JHEP

Published for SISSA by Springer

Received: *August 12, 2015*
Accepted: *October 27, 2015*
Published: *November 20, 2015*

JHEP11(2015)136

Hedgehog bases for A_n cluster polylogarithms and an application to six-point amplitudes

Daniel E. Parker, Adam Scherlis, Marcus Spradlin and Anastasia Volovich

Department of Physics, Brown University,
Providence RI 02912, U.S.A.

E-mail: daniel_parker@berkeley.edu, scherlis@stanford.edu, marcus_spradlin@brown.edu, anastasia_volovich@brown.edu

Abstract: Multi-loop scattering amplitudes in $\mathcal{N} = 4$ Yang-Mills theory possess cluster algebra structure. In order to develop a computational framework which exploits this connection, we show how to construct bases of Goncharov polylogarithm functions, at any weight, whose symbol alphabet consists of cluster coordinates on the A_n cluster algebra. Using such a basis we present a new expression for the 2-loop 6-particle NMHV amplitude which makes some of its cluster structure manifest.

Keywords: Supersymmetric gauge theory, Scattering Amplitudes

ArXiv ePrint: 1507.01950

doi:10.1007/JHEP11(2015)136

Contents

JHEP11(2015)136

1 Introduction

In a series of recent papers following [1] it has been realized that (all known) multi-loop n-particle scattering amplitudes of planar $\mathcal{N} = 4$ super-Yang-Mills (SYM) theory possess special properties that are intimately connected to mathematical structures known as cluster algebras. The most basic aspect of this connection is that amplitudes are linear combinations of generalized polylogarithm functions whose symbol arguments are cluster coordinates on the $\mathrm{Gr}(4, n)$ Grassmannian cluster algebra.[1]

[1] This aspect is the focus of our paper, but other connections have been observed for particular amplitudes. For example, there is a tight connection between the cobracket of motivic 2-loop MHV amplitudes and the Poisson structure on the underlying cluster algebra, which has been explored in [1, 3, 4].

This connection between scattering amplitudes and cluster algebras is undoubtedly related to a similar cluster structure that has been observed at the level of integrands in [2], though the precise connection has yet to be made. Nevertheless, the observed cluster structure of integrated amplitudes has already helped to facilitate the computation of new expressions for various quantities associated to amplitudes (see for example [3–7]). In parallel, work by Dixon, Drummond, and collaborators has resulted in spectacular progress in determining 6-particle amplitudes via a bootstrap approach (see [8–12], or the review [13]) utilizing input from the OPE of null Wilson loops (see for example [14–19]).

Typically, results in SYM theory take the form of colossal linear combinations of generalized polylogarithm functions. These special functions satisfy a huge number of functional identities: shuffle identities, stuffle identities, the Abel identity, the trilogarithm identity of [1], and many others. These make generalized polylogarithms notoriously difficult to work with. Moreover, with so many identities, there are a multitude of possible ways to write the same formula. In general there is no "best" way to write a given expression, nor is it even clear how one ought to define "best" — perhaps the shortest expression, or one where certain physical or mathematical properties are manifest.

Large progress towards finding canonical bases for generalized polylogarithms has been made by Brown in [20] (see also [21] for some applications) and employed by Dixon et al. in their 6-particle bootstrap program. In this paper we demonstrate a natural way to "clusterize" Brown's basis of polylogarithm functions. Namely, we show how to generate, at any weight, a basis of generalized polylogarithm functions whose symbols are manifestly expressible in terms of cluster coordinates on the A_n cluster algebra. We call these "hedgehog" bases because they are naturally associated to certain spiny structures in the A_n exchange graph. Hedgehog bases provide an almost canonical way to write expressions for 6-particle MHV and NMHV amplitudes, presumably at any loop order. Compared to using other bases that have been considered in the literature, hedgehog bases have the theoretically-pleasing advantage of making some of the cluster structure of such amplitudes manifest, as well as the practical benefit of allowing notably shorter expressions. The latter feature echoes a common theme in the amplitudes program: identifying underlying mathematical structure and improving computational efficiency go hand in hand.

Section 2 briefly reviews the necessary mathematical technology of polylogarithms, cluster algebras, and scattering amplitudes. Section 3 introduces the idea of a "hedgehog" for a cluster algebra, and sketches the rigorous proof (with details relegated to an appendix) that they can be fashioned into a basis for polylogarithms. Section 4 presents, as an application of this technology, a construction of a hedgehog-basis representation for the the 2-loop 6-particle NMHV amplitude.[2]

2 Review

Polylogarithms and cluster algebras are each subjects unto themselves. Thus this section is not an all-encompassing review, but rather a brief reminder of some of the mathematical technology needed for the rest of the paper, together with citations where the curious

[2]Our result is included as an ancillary file with the arXiv submission.

JHEP11(2015)136

reader may find additional details. We also review the relevant aspects of the connection between Grassmannian cluster algebras and scattering amplitudes in SYM theory.

2.1 Generalized polylogarithms

Polylogarithms are a broad class of special functions that generalize the logarithm. More details on the material in this section may be found in the recent review [22].

Recall that the ordinary logarithm can be written as $\log z = \int_0^z \frac{dt}{t}$. Generalizing this to an iterated integral of the type first studied systematically by Chen [23] gives the *weight-k Goncharov polylogarithm* [24]:

$$G(a_1, \ldots, a_k; z) = \int_0^z \frac{dt}{t - a_1} G(a_2, \ldots, a_k; t)\,, \qquad G(z) = 1\,, \tag{2.1}$$

with the special case

$$G(\underbrace{0, \ldots, 0}_{k}; z) = \frac{1}{k!} \log^k z\,. \tag{2.2}$$

In general, $a_1, \ldots, a_k$ are valued in the complex numbers with $z \in \mathbb{C} \setminus \{a_1, \ldots, a_k\}$, and one should specify a contour of integration. We will see that for scattering amplitudes in a certain domain these variables are all real-valued, and there is a natural ordering which allows one to take the "naive" contour straight along the real axis. A large class of L-loop amplitudes in SYM theory, including at least all MHV and NMHV amplitudes, are expected to be expressible as linear combinations of weight-$2L$ polylogarithms. The classical polylogarithms $\mathrm{Li}_k(z) = -G(\underbrace{0, \ldots, 0}_{k-1}, 1; z)$ form a strict subset of the Goncharov polylogarithms.

As mentioned in the introduction, the bane of working with polylogarithms is the numerous functional identities they obey. Perhaps the most important of these is the *shuffle identity*. The product of two polylogarithms can be written as

$$G(a_1, \ldots, a_n; z) G(a_{n+1}, \ldots, a_{n+m}; z) = \sum_{\sigma \in S_{n,m}} G(a_{\sigma(1)}, \ldots, a_{\sigma(n+m)}; z) \tag{2.3}$$

where $S_{n,m}$ is the set of (n, m)-shuffles, i.e. permutations σ of length $n + m$ such that

$$\sigma^{-1}(1) < \sigma^{-1}(2) < \cdots < \sigma^{-1}(n) \quad \text{and} \quad \sigma^{-1}(n+1) < \sigma^{-1}(n+2) < \cdots < \sigma^{-1}(n+m)\,. \tag{2.4}$$

The name comes from riffle shuffling a deck of cards; shuffling two stacks of cards together interweaves them while leaving each stack in the same order.

Each polylogarithm has an associated object called its *symbol* (see for example [25, 26], and the review [27]). The symbol is a useful tool for converting the functional identities of polylogarithms into linear algebra, obviating many thorny problems. The symbol of a Goncharov polylogarithm admits a nice graphical interpretation as a sum over plane trivalent trees [28], and is given explicitly by the recursive formula

$$\begin{aligned} S\big(G(a_k, \ldots, a_1; a_{k+1})\big) = \sum_{i=1}^{k} & S\big(G(a_k, \ldots, \widehat{a}_i, \ldots, a_1; a_{k+1})\big) \otimes (a_i - a_{i+1}) \\ & - S\big(G(a_k, \ldots, \widehat{a}_i, \ldots, a_1; a_{k+1})\big) \otimes (a_i - a_{i-1})\,. \end{aligned} \tag{2.5}$$

JHEP11(2015)136

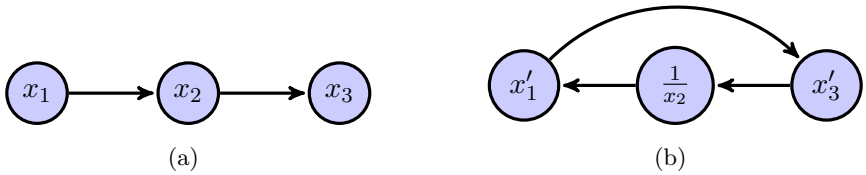


Figure 1. (a) The initial seed for the A_3 algebra. (b) The result of mutating on the vertex x_2.

Here $\widehat{a}_i$ denotes that the argument is omitted, and it is also understood that any term with 0 as a symbol entry (which can happen if some adjacent a's are equal) should simply be omitted. Symbols behave as if there were implicit "$d\log$'s" in front of each term: just as $d\log 1 = 0$ and $d\log\phi_1\phi_2 = d\log\phi_1 + d\log\phi_2$, symbols obey $(\cdots\otimes 1\otimes\cdots) = 0$ and

$$(\alpha\otimes\phi_1\phi_2\otimes\beta) = (\alpha\otimes\phi_1\otimes\beta) + (\alpha\otimes\phi_2\otimes\beta)\,. \tag{2.6}$$

The collection of ϕ_i which appear in the symbol of a given function is called its *symbol alphabet*.

2.2 Cluster algebras

Cluster algebras are a relatively new area of mathematics, introduced in 2002 by Fomin and Zelevinsky in [29, 30]. This section quickly reviews some salient facts about cluster algebras; the reader may consult [31, 32] for additional mathematical background and [1] for the amplitude perspective. A cluster algebra starts with a *seed* — a quiver where each vertex is labeled with a *cluster variable* (also called a *cluster coordinate*).[3] See figure 1(a) for an example of a seed.

A cluster algebra is generated from an initial seed through an iterative process. An operation called *mutation on a vertex* generates a new seed and new cluster variables, according to the formula given in eq. (B.2) (or see [33]). For example, mutating on the middle vertex on figure 1(a) gives figure 1(b) with the new cluster variables $x_1' = x_1(1+x_2)$ and $x_3' = \frac{x_2x_3}{1+x_2}$. Mutation is an involution, so applying the same mutation twice does nothing. The *cluster algebra* is the algebra generated by the set of all cluster variables which arise from repeatedly mutating the initial seed. Under certain conditions on the initial quiver, all possible repeated mutations will yield only finitely many seeds. Such cluster algebras are said to be of *finite type*.

A natural domain for a cluster algebra is the *positive domain*, where all cluster variables take positive real values. This property is preserved under mutation: if all variables in a given seed are positive-valued, then all possible cluster variables on the same algebra are also positive-valued.

The structure of the cluster algebra as a whole can be displayed as an *exchange graph*, where each vertex represents a seed, and *undirected* edges are drawn between seeds linked by a single mutation. (See figure 2(b).) Because applying a mutation will invert a single

[3]To be clear, throughout this paper the term "cluster variable" refers to the $\mathcal{X}$-coordinates of Fock and Goncharov [33], not to cluster $\mathcal{A}$-coordinates.

