Cluster Algebras in SYM Theory and QCD

Marcus Spradlin, Brown University

October 24 2025

LSA LEINWEBER INSTITUTE FOR THEORETICAL PHYSICS
UNIVERSITY OF MICHIGAN

Michigan Interdisciplinary Meeting on Amplitudes: Bridges between Physics & Mathematics

Introduction

In recent years we've seen remarkable progress on the problem of unlocking the hidden mathematical structure of scattering amplitudes in quantum field theory, both for its own sake (beauty) and for the desire to develop new methods of practical importance for comparison of theory to experiment (truth).

Particularly rich is the mathematical structure of the amplitudes in (planar) maximally supersymmetric Yang-Mills (SYM) theory, whose amplitudes are so highly constrained that they 'barely exist' given all the constraints they must satisfy.

The corner of this theory that is best understood is closely connected to cluster algebras, but it is clear that this is only the tip of a very big iceberg.

Introduction to Amplitudes

Amplitudes in planar SYM theory are indexed by

number of particles	$n = 4, 5, 6, \ldots$
helicity sector	$k = 0, 1, 2, \ldots, n-4$
	called MHV, NMHV, N^2MHV, etc.
loop order	$L = 0, 1, 2, \ldots$
	called tree-level, one-loop, two-loop, etc.

and are functions on $\mathrm{Gr}(4, n)/(\mathbb{C}^*)^{n-1}$, called the kinematic space.

A generic point in kinematic space can be represented by an $n \times 4$ matrix Z whose rows $Z_1, Z_2, \ldots, Z_n$ (called momentum twistors) can each be interpreted as the homogeneous coordinates of a point in $\mathbb{P}^{4-1}$.

It's fun to play with a generalization where "4" in the previous paragraph is replaced by a general parameter usually called "m".

Introduction to Amplitudes

Tree-level ($L = 0$) amplitudes are relatively simple, but still very rich. They are rational functions of the Plücker coordinates

$$\langle a\, b\, c\, d \rangle = \det(Z_a\, Z_b\, Z_c\, Z_d)$$

Moreover, they have poles only when $\langle a\, a{+}1\, b\, b{+}1 \rangle = 0$ for some a and b (indices are always understood mod n), and the residue of any n-particle amplitude on any of its poles is a sum of products of pairs of simpler (smaller-n) amplitudes.

Tree-level amplitudes can be computed using BCFW recursion, which expresses the n-particle $\mathrm{N}^k\mathrm{MHV}$ amplitude as sum of

$$\frac{1}{n-3}\binom{n-3}{k-1}\binom{n-3}{k-2}$$

rational functions of Plücker variables. This is the problem of "tiling the (tree-level) amplituhedron."

Introduction to Amplitudes

Loop amplitudes ($L > 0$) are more complicated; in particular they are multivalued functions, with very complicated structures of branch points.

The best-understood loop amplitudes belong to the class of functions that can be written as w-fold iterated integrals of the form

$$\int_{a_0}^{a_{w+1}} \frac{dt_1}{t_1 - a_1} \int_{a_0}^{t_1} \frac{dt_2}{t_2 - a_2} \cdots \int_{a_0}^{t_{w-1}} \frac{dt_w}{t_w - a_w}$$

that generalize the polylogarithm functions

$$\mathrm{Li}_1 = -\log(1 - z) \qquad \mathrm{Li}_{w+1}(z) = \int_0^z \frac{dt}{t} \, \mathrm{Li}_w(t)$$

The index $w$is called the weight of the function; L-loop amplitudes have weight $w = 2L$.

The Symbol Calculus for Generalized Polylogarithms

These are the best understood amplitudes because we have the most powerful mathematical tools for handling them.

Let $\mathcal{A} = \oplus_w \mathcal{A}_w$ be the algebra generated by the above iterated integrals (over $\mathbb{Q}$).

Goncharov showed that there is an associated coproduct $\Delta : \mathcal{A} \mapsto \mathcal{A} \times \mathcal{A}$ that is compatible with multiplication, giving $\mathcal{A}$ the structure of a Hopf algebra.

