A symbol of uniqueness: the cluster bootstrap for the 3-loop MHV heptagon

J.M. Drummond,[a,b,c] **G. Papathanasiou**[c] **and M. Spradlin**[d]

[a] *School of Physics & Astronomy, University of Southampton,*
Highfield, Southampton, SO17 1BJ, United Kingdom

[b] *Theory Division, Physics Department, CERN,*
CH-1211 Geneva 23, Switzerland

[c] *LAPTh, CNRS, Université de Savoie,*
F-74941 Annecy-le-Vieux Cedex, France

[d] *Department of Physics, Brown University,*
Providence, RI 02912, U.S.A.

E-mail: james.matthew.drummond@cern.ch, georgios@lapth.cnrs.fr, marcus_spradlin@brown.edu

ABSTRACT: Seven-particle scattering amplitudes in planar super-Yang-Mills theory are believed to belong to a special class of generalised polylogarithm functions called heptagon functions. These are functions with physical branch cuts whose symbols may be written in terms of the 42 cluster $\mathcal{A}$-coordinates on $\mathrm{Gr}(4,7)$. Motivated by the success of the hexagon bootstrap programme for constructing six-particle amplitudes we initiate the systematic study of the symbols of heptagon functions. We find that there is exactly one such symbol of weight six which satisfies the MHV last-entry condition and is finite in the 7 ∥ 6 collinear limit. This unique symbol is both dihedral and parity-symmetric, and remarkably its collinear limit is exactly the symbol of the three-loop six-particle MHV amplitude, although none of these properties were assumed a priori. It must therefore be the symbol of the three-loop seven-particle MHV amplitude. The simplicity of its construction suggests that the n-gon bootstrap may be surprisingly powerful for $n > 6$.

Contents

1 Introduction

A dream goal of the analytic S-matrix programme is to be able to construct expressions for the scattering amplitudes of a quantum field theory based on a few physical principles and a thorough knowledge of the analytic structure. In this work we are able to tie together recent advances in determining amplitudes by an analytic "bootstrap" procedure with discoveries about general classes of analytic functions which appear to play a central role. The theory we will study is the planar $\mathcal{N} = 4$ supersymmetric gauge (SYM) theory in four dimensions [1], where the greatest advances have been made in explicitly determining the scattering amplitudes.

The analytic structure of the S-matrices of general quantum field theories are notoriously complicated [2]. For the planar $\mathcal{N} = 4$ super Yang-Mills theory however, several simplifying features come into play which reduce the complexity sufficiently to allow conjectures to be made about which classes of functions describe the scattering amplitudes, at least in the simplest cases. The duality with Wilson loops [3–9] and the associated dual conformal symmetry [7, 10–13] of the planar theory mean that only amplitudes with six or more external legs are non-trivial. Moreover, for the six-particle case, the same amplitude/Wilson loop duality has allowed some explicit results [14] and beautifully simple expressions [15] to be obtained.

All of the above developments have led to an analytic bootstrap programme, so far focused on the six-particle ("hexagon") case [16–21]. The idea of the hexagon bootstrap programme is to declare that, order by order in perturbation theory, six-particle amplitudes can be determined in terms of a particular class of multiple polylogarithms called hexagon functions. Hexagon functions are polylogarithms associated to a natural nine-letter alphabet of singularities which can be identified with the nine multiplicatively independent cross-ratios one can form from six points in $\mathbb{CP}^1$. In addition, hexagon functions obey conditions on the locations of branch cuts, encoding the fact that amplitudes can have discontinuities only in certain kinematical regions.

In the case of maximally helicity-violating (MHV) amplitudes, the relevant piece (obtained by subtracting particular universal infrared-divergent terms [22, 23]) which is not fixed by dual conformal symmetry is called the remainder function. At L loops the n-particle remainder function $R_n^{(L)}$ should be a polylogarithm of weight $2L$ obeying additional criteria coming from various physical constraints. One constraint is that the remainder function should be fully dihedrally invariant, that is invariant under cyclic permutations $i \to i + 1$ and flips $i \to n + 1 - i$ of its particle labels. This is essentially because supersymmetry dictates that the MHV amplitudes are given by an overall supersymmetric [24, 25] Nair-Parke-Taylor factor at tree-level which then receives multiplicative quantum corrections. In addition the remainder function should approach $R_n^{(L)} \to R_{n-1}^{(L)}$ smoothly (i.e. with power suppressed corrections) in the limit where the momenta of two colour-adjacent particles become collinear.

Moreover, the relation to Wilson loops means that the remainder function must obey constraints on its discontinuities [26–29] and on the power-suppressed corrections [30–36] in the collinear limit coming from an operator product expansion (OPE) for light-like Wilson loops. The latter expansion is governed by the dynamics of an integrable colour-electric flux-tube, which in particular gives rise to all-loop integral formulas for individual power-suppressed terms. The success in systematically evaluating these in the weak coupling expansion for the first few terms [37, 38] gives hope that it may be even possible to resum the OPE (see [39] for a first step in this direction) to obtain full amplitudes. More generally, it is expected that the integrability of the theory [40] will play an instrumental role in determining its S-matrix, and apart from the collinear limit it has also led to all-loop expressions in the multi-Regge limit [41], another kinematical regime that has provided significant information on the remainder function [42–53].

Finally, the extension of dual conformal symmetry to dual superconformal symmetry [54], is expressed via the super-Wilson-loop correspondence [55–57] in terms of recursive differential equations [58, 59] which imply certain universal constraints on the total derivative of the remainder function. Including the original superconformal symmetry of the amplitudes, or equivalently invoking parity for the amplitudes, extends dual superconformal symmetry to its Yangian [60, 61], leading to further recursive equations, relevant for determining non-MHV amplitudes.

The hexagon bootstrap programme has yielded explicit expressions up to four loops for the MHV amplitudes [16, 18, 19] and three loops in the NMHV case [17, 21]. For higher multiplicities, explicit results in SYM theory so far have been confined (in general kinematics) to two loops [62–65]. With explicit two-loop results to hand, an important structural observation has been made [66]: the results are so far consistent with the conjecture that the relevant classes of functions are given by multiple polylogarithms whose singularities are dictated by a sequence of cluster algebras [67, 68]. This observation will be of central importance here because it will allow us to generalise the bootstrap programme to higher multiplicities, beginning with seven-particle ("heptagon") MHV amplitudes.

While one might expect that the bootstrap for heptagons would be a similar but more involved version of the bootstrap for hexagons, we are in fact led to a surprising and counterintuitive result. Up to three loops, to determine the symbol of the MHV remainder function, we need only construct the symbols of heptagon functions obeying the differential constraint coming from dual superconformal symmetry, and then demand that we have a linear combination of them which is finite in the collinear limit. Dihedral symmetry follows for free and no information coming from the OPE expansion of Wilson loops or the Regge limit of amplitudes is required at all. Moreover, the hexagon remainder can be obtained for free by taking the collinear limit of the heptagon remainder function. In this sense the heptagon bootstrap provides a conceptually more powerful framework for constructing even the hexagon amplitudes!

The plan of this paper is as follows. In section 2 we review the basic details needed to motivate the definition of heptagon functions. In section 3 we review a few of the simplest general properties of MHV amplitudes in SYM theory, which in the bootstrap programme are applied as constraints on the space of heptagon functions. Section 4 contains a discussion of algorithms for imposing the constraint of integrability, which is by far the most significant computational challenge in applying the bootstrap. The expert reader may wish to jump directly (after taking a peek at the heptagon alphabet shown in eq. (2.12)) to section 5, where our main results are discussed and summarised in table 1. In section 6 we attempt to formulate some explanation for why the heptagon (and higher-n) bootstrap is unexpectedly powerful.

Attached to the arXiv submission of this paper the reader may find data files containing: (1) the heptagon symbol alphabet shown in eq. (2.12), (2) the symbol of the remainder function $R_7^{(3)}$, and (3) the symbol of the other irreducible weight-6 heptagon function which satisfies the MHV last-entry condition (see subsection 5.2).

2 Heptagon functions

In this section we review some basic facts about generalised polylogarithms and symbols, leading up to our definition of heptagon functions which mirrors that of the hexagon functions studied in [16–20].

2.1 Symbols

Precise definitions and additional details may be found in [69–72] (see also [73] for a review), but here it is sufficient to recall the recursive definition according to which f_k is called a generalised polylogarithm function of weight (or transcendentality) k if its total differential may be written as a finite linear combination

$$df_k = \sum_\alpha f_{k-1}^{(\alpha)} \, d\log\phi_\alpha \tag{2.1}$$

over some set of ϕ_α, where the coefficients $f_{k-1}^{(\alpha)}$ are functions of weight $k-1$. Functions of weight 1 are defined to be finite linear combinations (with rational coefficients) of $\log\phi_\alpha$. By applying total derivatives d to each of the coefficient functions $f_{k-1}^{(\alpha)}$ and using property (2.1) recursively we arrive at a collection of rational numbers $f_0^{(\alpha_1,\alpha_2,\dots,\alpha_k)}$ characterising the original function f_k. The symbol $\mathcal{S}(f_k)$ encapsulates this data via the definition

$$\mathcal{S}(f_k) = \sum_{\alpha_1,\dots,\alpha_k} f_0^{(\alpha_1,\alpha_2,\dots,\alpha_k)} \left(\phi_{\alpha_1} \otimes \cdots \otimes \phi_{\alpha_k}\right). \tag{2.2}$$

Since $\log\phi_1\phi_2 = \log\phi_1 + \log\phi_2$, it is evident that the consistency of eqs. (2.1) and (2.2) requires symbols to satisfy

$$(\cdots \otimes \phi_1\phi_2 \otimes \cdots) = (\cdots \otimes \phi_1 \otimes \cdots) + (\cdots \otimes \phi_2 \otimes \cdots). \tag{2.3}$$

Moreover,

$$(\cdots \otimes c \otimes \cdots) = 0 \tag{2.4}$$

for any numerical constant c, since $d\log c = 0$. The collection of ϕ_α which appear in the symbol of a given function is called its *symbol alphabet*. A symbol alphabet is never uniquely defined because one can use the identity (2.3) to write symbols in various ways. Two alphabets $\{\phi_\alpha\}$, $\{\phi'_\alpha\}$ are considered equivalent if there exists a linear transformation

$$\log\phi_\alpha = \sum_\beta M_{\alpha\beta} \log\phi'_\beta \tag{2.5}$$

given by an invertible matrix M whose entries are rational numbers.