JHEP11(2015)136

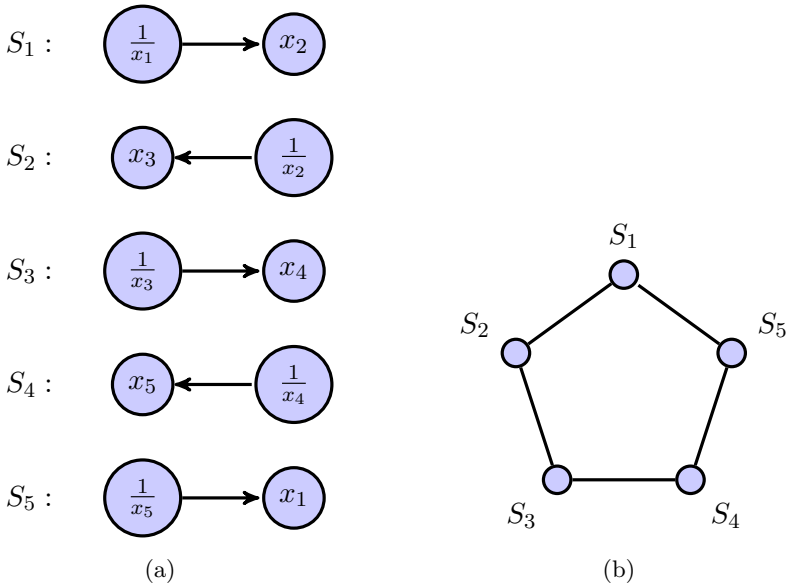
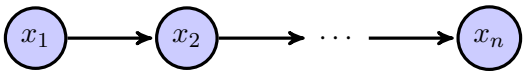


Figure 2. (a) The five seeds for the A_2 algebra, with each node labeled by its associated cluster variable. (b) The exchange graph, showing how to move from one seed to another by mutation.

cluster variable $x_i \mapsto \frac{1}{x_i}$, a *directed* edge of the exchange graph can be associated to a unique cluster variable x_i. The same edge with the opposite direction corresponds to $1/x_i$.

Later in this paper we will focus on the A_n family of cluster algebras, which start with the initial seed



for $n \geq 1$. For the special case of A_n algebras, which are all of finite type, there is a convenient alternative to representing clusters with quivers. This construction is reviewed in appendix A.

The A_2 algebra, for example, has exactly 5 distinct seeds and 10 cluster variables[4] given by

$$x_1\,, \quad x_2\,, \quad x_3 = \frac{1+x_2}{x_1}\,, \quad x_4 = \frac{1+x_1+x_2}{x_1 x_2}\,, \quad x_5 = \frac{1+x_1}{x_2} \tag{2.7}$$

and their reciprocals. These variables obey the recursive formula

$$x_{i+1} = \frac{1+x_i}{x_{i-1}}\,. \tag{2.8}$$

This paper makes extensive use of the A_3 cluster algebra, which starts with the initial seed in figure 1(a). It has 14 seeds and 30 cluster variables. We take the opportunity to enumerate in eq. (2.9) 15 of these cluster variables (the other 15 are their reciprocals) in four ways: (1) in terms of the names v_i, $x_i^{\pm}$ and e_i that these variables have been given in previous work (see in particular [1, 3]), (2) as rational functions of the variables x_1, x_2, x_3 in

[4] In this paper, we treat x and $1/x$ as two separate cluster variables, but they are sometimes conflated in the literature when it is useful to do so.

JHEP11(2015)136

the initial seed, (3) in terms of the $\{u, v, w, y_u, y_v, y_w\}$ variables used extensively by Dixon et al. in their study of 6-particle scattering amplitudes in SYM theory, (4) and in terms of Plücker coordinates on $\mathrm{Gr}(4,6)$ (the connection to Plücker coordinates is explained in the following subsection)[5]

$$
\begin{aligned}
v_1 &= \frac{(1+x_2)(1+x_3+x_2x_3+x_1x_2x_3)}{x_1x_2} &&= \frac{1-v}{v} &&= \frac{\langle 1246\rangle\langle 1345\rangle}{\langle 1234\rangle\langle 1456\rangle}\\
v_2 &= \frac{1+x_3}{x_2x_3} &&= \frac{1-w}{w} &&= \frac{\langle 1235\rangle\langle 2456\rangle}{\langle 1256\rangle\langle 2345\rangle}\\
v_3 &= (1+x_1)x_2 &&= \frac{1-u}{u} &&= \frac{\langle 1356\rangle\langle 2346\rangle}{\langle 1236\rangle\langle 3456\rangle}\\
x_1^+ &= \frac{1}{x_3} &&= \sqrt{\frac{vy_uy_vy_w}{uw}} &&= \frac{\langle 1456\rangle\langle 2356\rangle}{\langle 1256\rangle\langle 3456\rangle}\\
x_2^+ &= \frac{1+x_2+x_1x_2}{x_1} &&= \sqrt{\frac{wy_uy_vy_w}{uv}} &&= \frac{\langle 1346\rangle\langle 2345\rangle}{\langle 1234\rangle\langle 3456\rangle}\\
x_3^+ &= \frac{1+x_3+x_2x_3}{x_1x_2x_3} &&= \sqrt{\frac{uy_uy_vy_w}{vw}} &&= \frac{\langle 1236\rangle\langle 1245\rangle}{\langle 1234\rangle\langle 1256\rangle}\\
x_1^- &= x_1 &&= \sqrt{\frac{v}{uwy_uy_vy_w}} &&= \frac{\langle 1234\rangle\langle 2356\rangle}{\langle 1236\rangle\langle 2345\rangle}\\
x_2^- &= (1+x_2+x_1x_2)x_3 &&= \sqrt{\frac{w}{uvy_uy_vy_w}} &&= \frac{\langle 1256\rangle\langle 1346\rangle}{\langle 1236\rangle\langle 1456\rangle}\\
x_3^- &= \frac{1+x_3+x_2x_3}{x_2} &&= \sqrt{\frac{u}{vwy_uy_vy_w}} &&= \frac{\langle 1245\rangle\langle 3456\rangle}{\langle 1456\rangle\langle 2345\rangle}\\
e_1 &= \frac{1+x_3+x_2x_3+x_1x_2x_3}{(1+x_1)x_2} &&= \sqrt{\frac{(1-v)u}{v(1-u)y_uy_v}} &&= \frac{\langle 1246\rangle\langle 3456\rangle}{\langle 1456\rangle\langle 2346\rangle}\\
e_2 &= \frac{1}{(1+x_2)x_3} &&= \sqrt{\frac{v(1-w)y_vy_w}{(1-v)w}} &&= \frac{\langle 1235\rangle\langle 1456\rangle}{\langle 1256\rangle\langle 1345\rangle}\\
e_3 &= \frac{(1+x_1)x_2x_3}{1+x_3} &&= \sqrt{\frac{w(1-u)}{(1-w)uy_uy_w}} &&= \frac{\langle 1256\rangle\langle 2346\rangle}{\langle 1236\rangle\langle 2456\rangle}\\
e_4 &= \frac{1+x_2}{x_1x_2} &&= \sqrt{\frac{(1-v)uy_uy_v}{v(1-u)}} &&= \frac{\langle 1236\rangle\langle 1345\rangle}{\langle 1234\rangle\langle 1356\rangle}\\
e_5 &= \frac{x_1(1+x_3)}{1+x_3+x_2x_3+x_1x_2x_3} &&= \sqrt{\frac{v(1-w)}{(1-v)wy_vy_w}} &&= \frac{\langle 1234\rangle\langle 2456\rangle}{\langle 1246\rangle\langle 2345\rangle}\\
e_6 &= x_2 &&= \sqrt{\frac{w(1-u)y_uy_w}{(1-w)u}} &&= \frac{\langle 1356\rangle\langle 2345\rangle}{\langle 1235\rangle\langle 3456\rangle}
\end{aligned}
\tag{2.9}
$$

2.3 Scattering amplitudes and Grassmannian cluster algebras

The connection between scattering amplitudes in SYM theory and cluster algebras was first made in [1] and further explored in [3, 4, 34, 35]. The basic fact that allows for such

[5] The A_3 algebra has several seeds of the form shown in figure 1(a). We caution that the definition of x_1, x_2 and x_3 in eq. (2.9) reflects a convention chosen in [3] which differs from the particular x_1, x_2 and x_3 assigned to the initial seed in [31].

JHEP11(2015)136

a connection is that the kinematic domain for n-particle scattering in SYM theory, called $\mathrm{Conf}_n(\mathbb{P}^3)$, has, according to [33], the structure of a cluster Poisson variety associated to the $\mathrm{Gr}(4,n)$ Grassmannian cluster algebra. This fact is special to SYM theory in four dimensions because it relies on the dual conformal symmetry of the theory, discovered in [36–41].

An ordered scattering amplitude of n massless particles is a function of n null vectors in Minkowski space that sum up to zero due to energy-momentum conservation. Using the momentum twistor variables of Hodges [42], the space of such configurations can be realized as n ordered points in $\mathbb{P}^3$, or concretely as a $4\times n$ matrix $[Z_1 Z_2 \cdots Z_n]$ where each column Z_i is a four-component homogeneous coordinate on $\mathbb{P}^3$. In this presentation, dual conformal symmetry, which must leave all amplitudes invariant, acts as left-multiplication by $\mathrm{SL}(4,\mathbb{C})$. Passing to the quotient space we get a birational isomorphism (which means a bijection for generic points)

$$\mathrm{Gr}(4,n)/(\mathbb{C}^*)^{n-1} \xrightarrow{\sim} \mathrm{Conf}_n(\mathbb{P}^3)\,. \tag{2.10}$$

Thus, scattering amplitudes can be (essentially) regarded as complex-valued functions on Grassmannians, making it natural to use the $\mathrm{SL}(4,\mathbb{C})$ invariant Plücker coordinates $\langle ijk\ell\rangle = \det[Z_i Z_j Z_k Z_\ell]$, which are well-defined complex-valued functions on $\mathrm{Gr}(4,n)$. However, since the Z's are homogeneous coordinates, it is necessary to use ratios of Plücker coordinates (or, more generally, ratios of homogeneous polynomials of Plücker coordinates), with the same Z's appearing in the numerator and denominator, such as

$$\frac{\langle 5713\rangle\,\langle 5624\rangle}{\langle 4512\rangle\,\langle 3567\rangle}\,, \tag{2.11}$$

to get well-defined coordinates on $\mathrm{Gr}(4,n)/(\mathbb{C}^*)^{n-1}$. Scattering amplitudes in SYM theory are naturally written as functions of such cross-ratios.