The Symbol Calculus for Generalized Polylogarithms

Let $\mathcal{I}_w \in \mathcal{A}_w$ be an iterated integral of weight w. The symbol of $\mathcal{I}_w$, defined by

$$\mathcal{S}(I_w) = \Delta^w(\mathcal{I}_w)|_{\otimes^w \mathcal{A}_1}$$

was introduced to physics by Goncharov, Vergu, Volovich, MS and has enabled the calculation of many new loop amplitudes, because it (1) trivializes complicated functional identities and (2) efficiently encapsulates key information about the differential and monodromies of $\mathcal{I}_w$.

Some Alchemy of SYM Amplitudes

GSVV used the symbol calculus to help write the 2-loop 6-point MHV amplitude as

$$-\frac{1}{2}\,\mathrm{Li}_4\left(-\frac{\langle 1\,2\,4\,6\rangle\langle 1\,3\,4\,5\rangle}{\langle 1\,4\,5\,6\rangle\langle 1\,2\,3\,4\rangle}\right) + \text{“similar”}$$

It's symbol (7272 terms long) can be expressed entirely in terms of the 15 Plücker coordinates on $\mathrm{Gr}(4,6)$.

Some Alchemy of SYM Amplitudes

Caron-Huot subsequently computed the 2-loop MHV amplitudes for all n.

Staring at the results, you'll see things like

$$\mathrm{Li}_4\left(-\frac{\langle 1\,2\,5\,6\rangle\langle 2\,5\,7\,8\rangle(\langle 1\,2\,3\,7\rangle\langle 4\,5\,6\,8\rangle - \langle 1\,2\,3\,8\rangle\langle 4\,5\,6\,7\rangle)}{\langle 1\,2\,3\,7\rangle\langle 1\,2\,5\,8\rangle\langle 2\,4\,5\,6\rangle\langle 5\,6\,7\,8\rangle}\right)$$

but never

$$\mathrm{Li}_4\left(-\frac{\langle 1\,2\,5\,6\rangle\langle 2\,5\,7\,8\rangle(\langle 1\,2\,3\,7\rangle\langle 4\,5\,6\,8\rangle - \langle 1\,2\,3\,8\rangle\langle 4\,5\,6\,7\rangle)}{\langle 1\,2\,3\,8\rangle\langle 1\,2\,5\,7\rangle\langle 2\,4\,5\,6\rangle\langle 5\,6\,7\,8\rangle}\right)$$

Why? Cluster algebras! [WHY??? We don't know]

Cluster Algebras

Introduced by Fomin & Zelevinsky, cluster algebras are commutative algebras with distinguished generators called cluster variables grouped into collections called clusters, with a specific involution (called mutation) for generating a new cluster from a given one.

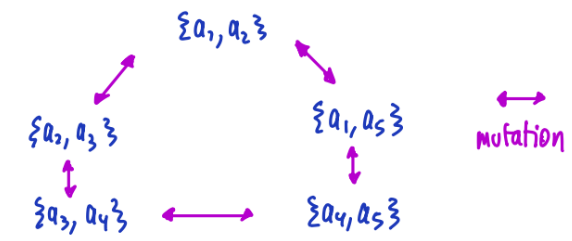
In general two cluster variables are called compatible if they appear in a cluster together.

Cluster Algebras

The simplest nontrivial example is the A_2 cluster algebra which has 5 clusters, thanks to the remarkable fact that the sequence

$$a_{i+1} = \frac{1 + a_i}{a_{i-1}}$$

is periodic with periodicity 5 (check it!). The 5 clusters are



In this example, a_i is compatible with $a_{i \pm 1 \bmod 5}$.

Aside: Cluster Algebras & Polylogarithms

The 5-term identity for the dilogarithm (Abel, 1881) is closely connected to the A_2 cluster algebra; it is schematically

$$\sum_{i=1}^{5} \mathrm{Li}_2(-a_i) + \log a_i \log a_{i+1} = 0$$

The study of amplitudes has stimulated the discovery of higher-weight generalizations; cluster algebras & polylogarithms are a natural match!

Cluster Algebras for SYM Theory

The Grassmannian $\text{Gr}(4, n)$ relevant to the kinematic space in SYM theory also admits a cluster structure.

