [6–8, 15–17]. The significance of this property is illustrated by the two-loop, six-particle remainder function, which encodes the MHV amplitude. Namely, this function can be put in the form

$$R_6^{(2)} = - \sum_{\text{cyclic}} \left[ \text{Li}_4 \left( -\frac{\langle 1234 \rangle \langle 3456 \rangle}{\langle 2345 \rangle \langle 1346 \rangle} \right) - \frac{1}{4} \text{Li}_4 \left( -\frac{\langle 1234 \rangle \langle 1456 \rangle}{\langle 1246 \rangle \langle 1345 \rangle} \right) \right] + \dots, \quad (2.31)$$

where the cyclic sum is over all rotations of the four-bracket indices  $i \rightarrow i + j$  for  $0 \leq j < 6$ , and the dots indicate this equality only holds up to products of lower-weight polylogarithms. This projection is well-defined and will be introduced, along with the  $n$ -particle remainder function, in section 3. Here we just emphasize the simplicity of this expression, which takes the form of classical polylogarithms with negative  $\mathcal{X}$ -coordinate arguments. (In particular, the argument of the first polylogarithm is the top node in figure 5b, while the argument of the second polylogarithm appears in the cluster generated by mutating on that node.) Moreover, the part of the expression we have dropped in (2.31) can also be expressed entirely in terms of products of classical polylogarithms with negative  $\mathcal{X}$ -coordinate arguments [10].

This surprising property—of being expressible as polylogarithms with negative  $\mathcal{X}$ -coordinate arguments—is enjoyed by the two-loop MHV amplitude at all  $n$ . However, for  $n > 6$  these amplitudes have a nonclassical component, so generalized polylogarithms with negative  $\mathcal{X}$ -coordinate arguments also appear [2]. Although this component represents the mathematically most complicated part of the remainder function, it was shown in [12] that it is decomposable into building blocks related to the  $A_2$  and  $A_3$  subalgebras of  $\text{Gr}(4, n)$ . This allows the all- $n$  symbol computed in [9] to be systematically upgraded to a function, as was done for seven particles in [10]. In the later sections of this paper, we demonstrate the existence of further subalgebra structure in the nonclassical part of the seven-particle MHV amplitude (leaving higher-point kinematics to future work [37]). The  $\mathcal{A}$ -coordinate and  $\mathcal{X}$ -coordinate seeds for the cluster algebra relevant to seven-particle scattering,  $\text{Gr}(4, 7)$ , are presented in figure 6.



**Figure 6:** The  $\mathcal{A}$ -coordinate seed quiver (a) and  $\mathcal{X}$ -coordinate seed quiver (b) for  $\text{Gr}(4, 7)$ .

Before turning to the remaining aspects of cluster algebras that we wish to develop, we note that plabic graphs—which are dual to the clusters we’ve been describing—encode a great deal more about planar  $\mathcal{N} = 4$  than we have had reason to touch on. We refer interested readers to the exposition of this rich structure given in [1].

## 2.5 Finite Cluster Algebras

The procedure of writing down an oriented quiver, dressing it with coordinates, and iteratively mutating on all non-frozen nodes using either the  $\mathcal{A}$ -coordinate or  $\mathcal{X}$ -coordinate mutation rule will always produce a cluster algebra. However, generic quivers give rise to exceedingly complicated cluster algebras—in fact, for a wide class of seeds, mutation will generate an infinite numbers of clusters. For the remainder of this paper we will mostly restrict our attention to finite cluster algebras, leaving the consideration of infinite algebras to future work.

Fortunately, all finite cluster algebras were classified in [51]. In particular, it was shown that a cluster algebra is of finite type if and only if the mutable part of at least one of its clusters takes the form of an oriented, simply-laced Dynkin diagram:  $A_n$ ,  $D_n$ , or  $E_{n \leq 8}$ . As we will primarily be interested in subalgebras of the cluster algebra on  $\text{Gr}(4, 7)$ , we here focus on the cases where  $n < 6$  (and on the case of  $E_6 \simeq \text{Gr}(4, 7)$  itself).

As mentioned above, cluster algebras of type  $A_n$  can be generated by the seed

$$x_1 \rightarrow x_2 \rightarrow \dots \rightarrow x_n, \quad (2.32)$$

which corresponds to the cluster algebra on  $\text{Gr}(2, n+3)$ . Each of the clusters in these algebras can be thought as triangulating an  $(n+3)$ -gon, where the  $\mathcal{A}$ -coordinates correspond to chords and the  $\mathcal{X}$ -coordinates to quadrilateral faces. This makes the counting easy: the number of clusters for  $A_n$  is given by the Catalan number  $C(n+1)$ , the number of distinct (mutable)  $\mathcal{A}$ -coordinates is  $\binom{n+3}{2} - n$ , and the number of distinct  $\mathcal{X}$ -coordinates is  $2\binom{n+3}{4}$ . Any smaller polygon embedded into the  $(n+3)$ -gon gives rise to a subalgebra; for example, there are  $56 = \binom{8}{5}$  pentagonal embeddings in an octagon, so there are 56  $A_2$  subalgebras in  $A_5$ .

Of particular interest is the cluster algebra generated by  $A_3 \simeq \text{Gr}(4, 6)$ , which describes six-particle scattering. By comparison with figure 5b, we see that the  $\mathcal{X}$ -coordinates in the quiver (2.32) act as coordinates on the space of momentum twistors, where they correspond to the functions

$$x_1 = \frac{\langle 1234 \rangle \langle 3456 \rangle}{\langle 2345 \rangle \langle 1346 \rangle}, \quad x_2 = \frac{\langle 2346 \rangle \langle 1456 \rangle}{\langle 3456 \rangle \langle 1246 \rangle}, \quad x_3 = \frac{\langle 1346 \rangle \langle 1256 \rangle}{\langle 1456 \rangle \langle 1236 \rangle}. \quad (2.33)$$

More generally, any  $\mathcal{X}$ -coordinate in  $\text{Gr}(4, 6)$  can be expressed in terms of the variables  $x_1$ ,  $x_2$ , and  $x_3$  by evaluating its four-brackets on the momentum twistor matrix

$$Z_{A_3} = \begin{pmatrix} 1 & 0 & 0 & 0 & -1 & -1 \\ 1 & 1 & 0 & 0 & x_1 & 0 \\ 0 & 1 & 1 & 0 & -x_1 x_2 & 0 \\ 0 & 0 & 1 & 1 & x_1 x_2 x_3 & 0 \end{pmatrix}. \quad (2.34)$$

The chief advantage of working directly in terms of cluster  $\mathcal{X}$ -coordinates such as  $x_1$ ,  $x_2$ , and  $x_3$  is that they trivialize all Plücker relations. Furthermore, cluster coordinates rationalize many of the square roots that appear when amplitudes and integrals are expressed in terms of dual-conformally-invariant cross ratios [21]. For instance, in this chart the dual conformal cross ratios commonly used to express the six-particle amplitude evaluate to

$$u = \frac{\langle 6123 \rangle \langle 3456 \rangle}{\langle 6134 \rangle \langle 2356 \rangle} = \frac{1}{1 + x_2 + x_2 x_3}, \quad (2.35)$$

$$v = \frac{\langle 1234 \rangle \langle 4561 \rangle}{\langle 1245 \rangle \langle 3461 \rangle} = \frac{x_1 x_2}{1 + x_1 + x_1 x_2}, \quad (2.36)$$

$$w = \frac{\langle 2345 \rangle \langle 5612 \rangle}{\langle 2356 \rangle \langle 4512 \rangle} = \frac{x_2 x_3}{(1 + x_1 + x_1 x_2)(1 + x_2 + x_2 x_3)}, \quad (2.37)$$

which rationalizes the well-known square root that appears in these cross ratios

$$\sqrt{(1 - u - v - w)^2 - 4uvw} = \frac{x_2(1 - x_1 x_3)}{(1 + x_1 + x_1 x_2)(1 + x_2 + x_2 x_3)}. \quad (2.38)$$

Note that the cluster coordinate expressions (2.35) are rotated compared to those given elsewhere in the literature (for example, [2, 52]) even though both arise from an  $\mathcal{X}$ -coordinate seed of the form (2.32); this reflects a differing convention for the seed of  $\text{Gr}(k, n)$ .

The first nondegenerate Dynkin diagram of type  $D_n$  is  $D_4$ , corresponding to the seed quiver

$$\begin{array}{c} & & x_3 \\ & \nearrow & \\ x_1 \longrightarrow x_2 & & \\ & \searrow & \\ & & x_4 \end{array} . \quad (2.39)$$

(Note here that the variables  $x_1$ ,  $x_2$ , and  $x_3$  are not the same as those defined by (2.33); we apologize for using the variables  $x_i$  ubiquitously, but hope their meaning is always clear from context.) This seed turns out to generate the same cluster algebra as  $\text{Gr}(3, 6)$ ; in particular, starting from the seed in (2.22) and mutating on the nodes initially labeled by  $f_{13}$  and then  $f_{23}$ , one arrives at the  $\mathcal{X}$ -coordinate quiver (2.39), where

$$x_1 = \frac{\langle 123 \rangle \langle 345 \rangle}{\langle 234 \rangle \langle 135 \rangle}, \quad x_2 = \frac{\langle 156 \rangle \langle 235 \rangle}{\langle 125 \rangle \langle 356 \rangle}, \quad x_3 = \frac{\langle 135 \rangle \langle 456 \rangle}{\langle 156 \rangle \langle 345 \rangle}, \quad x_4 = \frac{\langle 126 \rangle \langle 135 \rangle}{\langle 123 \rangle \langle 156 \rangle}. \quad (2.40)$$

The corresponding momentum twistor matrix is given by

$$Z_{D_4} = \begin{pmatrix} 1 & 0 & 0 & x_2 & x_2 & 1 + x_2 + x_2 x_4 \\ 0 & 1 & 0 & -(1 + x_1) & -1 & -(1 + x_4) \\ 0 & 0 & 1 & 1 + x_1 + x_1 x_2 + x_1 x_2 x_3 & 1 & x_4 \end{pmatrix}. \quad (2.41)$$

Conversely, the cluster algebra generated by  $D_5$ ,

$$\begin{array}{c} & & & x_4 \\ & & \nearrow & \\ x_1 \longrightarrow x_2 \longrightarrow x_3 & & & \\ & & \searrow & \\ & & & x_5 \end{array} , \quad (2.42)$$

is not equivalent to the cluster algebra on (the top cell of) any Grassmannian. However, it appears as a subalgebra of any  $\text{Gr}(k, n)$  with rank greater than five. The  $D_4$  cluster algebra consists of 50 clusters, 16  $\mathcal{A}$ -coordinates, and 104  $\mathcal{X}$ -coordinates, while  $D_5$  has 182 clusters, 25  $\mathcal{A}$ -coordinates, and 260  $\mathcal{X}$ -coordinates. More generally, there are  $\frac{n}{3}(n-1)(n^2+4n-6)$   $\mathcal{X}$ -coordinates in cluster algebras of type  $D_n$  [53].

Finally, the cluster algebra  $E_6$  is generated by the quiver

$$\begin{array}{c} & & x_4 \\ & & \uparrow \\ x_1 \longrightarrow x_2 \longrightarrow x_3 & \longleftarrow & x_5 \longleftarrow x_6 \end{array} . \quad (2.43)$$

This cluster algebra is equivalent to  $\text{Gr}(4, 7)$ , as can be seen by mutating the seed (2.22) on nodes  $f_{12}, f_{13}, f_{23}, f_{12}, f_{22}$ , and then  $f_{32}$ . By comparison with (2.43), we then have

$$\begin{aligned} x_1 &= \frac{\langle 1234 \rangle \langle 1267 \rangle}{\langle 1237 \rangle \langle 1246 \rangle}, & x_2 &= -\frac{\langle 1247 \rangle \langle 3456 \rangle}{\langle 4(12)(35)(67) \rangle}, \\ x_3 &= \frac{\langle 1246 \rangle \langle 5(12)(34)(67) \rangle}{\langle 1245 \rangle \langle 1267 \rangle \langle 3456 \rangle}, & x_4 &= -\frac{\langle 4(12)(35)(67) \rangle}{\langle 1234 \rangle \langle 4567 \rangle}, \\ x_5 &= -\frac{\langle 1267 \rangle \langle 1345 \rangle \langle 4567 \rangle}{\langle 1567 \rangle \langle 4(12)(35)(67) \rangle}, & x_6 &= \frac{\langle 1567 \rangle \langle 2345 \rangle}{\langle 5(12)(34)(67) \rangle}, \end{aligned} \quad (2.44)$$

where we have made use of the notation

$$\langle a(bc)(de)(fg) \rangle \equiv \langle abde \rangle \langle acfg \rangle - \langle abfg \rangle \langle acde \rangle. \quad (2.45)$$

Any  $\mathcal{X}$ -coordinate on  $\text{Gr}(4, 7)$  can be expressed in terms of these  $\mathcal{X}$ -coordinates using the momentum twistor matrix

$$Z_{E_6} = \begin{pmatrix} x_3 x_5 & x_3 x_5 x_6 & x_1 & 0 & 1 & 0 & -1 \\ 1 + x_5 & 1 + x_6 + x_5 x_6 & -x_1 & 0 & -1 + x_4 & x_4 & x_2 x_4 \\ 0 & 0 & x_1 & 0 & 1 & 1 & x_2 \\ 0 & 0 & -1 & 1 & 1 + x_2 + x_2 x_3 & 1 & 0 \end{pmatrix}. \quad (2.46)$$

The cluster polytope of  $\text{Gr}(4, 7) \simeq E_6$  was displayed in figure 4; it contains 833 clusters, 42  $\mathcal{A}$ -coordinates, and 770  $\mathcal{X}$ -coordinates. The subalgebras of  $E_6$ , as well as those of its subalgebras, are tabulated in appendix A.

## 2.6 Cluster Automorphisms

Cluster algebras come equipped with an automorphism group that maps the set of cluster coordinates (but not necessarily the set of clusters) back to itself. We introduce here only what we need to elucidate the automorphisms of the cluster algebras introduced in the last section, and refer the interested reader to [54] for a more thorough mathematical introduction. Note that we describe automorphisms in terms of  $\mathcal{X}$ -coordinates, whereas [54] works in the  $\mathcal{A}$ -coordinate language.

The simplest example of a cluster automorphism is what we call a direct automorphism. Let  $\mathcal{A}$  be a cluster algebra equipped with a mutation rule  $\mu(x_i, \mathbf{X})$  that mutates the cluster  $\mathbf{X}$  on node  $x_i$ . Then, we can define:

**Direct Automorphism:** The map  $f : \mathcal{A} \rightarrow \mathcal{A}$  is a direct automorphism of  $\mathcal{A}$  if

- (i) for every cluster  $\mathbf{X}$  of  $\mathcal{A}$ ,  $f(\mathbf{X})$  is also a cluster of  $\mathcal{A}$ ,



(ii)  $f$  respects mutations, i.e.  $f(\mu(x_i, \mathbf{X})) = \mu(f(x_i), f(\mathbf{X}))$ .

An example of a direct automorphism on  $A_2$  is given by

$$\sigma_{A_2} : \mathcal{X}_i \rightarrow \mathcal{X}_{i+1}, \quad (2.47)$$

where we are using the coordinates introduced in (2.15). This automorphism cycles the five clusters  $1/\mathcal{X}_i \rightarrow \mathcal{X}_{i+1}$  amongst themselves. The action of this automorphism can also be recast as

$$\sigma_{A_2} : x_1 \rightarrow \frac{1}{x_2}, \quad x_2 \rightarrow x_1(1 + x_2), \quad (2.48)$$

using the  $\mathcal{X}$ -coordinates  $x_1$  and  $x_2$  that appear in (2.32). This is of course equivalent to the cyclic symmetry of the pentagon.

Cluster algebras are also endowed with what we call indirect automorphisms, which respect mutations but do not map the set of clusters back to itself. Instead, indirect automorphisms map the clusters in  $\mathcal{A}$  to clusters in  $\mathcal{A}'$ , where  $\mathcal{A}'$  is constructed from  $\mathcal{A}$  by multiplicatively inverting all cluster  $\mathcal{X}$ -coordinates and reversing the direction of all quiver arrows. Then we have:

**Indirect Automorphism:** The map  $f : \mathcal{A} \rightarrow \mathcal{A}'$  is an indirect automorphism if

- (i) for every cluster  $\mathbf{X}$  of  $\mathcal{A}$ ,  $f(\mathbf{X})$  is a cluster of  $\mathcal{A}'$
- (ii)  $f$  respects mutations, i.e.  $f(\mu(x_i, \mathbf{X})) = \mu(f(x_i), f(\mathbf{X}))$ .

$A_2$  is also equipped with an indirect automorphism generated by

$$\tau_{A_2} : \mathcal{X}_i \rightarrow \mathcal{X}_{6-i}, \quad (2.49)$$

where indices are understood to be mod 5. This can be recast in term of  $x_1$  and  $x_2$  as

$$\tau_{A_2} : x_1 \rightarrow \frac{x_1 x_2}{1 + x_1}, \quad x_2 \rightarrow \frac{1 + x_1 + x_1 x_2}{x_2}. \quad (2.50)$$

To see that this is an indirect automorphism, consider  $\tau_{A_2}(1/\mathcal{X}_1 \rightarrow \mathcal{X}_2) = 1/\mathcal{X}_5 \rightarrow \mathcal{X}_4$ . Inverting the cluster coordinates on the right hand side and reversing the arrow, we get back to  $\mathcal{X}_5 \leftarrow 1/\mathcal{X}_4$ , which was one of the original clusters of  $A_2$ . The operation generated by  $\tau_{A_2}$  can be interpreted as the dihedral flip of the pentagon.

It is useful to think of indirect automorphisms as generating a “mirror” or “flipped” version of the original cluster algebra, where the total collection of  $\mathcal{X}$ -coordinates is the same, but their Poisson bracket has flipped sign. The existence of this flip then can be seen as resulting from the arbitrary choice of overall sign for the exchange matrix  $b_{ij}$ ;

picking the other sign would have generated the same cluster-algebraic structure, but with different labels for the nodes. Indirect automorphisms capture the superficiality of this notation change.