This is where cluster algebras enter: the Plücker coordinates of any Grassmannian form a cluster algebra [43], and the quotient $\mathrm{Conf}_n(\mathbb{P}^3)$ has the structure of a cluster Poisson variety [33], with cluster coordinates given by certain very special cross-ratios of the abovementioned type.[6] The physics interest in such cluster algebras stems from the fact that all known multi-loop amplitudes that have been explicitly computed to date in SYM theory (including [1, 9–12, 26, 44–47]) are generalized polylogarithms whose symbol alphabets are subsets of cluster coordinates on this $\mathrm{Gr}(4,n)$ cluster algebra.[7] The A_n family of cluster algebras reviewed in section 2.2 corresponds to the Grassmannian $\mathrm{Gr}(2,n+3)$, which overlaps with the sequence of algebras relevant to scattering amplitudes in the case $A_3 \cong \mathrm{Gr}(2,6) \cong \mathrm{Gr}(4,6)$ relevant to 6-particle amplitudes.

2.4 The cluster bootstrap

The main problem we address in this paper is simple to state: given a cluster algebra $\mathcal{A}$, with a set of cluster coordinates $\mathcal{X}_\mathcal{A}$, we would like to write down a basis for weight-k

[6] The positive domain for the $\mathrm{Gr}(4,n)$ algebra is defined by $\langle ijk\ell\rangle > 0\ \forall 1 \le i < j < k < \ell \le n$.

[7] A number of results in two-dimensional kinematics including [48–52] provide partial evidence in support this assertion, though the full $\mathrm{Gr}(4,n)$ structure necessarily collapses in two-dimensional kinematics. This has been studied in [34].

JHEP11(2015)136

polylogarithm functions whose symbols may be written in the alphabet $\mathcal{X}_{\mathcal{A}}$. We call such functions "cluster polylogarithm functions" or simply *cluster functions*[8] on $\mathcal{A}$.

To be explicit, let us note that A_2 cluster functions, for example, are those which can be written in the symbol alphabet consisting of the five x_i shown in eq. (2.7). Thanks to eq. (2.6), we can equivalently consider the A_2 symbol alphabet to be the set

$$\{x_1, x_2, 1+x_1, 1+x_2, 1+x_1+x_2\} \tag{2.12}$$

since each of the five x_i may be (uniquely) expressed as products of powers of elements of this set. For A_3 only 9 of the 30 cluster variables are multiplicatively independent, and it is evident from eq. (2.9) that the A_3 symbol alphabet may be taken as the set

$$\{x_1, x_2, x_3, 1+x_1, 1+x_2, 1+x_3, 1+x_2+x_1x_2, 1+x_3+x_2x_3, 1+x_3+x_2x_3+x_1x_2x_3\}\,. \tag{2.13}$$

Closely related symbol alphabets have appeared elsewhere, notably in Brown's work on polylogarithm functions on the moduli space $\mathfrak{M}_{0,m}$ of m marked points on the Riemann sphere [20]. For example, for the case $m = 6$, Brown's polylogarithms are based on the symbol alphabet

$$\{c_1, c_2, c_3, 1-c_1, 1-c_2, 1-c_3, 1-c_1c_2, 1-c_2c_3, 1-c_1c_2c_3\} \tag{2.14}$$

in cubical coordinates[9] or

$$\{t_1, t_2, t_3, 1-t_1, 1-t_2, 1-t_3, t_3-t_1, t_3-t_1, t_3-t_2\} \tag{2.15}$$

in simplicial coordinates. Neither alphabet is multiplicatively equivalent to eq. (2.13), but their relation will be uncovered in the following section. In fact, one way to express the central result of our paper is to say that we demonstrate how to construct explicit changes of variables between those of [20] on $\mathfrak{M}_{0,n+3}$ and the A_n cluster $\mathcal{X}$-coordinates, for any n, which render the corresponding symbol alphabets multiplicatively equivalent.

Let us conclude our review by briefly recalling that for finite symbol alphabets this problem admits a conceptually straightforward, if computationally intensive, brute force solution. If the symbol alphabet for $\mathcal{A}$ has s multiplicatively independent letters $\{\phi_1, \ldots, \phi_s\}$, then the symbol of any weight-k cluster function may be expressed as a unique vector (with rational components) in the s^k dimensional vector space $\mathcal{V}_k$ spanned by basis elements $\phi_{i_1} \otimes \cdots \otimes \phi_{i_k}$. Going the other way around, any vector in $\mathcal{V}_k$ which satisfies a set of linear integrability conditions (see for example [25]) corresponds to (the symbol of) some cluster function. Therefore, the problem of finding a basis for the (symbols of) weight-k cluster functions on $\mathcal{A}$ is the same as that finding a basis for the nullspace of a certain linear operator on $\mathcal{V}_k$.

[8] Functions of this type were called "cluster $\mathcal{A}$-functions" in [3] to distinguish them from a smaller set of functions with more special properties called "cluster $\mathcal{X}$-functions", but we do not explore these additional properties here.

[9] These cubical coordinates were called x_i in [20], but we use c_i in eq. (2.14) to distinguish them from our x_i cluster coordinates.

JHEP11(2015)136

The efficiency of this approach can be considerably enhanced by recycling lower-weight information at higher weight, and by exploiting the Hopf algebra structure of polylogarithms (discovered in [28], and nicely reviewed for a physics audience in [53]).[10] Collectively these "bootstrap" techniques have been implemented systematically by Dixon and collaborators for the 6-particle case (associated to the $\mathrm{Gr}(4,6) \cong A_3$ cluster algebra) to great effect in [10–12]. A slightly modified "weight-skipping" bootstrap based on a symbol alphabet of $\mathrm{Gr}(4,7)$ cluster coordinates allowed for the calculation of the symbol of the 3-loop 7-particle MHV amplitude in [7].

Finally, we note a fact we will use later: the classical polylogarithm functions Li_k (and products thereof) are known to span the space of all polylogarithm functions of weight $k \leq 3$, so it is trivial to write down a (vastly overcomplete) set of irreducible cluster functions at weights $k = 1, 2, 3$:

$$\log(\mathcal{X}_{\mathcal{A}})\,, \quad \mathrm{Li}_2(-\mathcal{X}_{\mathcal{A}})\,, \quad \mathrm{Li}_3(-\mathcal{X}_{\mathcal{A}}) \cup \mathrm{Li}_3(1+\mathcal{X}_{\mathcal{A}})\,. \tag{2.16}$$

The problem we address in this paper is that of finding bases for all weights, not just overcomplete sets of cluster functions.

3 Hedgehog bases

We tackle the problem of constructing bases of cluster functions in three steps. (1) First we discuss the set of Goncharov polylogarithms whose symbols may be written in the alphabet of cluster coordinates. (2) Next, we review the form a generating set should have, based on work of Brown [20] and Drummond [54]. (3) Lastly, we define "hedgehogs" and prove that they provide bases for the space of A_n cluster functions.

3.1 Good arguments for Goncharov polylogarithms

To construct suitable collections of functions there is no need to reinvent the wheel. We may attempt to solve this problem by using a nice set of polylogarithm functions we already have at our disposal: the Goncharov polylogarithms defined in eq. (2.1). Then it remains only to decide what kinds of variables we should allow as the arguments $a_1, \ldots, a_k; z$.

Let us write $G_k[Q]$ to denote the set of weight-k Goncharov polylogarithms whose arguments are drawn from some set Q:

$$G_k[Q] = \{G(q_1, \ldots, q_k; q_0) : q_i \in Q\}\,. \tag{3.1}$$

It is evident from eq. (2.5) that functions in $G_k[Q]$ have symbol entries of the form q_i as well as $q_i - q_j$, for $q_i, q_j \in Q$. We may try to follow the path of least resistance by considering what happens when Q is chosen simply to be some subset of $\mathcal{X}_{\mathcal{A}}$. Actually, although this doesn't matter at the level of symbols, for later convenience it will be better to consider subsets of $-\mathcal{X}_{\mathcal{A}}$ since this will help to naturally provide Goncharov polylogarithms that are manifestly free of branch cuts in the positive domain. (Henceforth we shall use

[10] The Hopf algebra structure makes SYM theory an ideal setting in which to study *motivic amplitudes*, as proposed a decade ago in [28] (see in particular section 7).

JHEP11(2015)136

$x_i \in \mathcal{X}_\mathcal{A}$ to denote cluster coordinates and $q_i = -x_i$ to denote negative cluster coordinates.) Unfortunately, for two generic $q_i, q_j \in -\mathcal{X}_\mathcal{A}$, there is nothing particularly nice about the quantity $q_i - q_j$; it may not even have definite sign in the positive domain, in which case it should never appear in the symbol of a cluster function.

One approach to construct bases of cluster functions would use special linear combinations of Goncharov polylogarithms for which all "bad" letters cancel out at the level of symbols. Several examples of such functions have been studied in the literature. For the particular case of $\mathcal{A} = A_3$, Dixon et al. have constructed Goncharov polylogarithm representations for bases of "hexagon functions" through weight at least 8. These are cluster functions satisfying an additional important physical constraint (the first-entry condition), which we do not address here. The construction of these bases, and several impressive applications to 6-particle scattering amplitudes in SYM theory, are discussed in [9–13]. Also, the "cluster $\mathcal{X}$-functions" studied in [3, 6] for more general algebras can be expressed as suitable linear combinations of Goncharov polylogarithms with all "bad" symbol entries cancelling out. These functions also play a prominent role in SYM theory: in particular, it appears from the result of [4] that all 2-loop MHV amplitudes can be expressed in terms of classical polylogarithms and the single non-classical cluster function $K_{2,2}$ defined in [6].

In the present paper, we would like to explore a different approach to cluster functions. We explore the possibility of constructing Goncharov polylogarithms at any weight which are manifestly free of any "bad" letters, rather than having to rely on solving a (potentially computationally-challenging) linear algebra problem to ensure their cancellation. In light of the factorization property reviewed in eq. (2.6), it is evident that this will be the case if we can choose the set Q so that $q_i - q_j$ factors into a product of powers of cluster coordinates for all $q_i, q_j \in Q$. To be precise, let us define the *multiplicative span* of $\mathcal{X}_\mathcal{A}$ to be the set

$$M_\mathcal{A} = \left\{ \pm \prod_i x_i^{n_i} : x_i \in \mathcal{X}_\mathcal{A} \text{ and } n_i \in \mathbb{Z} \right\}. \tag{3.2}$$

(If $\mathcal{A}$ is an infinite algebra, then only finitely many of the n_i may be nonzero.) We say that a set Q *splits over* $\mathcal{X}_\mathcal{A}$ if $q_i - q_j \in M_\mathcal{A}$ for all $q_i \neq q_j \in Q$.

Then it is evident that $G_k[Q]$ is a set of cluster functions on $\mathcal{A}$ whenever $Q \subseteq -\mathcal{X}_\mathcal{A}$ splits over $\mathcal{X}_\mathcal{A}$. In fact, for any such Q we can get additional cluster functions "for free" by considering the enlarged set $G_k[\{0,1\} \cup Q]$. The inclusion of 0 is trivial, and 1 is allowed because of the property that $q - 1 = 1 + x \in M_\mathcal{A}$ for all $q \in -\mathcal{X}_\mathcal{A}$. A proof of this property, which played an important role in [1, 4, 6], is presented in appendix B.

We can conclude that

> If $Q \subseteq -\mathcal{X}_\mathcal{A}$ splits over $\mathcal{X}_\mathcal{A}$, then $G_k[\{0,1\} \cup Q]$ is a set of cluster functions on $\mathcal{A}$.