We consider the set of polylogarithm functions of weight k as a vector space over the rational numbers. Moreover, since the product of two functions of weights k_1 and k_2 is a function of weight $k_1 + k_2$, they constitute a graded algebra. The irreducible elements of this algebra are functions which cannot be written as products of lower-weight functions.

The algebra of generalised polylogarithm functions admits a coproduct Δ compatible with multiplication, rendering it a Hopf algebra [70] (see also [74] for a review aimed at

physicists). Moreover there is a cobracket δ which squares to zero (when acting on the quotient space of all functions modulo products of lower-weight functions), giving this algebra the structure of a Lie co-algebra. We will make no direct use of these higher mathematical structures in the present paper, but these tools have been very useful in elucidating the structure of two-loop MHV amplitudes [66, 75–77] and in particular the relation between their coproducts and the cluster Poisson structure [78] on the kinematical space on which they are defined.

2.2 Symbol alphabets

A fundamental assumption of the "cluster bootstrap" programme is that the symbol alphabet relevant for n-particle amplitudes in SYM theory consists of the special collection of functions called cluster $\mathcal{A}$-coordinates on the kinematical configuration space $\mathrm{Conf}_n(\mathbb{P}^3) = \mathrm{Gr}(4, n)/(\mathbb{C}^*)^{n-1}$. We describe the space as kinematical because the n points Z_i (known as momentum twistors [79]) in $\mathbb{P}^3$ define, after choosing a preferred bitwistor $I \in \mathbb{P}^3 \wedge \mathbb{P}^3$, a light-like polygonal contour in Minkowski space-time via

$$x_i \sim \frac{Z_{i-1} \wedge Z_i}{\langle Z_{i-1} Z_i I \rangle}, \tag{2.6}$$

where $\langle Z_i Z_j Z_k Z_l \rangle = \det(Z_i Z_j Z_k Z_l)$. The particle momenta can be identified with the null separations of neighbouring points,

$$p_i = x_{i+1} - x_i, \tag{2.7}$$

and the kinematical Mandelstam variables can be identified with the non-zero separations and hence related to the momentum twistors,

$$(p_i + p_{i+1} + \ldots p_{j-1})^2 = (x_i - x_j)^2 = \frac{\langle Z_{i-1} Z_i Z_{j-1} Z_j \rangle}{\langle Z_{i-1} Z_i I \rangle \langle Z_{j-1} Z_j I \rangle}. \tag{2.8}$$

There is a vast mathematical literature on cluster algebras; we refer the reader to [66] for an introduction focused on amplitudes. The cluster $\mathcal{A}$-coordinates relevant to n-particle amplitudes consist of the Plücker coordinates $\langle ijkl \rangle \equiv \langle Z_i Z_j Z_k Z_l \rangle$ which can be formed from the momentum twistors Z_i specifying the kinematics of the scattering particles, together with certain very particular homogeneous polynomials in Plücker coordinates which can be systematically constructed via an algorithm known as mutation.

The fact that amplitudes in SYM theory depend on the individual Z_i only through the (projective) SL(4) invariants $\langle ijkl \rangle$ is a consequence of dual conformal symmetry. The individual four-brackets are not invariant under projective transformations of the homogeneous coordinates Z_i on $\mathbb{P}^3$, so they must always appear in projectively invariant ratios.

The case $n = 6$ is the simplest, since mutation does not generate any $\mathcal{A}$-coordinates beyond the standard Plücker coordinates on $\mathrm{Gr}(4, 6)$. From the 15 individual four-brackets

we can form 9 invariant ratios, for example

$$
\begin{aligned}
u &= \frac{\langle 6123\rangle\langle 3456\rangle}{\langle 6134\rangle\langle 2356\rangle}\,, & v &= \frac{\langle 1234\rangle\langle 4561\rangle}{\langle 1245\rangle\langle 3461\rangle}\,, & w &= \frac{\langle 2345\rangle\langle 5612\rangle}{\langle 2356\rangle\langle 4512\rangle}\,, \\
1-u &= \frac{\langle 5613\rangle\langle 6234\rangle}{\langle 6134\rangle\langle 2356\rangle}\,, & 1-v &= \frac{\langle 6124\rangle\langle 1345\rangle}{\langle 1245\rangle\langle 3461\rangle}\,, & 1-w &= \frac{\langle 1235\rangle\langle 2456\rangle}{\langle 2356\rangle\langle 4512\rangle}\,, & (2.9) \\
y_u &= \frac{\langle 1345\rangle\langle 2456\rangle\langle 1236\rangle}{\langle 1235\rangle\langle 3456\rangle\langle 1246\rangle}\,, & y_v &= \frac{\langle 1235\rangle\langle 2346\rangle\langle 1456\rangle}{\langle 1234\rangle\langle 2456\rangle\langle 1356\rangle}\,, & y_w &= \frac{\langle 2345\rangle\langle 1356\rangle\langle 1246\rangle}{\langle 1345\rangle\langle 2346\rangle\langle 1256\rangle}\,.
\end{aligned}
$$

This particular choice of basis has been widely used in the literature, but as mentioned above any multiplicatively transformed set of ratios would serve just as well, if one is interested in working only at the level of symbols. The so-called hexagon bootstrap is predicated on the assumption that all L-loop six-particle amplitudes (both MHV and non-MHV) in SYM theory are generalised polylogarithm functions of weight $k = 2L$ whose symbols can be written in terms of the nine-letter alphabet shown in eq. (2.9). This hypothesis has been successfully tested for the MHV remainder function through four loops [16, 18, 19] and the NMHV ratio function through three loops [17, 21]. Further support for its validity comes from a particular "dlog" representation of the all-loop integrand [80], as well as the all-loop basis of harmonic polylogarithms found for the first few orders of these amplitudes in an expansion around the collinear limit [37, 38].

For $n > 7$ the cluster algebra associated to $\mathrm{Conf}_n(\mathbb{P}^3)$ has infinitely many $\mathcal{A}$-coordinates. This is not necessarily an obstacle to the cluster bootstrap programme as long as only a finite number of them appear at any finite order in perturbation theory. For example, it is known [65] that the two-loop n-particle MHV remainder function is written in terms of a symbol alphabet of precisely $\frac{3}{2}n(n-5)^2$ (projectively invariant) letters. It would be very interesting to determine whether, for example, the symbol of the three-loop eight-particle MHV remainder function may be written in terms of the same 108 letters which appear already at two loops, or whether it requires more exotic cluster $\mathcal{A}$-coordinates. This amplitude has been evaluated in in two-dimensional kinematics [81], but unfortunately this limit appears to be insufficient to decide the question.

In this paper we focus on the $n = 7$ Goldilocks zone, where the number of $\mathcal{A}$-coordinates is still finite, but in addition to the Plücker coordinates $\langle ijkl\rangle$ there are 14 $\mathcal{A}$-coordinates which are bilinears of the form

$$
\langle a(bc)(de)(fg)\rangle \equiv \langle abde\rangle\langle acfg\rangle - \langle abfg\rangle\langle acde\rangle\,. \tag{2.10}
$$

This notation emphasises the antisymmetry under exchange of any pair of indices inside parentheses, as well as antisymmetry under the exchange of the pairs amongst each other. In the mathematical literature on cluster algebras the n cyclic Plücker coordinates $\langle i\,i{+}1\,i{+}2\,i{+}3\rangle$ are usually treated differently and are sometimes called "coefficients" instead of "coordinates". With this terminology, there are precisely 42 cluster $\mathcal{A}$-coordinates for the case $n = 7$, given by

$$
\langle 2367\rangle\,, \quad \langle 2567\rangle\,, \quad \langle 2347\rangle\,, \quad \langle 2457\rangle\,, \quad \langle 1(23)(45)(67)\rangle\,, \quad \text{and} \ \langle 1(34)(56)(72)\rangle\,, \tag{2.11}
$$

together with their images under cyclic transformations $Z_i \to Z_{i+1}$.

Projectively invariant ratios can be formed by dressing each of these 42 coordinates with suitable powers of the $\langle i\, i+1\, i+2\, i+3 \rangle$ Plücker coordinates, as it is always possible to construct products of the latter with helicity weight at a single point, and combine them so as to cancel the excess weight of the points appearing in (2.11). We have found a convenient choice to be

$$
\begin{aligned}
a_{11} &= \frac{\langle 1234 \rangle \langle 1567 \rangle \langle 2367 \rangle}{\langle 1237 \rangle \langle 1267 \rangle \langle 3456 \rangle}\,, & a_{41} &= \frac{\langle 2457 \rangle \langle 3456 \rangle}{\langle 2345 \rangle \langle 4567 \rangle}\,, \\
a_{21} &= \frac{\langle 1234 \rangle \langle 2567 \rangle}{\langle 1267 \rangle \langle 2345 \rangle}\,, & a_{51} &= \frac{\langle 1(23)(45)(67) \rangle}{\langle 1234 \rangle \langle 1567 \rangle}\,, & (2.12) \\
a_{31} &= \frac{\langle 1567 \rangle \langle 2347 \rangle}{\langle 1237 \rangle \langle 4567 \rangle}\,, & a_{61} &= \frac{\langle 1(34)(56)(72) \rangle}{\langle 1234 \rangle \langle 1567 \rangle}\,,
\end{aligned}
$$

together with a_{ij} obtained from a_{i1} by cyclically relabeling $Z_m \to Z_{m+j-1}$. While the a_{ij} are multiplicatively independent, they are of course not algebraically independent: the dimension of $\mathrm{Conf}_7(\mathbb{P}^3)$ is only six, so one could choose to parameterise all 42 of the a_{ij} in terms of just 6 free variables if needed.