The automorphisms  $\sigma_{A_2}$  and  $\tau_{A_2}$  generate the complete automorphism group for  $A_2$ , namely the dihedral group  $\mathfrak{D}_5$  (we denote the dihedral group in this font throughout so as not to confuse with the Dynkin diagram  $D_n$ ). More generally, cluster algebras of type  $A_n$  have as their automorphism group the dihedral group  $\mathfrak{D}_{n+3}$ , which is generated by a cyclic generator

$$\sigma_{A_n} : \quad x_{k < n} \rightarrow \frac{x_{k+1}(1 + x_{1,\dots,k-1})}{1 + x_{1,\dots,k+1}}, \quad x_n \rightarrow \frac{1 + x_{1,\dots,n-1}}{\prod_{i=1}^n x_i}, \quad (2.51)$$

of length  $n + 3$  and a flip generator

$$\tau_{A_n} : \quad x_1 \rightarrow \frac{1}{x_n}, \quad x_2 \rightarrow \frac{1}{x_{n-1}}, \quad \dots, \quad x_n \rightarrow \frac{1}{x_1} \quad (2.52)$$

of length 2. (Note that these definitions don't exactly match (2.48) or (2.50) when  $n$  is 2, but produce a pair of equally valid generators.) In  $\sigma_{A_n}$  we have introduced the notation

$$x_{i_1,\dots,i_k} \equiv \sum_{a=1}^k \prod_{b=1}^a x_{i_b} = x_{i_1} + x_{i_1}x_{i_2} + \dots + x_{i_1} \cdots x_{i_k}, \quad (2.53)$$

which we will also use below. The operator  $\sigma_{A_n}$  generates a direct automorphism while  $\tau_{A_n}$  generates an indirect automorphism.

The cluster algebra  $D_4$  has automorphism group  $\mathfrak{D}_4 \times S_3$ . Both  $\mathfrak{D}_4$  and  $S_3$  come with a cyclic (direct automorphism) generator,

$$\begin{aligned} \sigma_{D_4}^{(\mathfrak{D}_4)} : \quad x_1 &\rightarrow \frac{x_2}{1 + x_{1,2}}, \quad x_2 \rightarrow \frac{(1 + x_1)x_1x_2x_3x_4}{(1 + x_{1,2,3})(1 + x_{1,2,4})}, \\ x_3 &\rightarrow \frac{1 + x_{1,2}}{x_1x_2x_3}, \quad x_4 \rightarrow \frac{1 + x_{1,2}}{x_1x_2x_4}, \end{aligned} \quad (2.54)$$

$$\sigma_{D_4}^{(S_3)} : \quad x_1 \rightarrow \frac{1}{x_3}, \quad x_2 \rightarrow \frac{x_1x_2(1 + x_3)}{1 + x_1}, \quad x_3 \rightarrow x_4, \quad x_4 \rightarrow \frac{1}{x_1},$$

where  $\sigma_{D_4}^{(\mathfrak{D}_4)}$  has length four, and  $\sigma_{D_4}^{(S_3)}$  has length three. Then there are two flip generators

$$\begin{aligned} \tau_{D_4}^{(\mathfrak{D}_4)} : \quad x_2 &\rightarrow \frac{1 + x_1}{x_1x_2(1 + x_3)(1 + x_4)}, \\ \tau_{D_4}^{(S_3)} : \quad x_3 &\rightarrow x_4, \quad x_4 \rightarrow x_3, \end{aligned} \quad (2.55)$$

where  $\tau_{D_4}^{(\mathfrak{D}_4)}$  generates an indirect automorphism, and  $\tau_{D_4}^{(S_3)}$  generates a direct automorphism.

The cluster algebras on  $D_{n>4}$  with defining quiver

$$x_1 \Rightarrow x_2 \Rightarrow \cdots \Rightarrow x_{n-2} \begin{matrix} \nearrow x_{n-1} \\ \searrow x_n \end{matrix}, \quad (2.56)$$

have the automorphism group  $\mathfrak{D}_n \times \mathbb{Z}_2$ , with generators  $\sigma_{D_n}$  (length  $n$ , direct),  $\tau_{D_n}$  (length 2, indirect), and  $\mathbb{Z}_{2,D_n}$  (length 2, direct). In the case of  $D_5$ , these generators can be chosen to be

$$\begin{aligned} \sigma_{D_5} : \quad & x_1 \rightarrow \frac{x_2}{1+x_{1,2}}, \quad x_2 \rightarrow \frac{(1+x_1)x_3}{1+x_{1,2,3}}, \quad x_3 \rightarrow \frac{x_1x_2x_3x_4x_5(1+x_{1,2})}{(1+x_{1,2,3,4})(1+x_{1,2,3,5})}, \\ & x_4 \rightarrow \frac{1+x_{1,2,3}}{x_1x_2x_3x_4}, \quad x_5 \rightarrow \frac{1+x_{1,2,3}}{x_1x_2x_3x_5}, \\ \tau_{D_5} : \quad & x_2 \rightarrow \frac{1+x_1}{x_1x_2(1+x_3x_5+x_{3,4,5})}, \quad x_3 \rightarrow \frac{x_3x_4x_5}{(1+x_{3,4})(1+x_{3,5})}, \\ & x_4 \rightarrow \frac{1+x_3x_5+x_{3,4,5}}{x_4}, \quad x_5 \rightarrow \frac{1+x_3x_5+x_{3,4,5}}{x_5}, \end{aligned} \quad (2.57)$$

$$\mathbb{Z}_{2,D_n} : \quad x_4 \rightarrow x_5, \quad x_5 \rightarrow x_4.$$

More generally, for  $D_n$  cluster algebras, the action of  $\mathbb{Z}_2$  is always realized by the exchange  $x_{n-1} \leftrightarrow x_n$ .

Finally, the automorphism group of  $E_6 \simeq \text{Gr}(4, 7)$  is the dihedral group  $\mathfrak{D}_{14}$ . This group has generators  $\sigma_{E_6}$  (length 7, direct),  $\tau_{E_6}$  (length 2, indirect), and  $\mathbb{Z}_{2,E_6}$  (length 2, direct). In the coordinates of the quiver (2.43), these can be chosen to be

$$\begin{aligned} \sigma_{E_6} : \quad & x_1 \rightarrow \frac{1}{x_6(1+x_{5,3,4})}, \quad x_2 \rightarrow \frac{1+x_{6,5,3,4}}{x_5(1+x_{3,4})}, \quad x_3 \rightarrow \frac{(1+x_{2,3,4})(1+x_{5,3,4})}{x_3(1+x_4)}, \\ & x_4 \rightarrow \frac{1+x_{3,4}}{x_4}, \quad x_5 \rightarrow \frac{1+x_{1,2,3,4}}{x_2(1+x_{3,4})}, \quad x_6 \rightarrow \frac{1}{x_1(1+x_{2,3,4})}, \\ \tau_{E_6} : \quad & x_1 \rightarrow \frac{x_5}{1+x_{6,5}}, \quad x_2 \rightarrow (1+x_5)x_6, \quad x_3 \rightarrow \frac{(1+x_{1,2})(1+x_{6,5})}{x_1x_2x_3x_5x_6(1+x_4)}, \\ & x_5 \rightarrow x_1(1+x_2), \quad x_6 \rightarrow \frac{x_2}{1+x_{1,2}}, \end{aligned} \quad (2.58)$$

$$\mathbb{Z}_{2,E_6} : \quad x_1 \rightarrow x_6, \quad x_2 \rightarrow x_5, \quad x_5 \rightarrow x_2, \quad x_6 \rightarrow x_1.$$

In the language of  $\text{Gr}(4, 7)$ , these generators correspond to cycling momentum twistor indices  $Z_i \rightarrow Z_{i+1}$ , flipping momentum twistor indices  $Z_i \rightarrow Z_{8-i}$ , and parity conjugation.

The polylogarithmic part of the  $n$ -particle MHV amplitude is invariant under parity transformations, as well as the dihedral group that represents Bose symmetry. These symmetries directly translate to automorphisms of the cluster algebra on  $\text{Gr}(4, n)$ . However, it turns out the nonclassical part of these amplitudes can also be decomposed into building blocks that respect the automorphism group of certain subalgebras of  $\text{Gr}(4, n)$ . Making this statement precise will be the focus of much of the remainder of this paper.

### 3 Cluster Polylogarithms and MHV Amplitudes

The BDS ansatz captures the infrared structure of planar  $\mathcal{N} = 4$  sYM to all orders in the coupling [42]. In four- and five-particle kinematics it also furnishes the complete finite part of the amplitude, while for six or more particles it must be corrected by a finite dual-conformally invariant function [43–45]. In the case of the MHV amplitude, this correction is often computed in the form of the  $n$ -particle remainder function  $R_n$ , defined by

$$\mathcal{A}_n^{\text{MHV}} = \mathcal{A}_n^{\text{BDS}} \times \exp(R_n), \quad (3.1)$$

where  $\mathcal{A}_n^{\text{BDS}}$  is the BDS ansatz for  $n$  particles [42]. Like the amplitude, the remainder function can be expanded in the coupling

$$R_n = g^4 R_n^{(2)} + g^6 R_n^{(3)} + g^8 R_n^{(4)} + \dots, \quad (3.2)$$

where  $g^2 = \frac{g_{\text{YM}}^2 N_c}{16\pi^2}$ . In this expansion we have used the fact that  $R_n^{(1)} = 0$ , since the BDS ansatz encodes the complete one loop MHV amplitude at all  $n$ .

The remaining  $L$ -loop contributions to the remainder function are expected to be expressible in terms of generalized polylogarithms [55–57] of uniform transcendental weight  $2L$ . This space is spanned by (products of) the functions

$$G(a_1, \dots, a_k; z) \equiv \int_0^z \frac{dt}{t - a_1} G(a_2, \dots, a_k; t), \quad G(\underbrace{0, \dots, 0}_k; z) \equiv \frac{\log^k z}{k!}, \quad (3.3)$$

where  $G(; z) \equiv 1$ , and the transcendental weight of each function corresponds to its number of indices  $k$ . In particular, the remainder function is expected to be a pure function of this type, meaning that its kinematic dependence appears in the indices and arguments  $a_i$  and  $z$ , but not in the rational prefactors multiplying these functions. This

is known to be true at two loops, due to an impressive all- $n$  computation that leveraged the superconformal symmetry of this theory [9], as well as through six loops in six-particle kinematics [4, 5, 7, 58] and through four loops in seven-particle kinematics [6, 8].

In addition to being generalized polylogarithms, loop-level contributions to the remainder function exhibit a great deal of cluster-algebraic structure. In particular, they are members of the space of ‘cluster polylogarithms’ studied in [12], indicating that their symbol is naturally expressible in terms of cluster  $\mathcal{A}$ -coordinates, while their Lie cobracket is naturally expressible in terms of cluster  $\mathcal{X}$ -coordinates (in a way that will be made precise below). Their cobracket, moreover, has been shown to be decomposable into simple functions associated with their  $A_2$  and  $A_3$  subalgebras [12]. As will be shown in the next section, these functions (the ‘ $A_2$  function’ and the ‘ $A_3$  function’) are invariant under the automorphism group of the algebras on which they are defined, up to an overall sign. This encodes the fact that these functions are well-defined under coordinate relabelings (or, are well-defined functions of oriented graphs). We correspondingly propose that the space of cluster polylogarithms be refined to include only functions that respect the automorphism group of the cluster algebra on which they are defined. In addition to this subalgebra structure, the symbols of these amplitudes have been found to satisfy a ‘cluster adjacency’ principle [14], and their cobracket takes a similarly restricted form [10]. The rest of this section is devoted to making these properties precise, for which purpose we first describe the motivic structure of polylogarithms.

### 3.1 The Symbol and Cobracket

The space of generalized polylogarithms defined by (3.3) is colossally overcomplete. This is because  $a_i$  and  $z$  are allowed to be arbitrarily complicated algebraic functions, and because these polylogarithms satisfy a shuffle and stuffle algebra. The shuffle algebra represents the fact that unordered integrations can be triangulated into a sum over iterated integrals [59, 60]. In general, this means that when two polylogarithms share an argument  $z$ , their product can be re-expressed as the sum of functions

$$G(a_1, \dots, a_{k_1}; z) G(a_{k_1+1}, \dots, a_{k_1+k_2}; z) = \sum_{\sigma \in \Sigma(k_1, k_2)} G(a_{\sigma(1)}, \dots, a_{\sigma(k_1+k_2)}; z), \quad (3.4)$$

where  $\Sigma(k_1, k_2)$  denotes the set of all shuffles between the sets of integers  $\{1, \dots, k_1\}$  and  $\{k_1 + 1, \dots, k_1 + k_2\}$  (that is, all ways of interleaving these two sets such that the ordering of the elements within each of the original sets is maintained). The stuffle algebra naturally arises when generalized polylogarithms are re-expressed as infinite

sums,

$$\begin{aligned} \text{Li}_{n_1, \dots, n_d}(z_1, \dots, z_d) &\equiv \sum_{0 < m_1 < \dots < m_d} \frac{z_1^{m_1} \dots z_d^{m_d}}{m_1^{n_1} \dots m_d^{n_d}} \\ &= (-1)^d G(\underbrace{0, \dots, 0}_{n_d-1}, \frac{1}{z_d}, \dots, \underbrace{0, \dots, 0}_{n_1-1}, \frac{1}{z_1 \dots z_d}; 1) \end{aligned} \quad (3.5)$$

where  $d$  is called the depth of the polylogarithm. Stuffle identities represent the freedom to split up unordered summation indices (arising from products of polylogarithms) into nested sums where these indices are ordered, as in (3.5).

This overcompleteness gives rise to a rich space of identities. This can already be seen at the level of classical polylogarithms, which correspond to the instances of (3.5) with depth one,

$$\text{Li}_k(z) = -G(\underbrace{0, \dots, 0}_{k-1}, 1; z). \quad (3.6)$$

For instance, classical polylogarithms satisfy Abel's identity, which itself can be expressed in the language of the  $A_2$  cluster algebra. Using the definition of the  $\mathcal{X}$ -coordinates  $\mathcal{X}_i$  given in eq. (2.16), we have

$$\sum_{i=1}^5 \text{Li}_2(-\mathcal{X}_i) + \log \mathcal{X}_i \log \mathcal{X}_{i+1} = -\frac{\pi^2}{2}. \quad (3.7)$$

Additional cluster-algebraic identities, including an identity involving  $\text{Li}_3$  evaluated on the cluster  $\mathcal{X}$ -coordinates of  $D_4$ , are discussed in [2, 61].

Fortunately, *all* identities between polylogarithms are (conjecturally) generated by shuffle and stuffle relations [62, 63]. These relations are trivialized (up to logarithmic identities) by the symbol map, which can be defined in terms of total derivatives. For instance, taking the total derivative of (3.3), we have

$$dG(a_1, \dots, a_k; z) = \sum_{i=1}^k G(a_1, \dots, \hat{a}_i, \dots, a_k; z) d \log \left( \frac{a_{k-i+1} - a_{k-i}}{a_{k-i+1} - a_{k-i+2}} \right), \quad (3.8)$$

where  $a_0 \equiv z$  and  $a_{k+1} \equiv 0$ , and the notation  $\hat{a}_i$  indicates this index should be omitted [64]. The symbol map is then defined recursively by [65]

$$\mathcal{S}(G(a_1, \dots, a_k; z)) \equiv \sum_{i=1}^k \mathcal{S}(G(a_1, \dots, \hat{a}_i, \dots, a_k; z)) \otimes \left( \frac{a_{k-i+1} - a_{k-i}}{a_{k-i+1} - a_{k-i+2}} \right). \quad (3.9)$$

The entries of the resulting  $k$ -fold tensor product are referred to as symbol letters. These symbol letters inherit the distributive properties of (arguments of) logarithms,

and can therefore be expanded into a multiplicatively independent basis of symbol letters (the ‘symbol alphabet’). Complicated polylogarithmic identities are thereby reduced to identities between logarithms, at the cost of losing information about the boundary of integration in (3.3).

The symbol also captures the analytic structure of polylogarithms, insofar as it encodes their (iterated) discontinuity structure. Namely, for generic indices  $a_i$  and argument  $z$ , these functions have nonzero monodromy only where the letters in the first entry of their symbol vanish or become infinite. These monodromies can themselves have branch cuts that the original function did not have, when new symbol letters appear in the second entry of the symbol (and similarly for iterated monodromies, when new symbol letters appear at higher weights). For special values of  $a_i$  and  $z$ , some of this information is lost due to the fact that higher-weight transcendental constants (such as Riemann  $\zeta$  values) are in the kernel of the symbol [57]. However, this information is retained by the coaction [66, 67], and can be recovered by specifying an integration constant at each weight. We defer further consideration of these transcendental constants to a follow-up paper [37].

In addition to the symbol, polylogarithms come equipped with a Lie cobracket structure [2]. The cobracket  $\delta$  can be calculated using the symbol projection operator

$$\rho(s_1 \otimes \cdots \otimes s_k) = \frac{k-1}{k} \left( \rho(s_1 \otimes \cdots \otimes s_{k-1}) \otimes s_k - \rho(s_2 \otimes \cdots \otimes s_k) \otimes s_1 \right), \quad (3.10)$$

where  $\rho(s_1) \equiv s_1$ . This projects onto the component of a symbol that cannot be written as a product of lower-weight polylogarithms (for instance, via the shuffle relations (3.4)). The action of  $\rho$  can be lifted to a projection on functions, up to terms proportional to transcendental constants; since we will not be concerned with these terms in what follows, we will abuse notation by applying  $\rho$  to functions directly. The cobracket  $\delta$  of a weight  $k$  polylogarithm  $f$  can then be calculated as

$$\delta(f) \equiv \sum_{i=1}^{k-1} (\rho_i \wedge \rho_{k-i}) \rho(f). \quad (3.11)$$

This notation indicates that the projection operator  $\rho$  is first applied to  $f$ , after which each term in the resulting sum is partitioned into a wedge product of weight  $i$  and weight  $k-i$  functions (either by splitting up the symbol into its first  $i$  and last  $k-i$  entries, or by taking the ‘ $i, k-i$  component’ of the coproduct); the projection operator is then applied to the entries in this wedge product separately. In general, the wedge product in (3.11) involves spaces of different weight. Without loss of generality, we can put the cobracket of any function into a form where the first factor of the wedge product has weight equal to or higher than that of the second factor (exchanging the order of

factors when needed, at the cost of a minus sign). We denote by  $\delta_{i,j}(f)$  the component of the cobracket of  $f$  that involves a wedge product of weight  $i$  and  $j$  functions, in that order—but we emphasize that this includes contributions from all terms in (3.11) that involve these weights in either order.

