Kinematic Hopf Algebra for Bern-Carrasco-Johansson Numerators in Heavy-Mass Effective Field Theory and Yang-Mills Theory

Andreas Brandhuber [1,*] Gang Chen [1,†] Henrik Johansson [2,3,‡] Gabriele Travaglini [1,§] and Congkao Wen [1,‖]

[1] Centre for Theoretical Physics, Department of Physics and Astronomy, Queen Mary University of London, Mile End Road, London E1 4NS, United Kingdom

[2] Department of Physics and Astronomy, Uppsala University, Box 516, 75120 Uppsala, Sweden

[3] Nordita, Stockholm University and KTH Royal Institute of Technology, Hannes Alfvéns väg 12, 10691 Stockholm, Sweden

(Received 8 December 2021; accepted 1 March 2022; published 25 March 2022)

We present a closed formula for all Bern-Carrasco-Johansson (BCJ) numerators describing D-dimensional tree-level scattering amplitudes in a heavy-mass effective field theory with two massive particles and an arbitrary number of gluons. The corresponding gravitational amplitudes obtained via the double copy directly enter the computation of black-hole scattering and gravitational-wave emission. Our construction is based on finding a kinematic algebra for the numerators, which we relate to a quasishuffle Hopf algebra. The BCJ numerators thus obtained have a compact form and intriguing features: gauge invariance is manifest, locality is respected for massless exchange, and they contain poles corresponding to massive exchange. Counting the number of terms in a BCJ numerator for $n-2$ gluons gives the Fubini numbers F_{n-3}, reflecting the underlying quasishuffle Hopf algebra structure. Finally, by considering an appropriate factorization limit, the massive particles decouple, and we thus obtain a kinematic algebra and all tree-level BCJ numerators for D-dimensional pure Yang-Mills theory.

DOI: 10.1103/PhysRevLett.128.121601

Introduction.—Quantum field theory holds many surprising discoveries, one of which is the Bern-Carrasco-Johansson (BCJ) duality between color and kinematics [1,2]. In addition to providing a field-theory underpinning of the Kawai-Lewellen-Tye (KLT) open-closed string relations [3], the duality hints at a hidden algebraic structure in a variety of gauge theories. Scattering amplitudes in these theories can be written as a sum of cubic diagrams, each one expressed as the product of a color and a kinematic factor. The color factors satisfy Jacobi relations inherited from the gauge-group Lie algebra, and the kinematic numerators satisfy corresponding kinematic Jacobi relations [1]. Through the double-copy construction, gravitational amplitudes can be obtained from the kinematic numerators.

A central question is to identify the hidden algebra behind the kinematic relations. In this Letter we provide an explicit construction, in two related contexts. First we will study the amplitudes in an effective theory of heavy particles coupled to gluons, or gravitons [4–8]. These theories, which we will refer to as HEFT (heavy-mass effective field theory), [9] are obtained from a Yang-Mills (YM) theory, or general relativity, by restricting to the leading-order term in an inverse mass expansion. This is an appropriate approximation for the dynamics of particles with momentum exchange much smaller than their masses. Astrophysical black-hole scattering in general relativity satisfies this, and the relevant gravitational amplitudes were recently studied through a gauge-invariant double copy [10]. The underlying gauge-theory factors are the central objects, and we will here unravel their algebraic structure, including that of pure YM theory after factorizing out the heavy particles.

The understanding of the kinematic algebra has so far only progressed in small steps. The first successful construction of the algebra was limited to the self-dual sector of YM theory [11]. In that case the algebra corresponds to area-preserving diffeomorphisms, and explicit representations of the generators were found. Self-dual YM is far from a complete theory, having vanishing tree amplitudes (apart from a single three-point amplitude for complex momenta) and a non-CPT invariant spectrum, yet it is the first confirmation of BCJ duality with explicit generators and cubic Feynman rules. Another example of the duality was found in the nonlinear sigma model [12], as realized in Ref. [13] using a cubic Lagrangian. The corresponding kinematic algebra was later [14] tied to that of higher-dimensional Poincaré symmetry [15].

Efforts to identify the kinematic algebra have recently been renewed for YM theory [16,17], and for HEFT [10]. The common idea is to realize the algebra with abstract vector and tensor currents, multiplied through a fusion product. A consistent fusion product was worked out for the maximally helicity-violating (MHV) and next-to-MHV sectors of YM theory [16,17] [18] The approach was then applied to HEFT in Ref. [10], giving explicit expressions for two heavy particles coupled to gluons or gravitons, and the fusion products were presented up to six particles. This approach is well adapted for gravitational physics, and was used to compute the black hole scattering angle in a post-Minkowskian (PM) expansion at 3PM order [19] (see also Refs. [20–26]).

In this Letter we construct a kinematic algebra for HEFT, and by factorization, infer that the same algebra also works for pure YM theory. In particular, we give a representation of all the generators, and all fusion products needed for computing tree-level HEFT amplitudes with two heavy particles and an arbitrary number of gluons or gravitons. Interestingly, the obtained fusion product has the same structure as the quasishuffle product, known from the mathematical literature, specifically in the context of combinatorial Hopf algebras of shuffles and quasishuffles [27–29]. The quasishuffle Hopf algebra generates all ordered partitions for a given set [27] (often called $\mathbb{SC}$—the linear species of set compositions, or ordered partitions). Mapping the generators to gauge-invariant expressions, we obtain a closed formula for all tree-level BCJ numerators relevant to the HEFT. The numerators are gauge invariant, manifestly crossing symmetric and factorize into lower-point numerators on the massive poles. The underlying quasishuffle Hopf algebra implies that the counting of the number of terms in a numerator with $n-2$ gluons gives the Fubini number F_{n-3}, which counts the number of ordered partitions of $n-3$ elements.