As noted in eq. (2.5) the choice of symbol alphabet is not unique or canonical. In contrast, it has been noted [66, 82] that the coproducts of two-loop MHV remainder functions $R_n^{(2)}$ involve only preferred cross-ratios known as cluster $\mathcal{X}$-coordinates on $\mathrm{Conf}_n(\mathbb{P}^3)$. None of the a_{ij} in eq. (2.12) are cluster $\mathcal{X}$-coordinates for $n = 7$. The latter, which have been tabulated in section 7.3 of [66], may be expressed as products of powers of the former. It would be interesting to understand if there is a connection between coproducts and $\mathcal{X}$-coordinates beyond two loops, or for non-MHV amplitudes.

It is also interesting to note that only 14 out of the 105 possible distinct $\langle a(bc)(de)(fg) \rangle$ objects appear in eq. (2.11). This is indicative of a qualitative difference between the cases $n = 6$ and $n > 6$. For $n = 6$ the set of $\mathcal{A}$-coordinates, as a whole, is invariant (up to overall signs, which are never a concern inside symbols) under arbitrary permutations of the particle labels, not just under cyclic permutations. However for $n > 6$, a non-cyclic permutation would actually change the symbol alphabet. For example, switching $1 \leftrightarrow 4$ would have no substantive effect on eq. (2.9) (it would rearrange the letters to an equivalent basis), but it would completely change the heptagon basis (2.12) by introducing genuinely new letters which are not cyclic rotations of those in eq. (2.11).

This dependence of the symbol alphabet on the choice of dihedral structure, i.e. on a particular ordering of the particles, is in fact natural. We recall that when we refer to "amplitude" we really mean the colour-ordered partial amplitude $\mathcal{A}(1, \dots, n)$ which produces the full amplitude upon summation over non-cyclic permutations σ,

$$
\mathcal{A}_{\text{full}} = \sum_\sigma \mathrm{Tr}(T^{a_{\sigma(1)}} \dots T^{a_{\sigma(n)}}) \, \mathcal{A}\big(\sigma(1), \dots, \sigma(n)\big)\,. \tag{2.13}
$$

Thus, while for six particles, each colour-ordered partial amplitude is described by one and the same class of polylogarithms, the general case requires different classes of polylogarithms for different colour-ordered partial amplitudes.

The *heptagon bootstrap* which we initiate in this paper is based on the hypothesis that all L-loop seven-particle amplitudes (whether MHV or non-MHV) are generalised

polylogarithm functions of weight $k = 2L$ whose symbols can be written in terms of the 42-letter alphabet shown in eq. (2.12).

2.3 Integrable words

Given a random symbol $\mathcal{S}$ of weight $k > 1$, there does not in general exist any function whose symbol is $\mathcal{S}$. A symbol of the form (2.2) is said to be *integrable*, (or, to be an *integrable word*) if it satisfies

$$\sum_{\alpha_1,\dots,\alpha_k} f_0^{(\alpha_1,\alpha_2,\dots,\alpha_k)} \underbrace{(\phi_{\alpha_1} \otimes \cdots \otimes \phi_{\alpha_k})}_{\text{omitting } \phi_{\alpha_j} \otimes \phi_{\alpha_{j+1}}} d\log \phi_{\alpha_j} \wedge d\log \phi_{\alpha_{j+1}} = 0 \qquad \forall j \in \{1,\dots,k-1\}. \tag{2.14}$$

These are necessary and sufficient conditions for a function f_k with symbol $\mathcal{S}$ to exist.

There are 42^k distinct symbols of weight k which can be written in the 42-letter symbol alphabet shown in eq. (2.12), but only certain linear combinations of these satisfy the integrability conditions (2.14). Determining these linear combinations is, in general, a computationally difficult problem which we discuss in detail in section 4 below. For weight $k = 1, 2, 3$, only $42, 1035, 19536$ linear combinations of the $42^k = 42, 1764, 74088$ available symbols are integrable, and hence correspond to actual functions.

It is relatively easy to tabulate these functions explicitly. We begin with the fact that any generalised polylogarithm function of weight 3 or less can be written in terms of the classical polylogarithm functions $\mathrm{Li}_k(x)$. Since

$$\mathcal{S}\big(\mathrm{Li}_k(x)\big) = -(1-x) \otimes \underbrace{x \otimes \cdots \otimes x}_{k-1 \text{ times}} \tag{2.15}$$

we can allow the argument x to be any product of powers of the a_{ij} with the property that $1 - x$ can also be expressed as a product of powers of a_{ij}'s. There are precisely 2310 distinct x's of this type.

At weight 2 not all 2310 of the $\mathrm{Li}_2(x)$'s are independent since there are many identities for the Li_2 function. These include $\mathrm{Li}_2(x) \approx -\mathrm{Li}_2(1/x) \approx -\mathrm{Li}_2(1-x)$ (where $\approx$ means modulo products of functions of lower weight, i.e. modulo $\mathcal{O}(\log^2)$ in this case), as well as the pentagon identity. It can be checked that only 132 out of the 2310 $\mathrm{Li}_2(x)$'s are linearly independent mod $\mathcal{O}(\log^2)$. Hence the vector space of irreducible weight-2 integrable words has dimension 132.

At weight 3 we have the identities $\mathrm{Li}_3(x) \approx \mathrm{Li}_3(1/x)$ and $\mathrm{Li}_3(x) + \mathrm{Li}_3(1-x) + \mathrm{Li}_3(1-1/x) \approx 0$, which leave $2310/3 = 770$ independent functions. There are also 22 linearly independent D_4 identities [66], so we see that there are precisely 748 linearly independent irreducible weight-3 integrable words.

Having determined that there are 42, 132, 748 irreducible functions at weight $k = 1, 2, 3$, it is simple to consider all possible ways of taking products of lower-weight functions to count the total number of functions $42, 1035, 19536$ given above. Let us stress that here we have counted all functions written from the heptagon alphabet, but we will only give the name "heptagon function" to the subset satisfying an important analytic constraint to which we now turn our attention.

2.4 Physical singularities

Most of the functions discussed in the previous section have no possible relevance to amplitudes. One simple criterion which eliminates many of them is locality, which imposes tight constraints on the analytic properties of any scattering amplitude. In particular, it is a basic consequence of locality that amplitudes may only have singularities when some intermediate particle goes on-shell. For planar colour-ordered amplitudes in massless theories this can only happen when some sum of cyclically adjacent momenta $p_i + p_{i+1} + \cdots + p_{j-1} = x_j - x_i$ becomes null. In terms of momentum twistors, this means that amplitudes may only have singularities when $\langle i-1\, i\, j-1\, j \rangle \to 0$ for some i, j. The Euclidean region, in which amplitudes must be free of branch points, corresponds to having all non-neighbouring separations $x_j - x_i$ space-like.

The effect of this branch cut condition on the symbol of a seven-particle amplitude is that only the seven a_{1j} are allowed to appear in the first entry. This is because the singularities of generalised polylogarithm functions are encoded in the first entry of their symbols: specifically, a letter ϕ appearing in the first entry indicates that the corresponding function has branch points at $\phi = 0$ and $\phi = \infty$. From eq. (2.12) we see that only the a_{1j} are built entirely out of Plücker coordinates of the form $\langle i-1\, i\, j-1\, j \rangle$; the other letters contain other brackets which could not possibly cancel between different additive terms in a symbol since the a_{ij} are multiplicatively independent. The restriction that only $\langle i-1\, i\, j-1\, j \rangle$ may appear in the first entry is referred to as the *first-entry condition*.

2.5 Heptagon functions

Following the definition of hexagon functions given in [18], we define a *heptagon function* of weight k to be a polylogarithm function of weight k whose symbol may be written in the alphabet (2.12) and which is free of branch points in the Euclidean region. As discussed in the previous subsection, such functions have symbols in which only the letters a_{1j} appear in the first entry.

We follow the standard convention of counting heptagon functions of a certain weight only modulo the addition of functions of lower weight (times numerical constants of the appropriate transcendental weight). Although in this paper we work entirely at the level of symbols, if we restrict to the definitions and conventions we have introduced so far, the counting of the heptagon functions and the counting of their symbols will coincide.

More generally however, it is important to note that when additional constraints are imposed, the number of heptagon functions satisfying them may be smaller than the number of their symbols. In particular, it can happen that a symbol which satisfies the first-entry condition and is well-defined in a collinear limit can be promoted to a function with physical branch cuts only by adding certain terms of lower weight, which may end up diverging in the collinear limit. An example of this phenomenon has already been seen at three loops in the MHV hexagon case [16, 18].

Bearing this caveat in mind, especially in light of the fact that we will be examining collinear limits in what follows, we will be careful to only interchange the terms "heptagon function" and "symbol of heptagon function" when the counting coincides, and otherwise

employ the term *heptagon symbol* to denote the latter in a more abbreviated fashion. Finally, it should be understood that we are really counting dimensions of vector spaces of symbols, not individual symbols, so when we say there is a unique symbol with certain properties, we mean unique up to an overall multiplicative factor.