Of course, additional functions of weight k may be constructed by taking products of functions of lower weight.

It may be helpful to visualize sets of cluster coordinates satisfying the required property with the assistance of what we call a *factorization graph*. For a given algebra $\mathcal{A}$, the factorization graph contains one vertex for each cluster coordinate $x \in \mathcal{X}_\mathcal{A}$ and two vertices

JHEP11(2015)136

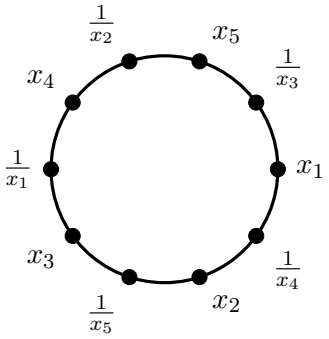


Figure 3. The factorization graph for A_2. Each vertex is one of the 10 cluster coordinates on the A_2 cluster algebra (see eq. (2.7)), and two vertices x_i, x_j are connected by an edge if $x_i - x_j$ factors into a product of cluster coordinates. Each of the 10 pairs of connected vertices, for example $\{1/x_2, x_5\}$, is a 2-clique.

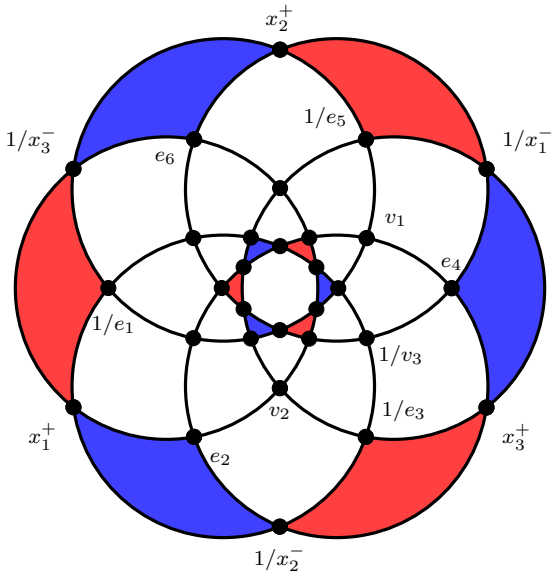


Figure 4. The factorization graph for A_3. Each vertex represents one of the 30 cluster coordinates on the A_3 cluster algebra, although to avoid clutter only 15 of the coordinates are labeled; the other 15 are reciprocals of the ones shown. Two vertices x_i, x_j are connected by an edge if $x_i - x_j \in M_{A_3}$. The six circles each pass through 10 vertices and indicate an A_2 subalgebra, as shown in figure 3. There are 12 subgraphs with the topology of a triangle, 6 around the outer edge and 6 around the inner edge; these are the 12 3-cliques.

x_i, x_j are connected if $x_i - x_j \in M_{\mathcal{A}}$. The factorization graphs for the A_2 and A_3 cluster algebras are shown in figures 3 and 4.

In mathematics, a complete subgraph (that is, a collection of vertices such that each pair is connected by an edge) is known as a *clique* (or an n-clique, if it has n vertices). It is evident from figures 3 and 4 that A_2 has 10 2-cliques and no higher cliques, while A_3 has 60 2-cliques, 12 3-cliques, and no higher cliques. Also note that the A_3 factorization

JHEP11(2015)136

graph is composed of 6 intersecting copies of the A_2 factorization graph. Therefore we can rephrase the above conclusion by saying that

> *Cliques give cluster functions.*
> If $-Q \subseteq \mathcal{X}_{\mathcal{A}}$ is a clique of the factorization graph of $\mathcal{A}$, then $G_k[\{0,1\} \cup Q]$ is a set of cluster functions on $\mathcal{A}$.

Since ordering will play a crucial role in what follows, this is the perfect opportunity for us to note the convenient fact that if $Q \subseteq -\mathcal{X}_{\mathcal{A}}$ splits over $\mathcal{X}_{\mathcal{A}}$, then there is a natural ordering on Q. Recalling that cluster coordinates are positive-valued everywhere in the interior of the positive domain, possibly taking value 0 or $+\infty$ only on the boundary of that domain, it is evident that for every pair $q_i \neq q_j \in Q$, the difference $q_i - q_j \in M_{\mathcal{A}}$ takes uniform sign inside the positive domain. Therefore, for each pair either $q_i < q_j$ or vice versa, so the natural ordering on Q is simply the true numerical order $q_1 < q_2 < \cdots < q_n$ of these coordinates in the positive domain. It will be convenient to choose the ordering on the set $\{0,1\} \cup Q$ to be $0, 1, q_1, \ldots, q_n$, even though this is not the true numerical ordering of these quantities (since the q's are negative in the positive domain).

3.2 Bases of cluster functions

So far, we have seen that elements of the set $G_k[\{0,1\} \cup Q]$ are cluster functions on $\mathcal{A}$, i.e. have symbols which can be written in the symbol alphabet $\mathcal{X}_{\mathcal{A}}$ of cluster coordinates on $\mathcal{A}$, whenever $-Q$ is a clique for the factorization graph of $\mathcal{A}$. We now want a basis for $\mathcal{A}_\bullet(A_n)$, the space of cluster functions on A_n. Let's first consider a simple case. Suppose we are only interested in the space of polylogarithms with a fixed last argument:

$$\mathcal{G}[S;z] = \bigoplus_{k=1}^{\infty} \operatorname{span} G_k[S;z] \quad \text{where} \quad G_k[S;z] = \{G(s_1,\ldots,s_k;z) : s_i \in S\} \tag{3.3}$$

where "span" denotes the vector space of $\mathbb{Q}$-linear combinations of the indicated functions. We use the notation $\mathcal{G}$ to carefully distinguish $G_k[S;z]$, which is a *set* of weight-k functions, from $\mathcal{G}[S;z]$, which is a *vector space* of functions of any weight.

Although $\mathcal{G}[S;z]$ is a vector space, because of eq. (2.3) it is more useful to consider it as a *shuffle algebra*. When dealing with such functions, it is more natural not to look for a vector space basis, but rather to find a minimal generating set for the algebra, such that each element of $\mathcal{G}[S;z]$ has a unique expression as a linear combination of *products of* elements of the minimal generating set.

For this, we use Radford's theorem (see [55]), which provides a minimal generating set for any free shuffle algebra in terms of *Lyndon words*. A Lyndon word of length k on an ordered set S is a sequence of k elements of S which is strictly smaller than all of its cyclic permutations with respect to the lexicographic order of S^k. (Several explicit examples will be presented in section 4.1.) Let $\text{Lyndon}_k(S)$ denote Lyndon words of length k on S. It is a consequence of Radford's theorem that $\mathcal{G}[S;z]$ has a minimal generating set

$$\bigcup_{k\in\mathbb{N}} G_k[\text{Lyndon}_k(S);z]\,. \tag{3.4}$$

JHEP11(2015)136

How can we use (3.4) to generate cluster functions? The answer to this question[11] is provided by Brown's extensive study of polylogarithms on the moduli spaces $\mathfrak{M}_{0,n+3}$[12] in [20] (see also [21] for some applications). The results of [20] were presented in various useful coordinate systems on $\mathfrak{M}_{0,n+3}$. One key result was that (essentially) the space of Goncharov polylogarithms on $\mathfrak{M}_{0,n+3}$ is the tensor product of n spaces of polylogarithms with fixed last arguments in a certain ordered set of variables S. The analysis in the previous subsection has revealed that choosing S to be a clique Q along with $\{0,1\}$ makes manifest the A_n cluster structure of these functions. And, by eq. (3.4), we have a generating set for each of those n spaces of polylogarithms. Combining these observations we arrive at:

> *For A_n, each n-clique gives a generating set for* all *cluster functions.*
> If $-Q \subseteq \mathcal{X}_{A_n}$ is an n-clique of the factorization graph of the A_n cluster algebra with an ordering $Q = \{q_1 < q_2 < \cdots < q_n\}$, then
>
> $$\bigcup_{k\in\mathbb{N}} \bigcup_{i=1}^{n} G_k[\text{Lyndon}_k\{0, 1, q_1, \ldots, q_{i-1}\}; q_i] \tag{3.5}$$
>
> is a minimal generating set for $\mathcal{A}_\bullet(A_n)$, the space of cluster functions on A_n. A vector space basis for $\mathcal{A}_\bullet(A_n)$ is given by all possible products of elements of the set (3.5).

We call the basis generated by eq. (3.5) a *hedgehog basis* for reasons that will become clear in the next section. A very nice feature of this basis is that, thanks to the natural ordering $q_1 < q_2 < \cdots < q_k < 0$ on the set Q discussed above, it is manifest from eq. (2.1) that each G function in eq. (3.5) is free of branch cuts everywhere in the interior of the positive domain, with possible branch cuts only on its boundary — with one important exception that we should note. The exception is that at weight 1, instead of $G(0; q_i)$ we should use the function

$$G(0; -q_i) = \log(-q_i)\,. \tag{3.6}$$

The feature of being free of branch cuts in the positive domain is a necessary feature for these functions to be useful in describing scattering amplitudes, but the analytic constraints on amplitudes are far stronger still: they must be singularity-free everywhere inside the larger Euclidean domain, with branch points allowed only on boundaries corresponding to multi-particle production thresholds. It is an outstanding problem of great importance to find an explicit basis for the subspace of cluster functions spanned by functions satisfying these tighter analytic constraints.

3.3 The hedgehog theorem for A_n

We have now reduced the problem of finding a basis for cluster functions on A_n to that of finding cliques Q of size n. In this section we show that there are precisely two such cliques

[11] We are grateful to J. Drummond for carefully explaining the application of Brown's results to the construction of functional bases.

[12] This is the space of configurations of $n+3$ distinct, ordered points on the Riemann sphere.

JHEP11(2015)136

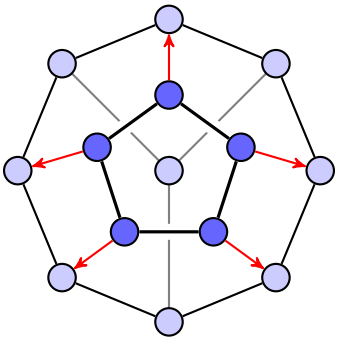


Figure 5. The A_3 algebra has six distinct hedgehogs (and six anti-hedgehogs). This figure shows the exchange graph for A_3, with one of its six pentagonal A_2 subalgebras highlighted. The "spines" of this $\mathcal{X}(A_3, A_2)$ hedgehog are the red edges connecting this A_2 to the rest of A_3. Specifically, this $\mathcal{X}(A_3, A_2)$ is the set of 3 $\mathcal{X}_{A_3}$ cluster coordinates associated to these 5 outward directed red edges.

for each A_{n-1} subalgebra of A_n. This correspondence can be visualized, at the level of the exchange graph, by collections of cluster variables that we call hedgehogs.