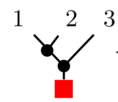
Finally, all the considerations in HEFT directly translate to pure YM theory. The pure-gluon BCJ numerators, and the corresponding expressions for the generators, are obtained from the natural on-shell factorization limit [10], which removes the two heavy particles and replaces them with an additional gluon (with label $n-1$). This is straightforward: replace the heavy-particle velocity v with the last polarization vector, $v\to\epsilon_{n-1}$, and impose the last on-shell condition $p^2_{1\ldots n-2}\to 0$. This operation does not modify the generator fusion rules, and hence YM theory admits the same kinematic algebra. The heavy-mass poles become spurious in this limit, and cancel out once the amplitude is assembled.

The HEFT kinematic algebra.—A novel color-kinematic duality and double copy for HEFT was obtained in Ref. [10], by four of us. Ignoring couplings, the YM and gravity tree amplitudes with two heavy particles and $n-2$ gluons or gravitons are

$$A(12\ldots n-2,v)=\sum_{\Gamma\in\rho}\frac{\mathcal{N}(\Gamma,v)}{d_\Gamma},$$
$$M(12\ldots n-2,v)=\sum_{\Gamma\in\tilde\rho}\frac{[\mathcal{N}(\Gamma,v)]^2}{d_\Gamma},\qquad(1)$$

where ρ ($\tilde\rho$) denotes all (un)ordered nested commutators of the particle labels $\{1,\ldots,n-2\}$, where the leftmost label is fixed to 1. The ordering is important since here we work with color-ordered YM amplitudes. Considering the set $\{1,2,3\}$, we have $\rho=\{[[1,2],3],[1,[2,3]]\}$ and $\tilde\rho=\{[[1,2],3],[[1,3],2],[1,[2,3]]\}$. In general, labels $1,\ldots,n-2$ are reserved for the gluons or gravitons and the heavy particles are assigned $n-1$ and n, and v is the velocity that characterizes the heavy particles.

The nested commutators are in one-to-one correspondence with cubic graphs (and hence BCJ numerators), and the corresponding massless scalarlike propagator denominators are denoted as d_Γ. For instance, the nested commutator $[[1,2],3]$ corresponds to the following cubic graph, associated BCJ numerator, and propagator denominator:



$$\leftrightarrow\ \mathcal{N}([[1,2],3],v)\,,\qquad d_{[[1,2],3]}=p^2_{12}p^2_{123}\,,\qquad(2)$$

where $p_{i_1\ldots i_r}\coloneqq p_{i_1}+\cdots+p_{i_r}$ and the red square denotes the heavy-particle source.

The BCJ numerator $\mathcal{N}(\Gamma,v)$ is a function of a nested set of labels Γ, and it has an expansion which parallels that of the commutator, e.g.,

$$\mathcal{N}([1,[2,3]],v)=\mathcal{N}(123,v)-\mathcal{N}(132,v)$$
$$-\mathcal{N}(231,v)+\mathcal{N}(321,v),\qquad(3)$$

and we refer to the object $\mathcal{N}(1\ldots n-2,v)$ as the *prenumerator*. In analogy with a Lie algebra, this quantity should be obtained by multiplying generators through an associative fusion product. Thanks to the nested commutator structure, the BCJ numerators will automatically satisfy kinematic Jacobi identities.

Explicit prenumerators can be obtained from the constraint imposed by requiring that they lead to correct amplitudes, and in Ref. [10] this was done up to six points. In the following, it will be crucial to find representations of the prenumerators where any nonlocality will correspond to a massive physical pole $\sim 1/(v\cdot P)$, where P is a sum of gluon momenta [30]. This linearized propagator arises because of the large-mass expansion. Our results will be an improvement compared to Ref. [10], since in that work additional spurious poles were present in the prenumerators. We find the following explicit new results up to five points:

$$\mathcal{N}(1,v)=v\cdot\varepsilon_1,$$
$$\mathcal{N}(12,v)=-\frac{v\cdot F_1\cdot F_2\cdot v}{2v\cdot p_1},$$
$$\mathcal{N}(123,v)=\frac{v\cdot F_1\cdot F_2\cdot F_3\cdot v}{3v\cdot p_1}-\frac{v\cdot F_1\cdot F_2\cdot V_{12}\cdot F_3\cdot v}{3v\cdot p_1 v\cdot p_{12}}$$
$$-\frac{v\cdot F_1\cdot F_3\cdot V_1\cdot F_2\cdot v}{3v\cdot p_1 v\cdot p_{13}},\qquad(4)$$

where $F_i^{\mu\nu} \coloneqq p_i^\mu \epsilon_i^\nu - \epsilon_i^\mu p_i^\nu$, and $V_\tau^{\mu\nu} \coloneqq v^\mu \sum_{j\in\tau} p_j^\nu = v^\mu p_\tau^\nu$. Note that gauge invariance is manifest except in the case of $\mathcal{N}(1,v)$, where it follows from three-point kinematics.