Using the algorithms described in section 4 below, we have found that the dimension of the space of heptagon functions is 7, 42, 237, 1288, 6763 for $k = 1, 2, 3, 4, 5$. These numbers, and the dimensions of various physically interesting subspaces, are tabulated in table 1. It follows from this counting that the vector space of irreducible heptagon functions has dimension 7, 14, 55, 196, 708 for $k = 1, 2, 3, 4, 5$.

3 MHV constraints

We believe that all seven-particle amplitudes in SYM theory are heptagon functions as defined in the previous section. In this section we discuss some of the additional properties special to MHV amplitudes, which will be the focus of most of the remainder of the paper.

3.1 The $\bar{Q}$ equation

It has been argued in [65], and subsequently shown to be a consequence of a proposed anomaly equation for the $\bar{Q}$ dual superconformal symmetry generators [59], that the extended superconformal symmetry of SYM theory implies that the differential of any MHV amplitude can be written as a linear combination of $d \log \langle i\, j{-}1\, j\, j{+}1 \rangle$. Evidently, from eqs. (2.1) and (2.2), this implies that only the Plücker coordinates $\langle i\, j{-}1\, j\, j{+}1 \rangle$ may appear in the last entry of the symbol of any MHV amplitude. This is called the *last-entry condition*. For the case $n = 7$, we see from eq. (2.12) that in our basis, only the 14 letters a_{2j} and a_{3j} may appear in the last entry of the symbol of the seven-particle MHV amplitude.

3.2 The collinear limit

MHV amplitudes have particularly simple behavior under collinear limits. It is baked into the definition of the BDS-subtracted n-particle L-loop MHV remainder function [83, 84] that it should smoothly approach the corresponding $n{-}1$-particle function in any simple collinear limit:

$$\lim_{i+1 \| i} R_n^{(L)} = R_{n-1}^{(L)}\,. \tag{3.1}$$

Although we do not do so in the present paper, it would be interesting to also consider the constraints imposed by multi-collinear limits, under which MHV remainder functions have a more intricate behavior (see for example [8, 85]).

We can parameterise the $7 \parallel 6$ collinear limit as

$$Z_7 \to Z_6 + \epsilon \frac{\langle 1246 \rangle}{\langle 1245 \rangle} Z_5 + \epsilon\tau \frac{\langle 2456 \rangle}{\langle 1245 \rangle} Z_1 + \eta \frac{\langle 1456 \rangle}{\langle 1245 \rangle} Z_2\,, \tag{3.2}$$

where the limit $\eta \to 0$ is taken first, followed by $\epsilon \to 0$, leaving the parameter τ fixed. The ratios of four-brackets in eq. (3.2) could be absorbed into ϵ, η and τ, but these factors are useful for keeping track of twistor weight.

Under the replacement (3.2), the 42-letter heptagon symbol alphabet collapses into the 9-letter hexagon symbol alphabet shown in eq. (2.9) plus nine additional letters: the vanishing letters ϵ and η, as well as the seven finite letters

$$
\begin{aligned}
&\tau \,, \\
&1 + \tau \,, \\
&\langle 1235 \rangle \langle 1246 \rangle + \tau \langle 1236 \rangle \langle 1245 \rangle \,, \\
&\langle 1245 \rangle \langle 3456 \rangle + \tau \langle 1345 \rangle \langle 2456 \rangle \,, \\
&\langle 1246 \rangle \langle 2356 \rangle + \tau \langle 1236 \rangle \langle 2456 \rangle \,, \\
&\langle 1246 \rangle \langle 3456 \rangle + \tau \langle 1346 \rangle \langle 2456 \rangle \,, \\
&\langle 1235 \rangle \langle 1246 \rangle \langle 3456 \rangle + \tau \langle 1236 \rangle \langle 1345 \rangle \langle 2456 \rangle \,.
\end{aligned}
\tag{3.3}
$$

A function has a well-defined $7 \parallel 6$ collinear limit only if its symbol is independent of all nine of these letters. We can parameterise other $i+1 \parallel i$ simple collinear limits by appropriately relabeling eq. (3.2) cyclically.

3.3 Discrete symmetries

MHV amplitudes must satisfy several discrete symmetries. They are invariant under the n-particle dihedral group generated by cyclic transformations $Z_i \to Z_{i+1}$ as well as the flip (orientation reversal operation) $Z_i \to Z_{n+1-i}$. These discrete symmetries act simply on the a_{ij}, taking each heptagon letter to some other, as may be read off from eq. (2.12).

A less trivial symmetry of MHV amplitudes is spacetime parity, which in momentum twistor space is generated by the involution

$$
Z_i \to W_i \equiv \langle * \, i{-}1 \, i \, i{+}1 \rangle \,.
\tag{3.4}
$$

This notation is meant to indicate that W_i is a vector orthogonal to the hyperplane spanned by Z_{i-1}, Z_i and Z_{i+1}. Under parity the letters a_{1i} and a_{6i} are invariant, while the others obey:

$$
a_{21} \longleftrightarrow a_{37} \,, \qquad a_{41} \longleftrightarrow a_{51} \,,
\tag{3.5}
$$

and cyclically related transformations.

4 Methods for constructing integrable words

The problem of enumerating all integrable words of length k written in a given alphabet is computationally challenging in general. An exception is when the symbol alphabet consists of cluster coordinates on $\mathrm{Gr}(2,n)$, corresponding to iterated integrals [86] on a Riemann sphere with n marked points, in which case the functions may be explicitly enumerated [71].

When the symbol alphabet is finite, as is the case for the 42-letter heptagon alphabet, at least it is a finite problem. Beginning with the vector space spanned by all 42^k (or fewer, if other conditions have been imposed) length-k words, one needs simply to determine how many linear combinations satisfy the integrability constraints (2.14). Since these are linear constraints, the problem of enumerating all integrable words is ultimately one of linear

algebra: it is the problem of finding a basis for the kernel of the matrix of the integrability constraints.

The calculation may be organised in a couple of different ways, which have various advantages and disadvantages as we now discuss.

4.1 A stepwise approach

For low weights we can use a standard recursive method of iteratively constructing integrable words. First we make an ansatz for words of length k by adjoining one extra letter in all possible ways to integrable words of length $k-1$ and then we directly impose integrability on the last two slots. For the final step of imposing integrability it is convenient to calculate once, and store the value of, all possible combinations $\omega_{\alpha\beta} = d \log \phi_\alpha \wedge d \log \phi_\beta$ as explicit two-forms expressed in some choice of variables.

In general the two-forms $\omega_{\alpha\beta}$ will be non-trivial functions of the ϕ's. The condition that eq. (2.14) should vanish identically may be translated into a collection of linear equations by evaluating the equation at sufficiently many randomly selected points. The nullspace of this linear system is the vector space of integrable words. There is never any concern that an accidentally poor choice of random points may lead to an erroneously large nullspace (i.e., to mistakenly conclude that there are more integrable words than actually exist) because while solving eq. (2.14) is difficult, it is completely straightforward to check whether or not any putative solution is valid.

4.2 A bootstrap

For higher weights we have found an alternative recursive method preferable. Let $\mathcal{A}$ denote the symbol alphabet, let $\mathcal{W}_k$ be the vector space of integrable words of length k written in $\mathcal{A}$, and let $\{w_i^{(k)}\}$ be a basis for this space, where $i = 1, \ldots, d_k = \dim(\mathcal{W}_k)$. Suppose that we have determined such a basis for all weights up to some value k. Then we can expand each basis element $w_i^{(k)}$ as a linear combination of words of the form $\mathcal{W}_{k-1} \otimes \mathcal{A}$ in order to make the last entry in each term explicit, i.e.

$$w_l^{(k)} = \sum_{i=1}^{d_{k-1}} \sum_\alpha A_{li\alpha}^{(k)} \, (w_i^{(k-1)} \otimes \phi_\alpha) \tag{4.1}$$

for some rational coefficients A. Similarly, we can make the first entry in each term explicit by expanding in $\mathcal{A} \otimes \mathcal{W}_{k-1}$,

$$w_m^{(k)} = \sum_{j=1}^{d_{k-1}} \sum_\alpha B_{mj\alpha}^{(k)} \, (\phi_\alpha \otimes w_j^{(k-1)}) \,. \tag{4.2}$$

The A and B coefficients may be easily computed once bases for $\mathcal{W}_k$ and $\mathcal{W}_{k-1}$ are known.

Now let $1 < k_1, \, k_2 \leq k$. We may then write an ansatz for words of length $k_1 + k_2 - 1$ as a linear combination of the form

$$\sum_{l=1}^{d_{k_1}} \sum_{j=1}^{d_{k_2-1}} C_{lj}^{(1)} \, (w_l^{(k_1)} \otimes w_j^{(k_2-1)}) \tag{4.3}$$

for some rational coefficients $C^{(1)}$. This ansatz is manifestly integrable in the first k_1 entries, as well as in the last $k_2 - 1$ entries, so the coefficients $C^{(1)}$ are to be determined by imposing integrability only between entries k_1 and $k_1 + 1$. On the other hand we may write an alternative ansatz of the form

$$\sum_{i=1}^{d_{k_1-1}} \sum_{m=1}^{d_{k_2}} C_{im}^{(2)} (w_i^{(k_1-1)} \otimes w_m^{(k_2)}) \tag{4.4}$$

where integrability is manifest everywhere except between entries $k_1 - 1$ and k_1.