In the context of two-loop amplitudes, the salient property of the cobracket is that it isolates the component of weight four polylogarithms that cannot be written in terms of classical polylogarithms. Under the action of  $\rho$ , classical polylogarithms are mapped to elements of the Bloch group  $B_k$  [68, 69], namely the algebra of polylogarithms modulo identities between classical polylogarithms. Following [2], we denote these elements by

$$\{z\}_k \equiv \rho(-\text{Li}_k(-z)) \in B_k, \quad k > 1, \quad (3.12)$$

$$\{z\}_1 \equiv \rho(\log(z)) \in B_1. \quad (3.13)$$

For instance, Abel’s identity, eq. (3.7), can be expressed (and easily checked) when written in terms of Bloch group elements:

$$\sum_{i=1}^5 \{\mathcal{X}_i\}_2 = 0. \quad (3.14)$$

In this language, the action of the cobracket on classical polylogarithms is given by

$$\delta(\text{Li}_k(-z)) = -\{z\}_{k-1} \wedge \{z\}_1, \quad k > 2, \quad (3.15)$$

$$\delta(\text{Li}_2(-z)) = -\{1+z\}_1 \wedge \{z\}_1. \quad (3.16)$$

Note that, for the first time at weight four, there exists a component of the cobracket that is not mapped to by classical polylogarithms—namely,  $\delta_{2,2}(\text{Li}_4(-z)) = 0$ . This is not true of weight four polylogarithms in general, and in particular  $\delta_{2,2}(R_n^{(2)})$  is nonzero for  $n \geq 7$ . However, it has been shown that any weight four function that is annihilated by  $\delta_{2,2}$  can be written in terms of classical polylogarithms (with potentially complicated arguments) [61, 70–73]. The converse is worth stating as well: any function with non-zero  $\delta_{2,2}$  must involve a nonclassical polylogarithm of weight 4.

Lastly, it is often quite useful to employ the fact that the trivial cohomology of  $\delta$  gives us an integrability condition,

$$\delta^2(f) = 0, \quad (3.17)$$

for any polylogarithm  $f$ . This condition implies an intricate relationship between the arguments of the Bloch group elements appearing in  $\delta(f)$ . For example, at weight 4 the relation (3.17) translates to

$$\delta(\delta_{2,2}(f)) + \delta(\delta_{3,1}(f)) = 0, \quad (3.18)$$



where

$$\delta(\{x\}_3 \wedge \{y\}_1) = \{x\}_2 \wedge \{x\}_1 \wedge \{y\}_1 \quad (3.19)$$

$$\begin{aligned} \delta(\{x\}_2 \wedge \{y\}_2) &= \{y\}_2 \wedge \{1+x\}_1 \wedge \{x\}_1 \\ &\quad - \{x\}_2 \wedge \{1+y\}_1 \wedge \{y\}_1. \end{aligned} \quad (3.20)$$

So, while the  $\delta_{2,2}$  component alone captures the nonclassical contribution to a function, this nonclassical contribution is linked to the classical contribution through its relation with the  $\delta_{3,1}$  component. Cobracket integrability is discussed in much greater detail in [12].

The symbol and cobracket naturally stratify the study of two-loop amplitudes and integrals that can be expressed as polylogarithms. The operator  $\delta_{2,2}$  isolates the nonclassical component of these functions, while the symbol captures their analytic structure up to terms proportional to transcendental constants. In the case of MHV amplitudes in planar  $\mathcal{N} = 4$  sYM theory, both objects turn out to distill intriguing cluster-algebraic structure that would otherwise be hard to see at the level of full functions. It is to this structure that we now turn.

### 3.2 Cluster-Algebraic Structure at Two Loops

As outlined in the introduction, the two-loop MHV amplitudes of this theory exhibit different forms of cluster-algebraic structure at the level of their cobracket, their symbol, and as full functions [2, 10, 12, 13]. The first facet of this structure concerns the building blocks that appear at each level, which are found to lie within restricted classes:

- The cobracket of  $R_n^{(2)}$  can be written in terms of elements of the Bloch group taking the form  $\{x_i\}_k$ , where  $x_i$  is an  $\mathcal{X}$ -coordinate on  $\text{Gr}(4, n)$
- The symbol of  $R_n^{(2)}$  can be expressed in terms of symbol letters drawn from the set of  $\mathcal{A}$ -coordinates on  $\text{Gr}(4, n)$
- The function  $R_n^{(2)}$  can be expressed entirely in terms of (products of) polylogarithms taking the form  $\text{Li}_{n_1, \dots, n_d}(-x_i, \dots, -x_j)$ , where each  $x_p$  is again an  $\mathcal{X}$ -coordinate on  $\text{Gr}(4, n)$

The physical meaning of the symbol alphabet restriction is clear, if unilluminating—the kinematic configurations in which these amplitudes are singular (due to internal propagators going on-shell in the Feynman diagram expansion) coincide with the vanishing loci of certain cluster  $\mathcal{A}$ -coordinates. Even so, it is not clear how (or if) this restriction follows from physical principles. In this respect, the positive Grassmannian formulation of this theory is suggestive, insofar as the integrands of these amplitudes

are seen to develop physical singularities where certain  $\mathcal{A}$ -coordinates vanish or become infinite [1]. However, this does not preclude the emergence of new singularities during integration that are not rationally expressible in terms of  $\mathcal{A}$ -coordinates, as has already been explicitly observed in Feynman integrals contributing to non-MHV amplitudes in this theory [21, 22, 24]. The physical meaning of the restricted functional form and cobracket structure exhibited by these amplitudes remains even more obscure.

In the polylogarithmic amplitudes where cluster coordinates do prove sufficient as a symbol alphabet (in particular, in all two-loop MHV amplitudes), cluster algebras also seem to play a role in how these building blocks are assembled. In particular, it was recently observed that—when these amplitudes are normalized appropriately—the only cluster  $\mathcal{A}$ -coordinates that appear in adjacent entries of their symbol are those that appear together in at least one cluster of  $\text{Gr}(4, n)$  [14]. This ‘cluster adjacency’ principle is not enjoyed by the remainder function (3.1), but by BDS-like normalized amplitudes  $\mathcal{E}_n$  [7, 8, 74, 75]. These are defined by

$$\mathcal{A}_n^{\text{MHV}} = \mathcal{A}_n^{\text{BDS-like}} \times \mathcal{E}_n, \quad (3.21)$$

where  $\mathcal{A}_n^{\text{BDS-like}}$  is related to  $\mathcal{A}_n^{\text{BDS}}$  by the cusp anomalous dimension [76]

$$\Gamma_{\text{cusp}} = 4g^2 - 8\zeta_2 g^4 + \mathcal{O}(g^6) \quad (3.22)$$

and a simple weight two polylogarithm  $Y_n$  via

$$\mathcal{A}_n^{\text{BDS-like}} = \mathcal{A}_n^{\text{BDS}} \times \exp\left(\frac{\Gamma_{\text{cusp}}}{4} Y_n\right). \quad (3.23)$$

The function  $Y_n$  corresponds to the part of the one-loop MHV amplitude that depends on three- and higher-particle Mandelstam invariants, where these invariants have been assembled into dual-conformally-invariant cross ratios (with the help of two-particle invariants). The BDS-like ansatz that remains only depends on two-particle invariants, yet accounts for the full infrared structure of these amplitudes.

The motivation for switching to the BDS-like normalization is precisely this restricted kinematic dependence. Since the BDS-like ansatz depends only on two-particle invariants, the functions  $\mathcal{E}_n$  directly inherit the Steinmann relations between three- and higher-particle invariants that are obeyed by the full amplitude [7, 8, 77–79]. These relations tell us that

$$\left. \begin{aligned} \text{Disc}_{s_j, \dots, j+p+q} [\text{Disc}_{s_i, \dots, i+p}(\mathcal{E}_n)] &= 0, \\ \text{Disc}_{s_i, \dots, i+p} [\text{Disc}_{s_j, \dots, j+p+q}(\mathcal{E}_n)] &= 0, \end{aligned} \right\} \quad 0 < j-i \leq p \text{ or } q < i-j \leq p+q, \quad (3.24)$$

for positive  $p$  and nonnegative  $q$ . Formulated in terms of symbol entries, this implies that the cluster  $\mathcal{A}$ -coordinates  $\langle j-1, j, j+p+q-1, j+p+q \rangle$  and  $\langle i-1, i, i+p+q-1, i+p+q \rangle$  cannot appear in adjacent entries of the symbol.

$p-1, i+p\rangle$  never appear next to each other in the first two entries of the symbol (when the conditions in (3.24) are met). In fact, it is believed that these constraints can be applied at all depths in the symbol, as these letters are never seen to appear next to each other [7, 8, 23, 58]. These generalized constraints have been termed the extended Steinmann relations, as they amount to applying the relations (3.24) to all discontinuities of the amplitude in addition to the amplitude itself.

The fact that the functions  $\mathcal{E}_n$  also obey the cluster adjacency principle in all known cases is far more surprising. The constraints that follow from this principle take a form similar to the extended Steinmann relations (insofar as they restrict which symbol letters can appear in adjacent entries), and in fact turn out to be equivalent in six-particle kinematics when applied to functions with physical branch cuts. It is therefore tempting to believe cluster adjacency follows from some set of physical principles that includes the extended Steinmann relations. However, it is not yet known whether these two conditions are equivalent at all  $n$ , even in polylogarithmic cases where algebraic roots are absent. (Once functions more complicated than polylogarithms (such as elliptic polylogarithms [80, 81]) start to appear, it is not even known whether anything resembling cluster adjacency can be formulated.)

Seeming to complicate this question of equivalence is the fact that the BDS-like ansatz (3.23) only exists when  $n$  is not a multiple of four. This is because no function satisfying the above description of  $Y_n$  exists for these particle multiplicities [8, 82]. (Such a function not only exists for all other  $n$ , but is uniquely picked out by this description.) Stated another way, when  $n$  is a multiple of four, any normalization that absorbs the infrared-divergent part of the amplitude either depends on some set of three- or higher-particle Mandelstam invariants, or spoils the dual conformal invariance of the resulting normalized amplitude. However, this shows this complication is superficial—for the purpose of understanding the relationship between the extended Steinmann relations and the cluster adjacency principle, we need merely choose the latter horn of this dilemma, and give up dual conformal invariance.

This issue first arises in eight-particle kinematics. There it can be seen—by direct computation—that both the extended Steinmann and cluster adjacency conditions are obeyed when the amplitude is normalized by a ‘generalized BDS-like ansatz’ whose kinematic dependence is restricted to two-particle Mandelstam invariants [37]. (It is always possible to formulate such a normalization, as only two-particle Mandelstams appear in the infrared-divergent terms of the one-loop amplitude [83].) This provides further evidence that the cluster adjacency principle and the extended Steinmann relations (when combined with physical branch cuts) are equivalent, in the cases where cluster adjacency can be unambiguously applied.

While cluster adjacency was first observed in  $\mathcal{A}$ -coordinates, it can also be for-

mulated in terms of  $\mathcal{X}$ -coordinates. Unlike  $\mathcal{A}$ -coordinates, which are multiplicatively independent,  $\mathcal{X}$ -coordinates satisfy numerous multiplicative identities. Thus, there will exist many representations of the same symbol in terms of  $\mathcal{X}$ -coordinates, whereas its representation in terms of  $\mathcal{A}$ -coordinates is unique. In general, only a subset of these  $\mathcal{X}$ -coordinate representations will satisfy cluster adjacency (even if the  $\mathcal{A}$ -coordinate representation does). Moreover, while  $\mathcal{X}$ -coordinate adjacency trivially implies  $\mathcal{A}$ -coordinate adjacency by the relation (2.10), the converse can not be true in general. This is because  $\mathcal{A}$ -coordinates can express a larger class of functions than  $\mathcal{X}$ -coordinates, insofar as the latter necessarily respect dual conformal symmetry while the former do not. However, it is easy to check in simple symbol alphabets, for example  $\text{Gr}(4, 6)$ , that dual conformal invariance and  $\mathcal{A}$ -coordinate adjacency together imply the existence of at least one  $\mathcal{X}$ -coordinate adjacent representation. We conjecture that this remains true for all dual-conformally-invariant symbols constructed on  $\text{Gr}(k, n)$ , for general  $k$  and  $n$ .

The cobracket  $\delta(R_n^{(2)})$  also satisfies its own form of cluster adjacency [13]. Namely,  $\delta(R_n^{(2)})$  can be expressed as a linear combination of terms  $\{x_i\}_2 \wedge \{x_j\}_2$  and  $\{x_k\}_3 \wedge \{x_l\}_1$  where  $x_i$  and  $x_j$  appear together in a cluster of  $\text{Gr}(4, n)$ , and similarly for  $x_k$  and  $x_l$ . For example,  $\delta_{2,2}(R_n^{(2)})$  can be expressed as a sum of terms of the form  $\{v_{ijk}\}_2 \wedge \{z_{pqr}^\pm\}_2$ , where (using the notation defined in (2.45))

$$v_{ijk} = -\frac{\langle i+1(i, i+2)(j, j+1)(k, k+1) \rangle}{\langle i, i+1, k, k+1 \rangle \langle i+1, i+2, j, j+1 \rangle}, \quad (3.25)$$

and

$$\begin{aligned} z_{ijk}^+ &= \frac{\langle i, j-1, j, j+1 \rangle \langle i+1, k-1, k, k+1 \rangle - \langle i+1, j-1, j, j+1 \rangle \langle i, k-1, k, k+1 \rangle}{\langle i, k-1, k, k+1 \rangle \langle i+1, j-1, j, j+1 \rangle}, \\ z_{ijk}^- &= \frac{\langle i, i+1, j, k \rangle \langle i-1, i, i+1, i+2 \rangle}{\langle i-1, i, i+1, k \rangle \langle i, i+1, i+2, j \rangle}. \end{aligned} \quad (3.26)$$

As long as  $i < j < k$  (considered mod  $n$ ), the quantities  $v_{ijk}$ ,  $z_{ijk}^+$ , and  $z_{ijk}^-$  each constitute  $\mathcal{X}$ -coordinates, where  $z_{ijk}^+$  and  $z_{ijk}^-$  are also parity conjugate to each other. The coordinates  $v_{ijk}$  were originally motivated by consideration of the location of physical branch cuts, while the  $z_{ijk}^\pm$  were motivated by consideration of the final symbol entry of the remainder function, which is constrained to take the form  $\langle i-1, i, i+1, j \rangle$  by dual superconformal symmetry [15]. However, for the present discussion, the pertinent point is that  $\delta_{2,2}(R_n^{(2)})$  can be expressed as a sum of terms  $\{v_{ijk}\}_2 \wedge \{z_{pqr}^\pm\}_2$  in which  $v_{ijk}$  and  $z_{pqr}^\pm$  also occur together in a cluster. Importantly, this ‘cobracket-level cluster adjacency’ isn’t implied by the cluster  $\mathcal{A}$ -coordinate adjacency observed in [14], as can be seen in the  $A_2$  function we will define in the next subsection—the symbol of

this function satisfies cluster  $\mathcal{A}$ -coordinate adjacency, but the  $\mathcal{X}$ -coordinates appearing in its  $\delta_{2,2}$  cobracket component cannot be chosen to satisfy cobracket-level cluster adjacency.

While this cobracket-level structure was originally observed in the remainder function, the same characterization carries over to the BDS-like normalized amplitude. This can be seen by combining equations (3.1), (3.21), and (3.23) to express  $\mathcal{E}_n$  in terms of  $R_n$ :

$$\begin{aligned}\mathcal{E}_n &= \exp\left(R_n - \frac{\Gamma_{\text{cusp}}}{4}Y_n\right) \\ &= 1 - Y_n g^2 + \left(R_n^{(2)} + 2\zeta_2 Y_n + \frac{1}{2}Y_n^2\right)g^4 + \mathcal{O}(g^6).\end{aligned}\tag{3.27}$$

Thus,  $\rho(\mathcal{E}_n^{(2)}) = \rho(R_n^{(2)})$ , since these two functions differ only by products and terms involving transcendental constants, and the remainder function and the BDS-like normalized amplitude have identical cobrackets. At both the level of its cobracket and symbol,  $\mathcal{E}_n^{(2)}$  can therefore be expressed in terms of the same restricted building blocks as  $R_n^{(2)}$ . But in the case of  $\mathcal{E}_n^{(2)}$  we can now add:

- The cobracket  $\delta(\mathcal{E}_n^{(2)})$  can be expressed as a linear combination of terms  $\{x_i\}_2 \wedge \{x_j\}_2$  and  $\{x_k\}_3 \wedge \{x_l\}_1$  where  $x_i$  and  $x_j$  appear together in a cluster of  $\text{Gr}(4, n)$ , and similarly for  $x_k$  and  $x_l$ .
- Pairs of  $\mathcal{A}$ -coordinates only appear in adjacent entries of the symbol  $\mathcal{S}(\mathcal{E}_n^{(2)})$  when they also appear together in at least one cluster of  $\text{Gr}(4, n)$ .

As discussed above, the last statement can also be applied at  $n$  that are multiples of four by going to a generalized BDS-like normalization in which only two-particle invariants appear in the normalizing function.

As it turns out, there exists yet more structure in  $\delta_{2,2}(R_n^{(2)}) = \delta_{2,2}(\mathcal{E}_n^{(2)})$ . In particular, it was shown in [12] that this cobracket component can be decomposed into a sum over various  $A_2$  subalgebras of  $\text{Gr}(4, n)$ , by defining an  $A_2$  function that can be evaluated on each of these subalgebras. Moreover, this  $A_2$  function can be assembled into an  $A_3$  function, in terms of which this cobracket component can similarly be decomposed. As this subalgebra decomposability will play a central role in what follows, we devote the next subsection to its description.

### 3.3 Subalgebra Structure and Cluster Polylogarithms

As we saw in section 2.3, cluster algebras are endowed with subalgebras that can be generated by mutating on restricted sets of nodes. This motivates looking for physically

relevant cluster polylogarithms on algebras other than  $\text{Gr}(4, n)$ , when these algebras appear as subalgebras of the latter. Before we do so, let us return to the definition of these objects, which we are now in a position to make precise. Following [12], we define cluster polylogarithms (at least through weight four) to have the following properties:

**Cluster Polylogarithm:** A generalized polylogarithm  $f$  is a cluster polylogarithm on a cluster algebra  $\mathcal{A}$  if

- (i) the symbol alphabet of  $f$  is composed of only  $\mathcal{A}$ -coordinates on  $\mathcal{A}$ ,
- (ii) the cobracket of  $f$  can be expressed in terms of Bloch group elements  $\{x_i\}_k$ , where  $x_i$  is an  $\mathcal{X}$ -coordinate of  $\mathcal{A}$ ,
- (iii) the function  $f$  is invariant under the automorphisms of  $\mathcal{A}$ , up to a sign.