Following Refs. [10,16,17], the prenumerators are presumed to be constructible in an algebraic fashion, by multiplying abstract generators of the kinematic algebra via a fusion product,

$$\mathcal{N}(12\ldots n-2, v) \coloneqq \langle T_{(1)} \star T_{(2)} \star \cdots \star T_{(n-2)} \rangle, \tag{5}$$

where the $T_{(i)}$ s are generators carrying the gluon label i, and $\star$ denotes the bilinear and associative fusion product. The angle bracket represents a linear map from the abstract generators to gauge- and Lorentz-invariant functions. It preserves the multilinearity with respect to the polarization vectors and the linear scaling in the velocity v of the heavy particles.

The starting point of the construction is $\langle T_{(i)} \rangle = v \cdot \varepsilon_i$, which is the unique choice that respects all the properties listed above, and furthermore generates the correct three-point amplitude. We can then combine two generators to obtain

$$\mathcal{N}(12, v) \coloneqq \langle T_{(1)} \star T_{(2)} \rangle = -\langle T_{(12)} \rangle, \tag{6}$$

where we choose $\langle T_{(12)} \rangle = (v\cdot F_1 \cdot F_2 \cdot v)/(2v\cdot p_1)$ to reproduce Eq. (4) [31]. Similarly, at five points one finds

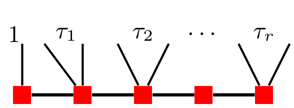
$$T_{(12)} \star T_{(3)} = -T_{(123)} + T_{(12),(3)} + T_{(13),(2)}, \tag{7}$$

with

$$\begin{aligned} \langle T_{(123)} \rangle &= \frac{v\cdot F_1 \cdot F_2 \cdot F_3 \cdot v}{3v\cdot p_1}, \\ \langle T_{(12),(3)} \rangle &= \frac{v\cdot F_1 \cdot F_2 \cdot V_{12} \cdot F_3 \cdot v}{3v\cdot p_1 v \cdot p_{12}}, \\ \langle T_{(13),(2)} \rangle &= \frac{v\cdot F_1 \cdot F_3 \cdot V_1 \cdot F_2 \cdot v}{3v\cdot p_1 v\cdot p_{13}}. \end{aligned} \tag{8}$$

The particular index assignments in the obtained generators are consistent with a general formula, which we find to work to any number of points,

$$\langle T_{(1\tau_1),(\tau_2),\ldots,(\tau_r)} \rangle \coloneqq$$



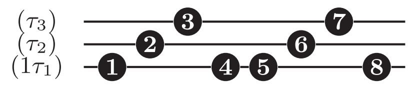
$$= \frac{v\cdot F_{1\tau_1} \cdot V_{\Theta(\tau_2)} \cdot F_{\tau_2} \cdots V_{\Theta(\tau_r)} \cdot F_{\tau_r} \cdot v}{(n-2) v\cdot p_1 v\cdot p_{1\tau_1} \cdots v\cdot p_{1\tau_1\tau_2\cdots\tau_{r-1}}}. \tag{9}$$

The τ_i's are ordered nonempty sets such that $\tau_1 \cup \tau_2 \cup \cdots \cup \tau_r = \{2,3,\ldots,n-2\}$ and $\tau_i \cap \tau_j = \varnothing$, i.e., they constitute a partition. The set $\Theta(\tau_i)$ consists of all indices to the left of τ_i and smaller than the first index in τ_i; that is $\Theta(\tau_i) = (\{1\} \cup \tau_1 \cup \cdots \cup \tau_{i-1}) \cap \{1,\ldots,\tau_{i[1]}\}$. Note that the denominators in Eq. (9) are the advertised massive propagators. For convenience, we also define F_{τ_i} as the ordered contraction of several linearized field strengths $F_j^{\mu\nu}$ with indices in τ_i, e.g., $F_{12}^{\mu\nu} = F_1^{\mu\alpha} F_{2\alpha}^\nu$.

To clarify the formula, consider a nontrivial example, $T_{(1458),(26),(37)}$, that is mapped to

$$\langle T_{(1458),(26),(37)} \rangle = \frac{v\cdot F_{1458} \cdot V_1 \cdot F_{26} \cdot V_{12} \cdot F_{37} \cdot v}{8v\cdot p_1 v\cdot p_{1458} v\cdot p_{124568}}. \tag{10}$$

We may further clarify the $\Theta(\tau_i)$s by drawing a "musical diagram," where the gluon labels (notes) are filled in progressively from left to right and each horizontal line indicates which set in the partition they belong to:



(11)

A given $\Theta(\tau_i)$ is associated with the first gluon on the horizontal line τ_i, and the set includes all labels "south-west" of this gluon. Specifically, in this example, the relevant sets used in Eq. (10) are $\Theta(26) = \{1\}$, and $\Theta(37) = \{1,2\}$. Furthermore, the contraction of field strengths can be read out by following each horizontal τ_i line in this musical diagram. A horizontal line can be thought of as the fundamental representation of the Lorentz group, and the linearized field strengths as Lorentz generators acting in this space.