Now any integrable word of length $k_1 + k_2 - 1$ must of course admit an expansion of both types simultaneously, so we can impose full integrability by equating the two forms of the ansatz. Using the basis decompositions (4.1), (4.2) to expose the intermediate letter in slot k_1 lets us express the compatibility conditions as

$$\sum_{i,j,l,\alpha} A_{li\alpha}^{(k_1)} C_{lj}^{(1)} (w_i^{(k_1-1)} \otimes \phi_\alpha \otimes w_j^{(k_2-1)}) = \sum_{i,j,m,\alpha} B_{mj\alpha}^{(k_2)} C_{im}^{(2)} (w_i^{(k_1-1)} \otimes \phi_\alpha \otimes w_j^{(k_2-1)}). \tag{4.5}$$

We may express these equations more simply in matrix form: we have a $d_{k_1} \times d_{k_2-1}$ matrix $C^{(1)}$, a $d_{k_1-1} \times d_{k_2}$ matrix $C^{(2)}$, and, for each value α (i.e., for each letter in the alphabet) a $d_{k_1} \times d_{k_1-1}$ matrix $A_\alpha^{(k_1)}$ and a $d_{k_2} \times d_{k_2-1}$ matrix $B_\alpha^{(k_2)}$, subject to the $d_{k_1-1} \times d_{k_2-1}$ matrix relations

$$(A_\alpha^{(k_1)})^{\mathrm{T}} C^{(1)} = C^{(2)} B_\alpha^{(k_2)} \qquad \forall \alpha. \tag{4.6}$$

Given the A's and B's constructed as described above, any solution $(C^{(1)}, C^{(2)})$ to this linear system determines an integrable word of length $k_1 + k_2 - 1$. Note that all of these matrices should have rational entries.

Although we have phrased it here in a general manner, this construction lets us impose the first- and/or last-entry conditions in a very straightforward way. For example, to impose the first-entry condition we simply restrict all of the above formulas from the d_{k_1-1}, d_{k_1} dimensional spaces of all integrable words to the $\tilde{d}_{k_1-1}, \tilde{d}_{k-1}$ dimensional subspaces satisfying the first-entry condition. The same relation (4.6) holds, but with significantly smaller $A_\alpha^{(k_1)}$, $C^{(1)}$ and $C^{(2)}$ matrices. Imposing the last-entry restriction reduces the size of $C^{(1)}$ and $C^{(2)}$ further, and also the size of $B_\alpha^{(k_2)}$.

4.3 Comparison of the two methods

The only notable disadvantage of the new recursive approach is that the equations (4.6) to solve involve considerably more free variables. To illustrate this point, let us describe the construction of integrable words of length 6 satisfying both the first and last-entry conditions. The traditional approach of subsection 4.1 involves an ansatz with $6763 \times 14 = 94682$ free parameters, corresponding (see table 1) to the number of weight-5 heptagon functions, tensored with the 14 allowed last entries. The number of equations depends on the number of random kinematical points at which we evaluate eq. (2.14), but should be at least comparable to the number of free variables.

In the new approach of subsection 4.2 we use eq. (4.6) with $k_1 = 4$ and $k_2 = 3$. From table 1 we see that there are $237\,(1288)$ heptagon functions at weight $3\,(4)$, i.e. integrable words of length $3\,(4)$ satisfying the first-entry condition. Meanwhile one can check that there are $146\,(1364)$ integrable words of length $2\,(3)$ satisfying the last-entry condition. Therefore, applying eq. (4.6) to find the space of heptagon functions satisfying the last-entry condition requires solving $237 \times 146 \times 42$ equations for the 1288×146 matrix $C^{(1)}$ and the 237×1364 matrix $C^{(2)}$, i.e. a total of about one and a half million equations for over half a million free variables! (As discussed in the next section, the solution space of this system turns out, amazingly, to have dimension four.)

However, we have found this disadvantage to be more than compensated by two significant advantages. The first is that if the bases for the $\mathcal{W}_k$ are chosen with reasonable care, the matrices A_α and B_α can be made quite sparse. By solving the simplest conditions, namely those that force a single free parameter to vanish, or express one parameter in terms of exactly one another, we may quickly reduce the size of the linear system. This contrasts to the traditional iterative construction described at the beginning of this section, since the $d \log \phi_\alpha \wedge d \log \phi_\beta$ two-forms generally have no, or only a few, vanishing elements. For the length 6 linear system we discussed in the previous paragraphs, solving the equations of length 1 or 2 and then partially solving some more of the shorter constraints leads to 63557 equations for the remaining 15979 free variables, a significant reduction of the size of the linear system compared to the traditional approach.

The second great advantage of the method of subsection 4.2 is also illustrated in the length 6 example we have used: if bases are known for all weights less than or equal to some value k, we may immediately bootstrap directly to a basis of integrable words at weight $2k - 1$, without having to recursively construct bases at weights $k + 1, k + 2, \ldots$ one step at a time, as with the traditional approach. (Because of this weight-skipping power of the bootstrap we had in fact found the symbol of the 3-loop MHV heptagon long before determining the total number of weight-5 heptagon functions.)

4.4 Solving the integrability constraints

Even with the improved method of subsection 4.2, starting at weight 5 the size of the linear system encoding the integrability constraints grows to such extent that its solution becomes the most important computational challenge of the bootstrap programme. Let us now discuss the strategy we adopted for addressing this challenge, which required an efficiency beyond the capabilities of standard scientific software such as `Mathematica` or `Matlab`.

After (or even before) partially reducing the integrability constraints in the form of subsection 4.2 according to the discussion of subsection 4.3, we may bring them to a more standard form by grouping all elements of the matrices of unknown coefficients $C^{(1)}, C^{(2)}$ into a column vector X, such that eq. (4.6) becomes

$$A \cdot X = 0 \,. \tag{4.7}$$

By virtue of eq. (4.3) or (4.4), the set of all integrable words of a given weight will thus be given by all linearly independent solutions of eq. (4.7), or in other words by the *right nullspace* of the matrix A.

A systematic procedure for computing the nullspace is Gaussian elimination, whereby A is brought into a column echelon form H by a transformation U,

$$A \cdot U = H \,, \tag{4.8}$$

$$U = (\underbrace{U_1}_{r} |N) \,, \qquad H = (\underbrace{H_1}_{r} |\mathbf{0}) \,. \tag{4.9}$$

In the last line we have written out the two matrices in block matrix form, where r denotes the rank of A, and the first nonzero element at each column of the invertible matrix H_1 is strictly below the corresponding element of the column at its left. Clearly, the submatrix N will form a basis for the right nullspace of A.

Even though standard Gaussian elimination can be completed in a number of arithmetic operations that depends polynomially on the size of the system, a major complication that arises when applying it to matrices with exactly represented rational entries like A is *intermediate expression swell*: generically, the size of the entries (in bits) doubles at each step, so that each operation takes longer and longer time, eventually leading to runtimes (and intermediate storage required) depending exponentially on the size of the system. A review of these well-established computer algebra results may be found in [87].

The key idea for avoiding this complication is to transform our matrix from rational to integer, for which there exist fraction-free variants of Gaussian elimination that bound the size of intermediate expressions by virtue of Hadamard's inequality, see [88] and references therein. Fortunately, there already exists an efficient C library implementation of such a variant, the `Integer Matrix Library` (IML) [89]. This implementation also builds on the use of modular arithmetic to further improve the size of intermediate computations. Finally, it reduces row and column operations to matrix multiplications, which can be done very fast with the help of other well-known algorithms, for example [90].

First starting with the transformation of A to an integer matrix A', we have found that a minimal increase in the size of its entries can be achieved by dividing each column of A with the greatest common divisor (GCD) of all its nonzero elements (as opposed to doing this for the rows or even worse for the entire matrix). In fact, in this way we may also track down free variables which don't appear at all in the equations, as their columns will have zero GCD. These will correspond to the simplest nullspace vectors, which we can immediately construct and remove from the linear system, in order to reduce its size. If D is the diagonal matrix whose diagonal elements are the inverses of the aforementioned GCDs, then A, A' will be related by $A' = A \cdot D$, and we may obtain the nullspace of the former from the one of the latter,

$$A' \cdot N' = 0 = A \cdot N \quad \Rightarrow \quad N = D \cdot N'. \tag{4.10}$$

(Alternatively we may absorb the transformation into a rescaling of the unknown coefficients, $X' = D \cdot X$.)

Once we have produced A' in this manner, we feed it as input into a custom C programme using the function `nullspaceLong` of the aforementioned IML library, which is optimised for matrices whose elements have absolute values smaller than 2^{31} (like the ones

we have encountered), and computes a nullspace N' with integer entries. Particularly for the case of the weight 6 hexagon functions obeying last-entry conditions, we also found advantageous to apply this procedure not to A' directly, but to its Gram matrix $A'^T \cdot A'$, exploiting the fact that the two matrices have the same nullspace,

$$A^T \cdot A \cdot X = 0 \Rightarrow X^T \cdot A^T \cdot A \cdot X = |A \cdot X|^2 = 0 \Rightarrow A \cdot X = 0. \qquad (4.11)$$

In this manner we traded a 63557×15979 matrix with a much smaller square 15979 matrix (albeit with larger entries), and in fact with slightly fewer nonzero entries (corresponding to a fill-in of 1.6% and 6.2% approximately).

Finally it is worth mentioning that after obtaining N, whose columns span a basis of solutions for our linear system (4.7), we can further simplify this basis with the help of the Lenstra-Lenstra-Lovász (LLL) algorithm, see [91] for a more recent, improved version. The latter, which has also found applications in the computation of anomalous dimensions in SYM and QCD (see [92] and references therein), is an algorithm for finding a short, nearly orthogonal basis for a d-dimensional integer lattice embedded in m-dimensional space, $d \leq m$. The integer matrix N has precisely the form of such a lattice, where d is the dimension of the nullspace, and m the number of components of its vectors. In addition, it is evident from eq. (4.7) that any rescaling of the columns of N will also be a nullspace basis. We can thus simplify our basis further by repeating a cycle of division of its vectors by the GCD of their nonzero elements, followed by an LLL reduction, until a new cycle leaves the basis unchanged. The final set of solutions to the integrability constraints has up to 3 times fewer nonzero coefficients than the initial set, leading to considerably shorter expressions for the corresponding integrable words (4.3) or (4.4).