Let us start by defining hedgehogs. Suppose $\mathcal{A}$ is a cluster algebra of rank r and $\mathcal{B}$ is a subalgebra of rank $r-1$. The exchange graph of $\mathcal{A}$ is an r-regular graph (each vertex has valence r), and the exchange graph for $\mathcal{B}$ is an embedded $(r-1)$-regular subgraph. Therefore, each vertex of $\mathcal{B}$ is incident to $r-1$ edges leading to other vertices of $\mathcal{B}$ and to one edge leading to a vertex of $\mathcal{A} \setminus \mathcal{B}$. In other words, each vertex of $\mathcal{B}$ has an edge which goes "out of" $\mathcal{B}$ and "into" $\mathcal{A} \setminus \mathcal{B}$. Recall from section 2.2 that a directed edge of the exchange graph can be associated with a cluster coordinate $x \in \mathcal{X}_{\mathcal{A}}$. Let the *hedgehog* $\mathcal{X}(\mathcal{A}, \mathcal{B}) \subseteq \mathcal{X}_{\mathcal{A}}$ be the set of cluster coordinates associated to the edges going out of $\mathcal{B}$ into $\mathcal{A} \setminus \mathcal{B}$.

Example hedgehogs for $\mathcal{X}(A_2, A_1)$ and $\mathcal{X}(A_3, A_2)$ are shown in figures 6(a) and 5 respectively. As can be seen from the pictures, the edge variables in the set $\mathcal{X}(\mathcal{A}, \mathcal{B})$ radiate outwards — just like the spines of a hedgehog. We might also consider the set of cluster coordinates associated to inward directed edges, which just gives the "anti-hedgehog"

$$\mathcal{X}^{-1}(\mathcal{A}, \mathcal{B}) = \{1/x : x \in \mathcal{X}(\mathcal{A}, \mathcal{B})\} . \tag{3.7}$$

We are now in a position to state the main result of this paper:

> *The hedgehog theorem for A_n: hedgehogs are n-cliques.*
> Let $\mathcal{X}(A_n, A_{n-1})$ be any hedgehog (or anti-hedgehog). Then $Q = -\mathcal{X}(A_n, A_{n-1})$ is an n-clique of the factorization graph for A_n. In particular, eq. (3.5) generates a basis for the set of all cluster functions on A_n.

The details of the proof of this theorem are presented in appendix D, using the machinery of triangulations reviewed in appendix A. Here we will be content to use the notation of the latter appendix to provide explicit formulas for all A_{n-1} hedgehogs of A_n, and to check that they are cliques.

Let us note that the symbol alphabet of the cluster functions generated by eq. (3.5), which consists of letters of the form q_i, $1-q_i$, or q_i-q_i, has exactly the same form as that of

JHEP11(2015)136

the polylogarithm functions studied by Brown [20] in what he calls simplicial coordinates, t_i. We are therefore able to conclude that the two sets of functions can be related to each other by the identification $t_i = -q_i = x_i$ between simplicial coordinates t_i on $\mathfrak{M}_{0,n+3}$ and the cluster coordinates x_i of any hedgehog $\mathcal{X}(A_n, A_{n-1})$ or anti-hedgehog $\mathcal{X}^{-1}(A_n, A_{n-1})$.

As reviewed in appendix A, there are precisely $2\binom{n+3}{4}$ cluster variables on A_n (counting x and $1/x$ separately); half of these can be enumerated explicitly as cross-ratios

$$r(i,j,k,\ell) = \frac{\langle ij\rangle \langle k\ell\rangle}{\langle jk\rangle \langle i\ell\rangle}, \qquad 1 \le i < j < k < \ell \le n+3 \tag{3.8}$$

of $n+3$ points in $\mathbb{P}^1$, while the other half are their reciprocals $1/r(i,j,k,\ell) = r(j,k,\ell,i)$. The $\mathrm{Gr}(2,n)$ Plücker coordinates $\langle ij\rangle$ used here may be related, in the case $n=3$, to the $\mathrm{Gr}(4,n)$ coordinates used in section 2 by

$$\langle ij\rangle = \frac{1}{4!}\epsilon_{ijk\ell mn}\langle k\ell mn\rangle . \tag{3.9}$$

The A_n cluster algebra has $n+3$ subalgebras of type A_{n-1}, so there are $n+3$ hedgehogs. In appendix D we show that these hedgehogs are given by sets of the form

$$\big\{r(k,k+1,k+2,i) : i \notin \{k,k+1,k+2\}\big\} \tag{3.10}$$

where $k+1$ and $k+2$ are taken mod n. It is easy to verify that these are n-cliques by taking two variables $r(k,k+1,k+2,i)$, $r(k,k+1,k+2,j)$ in this hedgehog and looking at their difference,

$$\begin{aligned} &r(k,k+1,k+2,i) - r(k,k+1,k+2,j) \\ &\quad= \frac{\langle k(k+1)\rangle}{\langle (k+1)(k+2)\rangle \langle ki\rangle \langle kj\rangle}\big(\langle (k+2)i\rangle \langle kj\rangle - \langle (k+2)j\rangle \langle ki\rangle\big) \\ &\quad= \frac{\langle k(k+1)\rangle}{\langle (k+1)(k+2)\rangle \langle ki\rangle \langle kj\rangle}\langle k(k+2)\rangle \langle ji\rangle \\ &\quad= r(k,k+1,k+2,i) r(k,k+2,i,j) \in M_{A_n} \end{aligned} \tag{3.11}$$

where the second equality is from a Plücker relation.

To summarize, for the A_n cluster algebras, hedgehogs are cliques of size n. There are $n+3$ hedgehogs, and $n+3$ anti-hedgehogs, related by the dihedral symmetry of the $n+3$-gon. This provides, via eq. (3.5) and the hedgehog theorem, $2(n+3)$ distinct, but equivalent, bases for cluster functions on A_n.

3.4 Comments on other algebras

Our problem was to write down a basis of cluster functions on a cluster algebra $\mathcal{A}$, and for $\mathcal{A} = A_n$ we have found that eq. (3.5) gives such a basis whenever $-Q$ is a hedgehog (or anti-hedgehog) in A_n. The algebras of most relevance to SYM theory, however, are the $\mathrm{Gr}(4,n)$ algebras (see [1]). Happily the one overlapping case $A_3 = \mathrm{Gr}(2,6) = \mathrm{Gr}(4,6)$ underlies the structure of 6-particle scattering amplitudes. We present an application of our results to this case in the following section.

JHEP11(2015)136

For more general algebras $\mathcal{A}$, the definition of hedgehog given above still makes sense, but it doesn't appear to be useful. In particular, it is straightforward to check, for example, that there is no $A_3 \subset D_4$, nor $A_5 \subset E_6$, such that $\mathcal{X}(D_4, A_3)$ or $\mathcal{X}(E_6, A_5)$ are cliques. For such hedgehogs, a set of functions of the type shown in eqs. (3.1) or (3.5) are still perfectly fine sets of polylogarithm functions, but they are not cluster functions: their symbols contain non-cluster coordinates as entries.

One could, of course, look at smaller hedgehogs, associated to $A_n \subset \mathcal{A}$ subalgebras, which are known to be cliques due to the hedgehog theorem. For example, D_4 has 12 distinct A_3 subalgebras, each of which has 6 A_2 subalgebras, so in all there exist 72 $Q(A_3, A_2)$ hedgehogs sitting inside D_4. However it is easy to check that no individual hedgehog furnishes enough functions to provide a basis for all cluster functions on D_4. We know this because we can compare with the dimension of the spanning sets for weight ≤ 3 described in eq. (2.16). The same comment holds for E_6, which has seven A_5 subalgebras, each of which in turn has eight A_4 subalgebras, for a total of 56 $\mathcal{X}(A_5, A_4)$ hedgehogs. On the other hand, at least for the D_4 case we have checked that the union of all cluster functions over these various hedgehogs provides a vastly overcomplete set of cluster functions at weight ≤ 3, but we do not have a collection of hedgehogs which exactly spans to provide a basis. It may be that, just as eq. (3.5) gives a basis for cluster functions on $A_n \cong \mathrm{Gr}(2, n)$ by gluing together sets of the form (3.4) in a certain pattern, some different pattern of gluing might work for other algebras including the cases $\mathrm{Gr}(4, n)$ of relevance to scattering amplitudes.

4 An application to the 2-loop 6-particle NMHV amplitude

All evidence available to date (including [1, 9–12, 26, 44, 45]) supports the hypothesis that all 6-particle scattering amplitudes in SYM theory can be expressed in terms of cluster functions on the A_3 cluster algebra. As an application of the hedgehog theorem, we discuss in this section how to express the 2-loop 6-particle NMHV amplitude in a hedgehog basis. This amplitude was originally computed in [9] and written (see eq. (2.27) of that paper) as $[12345](V + \tilde{V}) + \text{cyclic}$, where $[12345]$ is an R-invariant and $X \equiv 8(V + \tilde{V})$ is a weight-4 polylogarithm function. The exercise of rewriting X in a hedgehog basis has some practical benefit in that it produces a formula which is notably shorter than results previously available in the literature. But from our perspective a greater benefit of working with a hedgehog basis is that it makes some of the cluster structure of the amplitude manifest.

To highlight this point, let us note that in the presentation of [9], the amplitude X is written as a linear combination of various generalized polylogarithm functions whose symbols may be written in the 10-letter alphabet

$$\{u, v, w, 1-u, 1-v, 1-w, y_u, y_v, y_w, 1-y_u y_v y_w\}\,. \tag{4.1}$$

The relation between these variables and ours may be read off from eq. (2.9). The tenth letter $1 - y_u y_v y_w$ is not "clustery" — that is, it cannot be expressed as a product of A_3 cluster coordinates, so it should never appear in the symbol of anything we would call a cluster function. Indeed the full amplitude (like all 6-particle amplitudes) has the property that when all of the individual contributing polylogarithm functions are added up, this

JHEP11(2015)136

tenth letter cancels out of the symbol of the full amplitude. This is suggestive: if all these terms cancel out in the end, it seems desirable to express the amplitude in such a way that they never arise in the first place. This is exactly what an A_3 hedgehog basis does.

4.1 The hedgehog basis for A_3

Let's look at the hedgehog basis more concretely in the A_3 case. For $n = 3$, eq. (3.5) tells us that we can list basis elements for A_3 cluster functions by enumerating Lyndon words on the sets $\{0, 1\}$, $\{0, 1, q_1\}$, and $\{0, 1, q_1, q_2\}$.

At weight 1 there are respectively 2, 3, 4 Lyndon words on these three sets, which together provide the 9 weight-1 functions in the A_3 basis:

$$\begin{aligned}&\{G(0;-q_1), G(1;q_1),\\ &\ G(0;-q_2), G(1;q_2), G(q_1;q_2),\\ &\ G(0;-q_3), G(1;q_3), G(q_1;q_3), G(q_2;q_3)\}\,.\end{aligned} \tag{4.2}$$

(Here we recall that the three $G(0; z)$ functions are to be treated as explained in eq. (3.6).) At weight 2 there are respectively 1, 3, 6 Lyndon words on the three sets, which together provide the 10 pure weight-2 functions:

$$\begin{aligned}&\{G(0,1;q_1),\\ &\ G(0,1;q_2), G(0,q_1;q_2), G(1,q_1;q_2),\\ &\ G(0,1;q_3), G(0,q_1;q_3), G(0,q_2;q_3), G(1,q_1;q_3), G(1,q_2;q_3), G(q_1,q_2;q_3)\}\,.\end{aligned} \tag{4.3}$$

An additional 45 functions of weight 2 may be obtained by taking products of pairs of the weight-1 functions shown in eq. (4.2), so the total space of weight-2 functions on A_3 has dimension 55.