Let us return to the algebra of the abstract generators. The prenumerators can be recursively constructed from only knowing the following fusion product:

$$T_{(1\tau_1),(\tau_2),\ldots,(\tau_r)} \star T_{(j)}. \tag{12}$$

We assume that the possible outcome of this fusion product maintains the relative order of the labels in the left and right generator. Then by assuming we have a complete set of generators, we can only produce the terms

$$\begin{aligned} &T_{(1j),(\tau_1),(\tau_2),\ldots,(\tau_r)}, \qquad T_{(1\tau_1),(\tau_2),\ldots,(\tau_i),(j),(\tau_{i+1}),\ldots,(\tau_r)}, \\ &T_{(1\tau_1),(\tau_2),\ldots,(\tau_i j),\ldots,(\tau_r)}, \quad \text{where } i \in \{1,\ldots,r\}. \end{aligned} \tag{13}$$

By writing up a general ansatz, and fixing the free coefficients by comparing to the correct amplitudes via the map in Eq. (9), we find a simple all-multiplicity solution. The fusion product is captured by the general formula

$$\begin{aligned} T_{(1\tau_1),\ldots,(\tau_r)} \star T_{(j)} = &\sum_{\sigma \in \{(\tau_1),\ldots,(\tau_r)\} \sqcup \{(j)\}} T_{(1\sigma_1),\ldots,(\sigma_{r+1})} \\ &- \sum_{i=1}^{r} T_{(1\tau_1),\ldots,(\tau_{i-1}),(\tau_i j),(\tau_{i+1}),\ldots,(\tau_r)}, \end{aligned} \tag{14}$$

where ⧢ denotes the shuffle product between two sets, e.g., $\{A,B\}⧢\{C\}=\{ABC,ACB,CAB\}$. A proof for Eq. (14) will be given in the next section; here we will study examples. For $n=4$, 5, Eqs. (6) and (7) are recovered, and at six points, the fusion products are

$$
\begin{aligned}
T_{(123)}\star T_{(4)} &= -T_{(1234)} + T_{(123),(4)} + T_{(14),(23)} \\
T_{(12),(3)}\star T_{(4)} &= -T_{(12),(34)} - T_{(124),(3)} \\
&\quad + T_{(12),(3),(4)} + T_{(12),(4),(3)} + T_{(14),(2),(3)} \\
T_{(13),(2)}\star T_{(4)} &= -T_{(13),(24)} - T_{(134),(2)} + T_{(13),(2),(4)} \\
&\quad + T_{(13),(4),(2)} + T_{(14),(3),(2)},
\end{aligned}
\tag{15}
$$

leading to the six-point prenumerator

$$
\begin{aligned}
\mathcal{N}(1234,v) = \langle & -T_{(12),(3),(4)} - T_{(12),(4),(3)} - T_{(14),(2),(3)} \\
& - T_{(14),(3),(2)} - T_{(13),(2),(4)} - T_{(13),(4),(2)} \\
& + T_{(123),(4)} + T_{(124),(3)} + T_{(134),(2)} \\
& + T_{(12),(34)} + T_{(13),(24)} + T_{(14),(23)} - T_{(1234)} \rangle.
\end{aligned}
\tag{16}
$$

As already advertised, the algebra defined by the fusion product in Eq. (14) is known in the context of combinatorial Hopf algebras of shuffles and quasishuffles [27–29]. Specifically, our fusion product defines a quasishuffle Hopf algebra that generates all ordered partitions for a given set [27]. Indeed, the subscripts of the T's are precisely all possible ordered partitions of $\{2,3,\ldots,n-2\}$. This is also interpreted in Ref. [29] as a Hopf monoid in the category of coalgebra species. These Hopf algebras are endowed with a product that is commutative and associative [27,32,33], with a coproduct, counit, and antipode [27] (see the Supplemental Material [34] for more details).

We have thus found a realization of the kinematic algebra for HEFT by mapping it to a quasi-shuffle Hopf algebra. Note that the associativity of the fusion product is a natural property—for example, we can construct a BCJ numerator either as $((T_{(1)}\star T_{(2)})\star T_{(3)})\star\cdots$ or $\cdots\star(T_{(n-4)}\star(T_{(n-3)}\star T_{(n-2)}))$. To complete the story, we must also give the fusion product for the most general generators. Assuming the fusion product is associative and preserves the relative order for the left and right generators, one obtains a unique result [28],

$$
\begin{aligned}
&T_{(1\tau_1),\ldots,(\tau_r)}\star T_{(\omega_1),\ldots,(\omega_s)} \\
&\quad = \sum_{\substack{\sigma|_{\{\tau\}}=\{(\tau_1),\ldots,(\tau_r)\} \\ \sigma|_{\{\omega\}}=\{(\omega_1),\ldots,(\omega_s)\}}} (-1)^{t-r-s} T_{(1\sigma_1),(\sigma_2),\ldots,(\sigma_t)},
\end{aligned}
\tag{17}
$$

where τ_i and ω_j do not contain the label 1, as this index is always fixed to be the leftmost index of any expression, and thus it is inert to the algebra. The fusion product of two generators, neither containing label 1, is also given by Eq. (17) after dropping the 1. We use $\{\tau\}$ and $\{\omega\}$ to denote the total set of labels in the τ_i and ω_i, respectively. By $\sigma|_{\{\tau\}}$ we mean a restriction to the elements in $\{\tau\}$, e.g., $\{(235),(4),(67)\}|_{\{2,3,4\}}=\{(23),(4)\}$.

The number of ordered partitions of $\{2,3,\ldots,n-2\}$ are known as the Fubini numbers [35]

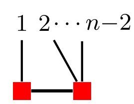
$$
\mathsf{F}_{n-3} = \sum_{r=1}^{n-3} r! \left\{ \begin{matrix} n-3 \\ r \end{matrix} \right\},
\tag{18}
$$

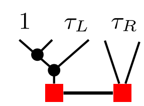
which therefore also counts the number of terms in the prenumerator of an n-point HEFT amplitude. Here, $\{{n \atop k}\}$ denotes the number of k partitions on n objects (also known as Stirling partition number of the second kind). The Fubini numbers also give the Hilbert series of $\mathbb{SC}$ [29].