The cluster variables of this algebra are certain homogeneous polynomials in Plücker coordinates; for example

$$\langle 1\,2\,4\,5\rangle\langle 2\,3\,6\,7\rangle - \langle 1\,2\,6\,7\rangle\langle 2\,3\,4\,5\rangle$$

is a cluster variable of $\text{Gr}(4, 7)$, but

$$\langle 1\,2\,4\,7\rangle\langle 3\,5\,6\,7\rangle - \langle 1\,3\,5\,7\rangle\langle 2\,4\,6\,7\rangle$$

is not.

Some Clustery Observations about SYM Amplitudes

For $n < 8$, all known amplitudes $\mathcal{A}$

(1) have branch points only when some cluster variable of $\mathrm{Gr}(4, n)$ vanishes.

(2) Moreover, if a_1 and a_2 are incompatible cluster variables, then

$$M_{a_1} M_{a_2} \mathcal{A} = 0$$

(where M_a denotes the monodromy around $a = 0$). This is an aspect of what we call cluster adjacency.

The hypothesis that this is true for all such amplitudes, at any loop order, has enabled a bootstrap program to compute them to very high loop order (8 loops for $n = 6$ and 4 loops for $n = 7$); see arXiv:2005.06735 for a review.

And Some Non-Clustery Aspects

Starting with the 2-loop 8-particle N^2MHV amplitude (He, Li, Zhang) it is known that there are amplitudes with branch points that do not correspond to the vanishing of cluster variables; for example, at the locus

$$A^2 - 4B = 0$$

where

$$A = \langle 1\,2\,5\,6\rangle\langle 3\,4\,7\,8\rangle - \langle 1\,2\,7\,8\rangle\langle 3\,4\,5\,6\rangle - \langle 1\,2\,3\,4\rangle\langle 5\,6\,7\,8\rangle$$
$$B = \langle 1\,2\,3\,4\rangle\langle 3\,4\,5\,6\rangle\langle 5\,6\,7\,8\rangle\langle 1\,2\,7\,8\rangle$$

Neither A nor $A^2 - 4B$ is a cluster variable of $\text{Gr}(4, 8)$.

Also, for $n > 7$ amplitudes generically have branch points at the vanishing of algebraic functions of Plücker coordinates, for example at

$$A - \sqrt{A^2 - 4B} = 0$$

Status of Known Symbol Alphabets at $n = 8$

The number of letters of all currently known $(k, L, 8)$ particle symbol alphabets:

	$(0,2,8)$ 1105.5606	$(1,2,8)$ 1911.01290	$(0,3,8)$ 2110.00350	$(2,2,8)$ 2209.10856	# available in $G(4,8)$
Plücker	68	68	68	68	70
Quadratic	48	96	104	104	120
Cubic	0	16	32	32	174
Algebraic	0	18	18	18	
Polynomial Non-Cluster	0	0	0	2	

More limited information, limited to two loops, is known for $n > 8$.

Some Big Open Questions for Math & Physics

For fixed $n > 7$, the $\mathrm{Gr}(4, n)$ cluster algebra has infinitely many cluster variables.

(1) Does the number of symbol letters in the L-loop amplitude grow without bound as L increases, or does it stabilize at some finite subset?

(2) If the former, do we eventually encounter <u>all</u> cluster variables of $\mathrm{Gr}(4, n)$, or only a proper subset?

(3) Is there a mathematically natural characterization of that subset, as well as of the apparently very special algebraic and non-cluster variable symbol letters?

Some Big Open Questions for Math & Physics

Several approaches have been proposed to build extra structure around the $\mathrm{Gr}(4, n)$ cluster algebra in order to describe these phenomena, including

Drummond, Foster, Kalousios, Gürdoğan
Henke, Papathanasiou
Herderschee
Mago, Schreiber, Yelleshpur Srikant, Volovich, MS
Arkani-Hamed, Lam, MS
Ren, Volovich, MS
Yang; He, Jiang, Liu, Yang
He, Li

but rather than reviewing all of this, let me make 2 observations.