Property (iii) can be thought of as the requirement that cluster polylogarithms be well-defined functions on the cluster algebra, or more specifically of the oriented graph representing that cluster algebra. For instance, if we wish to define a function on the  $A_2$  cluster algebra, as we shall do shortly, it should satisfy the property that  $f_{A_2}(1/\mathcal{X}_1 \rightarrow \mathcal{X}_2) = \pm f_{A_2}(1/\mathcal{X}_2 \rightarrow \mathcal{X}_3)$ , and similarly when the function is evaluated on the other automorphic images of this cluster. The ambiguity in the overall sign, which will be discussed below, reflects the fact that some automorphisms flip the orientation of cluster algebras while others do not.

The first nontrivial cluster polylogarithm is found on the rank-two cluster algebra associated with  $A_2$ . Given that  $A_2$  subalgebras are generated by any pair of connected nodes, they appear ubiquitously in  $\text{Gr}(4, n)$ . A central fact about the ‘ $A_2$  function’ is that it is uniquely determined by the cluster polylogarithm conditions, up to products of classical polylogarithms with weight  $\leq 3$  (see [12] for an in-depth discussion of the  $A_2$  function). However, its analytic properties can be tuned by adding and subtracting these products of lower-weight polylogarithms (which should only be done in a way that respects the automorphisms of  $A_2$ ). We choose to define the  $A_2$  function (which can be thought of as a ‘function on an oriented graph’) as

$$\begin{aligned}
 f_{A_2}(x_1 \rightarrow x_2) = \sum_{\text{skew-dihedral}} & \left[ \text{Li}_{2,2} \left( -\frac{1}{\mathcal{X}_{i-1}}, -\frac{1}{\mathcal{X}_{i+1}} \right) - \text{Li}_{1,3} \left( -\frac{1}{\mathcal{X}_{i-1}}, -\frac{1}{\mathcal{X}_{i+1}} \right) \right. \\
 & - 6 \text{Li}_3(-\mathcal{X}_{i-1}) \log(\mathcal{X}_{i+1}) - \frac{1}{2} \log(\mathcal{X}_{i-2}) \log^2(\mathcal{X}_i) \log(\mathcal{X}_{i+1}) \\
 & \left. + \text{Li}_2(-\mathcal{X}_{i-1}) \left( 3 \log(\mathcal{X}_{i-1}) \log(\mathcal{X}_{i+1}) + \log(\mathcal{X}_{i-2}/\mathcal{X}_{i+2}) \log(\mathcal{X}_{i+2}) \right) \right],
 \end{aligned} \tag{3.28}$$

where the  $\mathcal{X}_i$  are defined in terms of  $x_1$  and  $x_2$  as in eq. (2.15), and the skew-dihedral sum indicates subtracting the dihedral flip  $\mathcal{X}_i \rightarrow \mathcal{X}_{6-i}$  and summing over  $i$  from 1 to 5.

The function defined in (3.28) necessarily has the same cobracket as the  $A_2$  function considered in [12],

$$\delta(f_{A_2}) = - \sum_{\text{skew-dihedral}} \left[ \{\mathcal{X}_{i-1}\}_2 \wedge \{\mathcal{X}_{i+1}\}_2 + 3\{\mathcal{X}_i\}_2 \wedge \{\mathcal{X}_{i+1}\}_2 + 10\{\mathcal{X}_i\}_3 \wedge \{\mathcal{X}_{i+1}\}_1 \right], \quad (3.29)$$

but differs from the function considered there in some salient respects. Recalling that the clusters of  $A_2$  all take the form  $1/\mathcal{X}_i \rightarrow \mathcal{X}_{i+1}$ , it is easy to see that the symbol

$$\begin{aligned} \mathcal{S}(f_{A_2}) = \sum_{\text{skew-dihedral}} & \left[ 2 \mathcal{X}_i \otimes \mathcal{X}_{i+1} \otimes \mathcal{X}_i \otimes \mathcal{X}_{i+1} + 2 \mathcal{X}_i \otimes \mathcal{X}_{i+1} \otimes \mathcal{X}_{i+2} \otimes \mathcal{X}_{i+1} \right. \\ & + \mathcal{X}_{i+1} \otimes \mathcal{X}_i \otimes \mathcal{X}_{i+1} \otimes \mathcal{X}_{i+2} - \mathcal{X}_i \otimes \mathcal{X}_i \otimes \mathcal{X}_{i+1} \otimes \mathcal{X}_{i+1} \\ & \left. - 2 \mathcal{X}_i \otimes \mathcal{X}_{i+1} \otimes \mathcal{X}_{i+1} \otimes \mathcal{X}_{i+2} - \mathcal{X}_i \otimes \mathcal{X}_{i+1} \otimes \mathcal{X}_{i+1} \otimes \mathcal{X}_i \right], \end{aligned} \quad (3.30)$$

satisfies cluster  $\mathcal{X}$ -coordinate adjacency, and therefore also  $\mathcal{A}$ -coordinate adjacency (recall that  $\mathcal{X}_i$  and  $1/\mathcal{X}_i$  can always be exchanged in the symbol at the cost of a minus sign). As a function,  $f_{A_2}$  is also smooth and real-valued in the positive domain  $x_1, x_2 > 0$ . The  $A_2$  cluster algebra plays a crucial role in endowing  $f_{A_2}$  with this analytic behavior, as  $\text{Li}_{2,2}(x, y)$  and  $\text{Li}_{1,3}(x, y)$  have branch cuts in three locations, namely  $x = 1$ ,  $y = 1$ , and  $xy = 1$ . The first two branch cuts are trivially avoided as  $-1/\mathcal{X}_i < 0$  for  $x_1, x_2 > 0$ . However, the last branch cut is also avoided because of the  $A_2$  exchange relation:

$$0 < \left( -\frac{1}{\mathcal{X}_{i-1}} \right) \left( -\frac{1}{\mathcal{X}_{i+1}} \right) = \frac{1}{1 + \mathcal{X}_i} < 1. \quad (3.31)$$

Note that any sum of  $A_2$  functions evaluated on the subalgebras of some larger cluster algebra will also inherit these cluster adjacency and smoothness properties.

Remarkably, it was shown in [12] that all of the information contained in  $\delta_{2,2}(R_n^{(2)})$  is encoded in the  $A_2$  function, when this function is evaluated on some collection of the  $A_2$  subalgebras of  $\text{Gr}(4, n)$ . That is,

$$\delta_{2,2}(R_n^{(2)}) = \sum_{(x_i \rightarrow x_j) \subset \text{Gr}(4, n)} c_{ij} \delta_{2,2}(f_{A_2}(x_i \rightarrow x_j)) \quad (3.32)$$

for some rational coefficients  $c_{ij}$ . Moreover, the terms in this sum can themselves be arranged into  $A_3$  subalgebras, giving rise to a natural  $A_3$  function of the form

$$f_{A_3}(x_1 \rightarrow x_2 \rightarrow x_3) \sim \sum_{(x_i \rightarrow x_j) \subset A_3} d_{ij} f_{A_2}(x_i \rightarrow x_j), \quad (3.33)$$

where the  $d_{ij}$  are some rational coefficients that we treat in full detail in the next section. For now, we merely highlight the fact that these coefficients can be chosen in

such a way that  $\delta_{2,2}(f_{A_3})$  contains only terms  $\{x_i\}_2 \otimes \{x_j\}_2$  in which  $x_i$  and  $x_j$  appear together in at least one cluster of  $A_3$ . That is, there exists a decomposition

$$\delta_{2,2}(R_n^{(2)}) = \sum_{(x_i \rightarrow x_j \rightarrow x_k) \subset \text{Gr}(4,n)} c_{ijk} \delta_{2,2}(f_{A_3}(x_i \rightarrow x_j \rightarrow x_k)) \quad (3.34)$$

that makes the cobracket-level cluster adjacency enjoyed by these amplitudes manifest term-by-term.

In the following sections we analyze these types decompositions systematically, and find they can be extended to much larger subalgebras. As in the  $A_3$  decomposition, the  $A_2$  function will continue to play a privileged role.

## 4 Nonclassical Cluster Polylogarithms

We now turn to a more systematic exploration of the space of weight-four nonclassical cluster polylogarithms. We restrict our attention to functions that can be defined on  $\text{Gr}(4,7) \simeq E_6$  and its subalgebras, as this space is complex enough to give rise to an interesting collection of functions, yet can be explored exhaustively (this space was also explored in [32], using a slightly different approach). The techniques we utilize can also be applied to infinite cluster algebras (see for instance [12]), but we leave further exploration of this kind to future work [37]. Before exploring the properties of this space of functions on  $\text{Gr}(4,7)$ , we describe an efficient method for its generation using the  $A_2$  function (3.28). We then describe how this method can be extended to more restricted spaces of functions—those that are constructible out of cluster polylogarithms associated with the higher-rank subalgebras of a given cluster algebra. It is hoped that these subalgebra-constructible functions will prove to be of interest to physicists and mathematicians beyond the role they play in MHV amplitudes in planar  $\mathcal{N} = 4$  sYM theory.

### 4.1 $A_2$ Functions as a Basis

A remarkable (conjectured) property of  $f_{A_2}$  is that it forms a complete basis for the  $\delta_{2,2}$  component of any weight-four cluster polylogarithm [12]. (This property can be constructively realized for all cluster polylogarithms defined on  $E_6$  or its subalgebras, but has not been proven beyond these cases.<sup>2</sup>) That is, given a nonclassical weight-four cluster polylogarithm  $F$  defined on a cluster algebra  $\mathcal{A}$ , there always exists some decomposition

$$\delta_{2,2}(F) = \sum_{(x_i \rightarrow x_j) \subset \mathcal{A}} c_{ij} \delta_{2,2}(f_{A_2}(x_i \rightarrow x_j)) \quad (4.1)$$

---

<sup>2</sup>This was first checked in  $E_6$  by Daniel Parker and Adam Scherlis.



involving rational coefficients  $c_{ij}$ , where the sum ranges over all the  $A_2$  subalgebras of  $\mathcal{A}$ . Given the symbol (or cobracket) of  $F$ , this decomposition is simple to compute (but may not be unique). This decomposition can also be phrased as

$$F = \sum_{(x_i \rightarrow x_j) \subset \mathcal{A}} c_{ij} f_{A_2}(x_i \rightarrow x_j) + \dots, \quad (4.2)$$

where the dots indicate a residual, purely classical contribution to this function (also expressible in terms of cluster coordinates on  $\mathcal{A}$ ). Since the  $f_{A_2}$  function is smooth, real-valued, and satisfies cluster adjacency, this classical contribution will inherit these properties whenever they are also respected by  $F$ . In other words, this  $f_{A_2}$  decomposition cleanly separates the nonclassical from the classical components without disrupting any of the other salient properties of  $F$ .

We can glean some intuition for the meaning of this decomposition from the cobracket integrability condition (3.17). This condition implies specific algebraic relationships between the arguments of the  $\delta_{2,2}$  and  $\delta_{3,1}$  cobracket components. In particular, the  $A_2$  exchange relation

$$1 + \mathcal{X}_i = \mathcal{X}_{i-1} \mathcal{X}_{i+1} \quad (4.3)$$

is sufficient to generate one solution to integrability, namely  $f_{A_2}$ . The decomposition (4.1) thus tells us that *all* solutions to the cobracket integrability relation can be interpreted as linear combinations of exchange relations on  $A_2$  subalgebras. This hardly comes as a surprise, given that the only algebraic relations between cluster coordinates are generated by mutation (assuming that the seed cluster is composed of algebraically independent coordinates), and that an  $A_2$  subalgebra can be generated on any pair of connected nodes.

In practice, then, the existence (and uniqueness) of  $f_{A_2}$  solves the problem of writing down weight-four cluster polylogarithms. To illustrate this, let us consider the space of nonclassical  $A_3$  cluster polylogarithms. This space could be computed by forming an ansatz out of all possible cobracket components constructed from cluster  $\mathcal{X}$ -coordinates, and then imposing the integrability condition (3.17). However, the decomposability conjecture (4.1) reduces this computation to the simpler question: how many cluster polylogarithms can be constructed out of the  $A_2$  subalgebras of  $A_3$ ?

Recall from section 2.3 that there are six  $A_2$  subalgebras in  $A_3$ . To label these subalgebras, we choose a representative cluster from each one; in terms of  $A_3$  coordinates (defined by the initial seed  $x_1 \rightarrow x_2 \rightarrow x_3$ ), we take the set

$$x_1 \rightarrow x_2, \quad x_2 \rightarrow x_3, \quad \frac{x_2}{1 + x_{1,2}} \rightarrow \frac{(1 + x_1)x_3}{1 + x_{1,2,3}}, \quad (4.4)$$

$$\frac{x_1 x_2}{1 + x_1} \rightarrow x_3, \quad x_1(1 + x_2) \rightarrow \frac{x_2 x_3}{1 + x_2}, \quad x_1 \rightarrow x_2(1 + x_3). \quad (4.5)$$

We then construct an ansatz for the space of nonclassical  $A_3$  functions by considering the span of  $f_{A_2}$  evaluated on each of these subalgebras:

$$f_{A_3}(x_1 \rightarrow x_2 \rightarrow x_3) = c_1 f_{A_2}(x_1 \rightarrow x_2) + \dots + c_6 f_{A_2}(x_1 \rightarrow x_2(1+x_3)). \quad (4.6)$$

All that remains is to find values for the coefficients  $c_i$  such that  $f_{A_3}$  is invariant (up to an overall sign) under the automorphisms of  $A_3$ . In this case, there are two automorphism group generators,  $\sigma_{A_3}$  and  $\tau_{A_3}$ , which were defined in (2.51) and (2.52).

We first consider the case where  $f_{A_3}$  is invariant under both  $A_3$  automorphism generators, namely

$$\begin{aligned} \sigma_{A_3}(f_{A_3}(x_1 \rightarrow x_2 \rightarrow x_3)) &= f_{A_3}(x_1 \rightarrow x_2 \rightarrow x_3), \\ \tau_{A_3}(f_{A_3}(x_1 \rightarrow x_2 \rightarrow x_3)) &= f_{A_3}(x_1 \rightarrow x_2 \rightarrow x_3). \end{aligned} \quad (4.7)$$

In practice, this means we take the ansatz for  $f_{A_3}$  in eq. (4.6) and solve the constraints

$$\begin{aligned} f_{A_3}\left(\frac{x_2}{1+x_{1,2}} \rightarrow \frac{x_3(1+x_1)}{1+x_{1,2,3}} \rightarrow \frac{1+x_{1,2}}{x_1 x_2 x_3}\right) &= f_{A_3}(x_1 \rightarrow x_2 \rightarrow x_3), \\ f_{A_3}\left(\frac{1}{x_3} \rightarrow \frac{1}{x_2} \rightarrow \frac{1}{x_1}\right) &= f_{A_3}(x_1 \rightarrow x_2 \rightarrow x_3). \end{aligned} \quad (4.8)$$

It turns out there is no nontrivial solution to this set of constraints. Conversely, allowing  $f_{A_3}$  to pick up an overall minus sign when acted upon by  $\tau_{A_3}$ , we solve the constraints

$$\begin{aligned} \sigma_{A_3}(f_{A_3}(x_1 \rightarrow x_2 \rightarrow x_3)) &= f_{A_3}(x_1 \rightarrow x_2 \rightarrow x_3), \\ \tau_{A_3}(f_{A_3}(x_1 \rightarrow x_2 \rightarrow x_3)) &= -f_{A_3}(x_1 \rightarrow x_2 \rightarrow x_3), \end{aligned} \quad (4.9)$$

and find the solution

$$c_i = 1 \quad (4.10)$$

(which can always be rescaled by an overall constant). We label this particular solution

$$\begin{aligned} f_{A_3}^{+-}(x_1 \rightarrow x_2 \rightarrow x_3) &= f_{A_2}(x_1 \rightarrow x_2) + \dots + f_{A_2}(x_1 \rightarrow x_2(1+x_3)) \\ &= \sum_{i=1}^6 \sigma_{A_3}^i(f_{A_2}(x_1 \rightarrow x_2)), \end{aligned} \quad (4.11)$$

where the superscripts on  $f_{A_3}^{+-}$  label its behavior under  $\sigma$  and  $\tau$ , respectively. (The superscript on  $\sigma_{A_3}^i$  indicates how many times the operator  $\sigma_{A_3}$  should be applied.) There are two remaining sign choices to check:  $\sigma_{A_3}^- \tau_{A_3}^+$  and  $\sigma_{A_3}^- \tau_{A_3}^-$ . We find only the trivial solution in the first case, while

$$f_{A_3}^{--}(x_1 \rightarrow x_2 \rightarrow x_3) = \sum_{i=1}^6 (-1)^i \sigma_{A_3}^i(f_{A_2}(x_1 \rightarrow x_2)) \quad (4.12)$$

Nonclassical $A_n$ Polylogarithms				
	$\sigma^+\tau^+$	$\sigma^+\tau^-$	$\sigma^-\tau^+$	$\sigma^-\tau^-$
$A_2$	0	1 (0)	0	0
$A_3$	0	1 (0)	0	1 (1)
$A_4$	0	3 (0)	0	0
$A_5$	2 (1)	5 (1)	2 (0)	5 (3)

**Table 1:** The number of nonclassical weight-four cluster polylogarithms on the  $A_{n \leq 5}$  cluster algebra, for each of the four possible automorphism signatures. The number of functions that additionally respect cobracket-level cluster adjacency is given in parentheses.

turns out to be the unique  $A_3$  function that picks up a minus sign under the action of both  $A_3$  automorphism generators.

Therefore, unlike the case of  $A_2$ , there are two functions one can associate with the  $A_3$  cluster algebra:  $f_{A_3}^{+-}$  and  $f_{A_3}^{--}$ . These functions arise purely from the interplay between the overall symmetries of the  $A_3$  cluster algebra and the structure of the  $A_2$  subalgebras in  $A_3$ , i.e. there has been no physics input so far. However, motivated by the properties of the two-loop MHV amplitudes, we can check one further aspect of these functions—whether or not they respect cobracket-level cluster adjacency. Recall that the function  $f_{A_2}$  does not satisfy cobracket-level cluster adjacency, while it was stated around equation (3.33) that there exists an  $f_{A_3}$  that does have this property—in fact, it was pointed out in [12] that there exists only one such function. It is easy to check that only  $f_{A_3}^{--}$  can be expressed in terms of cluster-adjacent cobrackets, and that it matches the function reported in [12].