From the kinematic algebra, the closed form of the prenumerator is directly obtained as

$$
\mathcal{N}(1\ldots n-2,v) = \sum_{r=1}^{n-3} \sum_{\tau\in\mathbf{P}^{(r)}_{\{2,\ldots,n-2\}}} (-1)^{n+r} \langle T_{(1\tau_1),\ldots,(\tau_r)} \rangle,
\tag{19}
$$

where $\langle T_{(1\tau_1),\ldots,(\tau_r)}\rangle$ is defined in Eq. (9) and $\mathbf{P}^{(r)}_{\{2,\ldots,n-2\}}$ denotes all the ordered partitions of $\{2,3,\ldots,n-2\}$ into r subsets. This closed-form expression automatically induces a recursion relation for the prenumerator:

$\mathcal{N}(12\ldots n-2,v) =$ 

$+\sum_{\tau_R}$ 

$$
\begin{aligned}
&= (-1)^n \frac{v\cdot F_{12\ldots n-2}\cdot v}{(n-2)v\cdot p_1} \\
&\quad - \sum_{\tau_R\subset\{2,\cdots,n-2\}} (-1)^{n_R} \mathcal{N}(1\tau_L,v) \frac{(n-2-n_R)H_{\Theta_{\tau_R},\tau_R}}{(n-2)v\cdot p_{1\tau_L}},
\end{aligned}
\tag{20}
$$

where $\tau_L\cup\tau_R=\{2,3,\ldots,n-2\}$, $\tau_L,\tau_R\neq\varnothing$, and we have defined

$$
H_{\sigma,\tau} \coloneqq p_\sigma\cdot F_\tau\cdot v.
\tag{21}
$$

Here n_R denotes the number of indices in τ_R. From Eq. (20), we can see that the number of terms satisfies the recursion relation

$$
\mathsf{F}_{n-3} = \sum_{i=0}^{n-4} \binom{n-3}{i} \mathsf{F}_i,
\tag{22}
$$

where $\mathsf{F}_0=1$. This is the well-known recursion relation for the Fubini numbers [36]. To illustrate the simplicity of the prenumerator, we quote the fairly modest number of terms up to ten points:

n	3	4	5	6	7	8	9	10
F_{n-3}	1	1	3	13	75	541	4683	47 293

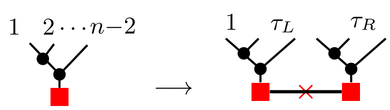
In the next section we present a general proof of our construction of the BCJ numerators.

Proof of the form of the prenumerator.—Here we give the proof of the BCJ numerators. Using $\mathcal{N}([12\ldots n-2],v) := \mathcal{N}([\ldots[[1,2],3]\ldots,n-2],v)$, we show that they give correct amplitudes, as obtained from the prenumerator in Eq. (19). In addition, we will show that the following simple relation holds, valid in the HEFT:

$$\mathcal{N}([12\ldots n-2],v) = (n-2)\mathcal{N}(12\ldots n-2,v). \quad (23)$$

The outline of the proof is as follows. Starting from the factorization properties on massive poles of our HEFT numerators as derived from the KLT formula [37], we prove that the quantity $(n-2)\mathcal{N}(12\ldots n-2,v)$ has the same factorization. We will then consider the difference between $(n-2)\mathcal{N}(12\ldots n-2,v)$ and the BCJ numerator (as derived from KLT), which is free of poles. Using arguments similar to those of Refs. [38,39], we will then show that gauge invariance ensures that this difference vanishes.

The starting point is the factorization on massive poles of BCJ numerators in HEFT. Using KLT relations, one can easily show that in the on-shell limit $v\cdot p_{1\tau_L}\to 0$



$$\mathcal{N}_{\rm KLT}([1\ldots n-2],v) \longrightarrow p_{\Theta(\tau_R)}\cdot p_{\tau_{R[1]}} \times \mathcal{N}_{\rm KLT}([1\tau_L],v)\mathcal{N}_{\rm KLT}([\tau_R],v)\,, \quad (24)$$

where $\tau_L\cup\tau_R=\{2,3,\ldots,n-2\}$ and $\tau_{R[1]}$ denote the first index in τ_R. We also called n_L (n_R) the number of gluons in τ_L (τ_R). The red cross denotes the cut on the physical pole. The derivation of this formula is given in the Supplemental Material [34], also making use of the results of Refs. [40–46].

The next step is to prove that the prenumerator in Eq. (19) has the same factorization as Eq. (24), which we will now do inductively. The seed of the induction is the factorization where the right-hand side of Eq. (24) only contains one gluon, that is, $n_R=1$, and we focus on the massive pole $1/(v\cdot p_{1\tau_L})$. The factorization is then immediately read off from Eq. (20): only one term in the second diagram in that equation contributes, with the residue given by

$$(n-3)\mathcal{N}(1i_2\ldots i_{n-3},v)p_{1i_2\ldots i_{n-3}}\cdot p_{i_{n-2}}\mathcal{N}(i_{n-2},v). \quad (25)$$

In the next step, we assume that factorization for $n_R=j-1$ at the massive pole $1/(v\cdot p_{1\tau_L})$ has the same form as Eq. (24), and we then derive that for $n_R=j$. According to Eq. (20), in this channel the residue of $(n-2)\mathcal{N}$ at the massive pole is

$$(n_L+1)\mathcal{N}(1\tau_L,v)\sum_{r=1}^{n_R}\sum_{\sigma\in\mathbf{P}^{(r)}_{\tau_R}}\frac{H_{\Theta(\sigma_1),\sigma_1}\cdots H_{\Theta(\sigma_r),\sigma_r}}{v\cdot p_{\sigma_1}\cdots v\cdot p_{\sigma_1\sigma_2\cdots\sigma_{r-1}}}. \quad (26)$$