It is a simple exercise to continue enumerating Lyndon words in this manner to higher weight. We find a total of 285 functions of weight 3 and 1351 functions of weight 4, which is as far as we need to go for the purpose of expressing the 2-loop amplitude X. Symbols of functions in this hedgehog basis can be expressed in the 9-letter "q" alphabet

$$\{q_1, q_2, q_3, 1-q_1, 1-q_2, 1-q_3, q_1-q_2, q_2-q_3, q_1-q_3\} \tag{4.4}$$

where $-Q = \{-q_1, -q_2, -q_3\}$ is any 3-clique of the A_3 factorization graph.

4.2 Hedgehogs for A_3

According to the hedgehog theorem, cliques for A_3 are given precisely by hedgehogs (or anti-hedgehogs), which are in one-to-one correspondence with $A_2 \subset A_3$ subalgebras. The hedgehogs for A_3 are triples of cluster coordinates associated with triangulations of a hexagon. In terms of the variables defined in eq. (2.9), the six hedgehogs are:

$$\begin{aligned}\text{Hedgehogs for } A_3 = \Big\{&\{1/x_3^-, 1/e_1, x_1^+\}, \{1/x_1^+, 1/e_2, x_2^-\}, \{1/x_2^-, 1/e_3, x_3^+\},\\ &\{1/x_3^+, 1/e_4, x_1^-\}, \{1/x_1^-, 1/e_5, x_2^+\}, \{1/x_2^+, 1/e_6, x_3^-\}\Big\}\,.\end{aligned} \tag{4.5}$$

Each triple $\{x_1, x_2, x_3\}$ here is listed in numerically increasing order in the positive domain (this can done consistently, see section 3.1), and we recall that for each hedgehog $\{x_1, x_2, x_3\}$ there is a corresponding anti-hedgehog $\{1/x_3, 1/x_2, 1/x_1\}$. These are highlighted in red and blue, respectively, in figure 4. Altogether, the twelve triples are related by dihedral transformations of the hexagon, or equivalently by dihedral transformations of the six scattering particles. Of course, the factorization and exchange graphs in figures 4 and 5 manifest this symmetry as well.

At this point, expressing the NMHV amplitude in such a hedgehog basis is an exercise in linear algebra. Since all 12 (anti-)hedgehogs give equivalent bases, we chose the basis from $\{1/x_3^+, e_3, x_2^-\}$ which provides the shortest representation of the pure weight 4 terms. That means we use the basis explained in the previous subsection with the ordered set

$$\begin{aligned} Q = \{q_1, q_2, q_3\} &= \{-x_2^-, -e_3, -1/x_3^+\} \\ &= \left\{-(1+x_2+x_1x_2)x_3, -\frac{(1+x_1)x_2x_3}{1+x_3}, -\frac{x_1x_2x_3}{1+x_3+x_2x_3}\right\} \\ &= \left\{-\sqrt{\frac{w}{uvy_uy_vy_w}}, -\sqrt{\frac{w(1-u)}{(1-w)uy_uy_w}}, -\sqrt{\frac{vw}{uy_uy_vy_w}}\right\}. \end{aligned} \tag{4.6}$$

A simple calculation using the second line quickly reveals that the difference of each pair lies in the multiplicative span of the symbol letters shown in eq. (2.13), and also that they are listed in increasing numerical order in the positive domain. These properties are less apparent from the third line.

Each of the first 9 terms of the "y" alphabet can be written as a product of elements of the "q" alphabet so, by means of the symbol rule (2.6), the NMHV amplitude X can be written in the "q" alphabet. Each element of the hedgehog basis can be expressed in the same alphabet and, because the symbol map is linear, the symbol of the amplitude can be written as a linear combination of the symbols of the basis vectors. To find the coefficients of this linear combination, it is convenient to work in the ambient 9^4 dimensional space of length 4 symbols in the "q" alphabet. The symbols of the hedgehog basis vectors, together with the amplitude, constitute 1352 linear combinations in this larger space with one linear relation. Calculating the null space of the 1352×9^4 matrix with the linear algebra library `SparseSuite` gives the appropriate linear combination. To summarize, the result of this calculation is a particular linear combination of 376 elements of the weight-4 hedgehog basis whose symbol matches that of the amplitude X exactly. To find a representation for the full amplitude we turn in the next section to the problem of fixing terms of the form (transcendental numerical coefficient) $\times$ (functions of weight less than four).

4.3 Fixing beyond-the-symbol terms

If the symbols of two functions are equal, then the functions are equal, modulo "beyond-the-symbol" terms of lower weight. So this 376-term expression is the highest-weight part of the NMHV amplitude. A priori, we might expect up to 65 possible terms of lower weight. These include 55 weight-2 functions times $\zeta(2)$, 9 weight-1 functions times $\zeta(3)$, and one overall additive constant proportional to $\zeta(4)$. The coefficients of these 65 terms can be fixed

JHEP11(2015)136

by numerically evaluating the amplitude and our 376-term highest-weight expression at 65 random points in the positive domain and performing a row reduction. All the coefficients turn out to be rational numbers with small denominators. Our final result[13] is a 416-term expression for the 2-loop, 6-particle NMHV amplitude X. The validity of our ansatz, and solution, for the lower-weight terms has been stringently tested by comparing our result to the known expression at high precision for additional random kinematic points.[14]

5 Outlook

Hedgehog bases give a natural way to express 6-particle amplitudes, since they make manifest that these amplitudes have symbols which can be expressed in terms of A_3 cluster coordinates. In practice, this may translate into more "compact" representations of amplitudes than might be otherwise achieved. It should be stressed again that this the hedgehog basis is a true basis for cluster functions, with no functional or linear relations between its elements.

However, hedgehog bases are clearly not the ultimate solution for representing scattering amplitudes. The most important reason is that amplitudes satisfy a stringent analytic constraint on the possible locations of their branch points, which translates into a condition that allows only certain letters to appear in the first entry of their symbols. For example, 6-particle amplitudes may only have the letters $\{u, v, w\}$ in the first entry of their symbols, whereas all nine letters of the A_3 symbol alphabet appear as first entries in the hedgehog basis. It would be extremely interesting, as well as of great practical utility, to see if there is a natural way to construct bases of cluster functions manifesting this additional property. It would also be very interesting, both mathematically and physically, to find an appropriate extension of the hedgehog theorem to algebras other than A_n.

Acknowledgments

We have benefitted from enormously valuable discussions with J. Drummond, correspondence with L. Dixon, and collaboration on closely related topics with J. Golden and A. Goncharov. This work was supported by the U.S. Department of Energy under contract DE-SC0010010 (MS) and the DE-FG02-11ER41742 Early Career Award (AV), by the Sloan Research Foundation (AV), and by Brown University UTRA Awards (DP and AS). MS and AV are also grateful to the CERN Theory Group for support during the completion of this work.

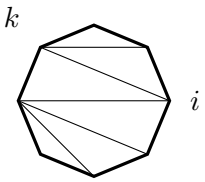
A Triangulations and A_n cluster algebras

Here we review from [30] the fact that in the special case case of A_n cluster algebras there is a convenient alternative to representing clusters with quivers: each cluster can instead

[13] Our result is included as an ancillary file with the arXiv submission.

[14] We are grateful to L.J. Dixon and A. McLeod for kindly providing an independent audit of our result in this manner.

be associated with a triangulation of an $(n+3)$-sided polygon. Beginning with a labeled $(n+3)$-gon, a triangulation is obtained by repeatedly adding non-crossing internal chords $\overline{ik}$ between nonadjacent vertices i, k until no further chords can be added. There are always n chords in a triangulation. For example, the five chords in a particular triangulation of an octagon are shown here:



Cluster coordinates are associated with these chords. Specifically, the chord $\overline{ik}$ shared between two triangles ijk and ikl is associated with the cluster variable[15]

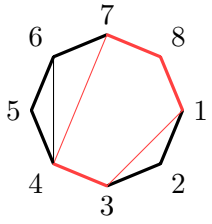
$$r(j,k,\ell,i) = \frac{1}{r(i,j,k,\ell)} := \frac{\langle jk\rangle \langle i\ell\rangle}{\langle ij\rangle \langle k\ell\rangle}\,, \tag{A.1}$$

where $\langle ij\rangle$ denotes the Plücker coordinate of two points z_i, z_j in $\mathbb{P}^1$.

In this representation, mutations are associated with chord-flips. To perform a chord-flip on $r(i,j,k,\ell)$, remove the chord $\overline{ik}$ and add the chord $\overline{jl}$. It is easy to see that the resulting variable $r(j,k,\ell,i)$ is indeed equal to $1/r(i,j,k,\ell)$; adjacent chords (those which lie on the same triangle) take the place of adjacent nodes in a quiver.

The added convenience of using triangulations over quivers comes from the fact that every triangulation is associated with a single cluster whose variables can be found explicitly via the formula above. Using this explicit formula, one can determine many useful facts about A_n: it's order is the Catalan number C_{n+1}, there are $2\binom{n+3}{4}$ cluster variables, and those variables are $r(i,j,k,\ell)$ for cyclically ordered i,j,k,ℓ.

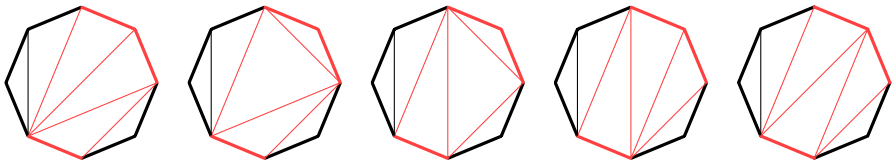
Triangulations also make it easy to enumerate and analyze subalgebras. Consider the case $n=5$. Clusters of A_5 correspond to triangulations of a labeled octagon. Selecting the vertices $1,3,4,7$, and 8, we can form a pentagon within our octagon:



[15]Here we use an inverse convention compared to [43] and parts of [1].

JHEP11(2015)136

Some triangulations of the octagon contain all of the edges of the pentagon as chords $(\overline{13}, \overline{34}, \overline{47}, \overline{48}, \overline{81})$. The subtriangulation obtained by discarding everything outside the pentagon is associated with a cluster of A_2. By flipping only the chords lying strictly within the pentagon, we can obtain other A_2 clusters, until we have an entire A_2 subalgebra:



By flipping the chords lying strictly outside the pentagon, or choosing a different pentagon to begin with, we can obtain different A_2 subalgebras. Note that the different subalgebras sharing the same pentagonal boundary must all have the same set of cluster variables; therefore, if we consider two subalgebras "equivalent" when they have the same variables; there are exactly $\binom{8}{5} = 56$ nonequivalent A_2 subalgebras of A_5.

This generalizes nicely: the clusters of A_m subalgebras of A_n correspond to $(m+3)$-vertex subtriangulations of $(n+3)$-gon triangulations. Up to equivalence, there are $\binom{n+3}{m+3}$ of these.