## 4.2 Constructing all Cluster Polylogarithms in $\text{Gr}(4, 7)$

It is a straightforward exercise to extend the construction outlined in the last subsection to any finite algebra (as well as to infinite algebras by specifying a set of  $A_2$  subalgebras). In the case of  $A_n$ , there are just the two automorphism generators  $\sigma_n$  and  $\tau_n$ , which were given in (2.51) and (2.52). The results are summarized in table 1, where each entry denotes the number of nonclassical weight-four cluster polylogarithms with a given automorphism signature, and the numbers in parentheses indicate how many of these functions additionally satisfy cobracket-level cluster adjacency. It is interesting to note that the automorphism signature  $\sigma^+\tau^-$  admits at least one solution for each of these  $A_n$ , and in particular that it gives rise to the only nontrivial solutions in both  $A_2$

Nonclassical  $D_4$  Polylogarithms

$\sigma_{\mathfrak{D}_4}^+ \tau_{\mathfrak{D}_4}^+$			$\sigma_{\mathfrak{D}_4}^+ \tau_{\mathfrak{D}_4}^-$			$\sigma_{\mathfrak{D}_4}^- \tau_{\mathfrak{D}_4}^+$			$\sigma_{\mathfrak{D}_4}^- \tau_{\mathfrak{D}_4}^-$		
	$\sigma_{S_3}^+$	$\sigma_{S_3}^-$		$\sigma_{S_3}^+$	$\sigma_{S_3}^-$		$\sigma_{S_3}^+$	$\sigma_{S_3}^-$		$\sigma_{S_3}^+$	$\sigma_{S_3}^-$
$\tau_{S_3}^+$	0	0	$\tau_{S_3}^+$	2 (0)	0	$\tau_{S_3}^+$	1 (0)	0	$\tau_{S_3}^+$	1 (0)	0
$\tau_{S_3}^-$	1 (0)	0	$\tau_{S_3}^-$	0	0	$\tau_{S_3}^-$	1 (1)	0	$\tau_{S_3}^-$	0	0

Nonclassical  $D_5$  Polylogarithms

$\sigma_{\mathfrak{D}_5}^+ \tau_{\mathfrak{D}_5}^+$		$\sigma_{\mathfrak{D}_5}^+ \tau_{\mathfrak{D}_5}^-$		$\sigma_{\mathfrak{D}_5}^- \tau_{\mathfrak{D}_5}^+$		$\sigma_{\mathfrak{D}_5}^- \tau_{\mathfrak{D}_5}^-$	
$\mathbb{Z}_2^+$	$\mathbb{Z}_2^-$	$\mathbb{Z}_2^+$	$\mathbb{Z}_2^-$	$\mathbb{Z}_2^+$	$\mathbb{Z}_2^-$	$\mathbb{Z}_2^+$	$\mathbb{Z}_2^-$
5 (2)	0	9 (2)	0	0	3 (1)	0	7 (5)

**Table 2:** The number of nonclassical weight-four cluster polylogarithms on the  $D_4$  and  $D_5$  cluster algebras, with each possible automorphism signature. The number of functions that additionally respect cobracket-level cluster adjacency is given in parentheses.

and  $A_4$ .

This procedure gets only slightly more complicated in the  $D_n$  algebras, due to their larger automorphism groups. As discussed in section 2.6, the automorphism group of  $D_4$  is a product of dihedral and symmetric groups,  $\mathfrak{D}_4 \times S_3$ . The dihedral group  $\mathfrak{D}_4$  is generated by a pair of operators  $\sigma_{D_4}^{(\mathfrak{D}_4)}$  and  $\tau_{D_4}^{(\mathfrak{D}_4)}$ , while the symmetric group is generated by a pair of operators  $\sigma_{D_4}^{(S_3)}$  and  $\tau_{D_4}^{(S_3)}$ . This gives rise to 16 possible automorphism signatures, which we impose on a general ansatz of  $A_2$  functions; the results are presented in table 2. In the same table we give the results for  $D_5$ , which has only the three automorphism generators  $\sigma_{D_5}$ ,  $\tau_{D_5}$ , and  $\mathbb{Z}_{2,D_5}$ , corresponding to the automorphism group  $\mathfrak{D}_5 \times \mathbb{Z}_2$ . In both  $D_4$  and  $D_5$ , we again see that the space of functions respecting automorphisms is remarkably constrained. There are no functions with odd signature under  $\sigma_{D_4}^{(\mathfrak{D}_4)}$ , and only a single  $D_4$  automorphism signature gives rise to more than one solution. In  $D_5$ , there are no functions that have opposite signature in  $\sigma_{D_5}$  and  $\mathbb{Z}_{2,D_5}$ .

Finally, we turn to  $E_6$ , which has the automorphism group  $\mathfrak{D}_{14}$ . This group has three generators— $\sigma_{E_6}$ ,  $\tau_{E_6}$ , and  $\mathbb{Z}_{2,E_6}$ .  $E_6$  is much larger than any of the cluster algebras considered above, with 504 distinct  $A_2$  subalgebras. Even so, the spaces of automorphic functions on it remains surprisingly small, as shown in table 3. It is especially surprising here that there are no odd solutions in  $\sigma_{E_6}$ . Of primary interest is the space with automorphism signature  $\sigma_{E_6}^+ \tau_{E_6}^+ \mathbb{Z}_{2,E_6}^+$ , as this space contains  $R_7^{(2)}$ , which

Nonclassical $E_6$ Polylogarithms							
$\frac{\sigma_{\mathfrak{D}_{14}}^+ \tau_{\mathfrak{D}_{14}}^+}{\mathbb{Z}_2^+ \mid \mathbb{Z}_2^-}$		$\frac{\sigma_{\mathfrak{D}_{14}}^+ \tau_{\mathfrak{D}_{14}}^-}{\mathbb{Z}_2^+ \mid \mathbb{Z}_2^-}$		$\frac{\sigma_{\mathfrak{D}_{14}}^- \tau_{\mathfrak{D}_{14}}^+}{\mathbb{Z}_2^+ \mid \mathbb{Z}_2^-}$		$\frac{\sigma_{\mathfrak{D}_{14}}^- \tau_{\mathfrak{D}_{14}}^-}{\mathbb{Z}_2^+ \mid \mathbb{Z}_2^-}$	
12 (6)	14 (6)	21 (6)	17 (9)	0	0	0	0

**Table 3:** The number of nonclassical weight-four cluster polylogarithms on the  $E_6$  cluster algebra, with each of eight possible automorphism signatures. The number of functions that additionally respect cobracket-level cluster adjacency is given in parentheses.

will be our primary object of interest in section 5. In particular,  $R_7^{(2)}$  must be a linear combination of the 6 cluster polylogarithms in this space that respect cobracket-level cluster adjacency.

We pause at this point to emphasize that the  $A_2$  constructibility of all nonclassical cluster polylogarithms has been verified on each of the finite cluster algebras considered above, and its completeness only becomes conjectural on infinite cluster algebras. While it is clearly not possible to form an ansatz out of all  $A_2$  subalgebras in an infinite cluster algebra such as  $\text{Gr}(4,8)$ , this method can still be used to generate all nonclassical polylogarithms that are constructible out of (the automorphic completion of) any finite set of  $A_2$  subalgebras. This proves sufficient in the case of the two-loop remainder function to all  $n$  [12], and may prove sufficient in other cases of physical interest.

### 4.3 Nested Cluster Polylogarithms

Given that the space of nonclassical cluster polylogarithms coincides with the space of  $A_2$ -constructible functions, it is natural to ask whether other (even more special) spaces of functions are constructible out of other cluster polylogarithms. In fact, this has already been shown to be the case, since the nonclassical part of all two-loop MHV amplitudes can be constructed out of  $f_{A_3}^{--}$  as defined in (4.12) [12]. This gives rise to an interesting nested structure, since  $f_{A_3}^{--}$  is itself constructible out of  $f_{A_2}$ . More generally, in the functions constructed in section 4.2 there will be many instances of  $A_2$  subalgebras assembling into larger subalgebras, and it remains an open question how intrinsically interesting such nested constructibility might be.

The procedure for constructing such spaces clearly proceeds just as in the case of  $A_2$  constructibility. There are many spaces one can consider constructing on each cluster algebra, corresponding to the inclusion of different sets of functions defined on (possibly) different subalgebras. We leave the exploration of compositely-constructible

$A_3^{--}$ -constructible $A_n$ Polylogarithms				
	$\sigma^+ \tau^+$	$\sigma^+ \tau^-$	$\sigma^- \tau^+$	$\sigma^- \tau^-$
$f_{A_4} \in \text{span}(f_{A_3}^{--})$	0	0	0	0
$f_{A_5} \in \text{span}(f_{A_3}^{--})$	1	1	0	3

$A_3^{--}$ -constructible $D_4$ Polylogarithms											
$\sigma_{\mathfrak{D}_4}^+ \tau_{\mathfrak{D}_4}^+$			$\sigma_{\mathfrak{D}_4}^+ \tau_{\mathfrak{D}_4}^-$			$\sigma_{\mathfrak{D}_4}^- \tau_{\mathfrak{D}_4}^+$			$\sigma_{\mathfrak{D}_4}^- \tau_{\mathfrak{D}_4}^-$		
	$\sigma_{S_3}^+$	$\sigma_{S_3}^-$		$\sigma_{S_3}^+$	$\sigma_{S_3}^-$		$\sigma_{S_3}^+$	$\sigma_{S_3}^-$		$\sigma_{S_3}^+$	$\sigma_{S_3}^-$
$\tau_{S_3}^+$	0	0	$\tau_{S_3}^+$	0	0	$\tau_{S_3}^+$	0	0	$\tau_{S_3}^+$	0	0
$\tau_{S_3}^-$	0	0	$\tau_{S_3}^-$	0	0	$\tau_{S_3}^-$	1	0	$\tau_{S_3}^-$	0	0

$A_3^{--}$ -constructible $D_5$ Polylogarithms											
$\sigma_{\mathfrak{D}_5}^+ \tau_{\mathfrak{D}_5}^+$		$\sigma_{\mathfrak{D}_5}^+ \tau_{\mathfrak{D}_5}^-$		$\sigma_{\mathfrak{D}_5}^- \tau_{\mathfrak{D}_5}^+$		$\sigma_{\mathfrak{D}_5}^- \tau_{\mathfrak{D}_5}^-$					
$\mathbb{Z}_2^+$	$\mathbb{Z}_2^-$	$\mathbb{Z}_2^+$	$\mathbb{Z}_2^-$	$\mathbb{Z}_2^+$	$\mathbb{Z}_2^-$	$\mathbb{Z}_2^+$	$\mathbb{Z}_2^-$				
2	0	2	0	0	1	0	5				

$A_3^{--}$ -constructible $E_6$ Polylogarithms											
$\sigma_{\mathfrak{D}_{14}}^+ \tau_{\mathfrak{D}_{14}}^+$		$\sigma_{\mathfrak{D}_{14}}^+ \tau_{\mathfrak{D}_{14}}^-$		$\sigma_{\mathfrak{D}_{14}}^- \tau_{\mathfrak{D}_{14}}^+$		$\sigma_{\mathfrak{D}_{14}}^- \tau_{\mathfrak{D}_{14}}^-$					
$\mathbb{Z}_2^+$	$\mathbb{Z}_2^-$	$\mathbb{Z}_2^+$	$\mathbb{Z}_2^-$	$\mathbb{Z}_2^+$	$\mathbb{Z}_2^-$	$\mathbb{Z}_2^+$	$\mathbb{Z}_2^-$				
6	6	6	9	0	0	0	0				

**Table 4:** The number of  $f_{A_3}^{--}$ -constructible cluster polylogarithms on the  $E_6$  cluster algebra and its subalgebras, with each possible automorphism signature.

spaces to future work. For now, we just consider the example of functions constructible out of  $f_{A_3}^{--}$ , since all  $R_n^{(2)}$  are believed to be in this class. We tabulate the space of  $f_{A_3}^{--}$ -constructible functions in table 4.

Since  $f_{A_3}^{--}$  satisfies cobracket-level cluster adjacency, all functions constructed out of it also have this property. As can be seen by comparing table 4 to tables 1-3, the converse is also true—all cluster polylogarithms that satisfy cobracket-level cluster adjacency are  $f_{A_3}^{--}$ -constructible. While we have checked this explicitly on  $E_6$  and all its subalgebras, it has been conjectured to hold more generally [12].

It is worth highlighting that the dimensions we have tabulated in this (and the

previous) section have taken into account polylogarithmic identities that reduce the difference between two nonclassical functions to something purely classical. For instance, in  $E_6$  there are identities relating different instances of  $f_{A_2}$  and  $f_{A_3}^{--}$  such as

$$\sum_{E_6^{+++}} \left[ \delta_{2,2} \left( 2f_{A_3}^{--} \left( x_1 \rightarrow x_2(1+x_3) \rightarrow \frac{x_3x_4x_5}{1+x_{5,3}} \right) \right. \right. \quad (4.13)$$

$$\left. \left. - f_{A_3}^{--} \left( x_1 \rightarrow x_2(1+x_3) \rightarrow \frac{x_3x_4x_5x_6}{1+x_{6,5,3}} \right) \right. \right.$$

$$\left. \left. - f_{A_3}^{--} \left( x_1 \rightarrow \frac{x_2(1+x_{6,5,3,4})}{1+x_{6,5}} \rightarrow \frac{x_3x_5}{(1+x_5)(1+x_{6,5,3})} \right) \right) \right] = 0,$$

where we have introduced the notation

$$\sum_{E_6^{+++}} f = \sum_{i=0}^6 \sum_{j=0}^1 \sum_{k=0}^1 \tau_{E_6}^k \circ \mathbb{Z}_{2,E_6}^j \circ \sigma_{E_6}^i(f). \quad (4.14)$$

Although we won't discuss these spaces of identities further, it would be interesting to study them (and their geometry on cluster polytopes) systematically.

## 5 Subalgebra Constructibility and $R_7^{(2)}$

Having seen the extent to which the  $E_6$  cluster algebra admits nontrivial functional embeddings, we turn to the study of  $R_7^{(2)}$ . Specifically, we explore the ways in which the nonclassical part of this function is subalgebra-constructible, thus probing the extent to which  $\mathcal{N} = 4$  sYM theory “knows” about the rich subalgebra structure of  $E_6$ . It is already known that  $R_7^{(2)}$  is  $A_2$ - and  $A_3$ -constructible in terms of the functions defined in (3.28) and (4.12) [12]. Using the techniques of section 4, we can ask this question systematically for any set of cluster polylogarithms defined on a subalgebra of  $E_6$ .

We will in particular be interested in the corank-one subalgebras of  $E_6$ —namely,  $D_5$  and  $A_5$ —as  $D_5$ - and  $A_5$ -constructible functions are not subject to the same amount of representational ambiguity as functions that are constructible out of smaller subalgebras. For instance, while there are 1071 subpolytopes of the  $E_6$  cluster polytope that correspond to  $A_2$  subalgebras, only 504 of these give rise to distinct cluster algebras. Moreover, there exist 56 identities between the instances of  $f_{A_2}$  evaluated on these 504 subalgebras [12]. Similarly, out of 476 distinct  $A_3$  subpolytopes, only 364 give rise to distinct cluster algebras, and there are 169 identities between the instances of  $f_{A_3}^{--}$  evaluated on these subalgebras. These redundancies make the representation of  $\delta_{2,2}(R_7^{(2)})$  in terms of either  $f_{A_2}$  or  $f_{A_3}^{--}$  far from unique. Correspondingly, it is hard to assign

a clear geometric interpretation to these decompositions on the  $E_6$  cluster polytope, since the  $A_2$  function associated with any specific subpolytope can be traded for  $A_2$  functions associated with other subpolytopes (and similarly for  $f_{A_3}^{--}$ ).

On the other hand, there are only 14  $D_5$  subpolytopes and 7  $A_5$  subpolytopes of the  $E_6$  cluster polytope, each of which gives rise to a distinct subalgebra. These collections of subalgebras respectively form complete orbits under the automorphism group of  $E_6$ , implying that all  $D_5$ - and  $A_5$ -constructible functions in  $E_6$  take the form

$$\sum_{D_5 \subset E_6} f_{D_5}(x_i \rightarrow \dots) = \sum_{i=0}^6 \sum_{j=0}^1 (\pm 1)^i (\pm 1)^j \mathbb{Z}_{2,E_6}^j \circ \sigma_{E_6}^i \left( f_{D_5}(x_i \rightarrow \dots) \right) \quad (5.1)$$

and

$$\sum_{A_5 \subset E_6} f_{A_5}(x_i \rightarrow \dots) = \sum_{i=0}^6 (\pm 1)^i \sigma_{E_6}^i \left( f_{A_5}(x_i \rightarrow \dots) \right), \quad (5.2)$$

where each sign choice determines the signature of the resulting  $E_6$  polylogarithm under the associated automorphism generator. In the case of  $R_7^{(2)}$ , the positive sign should be chosen in each case. Thus,  $D_5$  and  $A_5$  functions can only be embedded in  $E_6$  with the right automorphism signature a single way—making  $D_5$  and  $A_5$  decompositions of  $R_7^{(2)}$  (if they exist) canonical in a way that decompositions into smaller subalgebras are not.

We pause to note that the sum notation used in (5.1) and (5.2) differs from the notation used in (4.14). The two notations are in practice equivalent, since—as we require cluster polylogarithms to respect automorphisms (up to a sign)—summing over indirect automorphisms (with the appropriate sign) only rescales these functions by a multiplicative factor. We will adopt notation such as in (5.1) and (5.2) when it is useful to think of summing over all subalgebras corresponding to subpolytopes of a given type in the cluster polytope. The downside of this representation is that it obfuscates the behavior of the resulting object under the indirect automorphisms of the parent algebra. We will thus make use of notation similar to (4.14) when we deem it more important to make the behavior under all automorphisms explicit. Adding such a sum over indirect automorphisms amounts to summing over subalgebras that appear in the “flipped” cluster polytope as well as the original.