As we show in the Supplemental Material [34], in the limit $v\cdot p_{1\tau_L}\to 0$, the sum in Eq. (26) becomes precisely a BCJ numerator,

$$\sum_{r=1}^{n_R}\sum_{\sigma\in\mathbf{P}^{(r)}_{\tau_R}}\frac{H_{\Theta(\sigma_1),\sigma_1}\cdots H_{\Theta(\sigma_r),\sigma_r}}{v\cdot p_{\sigma_1}\cdots v\cdot p_{\sigma_1\sigma_2\cdots\sigma_{r-1}}} = p_{\Theta(\tau_R)}\cdot p_{\tau_{R[1]}}n_R\mathcal{N}(\tau_R,v). \quad (27)$$

This establishes that our proposed formula in Eq. (19) has the required factorization property of Eq. (24).

Next, we consider the difference

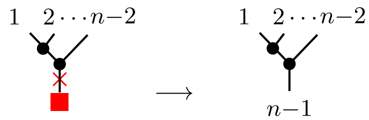
$$f=(n-2)\mathcal{N}(1\ldots n-2,v)-\mathcal{N}_{\rm KLT}([1\ldots n-2],v). \quad (28)$$

As the factorization on the heavy-mass poles is the same, and both contain only such poles, f must be a polynomial. An adaptation of the argument of Refs. [38,39] allows us to show that $f=0$. To this end, we note that the velocity v can appear in two possible ways. First, through the combination $p\cdot v$ with p being any of the momenta. This multiplies a polynomial function of dimension $n-4$ built from $n-2$ gluon momenta and $n-2$ polarizations. As is well known, and pointed out recently in Refs. [38,39], no such gauge-invariant function exists and hence it must vanish. Second, v can appear in the combination $v\cdot F$, which now multiplies a function of dimensions $n-4$, constructed from $n-2$ gluon momenta and $n-3$ polarizations. As before, such a function must vanish. Hence we conclude that $f=0$, and therefore

$$(n-2)\mathcal{N}(1\ldots n-2,v) = \mathcal{N}_{\rm KLT}([1\ldots n-2],v). \quad (29)$$

This completes the derivation of our BCJ numerator. As shown in the Supplemental Material [34] $\mathcal{N}_{\rm KLT}$ is crossing symmetric, hence $\mathcal{N}(1\ldots n-2,v)$ has the same property, which leads to Eq. (23). One may also verify Eq. (23) explicitly, e.g., at four points we have $\mathcal{N}([12],v)=\mathcal{N}(12,v)-\mathcal{N}(21,v)=(v\cdot F_1\cdot F_2\cdot v)/(2v\cdot p_1)-(v\cdot F_2\cdot F_1\cdot v)/(2v\cdot p_2)=2\mathcal{N}(12,v)$.

From HEFT to Yang-Mills.—It is straightforward to obtain the kinematic algebra, and the BCJ numerators, of pure YM theory from the HEFT construction, by exploiting the factorization property of the HEFT amplitude on a gluon pole [10]. The massive particles decouple on the pole, and we obtain the BCJ numerator for YM amplitudes:



$$\mathcal{N}^{\rm YM}([1\ldots n{-}1]) = \mathcal{N}([1\ldots n{-}2],v)\big|^{v\to\epsilon_{n-1}}_{p^2_{1\ldots n-2}\to 0}\,. \tag{30}$$

The same replacement should be performed on the expressions of the generators of the algebra given in Eq. (9), with no modification to the fusion rules. We have also explicitly verified Eq. (30) for D-dimensional YM amplitudes up to nine points.

The BCJ numerators thus obtained are manifestly gauge invariant and crossing symmetric for all gluons except the last one, $n-1$. The price to pay is that they also contain spurious poles of the form $1/(\epsilon_{n-1}\cdot P)$, which, however, can be eliminated in the complete amplitudes—in practice, one can use only independent variables (after imposing on-shell conditions and momentum conservation) in the amplitudes, then terms with spurious poles should cancel out, or we can simply drop them by hand.

As for the BCJ numerators, the spurious poles can also be eliminated explicitly. We take the MHV sector of the four-point case as an example to illustrate this idea. According to Eqs. (19) and (30), the corresponding numerator is

$$\left(v\cdot\epsilon_1 \frac{p_1\cdot\epsilon_2 p_{12}\cdot\epsilon_3 p_2\cdot v v\cdot p_3}{v\cdot p_1 v\cdot p_{12}} + v\cdot\epsilon_1 \frac{p_1\cdot\epsilon_3 p_1\cdot\epsilon_2 p_3\cdot v v\cdot p_2}{v\cdot p_1 v\cdot p_{13}} \right.$$
$$\left. - v\cdot\epsilon_1 \frac{p_1\cdot\epsilon_2 p_2\cdot\epsilon_3 p_3\cdot v}{v\cdot p_1}\right)\Bigg|_{v\to\epsilon_4} = \epsilon_1\cdot\epsilon_4 p_1\cdot\epsilon_2 p_{12}\cdot\epsilon_3, \tag{31}$$

which is in agreement with [16]. Similarly, we have verified this in the non-MHV sector in several examples.

A natural question arises as to how the known generalized gauge symmetry of the BCJ numerators in YM [1,2,47] manifests itself after taking the decoupling limit on the HEFT numerators. In this limit, the propagator matrix will become degenerate, which implies that one can add or subtract terms in its kernel and obtain a family of valid BCJ numerators.