Positivity

Above I defined

$$A = \langle 1\,2\,5\,6 \rangle \langle 3\,4\,7\,8 \rangle - \langle 1\,2\,7\,8 \rangle \langle 3\,4\,5\,6 \rangle - \langle 1\,2\,3\,4 \rangle \langle 5\,6\,7\,8 \rangle$$
$$B = \langle 1\,2\,3\,4 \rangle \langle 3\,4\,5\,6 \rangle \langle 5\,6\,7\,8 \rangle \langle 1\,2\,7\,8 \rangle$$

It is not obvious, but true, that $A > 0$ everywhere inside the positive Grassmannian, defined as the subset of the real Grassmannian where all (ordered) Plücker coordinates are positive.

In fact, we (Arkani-Hamed, Lam, MS) call A overpositive because

$$A > 2\sqrt{B} > 0$$

inside the positive Grassmannian (the "2" cannot be decreased), and hence

$$A^2 - 4B > 0$$

Positivity

In fact, all evidence available to date is consistent with the fundamental positivity hypothesis :

> Amplitudes in SYM theory have no singularities (neither poles–which is trivially true, nor branch points) when the momentum twistors Z lie inside the positive Grassmannian $\mathrm{Gr}^{>0}(4, n)$.

Yangian Invariants

Precisely the same 18 algebraic functions on $\mathrm{Gr}(4, 8)$ that appeared in the 2019 calculation of the 2-loop 8-particle NMHV amplitude are actually there in the literature all the way back in the 2012 paper by Arkani-Hamed et. al. on the positive Grassmannian, but in a seemingly different context.

Yangian Invariants

Mathematicians Even-Zohar, Lakrec, Parisi, Sherman-Bennett, Tessler & Williams define a tile to be an image of a cell $S \in \mathrm{Gr}^{\geq 0}(k, n)$ under the amplituhedron map

$$\mathrm{Gr}(k, n) \mapsto \mathrm{Gr}(k, k+4)$$

induced by multiplying by an $n \times k{+}m$ matrix Z, provided that the map is injective on S and the image is km-dimensional.

In physics we consider associate tiles with "Yangian invariants", but the meaning of the latter term is broader and allows several of the above conditions to be relaxed.

Algebraic Functions from Yangian Invariants

In particular, for $n = 8, m = 4, k = 2$ there is for the first time a $km = 8$-dimensional cell of $\mathrm{Gr}^{\geq 0}(2, 8)$ on which the amplituhedron map is not injective, but two-to-one.

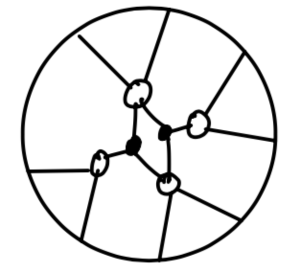
The corresponding "tile" has facets characterized by inequalities

$$A^2 - 4B > 0$$

and

$$A - \sqrt{A^2 - 4B} > 0$$

and similar, for precisely 9 of the 18 algebraic letters previously mentioned.



The plabic graph associated with this cell.

Cluster Adjacency for Yangian Invariants

Yangian invariants that do correspond to tiles exhibit a remarkable form of cluster adjacency; each one is a rational function of Plücker coordinates on $\mathrm{Gr}(4, n)$ whose denominator is a product of compatible cluster variables!

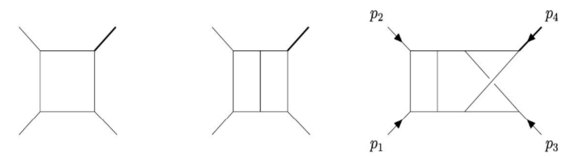
This was first observed "experimentally" by Drummond, Foster, Gürdoğan; Golden, McLeod, Volovich, MS and has been proven by Even-Zohar et. al.

It is a big open problem to understand how the notion of "compatibility" of cluster variables extends to the algebraic functions that we see in these "non-injective tiles".