### 5.1 The $D_5$ Constructibility of $R_7^{(2)}$

We begin with  $D_5$ , the largest subalgebra of  $E_6$  in terms of its number of clusters. For each of the automorphism signatures in  $D_5$  that admits a nontrivial space of (non-classical) cluster polylogarithms, we insert a general ansatz of such functions into the



sum (5.1) with positive signs chosen for both direct automorphism generators. Referring back to table 2, we see there are four nontrivial cases to consider. Note that we do not restrict our ansätze to  $D_5$  functions that respect cobracket-level cluster adjacency, because it is possible for functions that don't have this property to assemble into linear combinations that do when evaluated on the subalgebras of  $E_6$  (as happens when  $f_{A_3}^{--}$  is constructed out of  $A_2$  functions). We then check to see if the nonclassical part of  $R_7^{(2)}$  is in the span of any of these ansatz sums.

Amazingly, there exists precisely one  $D_5$  automorphism signature in terms of which  $R_7^{(2)}$  can be decomposed—the totally odd signature  $\sigma_{D_5}^- \tau_{D_5}^- \mathbb{Z}_{2,D_5}^-$ . Two free parameters remain in this decomposition, corresponding to two degrees of freedom that cancel in the sum (5.1). More concretely, we can decompose

$$\delta_{2,2}(R_7^{(2)}) = \frac{1}{20} \sum_{D_5 \subset \text{Gr}(4,7)} \delta_{2,2} \left( f_{D_5}^{--} \left( x_1 \rightarrow x_2 \rightarrow \frac{x_3 x_5}{1+x_5} \begin{array}{c} \nearrow x_4 \\ \searrow \frac{1}{x_5} \end{array} \right) \right), \quad (5.3)$$

where the sum over all  $D_5$  subalgebras is taken according to (5.1), and we define

$$\begin{aligned} f_{D_5}^{--} \left( x_1 \rightarrow x_2 \rightarrow x_3 \begin{array}{c} \nearrow x_4 \\ \searrow x_5 \end{array} \right) &\equiv \sum_{D_5^{--}} \left[ \frac{1}{2} c_1 f_{A_2}(x_1 \rightarrow x_2 (1+x_3)) \right. \\ &\quad - \left( \frac{1}{2} - \frac{c_1}{2} \right) f_{A_2} \left( \frac{x_1 x_2}{1+x_1} \rightarrow x_3 (1+x_4) \right) + \frac{1}{4} c_2 f_{A_2} \left( \frac{x_1 x_2 x_3}{1+x_{12}} \rightarrow x_4 \right) \\ &\quad + (c_1 - c_2) f_{A_2} \left( \frac{x_2 x_3}{1+x_2} \rightarrow x_4 \right) + (1 - c_1) f_{A_2}(x_2 \rightarrow x_3 (1+x_4)) \\ &\quad \left. + \left( \frac{1}{2} - c_1 + \frac{3c_2}{4} \right) f_{A_2} \left( x_2 (1+x_3) \rightarrow \frac{x_3 x_4}{1+x_3} \right) + \frac{1}{2} c_2 f_{A_2}(x_3 \rightarrow x_4) \right] \end{aligned} \quad (5.4)$$

using the notation

$$\sum_{D_5^{--}} f = \sum_{i=0}^4 \sum_{j=0}^1 \sum_{k=0}^1 (-1)^{i+j+k} \sigma_{D_5}^i \circ \tau_{D_5}^j \circ \mathbb{Z}_{2,D_5}^k (f). \quad (5.5)$$

Unlike in the sums (5.1) and (5.2)—where the sum was over all subalgebras of a given type in  $E_6$  (but not its flipped twin)—we now include a sum over the indirect automorphism  $\tau_{D_5}$  so as to make explicit the totally odd automorphism signature. The  $\mathcal{X}$ -coordinates in (5.3) should be understood to be the  $E_6$  coordinates defined in (2.44), while the  $\mathcal{X}$ -coordinates in (5.4) should be thought of as  $D_5$  coordinates that take different values when evaluated on the  $D_5$  subalgebras of  $E_6$ . (It is easy to see that the  $E_6$

seed (2.43) doesn't contain any  $D_5$  sub-quivers of the form (2.42), but that mutating on the  $x_5$  node will give rise to a quiver containing the  $D_5$  quiver seen in (5.3).) The coefficients  $c_1$  and  $c_2$  represent the two remaining degrees of freedom in our ansatz, which drop out of the sum (5.3).

It is natural to ask whether the parameters  $c_1$  and  $c_2$  can be chosen in such a way that  $f_{D_5}^{---}$  is endowed with additional nice properties. For instance,  $f_{D_5}^{---}$  does not satisfy cobracket-level cluster adjacency for generic values of  $c_1$  and  $c_2$ , but it can be given this property by choosing  $c_2 = -\frac{6}{5} + \frac{8}{5}c_1$ . As discussed in section 4.3, this necessarily makes  $f_{D_5}^{---}$  itself an  $A_3$ -constructible function:

$$\begin{aligned} f_{A_3 \subset D_5}^{---} = & -\frac{1}{10} \sum_{D_5^{---}} \left[ (6 - 3c_1) f_{A_3}^{--} \left( x_1 \rightarrow x_2(1 + x_3) \rightarrow \frac{x_3 x_4}{1 + x_3} \right) \right. \\ & + (-3 + 4c_1) f_{A_3}^{--} (x_1 \rightarrow x_2 \rightarrow x_3(1 + x_4)) \\ & + (2 - c_1) f_{A_3}^{--} (x_2 \rightarrow x_3(1 + x_4) \rightarrow x_5) \\ & + \left( \frac{1}{2} - \frac{3c_1}{2} \right) f_{A_3}^{--} \left( \frac{1}{x_4} \rightarrow x_3(1 + x_4) \rightarrow x_5 \right) \\ & \left. + \left( -1 + \frac{c_1}{2} \right) f_{A_3}^{--} \left( \frac{1}{x_4} \rightarrow \frac{x_2 x_3(1 + x_4)}{1 + x_2} \rightarrow x_5 \right) \right]. \end{aligned} \quad (5.6)$$

As in the (more general)  $f_{A_2}$  representation (5.4), we have made the full automorphism signature of the  $D_5$  function manifest at the level of the sum.

It turns out there is one additional way  $f_{D_5}^{---}$  can be decomposed. An  $A_4$  decomposition is obtainable if we choose  $c_1 = \frac{3}{5}$  and  $c_2 = 0$ . Moreover, although there exists no canonical decomposition such as (5.1) or (5.2) for  $A_4$ -constructible  $E_6$  polylogarithms, there does exist such a decomposition for  $A_4$ -constructible  $D_5$  polylogarithms. That is, all 10  $A_4$  subpolytopes of the  $D_5$  cluster polytope are in the same orbit of the automorphism group of  $D_5$ . Thus, such a decomposition must take the form

$$\sum_{A_4 \subset D_5} f_{A_4}(x_i \rightarrow \dots) = \sum_{i=0}^4 \sum_{j=0}^1 (\pm 1)^i (\pm 1)^j \mathbb{Z}_{2, D_5}^j \circ \sigma_{D_5}^i \left( f_{A_4}(x_i \rightarrow \dots) \right). \quad (5.7)$$

Choosing both signs to be negative to match the automorphism signature of  $f_{D_5}^{---}$ , we find the decomposition

$$f_{A_4 \subset D_5}^{---} = \sum_{A_4 \subset D_5} f_{A_4}^{+-} (x_1 \rightarrow x_2 \rightarrow x_3 \rightarrow x_4) \quad (5.8)$$

in terms of a new  $A_4$  function

$$f_{A_4}^{+-} = \frac{1}{10} \sum_{A_4^{+-}} \left( 3f_{A_2}(x_1 \rightarrow x_2(1 + x_3)) - f_{A_2}(x_2 \rightarrow x_3(1 + x_4)) \right), \quad (5.9)$$

where we have made use of the notation

$$\sum_{A_4^{+-}} f = \sum_{i=0}^6 \sum_{j=0}^1 (-1)^j \tau_{A_4}^j \circ \sigma_{A_4}^i (f). \quad (5.10)$$

The function  $f_{A_4}^{+-}$  itself is not  $A_3$ -constructible, as it does not satisfy cobracket-level adjacency. However, the  $f_{A_4 \subset D_5}^{---}$  representation of  $R_7^{(2)}$ ,

$$\delta_{2,2}(R_7^{(2)}) = \frac{1}{20} \sum_{D_5 \subset \text{Gr}(4,7)} \sum_{A_4 \subset D_5} \delta_{2,2}(f_{A_4}^{+-}(x_1 \rightarrow x_2 \rightarrow x_3 \rightarrow x_4)), \quad (5.11)$$

is still quite remarkable—by evaluating the simple function (5.9) on all  $A_4$  subalgebras of  $\text{Gr}(4,7)$  and arranging these functions into  $D_5$  polylogarithms (as uniquely dictated by the choice of a totally odd  $D_5$  automorphism signature),  $\delta_{2,2}(R_7^{(2)})$  is recovered by simply summing over these  $D_5$  polylogarithms. Note that every  $A_4$  subalgebra in  $\text{Gr}(4,7)$  appears at least once in this sum, however some instances of  $f_{A_4}^{+-}$  appear multiple times and with different sign, leaving only 56  $A_4$  subalgebras which contribute.

Having considered some cluster-algebraic motivations for choosing certain values of  $c_1$  and  $c_2$ , we now investigate whether collinear limits provide us with a natural choice for these parameters. In particular, we know that the nonclassical portion of  $R_7^{(2)}$  must vanish under the  $7 \rightarrow 6$  collinear limit, as  $R_6^{(2)}$  is purely classical. We can therefore ask whether the 14 instances of  $f_{D_5}^{---}$  in  $E_6$  can be made to separately vanish in the collinear limit. It turns out this is not possible—at best, by choosing  $c_1 = 0$  and  $c_2 = -\frac{6}{5}$ , one can get 10 of the 14  $f_{D_5}^{---}$  functions to separately vanish (at the nonclassical level). The nonclassical part of the remaining 4  $f_{D_5}^{---}$  functions then cancel off pairwise.

## 5.2 The $A_5$ Constructibility of $R_7^{(2)}$

We now turn our attention to  $A_5$  decompositions of  $R_7^{(2)}$ , using the ansatz sum (5.2). Intriguingly, there is (again) precisely one automorphism signature in terms of which  $\delta_{2,2}(R_7^{(2)})$  can be decomposed—and it is (again) the totally antisymmetric signature  $\sigma_{A_5}^- \tau_{A_5}^-$ . This time there is a single internal degree of freedom that cancels in the sum (5.2). This decomposition can be written explicitly as

$$\delta_{2,2}(R_7^{(2)}) = \frac{1}{20} \sum_{A_5 \subset E_6} \delta_{2,2} \left( f_{A_5}^{--} \left( x_1 \rightarrow x_2 \rightarrow \frac{x_3 x_5 x_6}{1 + x_{6,5}} \rightarrow \frac{1}{x_6(1 + x_5)} \rightarrow \frac{1 + x_{6,5}}{x_5} \right) \right), \quad (5.12)$$

where the  $A_5$  sum (5.2) is taken over the function

$$f_{A_5}^{--}(x_1 \rightarrow x_2 \rightarrow x_3 \rightarrow x_4 \rightarrow x_5) \equiv \sum_{A_5^{--}} \left[ \frac{1}{2} c_1 f_{A_2}(x_2 \rightarrow x_3(1+x_4)) \right. \\ \left. - (1+c_1) f_{A_2}(x_2 \rightarrow x_3(1+x_{4,5})) \right. \\ \left. - \left( \frac{1}{2} + c_1 \right) f_{A_2} \left( \frac{x_1 x_2}{1+x_1} \rightarrow x_3(1+x_4) \right) \right], \quad (5.13)$$

which makes use of the notation

$$\sum_{A_5^{--}} f = \sum_{i=0}^7 \sum_{j=0}^1 (-1)^{i+j} \sigma_{A_5}^i \circ \tau_{A_5}^j(f). \quad (5.14)$$

As in the last subsection, the  $\mathcal{X}$ -coordinates in (5.12) are the  $E_6$  coordinates defined in (2.44), while the  $\mathcal{X}$ -coordinates in (5.13) are  $A_5$  coordinates that take different values when evaluated on different  $A_5$  subalgebras of  $E_6$ . The initial  $A_5$  quiver in (5.12) can be generated by mutating on the nodes associated with  $x_5$  and then  $x_6$  in (2.43), and then freezing the node associated with  $x_4$ .

Let us first address  $f_{A_5}^{--}$ 's collinear limits. In this case, the nonclassical part of all 7 instances of  $f_{A_5}^{--}$  on  $E_6$  separately vanish in each  $7 \rightarrow 6$  collinear limit, for all values of  $c_1$ . This makes the collinear behavior of  $R_7^{(2)}$  considerably cleaner in the  $A_5$  decomposition than in the  $D_5$  decomposition. However, it also gives us no hints for how to choose  $c_1$ .

Luckily, we can find preferred values for  $c_1$  by requiring  $f_{A_5}^{--}$  to be either  $A_3$ - or  $A_4$ -constructible. It becomes  $A_3$ -constructible if we choose  $c_1 = -\frac{1}{2}$ , but this time in terms of the function  $f_{A_3}^{+-}$ . Specifically, we find the decomposition

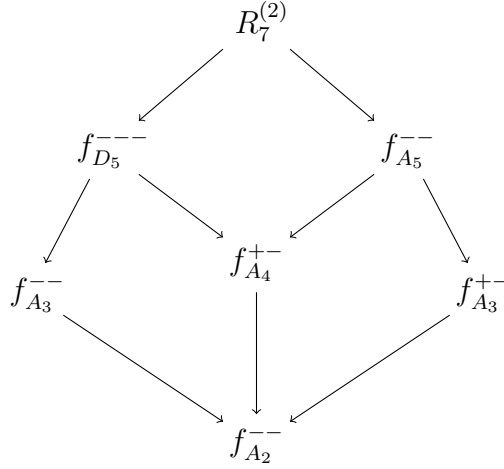
$$f_{A_3 \subset A_5}^{--} = -\frac{1}{8} \sum_{A_5^{--}} f_{A_3}^{+-} \left( x_2 \rightarrow x_3(1+x_4) \rightarrow \frac{x_4 x_5}{1+x_4} \right). \quad (5.15)$$

If we instead choose  $c_1 = -1$ ,  $f_{A_5}^{--}$  becomes  $A_4$ -constructible. Analogous to the  $D_5$  case, there is only a single orbit of  $A_4$  subalgebras in  $A_5$ , implying that this decomposition must take the form

$$\sum_{A_4 \subset A_5} f_{A_4}(x_i \rightarrow \dots) = \sum_{i=0}^7 (\pm 1)^i \sigma_{A_5}^i \left( f_{A_4}(x_i \rightarrow \dots) \right). \quad (5.16)$$

Choosing the minus sign in this sum, we find the unique decomposition

$$f_{A_4 \subset A_5}^{--} = - \sum_{A_4 \subset A_5} f_{A_4}^{+-}(x_1 \rightarrow x_2 \rightarrow x_3 \rightarrow x_4), \quad (5.17)$$



**Figure 7:** The various pathways along which  $R_7^{(2)}$  can be decomposed.

where  $f_{A_4}^{+-}$  is the *same* function that appeared in the  $A_4$  decomposition of  $f_{D_5}^{---}$ . Therefore we find an  $A_4 \subset A_5$  representation of  $R_7^{(2)}$  that bears a striking resemblance to the  $A_4 \subset D_5$  representation found in (5.11):

$$\delta_{2,2}(R_7^{(2)}) = -\frac{1}{20} \sum_{A_5 \subset \text{Gr}(4,7)} \sum_{A_4 \subset A_5} \delta_{2,2}(f_{A_4}^{+-}(x_1 \rightarrow x_2 \rightarrow x_3 \rightarrow x_4)). \quad (5.18)$$

In fact, the two sums are identical. Each of the seven  $A_5$  subalgebras contains eight  $A_4$  subalgebras, resulting in 56 instances of  $f_{A_4}^{+-}$ . These are the same 56 instances of  $f_{A_4}^{+-}$  that contributed in (5.11) (the overall sign difference between these two decompositions results from the orientations of the initial  $D_5$  and  $A_5$  subalgebras chosen in (5.3) and (5.12), respectively). We now turn to a discussion of these overlapping decompositions of  $R_7^{(2)}$ .

### 5.3 The Many Facets of $R_7^{(2)}$

The results of the previous sections are summarized in figure 7. The fact that the two most canonical decompositions of  $R_7^{(2)}$  involve functions that are themselves canonically decomposable in terms of the same  $A_4$  function is highly unexpected. In equation form, this interlinked structure can be summarized as

$$R_7^{(2)} = \frac{1}{20} \sum_{D_5 \subset \text{Gr}(4,7)} \sum_{A_4 \subset D_5} f_{A_4}^{+-}(x_1 \rightarrow x_2 \rightarrow x_3 \rightarrow x_4) + \dots \quad (5.19)$$

$$= -\frac{1}{20} \sum_{A_5 \subset \text{Gr}(4,7)} \sum_{A_4 \subset A_5} f_{A_4}^{+-}(x_1 \rightarrow x_2 \rightarrow x_3 \rightarrow x_4) + \dots \quad (5.20)$$

where the trailing dots indicate that this equality holds up to the contribution of purely classical polylogarithms, and we have left the signs in these sums implicit. As emphasized previously, the required classical contribution is the *same* in both lines, as strict equality holds between the two nested sums over  $f_{A_4}^{+-}$  (including the minus sign in the second line, for the orientations chosen above). This equality does not stem from any complicated polylogarithm identities, but is rather a direct consequence of the arrangement of these subalgebras within  $\text{Gr}(4, 7)$ . It is perhaps suggestive to interpret the equivalence of these two decompositions as coming from the fact that they are different subalgebraic “coverings” of the same object, namely  $\text{Gr}(4, 7)$ .

In addition to this tightly interlinked pair of decompositions, both  $f_{D_5}^{---}$  and  $f_{A_5}^{--}$  admit  $A_3$  decompositions. As discussed above, each of these decompositions makes some property of the remainder function manifest at the expense of others—for instance, the  $f_{A_5}^{--}$  representation makes the vanishing of the nonclassical component of this amplitude in collinear limits manifest term-by-term, whereas the  $A_2 \subset A_3 \subset D_5 \subset \text{Gr}(4, 7)$  representation manifestly satisfies cobracket-level cluster adjacency when phrased in terms of  $f_{A_3}^{--}$  or  $f_{A_3 \subset D_5}^{--}$ . While it might have been hoped that a single decomposition would exhibit all the nice mathematical features of the remainder function at once, it is perhaps better that  $R_7^{(2)}$  “requires” nearly the full breadth of the subalgebra structure of  $\text{Gr}(4, 7)$  in order to express all of its intricate behavior.