Finally, we highlight potential connections between our and other approaches in the literature, e.g., in Refs. [48–51]. For example, the construction of Ref. [51] also maintains gauge invariance and crossing symmetry of $n-1$ external legs, and contains linear spurious poles. Intriguingly, the numerator of Ref. [51] for n gluons has $2\mathsf{F}_{n-2}$ terms, while ours has F_{n-2}. We also note the appearance in Ref. [50] of Cayley's trees in the construction of BCJ numerators, and it is well known that Fubini numbers are related to such graphs. It would be interesting to explore the connections among these approaches.

Conclusions.—In this Letter we constructed a kinematic algebra that manifests BCJ color-kinematics duality in tree-level HEFT and YM theory, and showed that it can be mapped to a quasishuffle Hopf algebra. It is intriguing to note that Hopf algebras have already appeared in several different contexts in quantum field theory and string theory, e.g., renormalization theory [52], symbols and co-actions of loop integrals [53–56], harmonic sums [57], and string α' expansion [58,59]. The obtained kinematic algebra is very simple in terms of the abstract generators, and a nontrivial aspect of the construction is the map between these generators and the kinematic variables (momenta and polarizations), for which we find a simple closed formula that exhibits manifest gauge invariance (see, e.g., Refs. [37,51,60–62] for other all-multiplicity BCJ constructions).

Several questions remain open. First, it would be interesting to derive our BCJ numerators from a Lagrangian description, which may expose hidden symmetries or structures of the theory. The nonlocalities of the numerators are both mild and physical in HEFT, thus a Lagrangian approach seems feasible. It may also prove fruitful to try to find representations of the generators in the form of differential operators in kinematic variables, thus reintroducing kinematics in the fusion rules. On the mathematical side one may note that a Hopf algebra should have a coproduct and counit: what do these operations imply for the numerator and amplitude? The extension of our construction to loop amplitudes, as well as other theories, is an important avenue. Finally, it would be interesting to find a more direct construction of the pure YM kinematic algebra, without passing through the HEFT. We leave these questions for future investigation.

We thank Kays Haddad, Sanjaye Ramgoolam, Bill Spence, and Mao Zeng for useful discussions, and Maor Ben-Shahar, Lucia Garozzo, Fei Teng, and Tianheng Wang for collaborations on related topics. This work was supported by the Science and Technology Facilities Council (STFC) Consolidated Grants No. ST/P000754/1 *"String Theory, Gauge Theory and Duality"* and No. ST/T000686/1 *"Amplitudes, Strings and Duality"* and by the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie Grant agreement No. 764850 *"SAGEX"*. H. J. is supported by the Knut and Alice Wallenberg Foundation under Grants No. KAW 2018.0116 (*From Scattering Amplitudes to Gravitational Waves*) and No. KAW 2018.0162, the Swedish Research Council under Grant No. 621-2014-5722, and the Ragnar Söderberg Foundation (Swedish Foundations' Starting Grant). C. W. is supported by a Royal Society University Research Fellowship No. UF160350.

*a.brandhuber@qmul.ac.uk
†g.chen@qmul.ac.uk
‡henrik.johansson@physics.uu.se
§g.travaglini@qmul.ac.uk
‖c.wen@qmul.ac.uk