Cluster Algebras Beyond SYM Theory

The question of whether cluster algebras might be relevant for scattering beyond the tame world of SYM theory was answered in the affirmative by Chicherin, Henn, Papathanasiou and Aliaj, Papathanasiou who found that the symbol letters of certain $L =$ 1-, 2- and 3-loop Feynman integrals relevant to general massless scattering are described respectively by the A_2, C_2, G_2 cluster algebras.

$$a_1 = \frac{s - p_4^2}{t}, \qquad a_2 = \frac{s}{t}\frac{p_4^2 - s - t}{p_4^2}, \qquad a_{i+1} = \begin{cases} \frac{1+a_i}{a_{i-1}} & i \text{ odd} \\ \frac{1+a_i^L}{a_{i-1}} & i \text{ even} \end{cases}$$



A Cluster Algebra 5-Particle 2-Loop QCD

The full set of master integrals for two-loop five-particle massless planar (or non-planar) QCD involves a symbol alphabet with 26(or 31) letters. [Gehrmann, Henn, Lo Presti; Chicherin, Henn, Mitev]

These letters were shown by Bossinger, Drummond, Glew to coincide with those from a construction based on the D_4 cluster algebra.

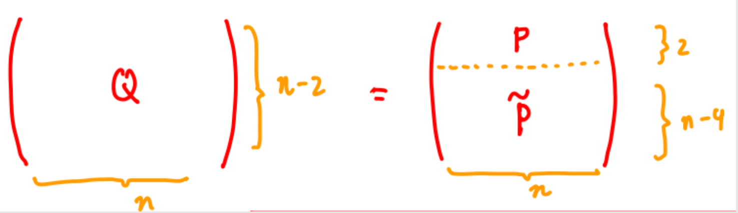
"Based on" is key, since D_4 only has 22 cluster variables!

I'll present their construction in a different but equivalent way that makes use of a cluster structure on partial flag varieties recently described by Bossinger & Li.

Of relevance to us is $F(2, n-2; n)$, whose initial cluster is related to that of $\mathrm{Gr}(n-2, 2n-4)$ by a certain sequence of mutations, freezings, and deletions.

A Cluster Algebra 5-Particle 2-Loop QCD

These are relevant to physics because using spinor helicity variables, there is an isomorphism between a configuration of n massless particles and a point in $F(2, n-2; n)$ represented by an $(n-2) \times n$ matrix [Maazouz, Pfister, Sturmfels]



The correspondence is

$$\langle ij \rangle = \det(P_i\ P_j)$$

$$[ij] = (-1)^{i+j+1} \det(Q_1\, Q_2 \cdots \hat{Q}_i \cdots \hat{Q}_j \cdots Q_n)$$

A Cluster Algebra Five-Particle Two-Loop QCD

In this notation the 22 cluster variables are

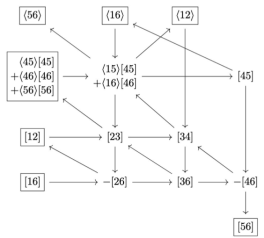
$$\{\underbrace{\langle i\, j\rangle}_{10}, \underbrace{[i\, j]}_{10}, \langle 2\, 3\rangle[2\, 3] - \langle 4\, 5\rangle[4\, 5], \langle 3\, 4\rangle[3\, 4] - \langle 1\, 2\rangle[1\, 2]\}$$

This set is not closed under the natural cyclic group acting on the five particles (because the identification described above breaks this symmetry!) but if we take the union under all permutations it produces exactly 30 of the 31 letters appearing in any two-loop five-particle integral; the sole exception is W_{31} which is not known to appear in any amplitude – it always drops out.

A Cluster Algebra for 6-Particle 2-Loop QCD?

Very recently the full symbol alphabet for all two-loop 6-particle master integrals has been worked out by
Abreu, Monii, Page, Usovitsch
Henn, Matijasic, Miczajka, Peraro, Xu, Zhang

Pokraka, Volovich, Weng, MS showed that most of the letters can be "explained" similarly to the five-particle case using a construction based on $F(2,4;6)$ – see also recent work by Bossinger, Drummond, Glew, Gürdoğan, Wright.



We are also exploring whether the "BCFW terms" associated to the momentum amplituhedron, in spinor helicity variables, exhibit flag cluster structure/adjacency like that described above – see He-Chen Weng's gong show.

Conclusion

The scattering amplitudes of SYM theory are a remarkable collection of functions with surprisingly rich mathematical properties.

They clearly exhibit deep connections to Grassmannian cluster algebras in the simplest cases, but that is only the tip of a very large iceberg.

There is a lot more to learn, which might be fun for mathematicians and possibly even useful for physicists who need new computational methods!