The only subalgebra of  $\text{Gr}(4, 7)$  we have not had reason to touch on yet is  $D_4$ . Interestingly, there exists exactly one  $D_4$ -constructible function in  $E_6$  that has signature  $\sigma_{E_6}^+ \tau_{E_6}^+ \mathbb{Z}_{2, E_6}^+$  and respects cobracket-level cluster adjacency. However, this function fails to vanish (or be well-defined) in the  $7 \rightarrow 6$  collinear limit, and so cannot be directly related to  $R_7^{(2)}$ . Still, the uniqueness of this function is itself intriguing and may merit further study.

While the mathematical features of each nonclassical decomposition are easy to check, their physical meaning remains more obscure. It would be interesting to find kinematic limits that cleanly isolate individual instances of the cluster polylogarithms these decompositions make use of, but—as the number of subalgebras in every decomposition is greater than the dimension of seven-particle kinematics—this would require intricate cancellations between the remaining cluster polylogarithms, making such limits hard to engineer. Soft and collinear limits are of particularly little use here, as the nonclassical part of the six-particle amplitude is zero. Moreover, while it is in principle no problem to consider singular limits, only the finite term will survive, as large logs will get projected out by the cobracket. (While we could also study these cluster polylogarithms as full functions, this doesn’t seem particularly meaningful since all the expressions presented in this paper only hold up to undetermined classical contributions.) We correspondingly defer the consideration of further kinematic limits to future

work involving the full analytic form of the eight-particle amplitude [37].

## 6 Conclusion

We have explored the space of nonclassical cluster polylogarithms on  $\text{Gr}(4, 7)$  and its subalgebras, and have shown that—while this space is relatively small—it admits a natural geometric construction that gives rise to interesting sequences of nested functional embeddings. The seven-particle two-loop MHV amplitude seems almost specially engineered to take advantage of this subalgebra structure, insofar as its nonclassical component can be decomposed into cluster polylogarithms defined on almost every type of subalgebra appearing in  $\text{Gr}(4, 7)$ , excluding only  $D_4$ . The decompositions associated with the corank-one subalgebras  $D_5$  and  $A_5$  are especially canonical, as are the further decompositions of the associated  $D_5$  and  $A_5$  cluster polylogarithms into their  $A_4$  subalgebras, as the form of each of these decompositions is uniquely dictated by the automorphism group of the parent algebra. These decompositions identify new cluster polylogarithms that are of special physical interest, supplementing the  $A_2$  and  $A_3$  cluster polylogarithms previously identified in [12]. It is natural to ask whether higher-point MHV amplitudes can also be decomposed into these new cluster polylogarithms, as is known to be the case for the  $A_2$  and  $A_3$  functions; we will take one step towards addressing this question in a forthcoming paper [37].

It would be interesting to try and extend this type of construction to the NMHV sector, where R-invariants also appear in the amplitude. There is a natural way to associate a subalgebra of  $\text{Gr}(4, n)$  to each R-invariant [19], allowing these subalgebras to (potentially) enter the decomposition of the polylogarithms these R-invariants multiply. More generally, although we have not attempted to explore (or even define) the space of cluster polylogarithms beyond weight four, it is hoped that some form of subalgebra constructibility can be extended to higher weight (and higher loops).

While we have almost entirely worked at the level of the nonclassical component of the cobracket, it can also be asked whether these amplitudes are decomposable into functions defined on their subalgebras before applying this projection. It is not hard to show that the symbol of  $R_6^{(2)}$  (as well as that of  $\mathcal{E}_6^{(2)}$ ) is not subalgebra-constructible in this way—the six-particle two-loop MHV amplitude is intrinsically an  $A_3$  polylogarithm. While it is possible that the  $A_2$  and  $A_1 \times A_1$  subalgebras of  $\text{Gr}(4, 6)$  are too simple to permit such a decomposition while larger subalgebras in  $\text{Gr}(4, n > 6)$  would prove sufficiently complex, it seems more likely that the cobracket projection is essential to distilling the part of the amplitude that can be decomposed in this way. The fact that it is the mathematically most complicated component of these amplitudes

that is easiest to compute to all  $n$  (as was done in [13]) still calls out for some physical explanation.

More speculatively, it can be hoped that the pervasiveness of cluster-algebraic structure exhibited by polylogarithmic amplitudes in planar  $\mathcal{N} = 4$  sYM theory points to the existence of some more general combinatorial structure underlying all amplitudes in this theory, including those that involve algebraic roots and/or functions beyond polylogarithms. Much of the cluster-algebraic structure seen in the polylogarithmic case is exposed by the coaction, which (especially in the guise of the symbol) distills these functions down to information about their integration kernels. Notably then, it is possible to formulate coactions on more general periods that appear in quantum field theory, such as elliptic polylogarithms [23, 80, 81, 84–87]. While much of the loop integration technology required for dealing with generic Feynman integrals remains to be developed, we can already begin to study the higher-genus and higher-dimensional varieties that appear in the integration contours contributing to planar  $\mathcal{N} = 4$  sYM theory (see for example [29–31, 88]). Hopefully, understanding these (seemingly always Calabi-Yau) geometries will provide salient hints for how the geometric and algebraic picture provided by cluster algebras can be generalized beyond the case of polylogarithms.

## Acknowledgements

We thank Jacob Bourjaily, Lance Dixon, James Drummond, Henriette Elvang, Matt von Hippel, Marcus Spradlin, and Matthias Wilhelm for illuminating discussion. We also thank Jacob Bourjaily, James Drummond, Andrea Orta, and Marcus Spradlin for comments on the manuscript. This project has received funding from an ERC Starting Grant (No. 757978), a grant from the Villum Fonden (AJM), and the support of a Van Loo Postdoctoral Fellowship (JG). The authors are grateful to the Kavli Institute for Theoretical Physics (National Science Foundation grant NSF PHY11-25915) for hospitality.



## A Counting Subalgebras of Finite Cluster Algebras

In this appendix we catalog the subalgebra structure of  $\text{Gr}(4, 7) \simeq E_6$  and all its finite connected subalgebras, namely  $A_2$ ,  $A_3$ ,  $A_4$ ,  $D_4$ ,  $A_5$ , and  $D_5$ . We include both the number of subalgebras and subpolytopes (of the cluster polytope) that occur in each algebra. We only consider subalgebras to be distinct when they differ in at least one of their mutable nodes, but count identical subalgebras as distinct subpolytopes when they are connected to different frozen nodes. Note also that we consider the  $\mathcal{X}$ -coordinates  $x$  and  $1/x$  to be distinct.

$A_2$

Clusters: 5 —  $\mathcal{A}$ -coordinates: 5 —  $\mathcal{X}$ -coordinates: 10

$A_3$

Clusters: 14 —  $\mathcal{A}$ -coordinates: 9 —  $\mathcal{X}$ -coordinates: 30

Type	Subpolytopes	Subalgebras
$A_2$	6	6
$A_1 \times A_1$	3	3

$A_4$

Clusters: 42 —  $\mathcal{A}$ -coordinates: 14 —  $\mathcal{X}$ -coordinates: 70

Type	Subpolytopes	Subalgebras
$A_2$	28	21
$A_1 \times A_1$	28	28
$A_3$	7	7
$A_2 \times A_1$	7	7
$A_1 \times A_1 \times A_1$	0	0

$D_4$

Clusters: 50 —  $\mathcal{A}$ -coordinates: 16 —  $\mathcal{X}$ -coordinates: 104

Type	Subpolytopes	Subalgebras
$A_2$	36	36
$A_1 \times A_1$	30	18
$A_3$	12	12
$A_2 \times A_1$	0	0
$A_1 \times A_1 \times A_1$	4	4

$A_5$

Clusters: 132 —  $\mathcal{A}$ -coordinates: 20 —  $\mathcal{X}$ -coordinates: 140

Type	Subpolytopes	Subalgebras
$A_2$	120	56
$A_1 \times A_1$	180	144
$A_3$	36	28
$A_2 \times A_1$	72	72
$A_1 \times A_1 \times A_1$	12	12
$D_4$	0	0
$A_4$	8	8
$A_3 \times A_1$	8	8
$A_2 \times A_2$	4	4
$A_2 \times A_1 \times A_1$	0	0
$A_1 \times A_1 \times A_1 \times A_1$	0	0

$D_5$

Clusters: 182 —  $\mathcal{A}$ -coordinates: 25 —  $\mathcal{X}$ -coordinates: 260

Type	Subpolytopes	Subalgebras
$A_2$	180	125
$A_1 \times A_1$	230	145
$A_3$	70	65
$A_2 \times A_1$	60	50
$A_1 \times A_1 \times A_1$	30	30
$D_4$	5	5
$A_4$	10	10
$A_3 \times A_1$	5	5
$A_2 \times A_2$	0	0
$A_2 \times A_1 \times A_1$	5	5
$A_1 \times A_1 \times A_1 \times A_1$	0	0

$E_6$

Clusters: 833 —  $\mathcal{A}$ -coordinates: 42 —  $\mathcal{X}$ -coordinates: 770

Type	Subpolytopes	Subalgebras
$A_2$	1071	504
$A_1 \times A_1$	1785	833
$A_3$	476	364
$A_2 \times A_1$	714	490
$A_1 \times A_1 \times A_1$	357	357
$D_4$	35	35
$A_4$	112	98
$A_3 \times A_1$	112	112
$A_2 \times A_2$	21	14
$A_2 \times A_1 \times A_1$	119	119
$A_1 \times A_1 \times A_1 \times A_1$	0	0
$D_5$	14	14
$A_5$	7	7
$D_4 \times A_1$	0	0
$A_4 \times A_1$	14	14
$A_3 \times A_2$	0	0
$A_3 \times A_1 \times A_1$	0	0
$A_2 \times A_2 \times A_1$	7	7
$A_2 \times A_1 \times A_1 \times A_1$	0	0
$A_1 \times A_1 \times A_1 \times A_1 \times A_1$	0	0

## B Cobracket Spaces in Finite Cluster Algebras

In this appendix we tabulate the number of independent weight-four cluster polylogarithms that have a nonzero Lie cobracket on  $\text{Gr}(4, 7) \simeq E_6$  and its suablagebras. In

Type	Nonclassical Cobrackets	Automorphism Signature			
		$\sigma^+\tau^+$	$\sigma^+\tau^-$	$\sigma^-\tau^+$	$\sigma^-\tau^-$
$A_2$	1 (0)	0	1 (0)	0	0
$A_3$	6 (1)	0	1 (0)	0	1 (1)
$A_4$	21 (6)	0	3 (0)	0	0
$D_4$	34 (9)				
$A_5$	56 (21)	2 (1)	5 (1)	2 (0)	5 (3)
$D_5$	116 (42)				
$E_6$	448 (195)				

$D_4$  Automorphism Signatures

	$\frac{\sigma_{\mathfrak{D}_4}^+ \tau_{\mathfrak{D}_4}^+}{\sigma_{S_3}^+ \quad \sigma_{S_3}^-}$	$\frac{\sigma_{\mathfrak{D}_4}^+ \tau_{\mathfrak{D}_4}^-}{\sigma_{S_3}^+ \quad \sigma_{S_3}^-}$	$\frac{\sigma_{\mathfrak{D}_4}^- \tau_{\mathfrak{D}_4}^+}{\sigma_{S_3}^+ \quad \sigma_{S_3}^-}$	$\frac{\sigma_{\mathfrak{D}_4}^- \tau_{\mathfrak{D}_4}^-}{\sigma_{S_3}^+ \quad \sigma_{S_3}^-}$
$\tau_{S_3}^+$	0	0	1 (0)	0
$\tau_{S_3}^-$	1 (0)	0	1 (1)	0

$D_5$  Automorphism Signatures

	$\frac{\sigma_{\mathfrak{D}_5}^+ \tau_{\mathfrak{D}_5}^+}{\mathbb{Z}_2^+ \quad \mathbb{Z}_2^-}$	$\frac{\sigma_{\mathfrak{D}_5}^+ \tau_{\mathfrak{D}_5}^-}{\mathbb{Z}_2^+ \quad \mathbb{Z}_2^-}$	$\frac{\sigma_{\mathfrak{D}_5}^- \tau_{\mathfrak{D}_5}^+}{\mathbb{Z}_2^+ \quad \mathbb{Z}_2^-}$	$\frac{\sigma_{\mathfrak{D}_5}^- \tau_{\mathfrak{D}_5}^-}{\mathbb{Z}_2^+ \quad \mathbb{Z}_2^-}$
	5 (2)	0	0	3 (1)
	9 (2)	0	0	7 (5)

$E_6$  Automorphism Signatures

	$\frac{\sigma_{\mathfrak{D}_{14}}^+ \tau_{\mathfrak{D}_{14}}^+}{\mathbb{Z}_2^+ \quad \mathbb{Z}_2^-}$	$\frac{\sigma_{\mathfrak{D}_{14}}^+ \tau_{\mathfrak{D}_{14}}^-}{\mathbb{Z}_2^+ \quad \mathbb{Z}_2^-}$	$\frac{\sigma_{\mathfrak{D}_{14}}^- \tau_{\mathfrak{D}_{14}}^+}{\mathbb{Z}_2^+ \quad \mathbb{Z}_2^-}$	$\frac{\sigma_{\mathfrak{D}_{14}}^- \tau_{\mathfrak{D}_{14}}^-}{\mathbb{Z}_2^+ \quad \mathbb{Z}_2^-}$
	12 (6)	14 (6)	21 (6)	17 (9)
	0	0	0	0

**Table 5:** The number of nonclassical weight-four cluster polylogarithms on various finite cluster algebras, prior to consideration of their automorphism group and after requiring specific automorphism signatures. The number of polylogarithms that can also be made to satisfy cobracket-level cluster adjacency is given in parentheses.

Type	Integrable Cobrackets	Automorphism Signature			
		$\sigma^+ \tau^+$	$\sigma^+ \tau^-$	$\sigma^- \tau^+$	$\sigma^- \tau^-$
$A_2$	6 (5)	1 (1)	1 (0)	0	0
$A_3$	21 (16)	2 (2)	2 (1)	0	3 (3)
$A_4$	56 (41)	4 (4)	4 (1)	0	0
$D_4$	86 (61)				
$A_5$	126 (91)	8 (7)	8 (4)	5 (3)	11 (9)
$D_5$	246 (172)				
$E_6$	833 (580)				

$D_4$ Automorphism Signatures							
$\frac{\sigma_{\mathfrak{D}_4}^+ \tau_{\mathfrak{D}_4}^+}{\sigma_{S_3}^+ \quad \sigma_{S_3}^-}$		$\frac{\sigma_{\mathfrak{D}_4}^+ \tau_{\mathfrak{D}_4}^-}{\sigma_{S_3}^+ \quad \sigma_{S_3}^-}$		$\frac{\sigma_{\mathfrak{D}_4}^- \tau_{\mathfrak{D}_4}^+}{\sigma_{S_3}^+ \quad \sigma_{S_3}^-}$		$\frac{\sigma_{\mathfrak{D}_4}^- \tau_{\mathfrak{D}_4}^-}{\sigma_{S_3}^+ \quad \sigma_{S_3}^-}$	
$\tau_{S_3}^+$	3 (3)	$\tau_{S_3}^+$	3 (1)	$\tau_{S_3}^+$	3 (2)	$\tau_{S_3}^+$	3 (2)
$\tau_{S_3}^-$	0	$\tau_{S_3}^-$	0	$\tau_{S_3}^-$	0	$\tau_{S_3}^-$	0
	1 (0)		1 (1)		2 (2)		0
$D_5$ Automorphism Signatures							
$\frac{\sigma_{\mathfrak{D}_5}^+ \tau_{\mathfrak{D}_5}^+}{\mathbb{Z}_2^+ \quad \mathbb{Z}_2^-}$		$\frac{\sigma_{\mathfrak{D}_5}^+ \tau_{\mathfrak{D}_5}^-}{\mathbb{Z}_2^+ \quad \mathbb{Z}_2^-}$		$\frac{\sigma_{\mathfrak{D}_5}^- \tau_{\mathfrak{D}_5}^+}{\mathbb{Z}_2^+ \quad \mathbb{Z}_2^-}$		$\frac{\sigma_{\mathfrak{D}_5}^- \tau_{\mathfrak{D}_5}^-}{\mathbb{Z}_2^+ \quad \mathbb{Z}_2^-}$	
	16 (13)		16 (9)		0		0
	0		0		5 (3)		13 (11)
$E_6$ Automorphism Signatures							
$\frac{\sigma_{\mathfrak{D}_{14}}^+ \tau_{\mathfrak{D}_{14}}^+}{\mathbb{Z}_2^+ \quad \mathbb{Z}_2^-}$		$\frac{\sigma_{\mathfrak{D}_{14}}^+ \tau_{\mathfrak{D}_{14}}^-}{\mathbb{Z}_2^+ \quad \mathbb{Z}_2^-}$		$\frac{\sigma_{\mathfrak{D}_{14}}^- \tau_{\mathfrak{D}_{14}}^+}{\mathbb{Z}_2^+ \quad \mathbb{Z}_2^-}$		$\frac{\sigma_{\mathfrak{D}_{14}}^- \tau_{\mathfrak{D}_{14}}^-}{\mathbb{Z}_2^+ \quad \mathbb{Z}_2^-}$	
	32 (26)		32 (17)		0		0
	25 (17)		30 (22)		0		0

**Table 6:** The number of weight-four cluster polylogarithms on various finite cluster algebras that have nonzero cobrackets, prior to consideration of their automorphism group and after requiring specific automorphism signatures. The number of polylogarithms that can also be made to satisfy cobracket-level cluster adjacency is given in parentheses.

table 5 we first record the number of cluster polylogarithms that have a nonzero  $\delta_{2,2}$  cobracket component, as considered in section 4. We tabulate the number of functions on each cluster algebra both before and after requiring a specific automorphism

signature, and include the number of functions that also respect cobracket-level cluster adjacency in parentheses. (Some of these numbers can also be found in [32].) In table 6 we record the same information, but for all weight-four cluster polylogarithms that have any nonzero cobracket component.