B A theorem on $1 + \mathcal{X}$ coordinates

In this appendix we prove that for all cluster coordinates x, the quantity $1 + x$ can be expressed as a product of cluster coordinates on the same algebra.

To be precise: suppose C is a cluster algebra of A, D, or E type whose quivers are connected with more than one node. Suppose $\mathcal{X}_C$ is its set of cluster coordinates. Then if $x_i \in \mathcal{X}_C$,

$$1 + x_i = \prod_{x_j \in \mathcal{X}} x_j^{n_j} \tag{B.1}$$

for some $n_j \in \{-1, 0, 1\}$.

The proof of this statement is straightforward. Pick some quiver of C that contains x_i. By connectedness, there exists some x_k connected to it. One of the properties of an A,D or E-type cluster algebra is that $|B_{ij}| \leq 1$ for all i, j. In particular, $B_{ik} \in \{-1, 0, 1\}$. Recall the mutation rule for cluster coordinates:

$$x_k' = \mu_i(x_k) = \begin{cases} x_i^{-1} & i = k \\ x_k\left(1 + x_i^{\operatorname{sgn} B_{ik}}\right)^{B_{ik}} & i \neq k\,. \end{cases} \tag{B.2}$$

If $B_{ik} = 1$, then

$$x_k' = x_k(1 + x_i^1)^1 \implies 1 + x_i = \frac{x_k'}{x_k}\,. \tag{B.3}$$

Otherwise, $B_{ik} = -1$, in which case

$$x_k' = x_k(1 + x_i^{-1})^{-1} = x_k\left(\frac{x_i}{1 + x_i}\right) \implies 1 + x_i = \frac{x_i x_k}{x_k'}\,. \tag{B.4}$$

JHEP11(2015)136

Since x_i and x_k are connected, $B_{ik} \neq 0$, so this is exhaustive. Thus $1 + x_i$ factors as a product of cluster coordinates.

Connectedness of a quiver is preserved by mutation, so if the initial quiver is connected, all quivers of an algebra are connected. Note also that the algebra A_1 as well as its derived algebras such as $A_1 \times A_2$ have disconnected quivers, hence expressions of the form $1 + x$ do not necessarily factor in this case.

Let us note here that a sort of converse statement, which has been stated and used for example in [1, 4, 5], remains a conjecture: if a and b are two elements of the multiplicative span $M_{\mathcal{A}}$ for some cluster algebra $\mathcal{A}$, and if $b = a + 1$, then precisely one element of the set $\{a, -1 - a, -1 - 1/a\}$ is a cluster coordinate.

C A cluster parameterization of 6-particle kinematics

We include here a parameterization of the positive domain of 6-particle scattering kinematics, in terms of momentum twistors, that we have found useful:

$$Z = \begin{pmatrix} -1 & 0 & 0 & 0 & 1 & 1 + x_1 \\ 0 & 1 & 0 & 0 & 1 + x_2 & 1 + x_2 + x_1 x_2 \\ 0 & 0 & -1 & 0 & 1 + x_3 + x_2 x_3 & 1 + x_3 + x_2 x_3 + x_1 x_2 x_3 \\ 0 & 0 & 0 & 1 & 1 & 1 \end{pmatrix}. \tag{C.1}$$

It is easily checked that this lies in the positive domain (that is, all minors $\langle ijkl \rangle > 0$ when $i < j < k < l$) whenever $x_1, x_2, x_3 > 0$, and that when plugged into the last column of eq. (2.9), it precisely reproduces the second column.

D Hedgehogs are cliques for A_n cluster algebras

Here we provide the details of the proof of the hedgehog theorem presented in section 3.3. First consider the case $n = 2$. The general case can be reduced to the $n = 2$ case, so it is worth doing in detail.

As reviewed in section 2.2, the mutation relations for $\mathcal{A} \cong A_2$ give ten cluster coordinates $\{x_1, \frac{1}{x_1}, \ldots, x_5, \frac{1}{x_5}\}$ related by

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

For $\mathcal{B} \cong A_1$ let us choose the subalgebra with coordinates $\{x_i, \frac{1}{x_i}\}$. So the relevant hedgehog is $\mathcal{X}(\mathcal{A}, \mathcal{B}) = \{x_{i+1}, \frac{1}{x_{i-1}}\}$. Pictorially, the hedgehog consists of the red and blue edges in the exchange graph shown in figure 6(a).

This is a clique because

$$x_{i+1} - \frac{1}{x_{i-1}} = \frac{1 + x_i}{x_{i-1}} - \frac{1}{x_{i-1}} = \frac{x_i}{x_{i+1}} \in M_{\mathcal{A}}. \tag{D.2}$$

Recasting this in terms of polygon triangulation is very illuminating. Recall that A_2 can be described in terms of pentagon triangulations. The red dashed lines in figure 6(b) are

JHEP11(2015)136

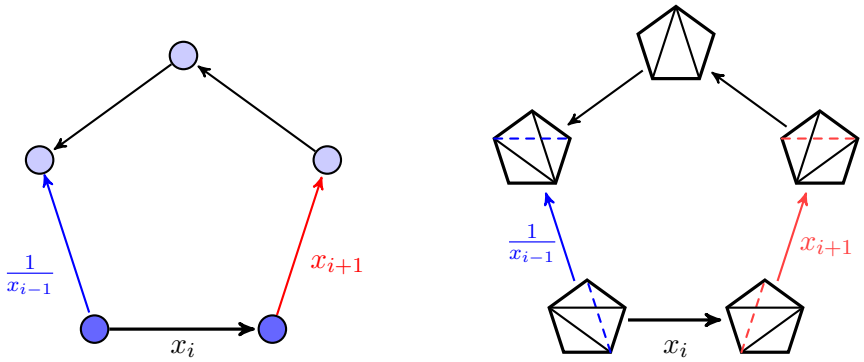
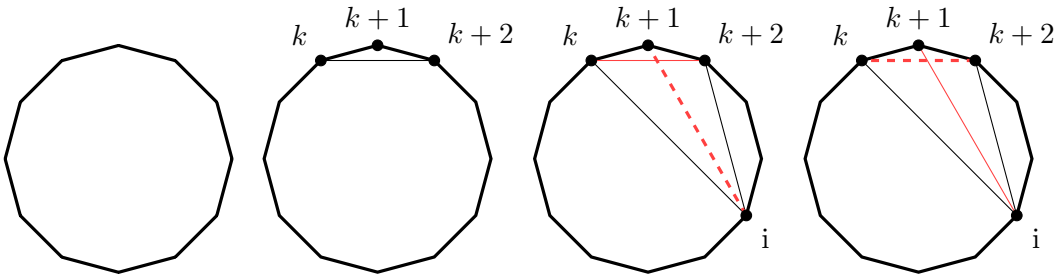


Figure 6. (a) The exchange graph for A_2, with the two vertices on the bottom row constituting an A_1 subalgebra. The hedgehog $\mathcal{X}(A_2, A_1)$ contains the two cluster variables $1/x_{i-1}$, x_{i+1} associated to the edges emanating away from the subalgebra. (b) The same exchange graph, but with each vertex showing the associated pentagon triangulation.

the chords that change as the red edge is traversed, and similarly for the blue. What we have shown then is that the difference between the cluster coordinates for the red and blue edges can be written in terms of products of cluster coordinates.

Now for the general case consider a $(n+3)$-gon. Choose three adjacent vertices $k, k+1, k+2$ and draw the chord $\overline{k(k+2)}$. This chord separates a triangle from an $(n+2)$-gon. The variable associated with this chord depends only on the triangles containing it, and not on the rest of the triangulation. The triangle on one side of the chord will always have the vertices $k, k+1, k+2$. The other triangle will have vertices $k, k+2, i$, where i is any of the n remaining vertices:



Therefore, there are n cluster variables that can be associated with $\overline{k(k+2)}$. These n variables are given by

$$\left\{r(i,k,k+1,k+2) : i \notin \{k,k+1,k+2\}\right\}. \tag{D.3}$$

Consider the subalgebra $B \cong A_{n-1}$ associated with the $(n+2)$-gon that excludes vertex $(k+1)$. Any triangulation containing $\overline{k(k+2)}$ will contain a triangulation of this polygon, and hence will be associated with a cluster in B. However, flipping $\overline{k(k+2)}$ will yield a triangulation that is not in B; therefore, the set of cluster variables associated with $\overline{k(k+2)}$ is a hedgehog of A_n! Because there are $n+3$ choices for k, all hedgehogs can be so obtained, and all will be of cardinality n.

JHEP11(2015)136

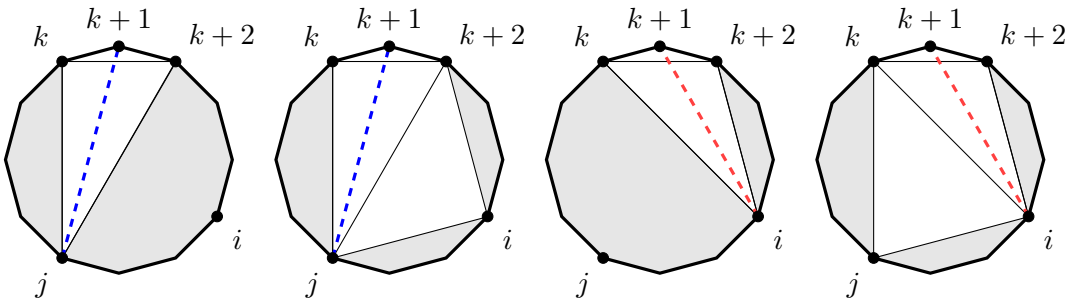


Figure 7. Partial triangulations of n-gons with an A_n hedgehog (red and blue edges). Notice that (b) and (d) reduce to the A_2 case of figure 6(b).

We can also obtain the anti-hedgehogs, which are associated with the result of any chord-flip of $\overline{k(k+2)}$:

$$\left\{r(k,k+1,k+2,i) : i \notin \{k,k+1,k+2\}\right\}. \tag{D.4}$$

Take x_i, x_j to be two arbitrary elements of the hedgehog, with (i,j,k) cyclically ordered. Then the mutation $x_i \mapsto 1/x_i$ corresponds to flipping $\overline{k(k+2)}$ to $\overline{(k+1)i}$ and $x_j \mapsto 1/x_j$ corresponds to flipping $\overline{k(k+2)}$ to $\overline{(k+1)j}$. These are indicated in figure 7 in (c) and (a) respectively.

Any other sub-triangulation of the gray region of (a) preserves x_j, so in particular one can choose the sub-triangulation with the pentagon $\{k,k+1,k+2,i,j\}$, shown in (b). Similarly, one can go from (c) to (d) and x_i will still be accessible by the red chord flip. But now notice that this is exactly the situation from the A_2 case! There is an embedded pentagon with exactly the same triangulations that appeared above. Therefore $x_i - x_j$ factors as a product of cluster coordinates, i.e. $x_i - x_j \in M_{A_n}$. We can also show this algebraically, making use of a Plücker relation, as displayed in eq. (3.11).