5 Heptagon symbols and their properties

Table 1 summarises the results of our partial analysis of the space of heptagon symbols through weight 6 (the question marks in the table indicate numbers that we have not yet explicitly determined). We remind the reader that in this paper we are working only at the level of symbols and that the counting of dimensions of spaces of functions obeying various constraints should be taken with this in mind. Following the conventions of subsection 2.5, in this section we continue to highlight this point by referring to *heptagon symbols* (or *hexagon symbols*) instead of the more cumbersome "symbols of heptagon functions".

We now discuss the results of table 1 in detail, beginning with the first three lines which contain, perhaps, no great qualitative surprises.

5.1 Collinear limits of heptagon symbols

The first line indicates the total number of heptagon symbols of a given weight, which we have already mentioned in section 2.5. The second line indicates the number of linear combinations of these which are finite in the collinear limit and independent of the "bad" letters shown in eq. (3.3). Many linear combinations are not only well-defined, but actually vanish in the $7 \parallel 6$ collinear limit; the number of these is indicated on the third line.

Weight $k =$	1	2	3	4	5	6
Number of heptagon symbols	7	42	237	1288	6763	?
well-defined in the 7 ∥ 6 limit	3	15	98	646	?	?
which vanish in the 7 ∥ 6 limit	0	6	72	572	?	?
well-defined for all $i+1 \parallel i$	0	0	0	1	?	?
with MHV last entries	0	1	0	2	1	4
with both of the previous two	0	0	0	1	0	1

Table 1. Heptagon symbols and their properties.

Weight $k =$	1	2	3	4	5	6
Number of hexagon symbols	3	9	26	75	218	643
well-defined (hence vanish) in the 6 ∥ 5 limit	0	2	11	44	155	516
well-defined (hence vanish) for all $i+1 \parallel i$	0	0	2	12	68	307
with MHV last entries	0	3	7	21	62	188
with both of the previous two	0	0	1	4	14	59

Table 2. Hexagon symbols and their properties.

For comparison with the hexagon bootstrap programme we include the analogous results for $n = 6$ in table 2. Here there is no distinction between the cases considered separately on lines 2 and 3 of table 1: if the 6 ∥ 5 collinear limit of a hexagon symbol is well-defined, then it necessarily vanishes in the limit, as there are no symbols for $n = 5$.

The collinear limit ties the two tables together in an interesting way, because the 7 ∥ 6 collinear limit of a heptagon symbol must be a hexagon symbol, whenever the limit is well-defined. Of course, by taking collinear limits of all possible heptagon symbol we cannot possibly find more hexagon symbols than actually exist. This criterion partially explains the third line of table 1. For example, at weights 1, 2, 3 we see by subtracting the third line from the second that there are 3, 9, 26 linearly independent hexagon symbols which can be obtained as collinear limits of heptagon symbols. These numbers match the top line of table 2. So for weight ≤ 3 we conclude that the space of all hexagon symbols is spanned by the collection of (well-defined) collinear limits of heptagon symbols.

Curiously this pattern breaks down at weight 4. Table 1 indicates that taking the 7 ∥ 6 collinear limit of heptagon symbols generates $646 - 572 = 74$ linearly independent hexagon symbols, but table 2 indicates that there exist 75 hexagon symbols. Therefore, there is a weight-4 hexagon symbol which is not the collinear limit of any heptagon symbol!

5.2 Symbols of uniqueness: MHV heptagons at 2 and 3 loops

The real surprises in table 1 lie in the last three lines, which stand out when compared to the last three lines of table 2. Although the total number of heptagon symbols at a given weight is much greater (asymptotically exponentially) than the number of hexagon symbols at the same weight, the entries on the last three lines of the heptagon table are small compared to the corresponding entries in the hexagon table. The discovery of this surprising fact is the unexpected "miracle" of our work.

Let us begin with the fourth line of table 1. A heptagon symbol may have a perfectly well-defined collinear limit as 7 ∥ 6 while being divergent, or just ill-defined, in a different collinear limit, 4 ∥ 3, say. (Note that nowhere in the two tables have we imposed cyclic symmetry.) MHV remainder functions must be finite and well-defined in all $i+1 \parallel i$ simple collinear limits. The number of heptagon symbols satisfying this criterion is indicated on the fourth line of table 1. There are no such symbols for weight less than 4, and precisely one such symbol at weight 4. Obviously, that symbol must be the symbol of the two-loop seven-particle MHV remainder function $R_7^{(2)}$! To recap:

> The symbol of the two-loop seven-particle MHV remainder function $R_7^{(2)}$ is the only weight-4 heptagon symbol which is well-defined in all $i+1 \parallel i$ collinear limits.

Let us emphasise that it is not necessary to assume dihedral symmetry, parity symmetry, or the last-entry condition. Nor is it necessary to use the expected collinear limit $R_6^{(2)}$ as an input to fix some remaining ambiguity (except for the overall multiplicative normalisation). All of these properties are automatically satisfied by the unique function described in the above box. Of course, we have checked that the symbol of the function obtained in this manner via the bootstrap programme indeed is proportional to the known symbol of $R_7^{(2)}$ found in [65].

It would be extremely interesting to see if this criterion continues to hold at weight 6, i.e. to see whether the question mark in the last column of the fourth line of table 1 is also 1, but we have not yet completed this calculation. Nevertheless we note that this criterion certainly could not work at arbitrary loop order; for example at weight 8 the square of $R_7^{(2)}$ and the four-loop seven-particle MHV remainder function $R_7^{(4)}$ are both well-defined in all simple collinear limits, and are distinct.

Let us now turn to the last two lines of table 1, where we impose the last-entry condition appropriate for MHV amplitudes, as discussed in section 3.1. In contrast to the general heptagon problem, where the complexity of the linear systems involved has forced us to leave some questions marks in the table, when we impose the last-entry condition the size of the linear systems becomes "small" enough that we have succeeded in a full classification through weight 6. This is certainly not to say that the calculation was easy — as explained in section 4, determining the number "4" in the last column of table 1 required finding the nullspace of a linear system with over half a million variables.

The number of heptagon symbols satisfying the last-entry condition is shown in the sixth line of table 1. Let us note right away that none of the numbers are multiples of 7, hence all of these functions are necessarily cyclically invariant, even though this was not an input to the calculation. Also it turns out (this is trivial at weights 2 and 5, where there is a single symbol, and is easily checked at weights 4 and 6) that they are all invariant under the full dihedral group, as well as under the parity operation shown in eq. (3.5). Again none of these discrete symmetries were imposed going into the calculation.

At weight 2 we find there is a unique heptagon symbol satisfying the last-entry condition. The corresponding heptagon function is written explicitly, and discussed in more

detail, in the following section. At weight 4 there are two functions: the square of the weight-2 function, and the two-loop MHV remainder function $R_7^{(2)}$. As we have already seen, the latter is the only one which is well-defined in all collinear limits. At weight 5 there is again a unique symbol satisfying the last-entry condition. This weight-5 symbol, like the weight-2 symbol, is not well-defined in the collinear limit, so these two symbols have no role to play in connection with MHV scattering amplitudes.

Let us now focus again on the surprising "4" in the last column of table 1. From our discussion so far we already know that there must be at least two weight-6 heptagon functions satisfying the last-entry condition: the cube of the weight-2 function discussed above, and the product of the weight-2 function with $R_7^{(2)}$. The surprise is that, in addition to the symbols of these two functions, we find only two irreducible symbols at weight 6. We find that there is a unique linear combination of these four symbols which is finite in the collinear limit (in this case it happens that it is sufficient to consider only the 7 ∥ 6 collinear limit since, as mentioned above, the symbols turn out to all be cyclically invariant anyway). We have checked that the collinear limit matches perfectly (up to an overall factor, which is not fixed by the bootstrap) the known symbol of the three-loop MHV hexagon [16, 59]. Therefore:

> The symbol of the three-loop seven-particle MHV remainder function $R_7^{(3)}$ is the only weight-6 heptagon symbol which satisfies the last-entry condition and which is finite in the 7 ∥ 6 collinear limit.

The only ambiguity which will be left when passing from the symbol of $R_7^{(3)}$ to an actual function is the addition of a rational linear combination of $\zeta_2 R_7^{(2)}$, ζ_3^2 and ζ_6. The collinear limit will fix all three coefficients (as well as the overall normalisation of $R_7^{(3)}$) uniquely.

Again we emphasise that the above conclusion does not rely on assuming that any of the discrete symmetries are satisfied; they all emerge as "accidental" (if there is such a thing in SYM theory) properties of the unique solution. Moreover, and even more surprisingly, the unique solution emerges without any free parameters which need to be tuned in order to match the correct value of the three-loop MHV hexagon in the collinear limit, let alone to match various terms in the Regge limit and/or OPE expansion around the collinear limit.

6 Speculations: the n-gon bootstrap at weight 2

Our results were completely surprising. Based on the hexagon bootstrap programme, we expected that even after imposing all discrete symmetries, there would likely be hundreds of free parameters in our heptagon ansatz which would need to be fit by comparison to various data in the literature.

The fact that none of this turned out to be necessary, and that the heptagon bootstrap turned out, in this sense, to be more powerful than the hexagon bootstrap, requires explanation. It is, after all, a basic tenet of amplitudeology that "accidents do not happen," especially in SYM theory.

Unfortunately we have only very little to offer at this time. In this section we make a few meager observations at weight 2, where it is simple to tabulate and explicitly analyze the relevant function spaces. Our observations here admittedly shed only a little light on the situation at higher weight, but perhaps they serve as a useful starting point.