## References

- [1] N. Arkani-Hamed, J. L. Bourjaily, F. Cachazo, A. B. Goncharov, A. Postnikov, and J. Trnka, *Grassmannian Geometry of Scattering Amplitudes*. Cambridge University Press, 2016.
- [2] J. Golden, A. B. Goncharov, M. Spradlin, C. Vergu, and A. Volovich, *Motivic Amplitudes and Cluster Coordinates*, *JHEP* **1401** (2014) 091, [[arXiv:1305.1617](#)].
- [3] J. L. Bourjaily, *Positroids, Plabic Graphs, and Scattering Amplitudes in Mathematica*, [arXiv:1212.6974](#).
- [4] L. J. Dixon, J. M. Drummond, M. von Hippel, and J. Pennington, *Hexagon functions and the three-loop remainder function*, *JHEP* **1312** (2013) 049, [[arXiv:1308.2276](#)].
- [5] L. J. Dixon, J. M. Drummond, C. Duhr, and J. Pennington, *The four-loop remainder function and multi-Regge behavior at NNLLA in planar  $\mathcal{N} = 4$  super-Yang-Mills theory*, *JHEP* **1406** (2014) 116, [[arXiv:1402.3300](#)].
- [6] J. M. Drummond, G. Papathanasiou, and M. Spradlin, *A Symbol of Uniqueness: The Cluster Bootstrap for the 3-Loop MHV Heptagon*, *JHEP* **03** (2015) 072, [[arXiv:1412.3763](#)].
- [7] S. Caron-Huot, L. J. Dixon, A. McLeod, and M. von Hippel, *Bootstrapping a Five-Loop Amplitude Using Steinmann Relations*, *Phys. Rev. Lett.* **117** (2016), no. 24 241601, [[arXiv:1609.00669](#)].
- [8] L. J. Dixon, J. Drummond, T. Harrington, A. J. McLeod, G. Papathanasiou, and M. Spradlin, *Heptagons from the Steinmann Cluster Bootstrap*, *JHEP* **02** (2017) 137, [[arXiv:1612.08976](#)].
- [9] S. Caron-Huot, *Superconformal symmetry and two-loop amplitudes in planar  $\mathcal{N} = 4$  super Yang-Mills*, *JHEP* **1112** (2011) 066, [[arXiv:1105.5606](#)].
- [10] J. Golden and M. Spradlin, *An analytic result for the two-loop seven-point MHV amplitude in  $\mathcal{N} = 4$  SYM*, *JHEP* **1408** (2014) 154, [[arXiv:1406.2055](#)].
- [11] J. Golden and M. Spradlin, *The differential of all two-loop MHV amplitudes in  $\mathcal{N} = 4$  Yang-Mills theory*, *JHEP* **1309** (2013) 111, [[arXiv:1306.1833](#)].
- [12] J. Golden, M. F. Paulos, M. Spradlin, and A. Volovich, *Cluster Polylogarithms for Scattering Amplitudes*, *J. Phys.* **A47** (2014), no. 47 474005, [[arXiv:1401.6446](#)].
- [13] J. Golden and M. Spradlin, *A Cluster Bootstrap for Two-Loop MHV Amplitudes*, *JHEP* **02** (2015) 002, [[arXiv:1411.3289](#)].
- [14] J. Drummond, J. Foster, and Ö. Gürdoğan, *Cluster Adjacency Properties of Scattering Amplitudes in  $N = 4$  Supersymmetric Yang-Mills Theory*, *Phys. Rev. Lett.* **120** (2018), no. 16 161601, [[arXiv:1710.10953](#)].

- [15] S. Caron-Huot and S. He, *Jumpstarting the All-Loop S-Matrix of Planar  $\mathcal{N} = 4$  Super Yang-Mills*, *JHEP* **1207** (2012) 174, [[arXiv:1112.1060](#)].
- [16] L. J. Dixon and M. von Hippel, *Bootstrapping an NMHV amplitude through three loops*, *JHEP* **1410** (2014) 65, [[arXiv:1408.1505](#)].
- [17] L. J. Dixon, M. von Hippel, and A. J. McLeod, *The four-loop six-gluon NMHV ratio function*, *JHEP* **01** (2016) 053, [[arXiv:1509.08127](#)].
- [18] L. J. Dixon, M. von Hippel, A. J. McLeod, and J. Trnka, *Multi-loop positivity of the planar  $\mathcal{N} = 4$  SYM six-point amplitude*, *JHEP* **02** (2017) 112, [[arXiv:1611.08325](#)].
- [19] J. Drummond, J. Foster, and Ö. Gürdoğan, *Cluster adjacency beyond MHV*, [arXiv:1810.08149](#).
- [20] J. M. Drummond, J. M. Henn, and J. Trnka, *New differential equations for on-shell loop integrals*, *JHEP* **04** (2011) 083, [[arXiv:1010.3679](#)].
- [21] J. L. Bourjaily, A. J. McLeod, M. von Hippel, and M. Wilhelm, *Rationalizing Loop Integration*, *JHEP* **08** (2018) 184, [[arXiv:1805.10281](#)].
- [22] J. Henn, E. Herrmann, and J. Parra-Martinez, *Bootstrapping two-loop Feynman integrals for planar  $N=4$  sYM*, [arXiv:1806.06072](#).
- [23] S. Caron-Huot, L. J. Dixon, M. von Hippel, A. J. McLeod, and G. Papathanasiou, *The Double Pentaladder Integral to All Orders*, *JHEP* **07** (2018) 170, [[arXiv:1806.01361](#)].
- [24] I. Prlina, M. Spradlin, J. Stankowicz, S. Stanojevic, and A. Volovich, *All-Helicity Symbol Alphabets from Unwound Amplituhedra*, *JHEP* **05** (2018) 159, [[arXiv:1711.11507](#)].
- [25] M. F. Paulos, M. Spradlin, and A. Volovich, *Mellin Amplitudes for Dual Conformal Integrals*, *JHEP* **08** (2012) 072, [[arXiv:1203.6362](#)].
- [26] S. Caron-Huot and K. J. Larsen, *Uniqueness of two-loop master contours*, *JHEP* **10** (2012) 026, [[arXiv:1205.0801](#)].
- [27] J. L. Bourjaily and J. Trnka, *Local Integrand Representations of All Two-Loop Amplitudes in Planar SYM*, *JHEP* **08** (2015) 119, [[arXiv:1505.05886](#)].
- [28] J. L. Bourjaily, E. Herrmann, and J. Trnka, *Prescriptive Unitarity*, *JHEP* **06** (2017) 059, [[arXiv:1704.05460](#)].
- [29] J. L. Bourjaily, A. J. McLeod, M. Spradlin, M. von Hippel, and M. Wilhelm, *Elliptic Double-Box Integrals: Massless Scattering Amplitudes beyond Polylogarithms*, *Phys. Rev. Lett.* **120** (2018), no. 12 121603, [[arXiv:1712.02785](#)].
- [30] J. L. Bourjaily, Y.-H. He, A. J. McLeod, M. Von Hippel, and M. Wilhelm, *Traintracks through Calabi-Yau Manifolds: Scattering Amplitudes beyond Elliptic Polylogarithms*, *Phys. Rev. Lett.* **121** (2018), no. 7 071603, [[arXiv:1805.09326](#)].



- [31] J. L. Bourjaily, A. J. McLeod, M. von Hippel, and M. Wilhelm, *A (Bounded) Bestiary of Feynman Integral Calabi-Yau Geometries*, [arXiv:1810.07689](#).
- [32] T. Harrington and M. Spradlin, *Cluster Functions and Scattering Amplitudes for Six and Seven Points*, *JHEP* **07** (2017) 016, [[arXiv:1512.07910](#)].
- [33] S. Fomin and A. Zelevinsky, *Cluster algebras I: Foundations*, *J. Am. Math. Soc.* **15** (2002), no. 2 497–529, [[math/0104151](#)].
- [34] N. Arkani-Hamed, J. L. Bourjaily, F. Cachazo, A. Postnikov, and J. Trnka, *On-Shell Structures of MHV Amplitudes Beyond the Planar Limit*, *JHEP* **06** (2015) 179, [[arXiv:1412.8475](#)].
- [35] V. V. Fock and A. B. Goncharov, *Cluster ensembles, quantization and the dilogarithm*, *Ann. Sci. Éc. Norm. Supér. (4)* **42** (2009), no. 6 865–930, [[math/0311245](#)].
- [36] M. Gekhtman, M. Shapiro, and A. Vainshtein, *Cluster algebras and Poisson geometry*, *Mosc. Math. J.* **3** (2003), no. 3 899, [[math/0208033](#)].
- [37] J. Golden, A. J. McLeod, in progress.
- [38] C. Vergu, *Polylogarithm identities, cluster algebras and the  $N=4$  supersymmetric theory*, 2015. [arXiv:1512.08113](#).
- [39] J. S. Scott, *Grassmannians and cluster algebras*, *Proc. Lond. Math. Soc., III. Ser.* **92** (2006), no. 2 345–380.
- [40] L. J. Dixon, *A brief introduction to modern amplitude methods*, in *Proceedings, 2012 European School of High-Energy Physics (ESHEP 2012): La Pommeraye, Anjou, France, June 06-19, 2012*, pp. 31–67, 2014. [arXiv:1310.5353](#).
- [41] H. Elvang and Y.-t. Huang, *Scattering Amplitudes*, [arXiv:1308.1697](#).
- [42] Z. Bern, L. J. Dixon, and V. A. Smirnov, *Iteration of planar amplitudes in maximally supersymmetric Yang-Mills theory at three loops and beyond*, *Phys. Rev.* **D72** (2005) 085001, [[hep-th/0505205](#)].
- [43] J. M. Drummond, J. Henn, G. P. Korchemsky, and E. Sokatchev, *Conformal Ward identities for Wilson loops and a test of the duality with gluon amplitudes*, *Nucl. Phys.* **B826** (2010) 337–364, [[arXiv:0712.1223](#)].
- [44] Z. Bern, L. Dixon, D. Kosower, R. Roiban, M. Spradlin, C. Vergu, and A. Volovich, *The Two-Loop Six-Gluon MHV Amplitude in Maximally Supersymmetric Yang-Mills Theory*, *Phys.Rev.* **D78** (2008) 045007, [[arXiv:0803.1465](#)].
- [45] J. Drummond, J. Henn, G. Korchemsky, and E. Sokatchev, *Hexagon Wilson loop = six-gluon MHV amplitude*, *Nucl.Phys.* **B815** (2009) 142–173, [[arXiv:0803.1466](#)].
- [46] J. M. Drummond, J. Henn, V. A. Smirnov, and E. Sokatchev, *Magic identities for conformal four-point integrals*, *JHEP* **01** (2007) 064, [[hep-th/0607160](#)].

- [47] Z. Bern, M. Czakon, L. J. Dixon, D. A. Kosower, and V. A. Smirnov, *The Four-Loop Planar Amplitude and Cusp Anomalous Dimension in Maximally Supersymmetric Yang-Mills Theory*, *Phys. Rev.* **D75** (2007) 085010, [[hep-th/0610248](#)].
- [48] Z. Bern, J. Carrasco, H. Johansson, and D. Kosower, *Maximally supersymmetric planar Yang-Mills amplitudes at five loops*, *Phys. Rev.* **D76** (2007) 125020, [[arXiv:0705.1864](#)].
- [49] L. F. Alday and J. M. Maldacena, *Gluon scattering amplitudes at strong coupling*, *JHEP* **0706** (2007) 064, [[arXiv:0705.0303](#)].
- [50] J. M. Drummond, J. Henn, G. P. Korchemsky, and E. Sokatchev, *Dual superconformal symmetry of scattering amplitudes in  $N=4$  super-Yang-Mills theory*, *Nucl. Phys.* **B828** (2010) 317–374, [[arXiv:0807.1095](#)].
- [51] S. Fomin and A. Zelevinsky, *Cluster algebras II: Finite type classification*, *Invent. Math.* **154** (2003), no. 1 63–121, [[math/0208229](#)].
- [52] D. Parker, A. Scherlis, M. Spradlin, and A. Volovich, *Hedgehog Bases for  $A_n$  Cluster Polylogarithms and An Application to Six-Point Amplitudes*, *JHEP* **11** (2015) 136, [[arXiv:1507.01950](#)].
- [53] M. Sherman-Bennett, *Combinatorics of  $\mathcal{X}$ -variables in finite type cluster algebras*, [arXiv:1803.02492](#).
- [54] W. Chang and B. Zhu, *Cluster automorphism groups of cluster algebras of finite type*, [arXiv:1506.01950](#).
- [55] K.-T. Chen, *Iterated path integrals*, *Bull. Amer. Math. Soc.* **83** (1977), no. 5 831–879.
- [56] F. C. Brown, *Multiple zeta values and periods of moduli spaces  $\overline{\mathfrak{M}}_{0,n}(\mathbb{R})$* , *Annales Sci. Ecole Norm. Sup.* **42** (2009) 371, [[math/0606419](#)].
- [57] A. B. Goncharov, *A simple construction of Grassmannian polylogarithms*, *Adv. Math.* **241** (2013) 79–102, [[arXiv:0908.2238](#)].
- [58] S. Caron-Huot, L. J. Dixon, M. von Hippel, A. J. McLeod, G. Papathanasiou, in progress.
- [59] C. Duhr, H. Gangl, and J. R. Rhodes, *From polygons and symbols to polylogarithmic functions*, *JHEP* **1210** (2012) 075, [[arXiv:1110.0458](#)].
- [60] C. Duhr, *Mathematical aspects of scattering amplitudes*, [arXiv:1411.7538](#).
- [61] H. Gangl, *Multiple polylogarithms in weight 4*, [arXiv:1609.05557](#).
- [62] F. Brown, *Mixed Tate motives over  $\mathbb{Z}$* , [arXiv:1102.1312](#).
- [63] F. Brown, *Feynman Amplitudes and Cosmic Galois group*, [arXiv:1512.06409](#).
- [64] A. B. Goncharov, *Multiple polylogarithms and mixed Tate motives*, [math/0103059](#).
- [65] A. B. Goncharov, M. Spradlin, C. Vergu, and A. Volovich, *Classical Polylogarithms for*

- Amplitudes and Wilson Loops*, *Phys.Rev.Lett.* **105** (2010) 151605, [[arXiv:1006.5703](#)].
- [66] F. Brown, *On the decomposition of motivic multiple zeta values*, [arXiv:1102.1310](#).
  - [67] C. Duhr, *Hopf algebras, coproducts and symbols: an application to Higgs boson amplitudes*, *JHEP* **1208** (2012) 043, [[arXiv:1203.0454](#)].
  - [68] S. J. Bloch, *Higher regulators, algebraic K-theory, and zeta functions of elliptic curves*. American Mathematical Society (AMS), 2000.
  - [69] A. Suslin, *K3 of a field and the bloch group*, *Proc. Stekov Inst. Math.* **183** (1990) 217–239.
  - [70] A. B. Goncharov, *Polylogarithms and motivic Galois groups*, in *Motives (Seattle, WA, 1991)*, vol. 55 of *Proc. Sympos. Pure Math.*, pp. 43–96. Amer. Math. Soc., Providence, RI, 1994.
  - [71] N. Dan, *Sur la conjecture de Zagier pour  $n=4$* , *ArXiv e-prints* (Sept., 2008) [[arXiv:0809.3984](#)].
  - [72] H. Gangl, *The Grassmannian complex and Goncharov’s motivic complex in weight 4*, *ArXiv e-prints* (Jan., 2018) [[arXiv:1801.07816](#)].
  - [73] A. B. Goncharov and D. Rudenko, *Motivic correlators, cluster varieties and Zagier’s conjecture on  $\zeta(F,4)$* , [arXiv:1803.08585](#).
  - [74] L. F. Alday, D. Gaiotto, and J. Maldacena, *Thermodynamic Bubble Ansatz*, *JHEP* **1109** (2011) 032, [[arXiv:0911.4708](#)].
  - [75] G. Yang, *A simple collinear limit of scattering amplitudes at strong coupling*, *JHEP* **03** (2011) 087, [[arXiv:1006.3306](#)].
  - [76] N. Beisert, B. Eden, and M. Staudacher, *Transcendentality and Crossing*, *J. Stat. Mech.* **0701** (2007) P01021, [[hep-th/0610251](#)].
  - [77] O. Steinmann, *Über den Zusammenhang zwischen den Wightmanfunktionen und der retardierten Kommutatoren*, *Helv. Physica Acta* **33** (1960) 257.
  - [78] O. Steinmann, *Wightman-Funktionen und retardierten Kommutatoren. II*, *Helv. Physica Acta* **33** (1960) 347.
  - [79] K. E. Cahill and H. P. Stapp, *Optical Theorems and Steinmann Relations*, *Annals Phys.* **90** (1975) 438.
  - [80] J. Broedel, C. Duhr, F. Dulat, and L. Tancredi, *Elliptic polylogarithms and iterated integrals on elliptic curves. Part I: general formalism*, *JHEP* **05** (2018) 093, [[arXiv:1712.07089](#)].
  - [81] J. Broedel, C. Duhr, F. Dulat, B. Penante, and L. Tancredi, *Elliptic Feynman integrals and pure functions*, [arXiv:1809.10698](#).

- [82] G. Yang, *Scattering amplitudes at strong coupling for  $4K$  gluons*, *JHEP* **12** (2010) 082, [[arXiv:1004.3983](#)].
- [83] Z. Bern, L. J. Dixon, D. C. Dunbar, and D. A. Kosower, *One-loop  $n$ -point gauge theory amplitudes, unitarity and collinear limits*, *Nucl.Phys.* **B425** (1994) 217–260, [[hep-ph/9403226](#)].
- [84] F. Brown, *Notes on Motivic Periods*, [arXiv:1512.06410](#).
- [85] S. Abreu, R. Britto, C. Duhr, and E. Gardi, *Algebraic Structure of Cut Feynman Integrals and the Diagrammatic Coaction*, *Phys. Rev. Lett.* **119** (2017), no. 5 051601, [[arXiv:1703.05064](#)].
- [86] S. Abreu, R. Britto, C. Duhr, and E. Gardi, *Diagrammatic Hopf algebra of cut Feynman integrals: the one-loop case*, *JHEP* **12** (2017) 090, [[arXiv:1704.07931](#)].
- [87] J. Broedel, C. Duhr, F. Dulat, B. Penante, and L. Tancredi, *Elliptic symbol calculus: from elliptic polylogarithms to iterated integrals of Eisenstein series*, *JHEP* **08** (2018) 014, [[arXiv:1803.10256](#)].
- [88] F. C. S. Brown, *On the periods of some Feynman integrals*, [arXiv:0910.0114](#).