[1] Z. Bern, J. J. M. Carrasco, and H. Johansson, Phys. Rev. D **78**, 085011 (2008).
[2] Z. Bern, J. J. M. Carrasco, and H. Johansson, Phys. Rev. Lett. **105**, 061602 (2010).
[3] H. Kawai, D. C. Lewellen, and S. H. H. Tye, Nucl. Phys. **B269**, 1 (1986).
[4] H. Georgi, Phys. Lett. B **240**, 447 (1990).
[5] M. E. Luke and A. V. Manohar, Phys. Lett. B **286**, 348 (1992).
[6] M. Neubert, Phys. Rep. **245**, 259 (1994).
[7] A. V. Manohar and M. B. Wise, *Heavy Quark Physics* (Cambridge University Press, Cambridge, England, 2000), Vol. 10.
[8] P. H. Damgaard, K. Haddad, and A. Helset, J. High Energy Phys. 11 (2019) 070.
[9] Note that in the phenomenological literature, HEFT denotes effective theories involving the BEH boson.
[10] A. Brandhuber, G. Chen, G. Travaglini, and C. Wen, J. High Energy Phys. 07 (2021) 047.
[11] R. Monteiro and D. O'Connell, J. High Energy Phys. 07 (2011) 007.
[12] G. Chen and Y.-J. Du, J. High Energy Phys. 01 (2014) 061.
[13] C. Cheung and C.-H. Shen, Phys. Rev. Lett. **118**, 121601 (2017).
[14] C. Cheung, G. N. Remmen, C.-H. Shen, and C. Wen, J. High Energy Phys. 04 (2018) 129.
[15] C. Cheung, C.-H. Shen, and C. Wen, J. High Energy Phys. 02 (2018) 095.
[16] G. Chen, H. Johansson, F. Teng, and T. Wang, J. High Energy Phys. 11 (2019) 055.
[17] G. Chen, H. Johansson, F. Teng, and T. Wang, J. High Energy Phys. 10 (2021) 042.
[18] By N^kMHV amplitude in Yang-Mills we mean one whose numerator has the schematic form $(\epsilon\cdot\epsilon)^{k+1}\prod(\epsilon\cdot p)$
[19] A. Brandhuber, G. Chen, G. Travaglini, and C. Wen, J. High Energy Phys. 10 (2021) 118.
[20] Z. Bern, C. Cheung, R. Roiban, C.-H. Shen, M. P. Solon, and M. Zeng, Phys. Rev. Lett. **122**, 201603 (2019).
[21] T. Damour, Phys. Rev. D **102**, 024060 (2020).
[22] G. Kälin, Z. Liu, and R. A. Porto, Phys. Rev. Lett. **125**, 261103 (2020).
[23] P. Di Vecchia, C. Heissenberg, R. Russo, and G. Veneziano, J. High Energy Phys. 07 (2021) 169.
[24] E. Herrmann, J. Parra-Martinez, M. S. Ruf, and M. Zeng, J. High Energy Phys. 10 (2021) 148.
[25] N. E. J. Bjerrum-Bohr, P. H. Damgaard, L. Planté, and P. Vanhove, J. High Energy Phys. 08 (2021) 172.
[26] Z. Bern, J. Parra-Martinez, R. Roiban, M. S. Ruf, C.-H. Shen, M. P. Solon, and M. Zeng, Phys. Rev. Lett. **126**, 171601 (2021).
[27] M. E. Hoffman, Journal of Algebraic Combinatorics **11**, 49 (2000).
[28] M. Aguiar and S. A. Mahajan, *Monoidal Functors, Species and Hopf Algebras* (American Mathematical Society, Providence, RI, 2010), Vol. 29.
[29] F. Fauvet, L. Foissy, and D. Manchon, Ann. l'Inst. Fourier **67**, 911 (2017).
[30] Note that $\sum_{i=1}^{n-2} v\cdot p_i = 0$.
[31] Note that on the massive pole $\langle T_{(12)}\rangle$ factorizes as $\langle T_{(12)}\rangle|_{v\cdot p_1\to 0} = -[(p_1\cdot p_2)/(2v\cdot p_1)]\langle T_{(1)}\rangle\langle T_{(2)}\rangle$.
[32] P. Cartier, Asterisque-Soc. Math. France **282**, 137 (2002).
[33] K. Ihara, M. Kaneko, and D. Zagier, Compos. Math. **142**, 307 (2006).
[34] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.128.121601 for a discussion of the factorization of BCJ numerators and details on quasishuffle Hopf algebras.
[35] I. Mező, *Combinatorics and Number Theory of Counting Sequences* (Chapman and Hall/CRC, London, 2019).
[36] O. A. Gross, Am. Math. Mon. **69**, 4 (1962).
[37] N. E. J. Bjerrum-Bohr, P. H. Damgaard, T. Sondergaard, and P. Vanhove, J. High Energy Phys. 01 (2011) 001.
[38] N. Arkani-Hamed, L. Rodina, and J. Trnka, Phys. Rev. Lett. **120**, 231602 (2018).
[39] L. Rodina, J. High Energy Phys. 09 (2019) 084.
[40] V. Del Duca, L. J. Dixon, and F. Maltoni, Nucl. Phys. **B571**, 51 (2000).
[41] R. Kleiss and H. Kuijf, Nucl. Phys. **B312**, 616 (1989).
[42] D. Vaman and Y.-P. Yao, J. High Energy Phys. 11 (2010) 028.
[43] N. E. J. Bjerrum-Bohr, P. H. Damgaard, B. Feng, and T. Sondergaard, Phys. Rev. D **82**, 107702 (2010).
[44] Z. Bern, L. J. Dixon, M. Perelstein, and J. S. Rozowsky, Nucl. Phys. **B546**, 423 (1999).
[45] J. J. M. Carrasco, C. R. Mafra, and O. Schlotterer, J. High Energy Phys. 06 (2017) 093.
[46] H. Frost, C. R. Mafra, and L. Mason, arXiv:2012.00519.
[47] Z. Bern, T. Dennen, Y.-t. Huang, and M. Kiermaier, Phys. Rev. D **82**, 065003 (2010).
[48] E. Bridges and C. R. Mafra, J. High Energy Phys. 10 (2019) 022.
[49] S. Mizera, Phys. Rev. Lett. **124**, 141601 (2020).
[50] S. He, L. Hou, J. Tian, and Y. Zhang, J. High Energy Phys. 08 (2021) 118.
[51] C. Cheung and J. Mangan, J. High Energy Phys. 11 (2021) 069.
[52] A. Connes and D. Kreimer, Commun. Math. Phys. **210**, 249 (2000).
[53] A. B. Goncharov, M. Spradlin, C. Vergu, and A. Volovich, Phys. Rev. Lett. **105**, 151605 (2010).
[54] C. Duhr, J. High Energy Phys. 08 (2012) 043.
[55] S. Abreu, R. Britto, C. Duhr, and E. Gardi, J. High Energy Phys. 10 (2014) 125.
[56] F. Brown, Commun. Num. Theor. Phys. **11**, 453 (2017).
[57] J. Blumlein, Comput. Phys. Commun. **159**, 19 (2004).
[58] J. Broedel, O. Schlotterer, and S. Stieberger, Fortschr. Phys. **61**, 812 (2013).
[59] C.-H. Fu and Y. Wang, J. High Energy Phys. 12 (2020) 106.
[60] C. R. Mafra, O. Schlotterer, and S. Stieberger, J. High Energy Phys. 07 (2011) 092.
[61] Y.-J. Du and F. Teng, J. High Energy Phys. 04 (2017) 033.
[62] A. Edison and F. Teng, J. High Energy Phys. 12 (2020) 138.