Let us define the cross-ratio

$$u_1 = \frac{\langle 1256\rangle\langle 2345\rangle}{\langle 1245\rangle\langle 2356\rangle} = \frac{a_{17}}{a_{13}a_{14}} \tag{6.1}$$

with six other u_i defined cyclically (sometimes u_1 is called u_{14} in the literature). According to the sixth line of table 1, there is a unique weight-2 heptagon function satisfying the last-entry condition. This function is

$$\sum_{i=1}^{7} \text{Li}_2(1 - 1/u_i) - \log u_i \log \frac{u_{i+3}}{u_{i-3}}. \tag{6.2}$$

Let us contrast this to the situation at $n = 6$, where there are three functions

$$\text{Li}_2(1 - 1/u), \qquad \text{Li}_2(1 - 1/v), \qquad \text{Li}_2(1 - 1/w) \tag{6.3}$$

which separately satisfy the last-entry condition.

Why does this happen? The functions shown in eq. (6.3) exist because of the identity

$$\frac{1 - u}{u} = \frac{\langle 1356\rangle\langle 2346\rangle}{\langle 1236\rangle\langle 3456\rangle} \tag{6.4}$$

(and two cyclic images). Note that all of the brackets appearing on the right-hand side are of the form $\langle i\, j{-}1\, j\, j{+}1\rangle$. Hence all three of $(1 - u)/u, (1 - v)/v, (1 - w)/w$ are valid MHV last entries.

Is there an analogue to the identity (6.4) for $n = 7$? That is, does there exist an identity which allows products of u_i's and $1 - u_i$'s to be rewritten only in terms of $\langle i\, j{-}1\, j\, j{+}1\rangle$'s? It is simple to check that there are precisely seven such identities:

$$\frac{1 - u_1}{u_1}\frac{1 - u_7}{u_7}\frac{1}{1 - u_4} = \frac{\langle 1235\rangle\langle 1247\rangle\langle 1345\rangle\langle 2456\rangle}{\langle 1234\rangle\langle 1257\rangle\langle 1456\rangle\langle 2345\rangle} \tag{6.5}$$

and its cyclic permutations. If it were not for the factor $1/(1 - u_4)$ on the left, then the functions $\text{Li}_2(1 - 1/u_i) + \text{Li}_2(1 - 1/u_{i+1})$ would satisfy both the first- and last-entry conditions for all i. Instead, only the particular linear combination shown in eq. (6.2) is allowed.

It is straightforward to extend this analysis to higher n. The total number of weight-2 n-gon functions grows very rapidly (in fact, as $\mathcal{O}(n^4)$) with n. How many of those functions satisfy the last-entry condition? Obviously, it is to be expected that there should be very strong interplay between the first- and last-entry conditions at weight 2, but we find an unexpectedly strong result: we find that there are precisely 3 n-gon functions at weight 2 satisfying the last-entry condition when $n = 6$ or when n is a multiple of four. For all other n, we find that there is only one such function!

As we go to higher weight we might expect the interplay between the first- and last-entry conditions to become less constraining. This expectation may or may not turn out to be true asymptotically at large weight, but tables 1 and 2 indicate little weakening of this interplay for k even as high as 6. Clearly it would be interesting to map out the space of n-gon functions at higher weight.

7 Discussion

We have found that the heptagon bootstrap for computing (symbols of) seven-point MHV amplitudes in SYM is unreasonably effective in comparison with the hexagon bootstrap, at least through three loops. In particular, the three-loop heptagon remainder function is the unique weight-6 heptagon function which satisfies the last-entry condition and which is finite in the $7 \parallel 6$ collinear limit. Evidently the conceptually simplest way of computing the three-loop hexagon remainder function is, somewhat perversely, to first compute the heptagon remainder and take its collinear limit.

Naturally, it would be very interesting to further explore the power of the heptagon bootstrap at higher loops or by relaxing the last-entry condition to those appropriate for NMHV amplitudes. It would also be interesting to explore the n-gon bootstrap for higher n. Our analysis at weight-2 in section 6 suggests that the first- and last-entry conditions are much tighter in combination than each is individually. It would be important to understand whether this is an accident at weight-2 (and whether the success of our heptagon bootstrap was similarly accidental), or whether there is some fundamental feature of the structure of n-gon functions which currently evades our understanding.

The cases $n = 6, 7$ are special because we believe that we know the appropriate symbol alphabets for amplitudes (both MHV and non-MHV) to all loop order, based on the fact that the associated cluster algebras have finitely many $\mathcal{A}$-coordinates. However starting at $n = 8$ their number is infinite, so there is the possibility that new, more exotic symbol letters could start appearing at each loop order (or even when we go from MHV to non-MHV at a given loop order). Anything we could learn about the pattern of symbol letters which appear at higher n and at higher weight would be very valuable.

In section 5 we found an indication that thinking about the collinear limits of n-gon functions may lead to a class of previously underappreciated constraints. Specifically we found that there exists a hexagon function at weight 4 which is not the collinear limit of any heptagon function. Similarly, it is natural to expect that there may be heptagon functions which are not the collinear limit of any octagon functions, that there are hexagon functions which are not the double-collinear limit of any octagon function, etc. In this way we see that the consistency of collinear limits places an entire infinite tower of potentially very powerful constraints on the bootstrap. Along these lines, it has recently been shown [82] that the collinear limit, together with dihedral symmetry and the first- and last-entry conditions, uniquely fixes the two-loop n-point MHV amplitude modulo classical polylogarithm functions for all n. The results of this paper suggest that even full symbols, if not full functions (which we have not addressed), may be surprisingly accessible via the n-gon bootstrap.

Finally, it is of course important to construct a complete functional representation for $R_7^{(3)}$. This would require first constructing a heptagon function of weight 6 with the correct symbol, obeying the differential $\bar{Q}$ constraint and finite in the collinear limit. After this there will be beyond-the-symbol ambiguities corresponding to the addition of a numerical constant (in particular a linear combination of ζ_3^2 and ζ_6) as well as a term proportional to $\zeta_2 R_7^{(2)}$. Both of these ambiguities will be uniquely fixed by the simple collinear limit. Having an explicit functional form for $R_7^{(3)}$ would not only allow for detailed checks against the available predictions for its behaviour in the collinear [31, 33] and multi-Regge limits [51–53], but would also shed light on yet unknown key quantities in these approaches. These include multi-particle scalar and fermion pentagon transitions, or higher BFKL eigenvalues, impact factors and central emission vertices. For example it would be great if our result could guide the generalisation of the all-loop formulas for the hexagon in the multi-Regge limit [41], to the heptagon. The continuation of this programme for $n = 8$ will have an even more interesting interplay with the BFKL approach, where a new bound state of three reggeised gluons first appears, and could moreover push forward our knowledge of the strong coupling behaviour of the amplitudes [93, 94].

Acknowledgments

References

[1] L. Brink, J.H. Schwarz and J. Scherk, *Supersymmetric Yang-Mills theories*, *Nucl. Phys.* **B 121** (1977) 77 [INSPIRE].

[2] R.J. Eden, P.V. Landshoff, D.I. Olive and J.C. Polkinghorne, *The analytic S-matrix*, Cambridge University Press, Cambridge U.K. (1966).

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

[4] J.M. Drummond, G.P. Korchemsky and E. Sokatchev, *Conformal properties of four-gluon planar amplitudes and Wilson loops*, *Nucl. Phys.* **B 795** (2008) 385 [arXiv:0707.0243] [INSPIRE].

[5] A. Brandhuber, P. Heslop and G. Travaglini, *MHV amplitudes in* $\mathcal{N} = 4$ *super Yang-Mills and Wilson loops*, *Nucl. Phys.* **B 794** (2008) 231 [arXiv:0707.1153] [INSPIRE].

[6] J.M. Drummond, J. Henn, G.P. Korchemsky and E. Sokatchev, *On planar gluon amplitudes/Wilson loops duality*, *Nucl. Phys.* **B 795** (2008) 52 [arXiv:0709.2368] [INSPIRE].

[7] 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].

[8] Z. Bern et al., *The two-loop six-gluon MHV amplitude in maximally supersymmetric Yang-Mills theory*, *Phys. Rev.* **D 78** (2008) 045007 [arXiv:0803.1465] [INSPIRE].

[9] J.M. Drummond, J. Henn, G.P. Korchemsky and E. Sokatchev, *Hexagon Wilson loop = six-gluon MHV amplitude*, *Nucl. Phys.* **B 815** (2009) 142 [arXiv:0803.1466] [INSPIRE].

[10] 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].

[11] 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].

[12] 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].

[13] L.F. Alday and J. Maldacena, *Comments on gluon scattering amplitudes via AdS/CFT*, *JHEP* **11** (2007) 068 [arXiv:0710.1060] [INSPIRE].

[14] 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].

[15] A.B. Goncharov, M. Spradlin, C. Vergu and A. Volovich, *Classical polylogarithms for amplitudes and Wilson loops*, *Phys. Rev. Lett.* **105** (2010) 151605 [arXiv:1006.5703] [INSPIRE].

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

[17] 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].

[18] 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].

[19] 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].

[20] 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].

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

[22] C. Anastasiou, Z. Bern, L.J. Dixon and D.A. Kosower, *Planar amplitudes in maximally supersymmetric Yang-Mills theory*, *Phys. Rev. Lett.* **91** (2003) 251602 [hep-th/0309040] [INSPIRE].

[23] Z. Bern, L.J. Dixon and V.A. Smirnov, *Iteration of planar amplitudes in maximally supersymmetric Yang-Mills theory at three loops and beyond*, *Phys. Rev.* **D 72** (2005) 085001 [hep-th/0505205] [INSPIRE].

[24] V.P. Nair, *A current algebra for some gauge theory amplitudes*, *Phys. Lett.* **B 214** (1988) 215 [INSPIRE].

[25] S.J. Parke and T.R. Taylor, *An amplitude for n gluon scattering*, *Phys. Rev. Lett.* **56** (1986) 2459 [INSPIRE].