References

[1] J. Golden, A.B. Goncharov, M. Spradlin, C. Vergu and A. Volovich, *Motivic amplitudes and cluster coordinates*, *JHEP* **01** (2014) 091 [arXiv:1305.1617] [INSPIRE].

[2] N. Arkani-Hamed et al., *Scattering amplitudes and the positive Grassmannian*, arXiv:1212.5605 [INSPIRE].

[3] J. Golden, M.F. Paulos, M. Spradlin and A. Volovich, *Cluster polylogarithms for scattering amplitudes*, *J. Phys.* **A 47** (2014) 474005 [arXiv:1401.6446] [INSPIRE].

[4] J. Golden and M. Spradlin, *A cluster bootstrap for two-loop MHV amplitudes*, *JHEP* **02** (2015) 002 [arXiv:1411.3289] [INSPIRE].

JHEP11(2015)136

[5] J. Golden and M. Spradlin, *The differential of all two-loop MHV amplitudes in $\mathcal{N} = 4$ Yang-Mills theory*, *JHEP* **09** (2013) 111 [arXiv:1306.1833] [INSPIRE].

[6] J. Golden and M. Spradlin, *An analytic result for the two-loop seven-point MHV amplitude in $\mathcal{N} = 4$ SYM*, *JHEP* **08** (2014) 154 [arXiv:1406.2055] [INSPIRE].

[7] 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] [INSPIRE].

[8] L.J. Dixon, J.M. Drummond and J.M. Henn, *Bootstrapping the three-loop hexagon*, *JHEP* **11** (2011) 023 [arXiv:1108.4461] [INSPIRE].

[9] L.J. Dixon, J.M. Drummond and J.M. Henn, *Analytic result for the two-loop six-point NMHV amplitude in $\mathcal{N} = 4$ super Yang-Mills theory*, *JHEP* **01** (2012) 024 [arXiv:1111.1704] [INSPIRE].

[10] L.J. Dixon, J.M. Drummond, M. von Hippel and J. Pennington, *Hexagon functions and the three-loop remainder function*, *JHEP* **12** (2013) 049 [arXiv:1308.2276] [INSPIRE].

[11] L.J. Dixon, J.M. Drummond, C. Duhr and J. Pennington, *The four-loop remainder function and multi-Regge behavior at NNLLA in planar $\mathcal{N} = 4$ super-Yang-Mills theory*, *JHEP* **06** (2014) 116 [arXiv:1402.3300] [INSPIRE].

[12] L.J. Dixon and M. von Hippel, *Bootstrapping an NMHV amplitude through three loops*, *JHEP* **10** (2014) 065 [arXiv:1408.1505] [INSPIRE].

[13] L.J. Dixon, J.M. Drummond, C. Duhr, M. von Hippel and J. Pennington, *Bootstrapping six-gluon scattering in planar $\mathcal{N} = 4$ super-Yang-Mills theory*, PoS(LL2014)077 [arXiv:1407.4724] [INSPIRE].

[14] B. Basso, A. Sever and P. Vieira, *Spacetime and flux tube S-matrices at finite coupling for $\mathcal{N} = 4$ supersymmetric Yang-Mills theory*, *Phys. Rev. Lett.* **111** (2013) 091602 [arXiv:1303.1396] [INSPIRE].

[15] B. Basso, A. Sever and P. Vieira, *Space-time S-matrix and flux tube S-matrix II. Extracting and matching data*, *JHEP* **01** (2014) 008 [arXiv:1306.2058] [INSPIRE].

[16] B. Basso, A. Sever and P. Vieira, *Space-time S-matrix and flux-tube S-matrix III. The two-particle contributions*, *JHEP* **08** (2014) 085 [arXiv:1402.3307] [INSPIRE].

[17] B. Basso, A. Sever and P. Vieira, *Space-time S-matrix and flux-tube S-matrix IV. Gluons and fusion*, *JHEP* **09** (2014) 149 [arXiv:1407.1736] [INSPIRE].

[18] A.V. Belitsky, *Nonsinglet pentagons and NMHV amplitudes*, *Nucl. Phys.* **B 896** (2015) 493 [arXiv:1407.2853] [INSPIRE].

[19] A.V. Belitsky, *Fermionic pentagons and NMHV hexagon*, *Nucl. Phys.* **B 894** (2015) 108 [arXiv:1410.2534] [INSPIRE].

[20] F.C.S. Brown, *Multiple zeta values and periods of moduli spaces $\bar{\mathfrak{M}}_{0,n}(R)$*, *Annales Sci. Ecole Norm. Sup.* **42** (2009) 371 [math.AG/0606419] [INSPIRE].

[21] C. Bogner and F. Brown, *Feynman integrals and iterated integrals on moduli spaces of curves of genus zero*, *Commun. Num. Theor. Phys.* **09** (2015) 189 [arXiv:1408.1862] [INSPIRE].

[22] C. Duhr, *Mathematical aspects of scattering amplitudes*, arXiv:1411.7538 [INSPIRE].

[23] K.T. Chen, *Iterated path integrals*, *Bull. Am. Math. Soc.* **83** (1977) 831.

JHEP11(2015)136

[24] A.B. Goncharov, *Multiple polylogarithms and mixed Tate motives*, math.AG/0103059 [INSPIRE].

[25] A.B. Goncharov, *A simple construction of Grassmannian polylogarithms*, *Adv. Math.* **241** (2013) 79 [arXiv:0908.2238] [INSPIRE].

[26] 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] [INSPIRE].

[27] C. Duhr, H. Gangl and J.R. Rhodes, *From polygons and symbols to polylogarithmic functions*, *JHEP* **10** (2012) 075 [arXiv:1110.0458] [INSPIRE].

[28] A.B. Goncharov, *Galois symmetries of fundamental groupoids and noncommutative geometry*, *Duke Math. J.* **128** (2005) 209 [math.AG/0208144] [INSPIRE].

[29] S. Fomin and A. Zelevinsky, *Cluster algebras. I: Foundations*, *J. Am. Math. Soc.* **15** (2002) 497 [math.RT/0104151].

[30] S. Fomin and A. Zelevinsky, *Cluster algebras. II: Finite type classification*, *Invent. Math.* **154** (2003) 63 [math.RA/0208229].

[31] M. Gekhtman, M. Shapiro and A. Vainshtein, *Cluster algebras and Poisson geometry*, American Mathematical Society (2010), http://www.ams.org/bookstore-getitem/item=surv-167.

[32] L.K. Williams, *Cluster algebras: an introduction*, arXiv:1212.6263.

[33] V.V. Fock and A.B. Goncharov, *Cluster ensembles, quantization and the dilogarithm*, *Ann. Sci. Éc. Norm. Supér.* **42** (2009) 865 [math.AG/0311245] [INSPIRE].

[34] M.A.C. Torres, *Cluster algebras in scattering amplitudes with special 2D kinematics*, *Eur. Phys. J.* **C 74** (2014) 2757 [arXiv:1310.6906] [INSPIRE].

[35] M.F. Paulos and B.U.W. Schwab, *Cluster algebras and the positive Grassmannian*, *JHEP* **10** (2014) 031 [arXiv:1406.7273] [INSPIRE].

[36] J.M. Drummond, J. Henn, G.P. Korchemsky and E. Sokatchev, *Dual superconformal symmetry of scattering amplitudes in $\mathcal{N} = 4$ super-Yang-Mills theory*, *Nucl. Phys.* **B 828** (2010) 317 [arXiv:0807.1095] [INSPIRE].

[37] 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.* **B 826** (2010) 337 [arXiv:0712.1223] [INSPIRE].

[38] Z. Bern, J.J.M. Carrasco, H. Johansson and D.A. Kosower, *Maximally supersymmetric planar Yang-Mills amplitudes at five loops*, *Phys. Rev.* **D 76** (2007) 125020 [arXiv:0705.1864] [INSPIRE].

[39] L.F. Alday and J.M. Maldacena, *Gluon scattering amplitudes at strong coupling*, *JHEP* **06** (2007) 064 [arXiv:0705.0303] [INSPIRE].

[40] Z. Bern, M. Czakon, L.J. Dixon, D.A. Kosower and V.A. Smirnov, *The four-loop planar amplitude and cusp anomalous dimension in maximally supersymmetric Yang-Mills theory*, *Phys. Rev.* **D 75** (2007) 085010 [hep-th/0610248] [INSPIRE].

[41] 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] [INSPIRE].

JHEP11(2015)136

[42] A. Hodges, *Eliminating spurious poles from gauge-theoretic amplitudes*, *JHEP* **05** (2013) 135 [arXiv:0905.1473] [INSPIRE].

[43] J.S. Scott, *Grassmannians and cluster algebras*, *Proc. London Math. Soc.* **92** (2006) 345 [math.CO/0311148].

[44] V. Del Duca, C. Duhr and V.A. Smirnov, *An analytic result for the two-loop hexagon Wilson loop in* $\mathcal{N} = 4$ *SYM*, *JHEP* **03** (2010) 099 [arXiv:0911.5332] [INSPIRE].

[45] V. Del Duca, C. Duhr and V.A. Smirnov, *The two-loop hexagon Wilson loop in* $\mathcal{N} = 4$ *SYM*, *JHEP* **05** (2010) 084 [arXiv:1003.1702] [INSPIRE].

[46] S. Caron-Huot, *Superconformal symmetry and two-loop amplitudes in planar* $\mathcal{N} = 4$ *super Yang-Mills*, *JHEP* **12** (2011) 066 [arXiv:1105.5606] [INSPIRE].

[47] S. Caron-Huot and S. He, *Jumpstarting the all-loop S-matrix of planar* $\mathcal{N} = 4$ *super Yang-Mills*, *JHEP* **07** (2012) 174 [arXiv:1112.1060] [INSPIRE].

[48] V. Del Duca, C. Duhr and V.A. Smirnov, *A two-loop octagon Wilson loop in* $\mathcal{N} = 4$ *SYM*, *JHEP* **09** (2010) 015 [arXiv:1006.4127] [INSPIRE].

[49] P. Heslop and V.V. Khoze, *Analytic results for MHV Wilson loops*, *JHEP* **11** (2010) 035 [arXiv:1007.1805] [INSPIRE].

[50] P. Heslop and V.V. Khoze, *Wilson loops @ 3-loops in special kinematics*, *JHEP* **11** (2011) 152 [arXiv:1109.0058] [INSPIRE].

[51] T. Goddard, P. Heslop and V.V. Khoze, *Uplifting amplitudes in special kinematics*, *JHEP* **10** (2012) 041 [arXiv:1205.3448] [INSPIRE].

[52] S. Caron-Huot and S. He, *Three-loop octagons and n-gons in maximally supersymmetric Yang-Mills theory*, *JHEP* **08** (2013) 101 [arXiv:1305.2781] [INSPIRE].

[53] C. Duhr, *Hopf algebras, coproducts and symbols: an application to Higgs boson amplitudes*, *JHEP* **08** (2012) 043 [arXiv:1203.0454] [INSPIRE].

[54] J.M. Drummond, private communication.

[55] D.E. Radford, *A natural ring basis for the shuffle algebra and an application to group schemes*, *J. Algebra* **58** (1979) 432.

JHEP11(2015)136