[26] L.F. Alday, D. Gaiotto, J. Maldacena, A. Sever and P. Vieira, *An operator product expansion for polygonal null Wilson loops*, *JHEP* **04** (2011) 088 [arXiv:1006.2788] [INSPIRE].

[27] D. Gaiotto, J. Maldacena, A. Sever and P. Vieira, *Bootstrapping null polygon Wilson loops*, *JHEP* **03** (2011) 092 [arXiv:1010.5009] [INSPIRE].

[28] D. Gaiotto, J. Maldacena, A. Sever and P. Vieira, *Pulling the straps of polygons*, *JHEP* **12** (2011) 011 [arXiv:1102.0062] [INSPIRE].

[29] A. Sever and P. Vieira, *Multichannel conformal blocks for polygon Wilson loops*, *JHEP* **01** (2012) 070 [arXiv:1105.5748] [INSPIRE].

[30] 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].

[31] 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].

[32] 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].

[33] 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].

[34] A.V. Belitsky, *Nonsinglet pentagons and NHMV amplitudes*, arXiv:1407.2853 [INSPIRE].

[35] A.V. Belitsky, *Fermionic pentagons and NMHV hexagon*, arXiv:1410.2534 [INSPIRE].

[36] B. Basso, J. Caetano, L. Cordova, A. Sever and P. Vieira, *OPE for all helicity amplitudes*, arXiv:1412.1132 [INSPIRE].

[37] G. Papathanasiou, *Hexagon Wilson loop OPE and harmonic polylogarithms*, *JHEP* **11** (2013) 150 [arXiv:1310.5735] [INSPIRE].

[38] G. Papathanasiou, *Evaluating the six-point remainder function near the collinear limit*, *Int. J. Mod. Phys.* **A 29** (2014) 1450154 [arXiv:1406.1123] [INSPIRE].

[39] J.M. Drummond and G. Papathanasiou, *Hexagon from collinear to multi-Regge kinematics*, to appear.

[40] N. Beisert et al., *Review of AdS/CFT integrability: an overview*, *Lett. Math. Phys.* **99** (2012) 3 [arXiv:1012.3982] [INSPIRE].

[41] B. Basso, S. Caron-Huot and A. Sever, *Adjoint BFKL at finite coupling: a short-cut from the collinear limit*, *JHEP* **01** (2015) 027 [arXiv:1407.3766] [INSPIRE].

[42] J. Bartels, L.N. Lipatov and A. Sabio Vera, *BFKL Pomeron, Reggeized gluons and Bern-Dixon-Smirnov amplitudes*, *Phys. Rev.* **D 80** (2009) 045002 [arXiv:0802.2065] [INSPIRE].

[43] J. Bartels, L.N. Lipatov and A. Sabio Vera, $\mathcal{N} = 4$ *supersymmetric Yang-Mills scattering amplitudes at high energies: the Regge cut contribution*, *Eur. Phys. J.* **C 65** (2010) 587 [arXiv:0807.0894] [INSPIRE].

[44] L.N. Lipatov and A. Prygarin, *Mandelstam cuts and light-like Wilson loops in* $\mathcal{N} = 4$ *SUSY*, *Phys. Rev.* **D 83** (2011) 045020 [arXiv:1008.1016] [INSPIRE].

[45] L.N. Lipatov and A. Prygarin, *BFKL approach and six-particle MHV amplitude in* $\mathcal{N} = 4$ *super Yang-Mills*, *Phys. Rev.* **D 83** (2011) 125001 [arXiv:1011.2673] [INSPIRE].

[46] J. Bartels, L.N. Lipatov and A. Prygarin, *MHV amplitude for* $3 \to 3$ *gluon scattering in Regge limit*, *Phys. Lett.* **B 705** (2011) 507 [arXiv:1012.3178] [INSPIRE].

[47] V.S. Fadin and L.N. Lipatov, *BFKL equation for the adjoint representation of the gauge group in the next-to-leading approximation at* $\mathcal{N} = 4$ *SUSY*, *Phys. Lett.* **B 706** (2012) 470 [arXiv:1111.0782] [INSPIRE].

[48] A. Prygarin, M. Spradlin, C. Vergu and A. Volovich, *All two-loop MHV amplitudes in multi-Regge kinematics from applied symbology*, *Phys. Rev.* **D 85** (2012) 085019 [arXiv:1112.6365] [INSPIRE].

[49] L. Lipatov, A. Prygarin and H.J. Schnitzer, *The multi-Regge limit of NMHV amplitudes in* $\mathcal{N} = 4$ *SYM theory*, *JHEP* **01** (2013) 068 [arXiv:1205.0186] [INSPIRE].

[50] L.J. Dixon, C. Duhr and J. Pennington, *Single-valued harmonic polylogarithms and the multi-Regge limit*, *JHEP* **10** (2012) 074 [arXiv:1207.0186] [INSPIRE].

[51] J. Bartels, A. Kormilitzin, L.N. Lipatov and A. Prygarin, *BFKL approach and* $2 \to 5$ *maximally helicity violating amplitude in* $\mathcal{N} = 4$ *super-Yang-Mills theory*, *Phys. Rev.* **D 86** (2012) 065026 [arXiv:1112.6366] [INSPIRE].

[52] J. Bartels, A. Kormilitzin and L. Lipatov, *Analytic structure of the* $n = 7$ *scattering amplitude in* $\mathcal{N} = 4$ *SYM theory in the multi-Regge kinematics: conformal Regge pole contribution*, *Phys. Rev.* **D 89** (2014) 065002 [arXiv:1311.2061] [INSPIRE].

[53] J. Bartels, A. Kormilitzin and L.N. Lipatov, *Analytic structure of the* $n = 7$ *scattering amplitude in* $\mathcal{N} = 4$ *theory in multi-Regge kinematics: conformal Regge cut contribution*, *Phys. Rev.* **D 91** (2015) 045005 [arXiv:1411.2294] [INSPIRE].

[54] 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].

[55] L.J. Mason and D. Skinner, *The complete planar S-matrix of* $\mathcal{N} = 4$ *SYM as a Wilson loop in twistor space*, *JHEP* **12** (2010) 018 [arXiv:1009.2225] [INSPIRE].

[56] S. Caron-Huot, *Notes on the scattering amplitude/Wilson loop duality*, *JHEP* **07** (2011) 058 [arXiv:1010.1167] [INSPIRE].

[57] T. Adamo, M. Bullimore, L. Mason and D. Skinner, *Scattering amplitudes and Wilson loops in twistor space*, *J. Phys.* **A 44** (2011) 454008 [arXiv:1104.2890] [INSPIRE].

[58] M. Bullimore and D. Skinner, *Descent equations for superamplitudes*, arXiv:1112.1056 [INSPIRE].

[59] 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].

[60] J.M. Drummond, J.M. Henn and J. Plefka, *Yangian symmetry of scattering amplitudes in* $\mathcal{N} = 4$ *super Yang-Mills theory*, *JHEP* **05** (2009) 046 [arXiv:0902.2987] [inSPIRE].

[61] J.M. Drummond and L. Ferro, *Yangians, Grassmannians and T-duality*, *JHEP* **07** (2010) 027 [arXiv:1001.3348] [inSPIRE].

[62] 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].

[63] 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].

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

[65] 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].

[66] 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].

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

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

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

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

[71] F.C.S. 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.AG/0606419] [inSPIRE].

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

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

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

[75] 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].

[76] 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].

[77] 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].

[78] V.V. Fock and A.B. Goncharov, *Cluster ensembles, quantization and the dilogarithm*, *Annales Sci. Ecole Norm. Sup.* **42** (2009) 865 [math.AG/0311245] [inSPIRE].

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

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

[81] 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].

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

[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.* **B 425** (1994) 217 [hep-ph/9403226] [INSPIRE].

[84] Z. Bern, J.S. Rozowsky and B. Yan, *Two loop four gluon amplitudes in* $\mathcal{N} = 4$ *super Yang-Mills*, *Phys. Lett.* **B 401** (1997) 273 [hep-ph/9702424] [INSPIRE].

[85] C. Anastasiou et al., *Two-loop polygon Wilson loops in* $\mathcal{N} = 4$ *SYM*, *JHEP* **05** (2009) 115 [arXiv:0902.2245] [INSPIRE].

[86] K.-T. Chen, *Iterated path integrals*, *Bull. Amer. Math. Soc.* **83** (1977) 831.

[87] M. Kauers, *Fast solvers for dense linear systems*, *Nucl. Phys. Proc. Suppl.* **183** (2008) 245 [INSPIRE].

[88] A. Storjohann, *Algorithms for matrix canonical forms*, Ph.D. Thesis, Swiss Federal Institute of Technology, ETH-Zurich (2000).

[89] Z. Chen and A. Storjohann, *A blas based C library for exact linear algebra on integer matrices*, in Proceedings of the *2005 International Symposium on Symbolic and Algebraic Computation (ISSAC '05)*, New York U.S.A. (2005), pp. 92–99.

[90] V. Strassen, *Gaussian elimination is not optimal*, *Numer. Math.* **13** (1969) 354.

[91] A. Storjohann, *Faster algorithms for integer lattice basis reduction*, Swiss Federal Institute of Technology (1996).

[92] V.N. Velizhanin, *Four loop anomalous dimension of the third and fourth moments of the non-singlet twist-2 operator in QCD*, arXiv:1411.1331 [INSPIRE].

[93] J. Bartels, V. Schomerus and M. Sprenger, *Heptagon amplitude in the multi-Regge regime*, *JHEP* **10** (2014) 067 [arXiv:1405.3658] [INSPIRE].

[94] J. Bartels, V. Schomerus and M. Sprenger, *The Bethe roots of Regge cuts in strongly coupled* $\mathcal{N} = 4$ *SYM theory*, arXiv:1411.2594 [INSPIRE].